

АКАДЕМІЯ НАУК РЕСПУБЛІКИ МОЛДОВА
ІНСТИТУТ ГЕОЛОГІЇ І СЕЙСМОЛОГІЇ
НАЦІОНАЛЬНА АКАДЕМІЯ НАУК УКРАЇНИ
ІНСТИТУТ ГЕОФІЗИКИ ІМ. С. І. СУББОТІНА

Кваліфікаційна наукова праця на
правах рукопису

БУРТІЄВ РАШІД ЗЕТОВИЧ

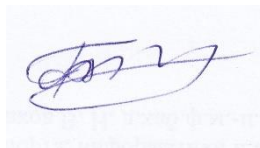
УДК 550.830

ДИСЕРТАЦІЯ
МЕТОДОЛОГІЯ ОЦІНКИ СЕЙСМІЧНОЇ НЕБЕЗПЕКИ
НА ОСНОВІ ЙМОВІРНІСНИХ МОДЕЛЕЙ СЕЙСМІЧНОСТІ

04.00.22 – геофізика

Подається на здобуття наукового ступеня доктора фізико-математичних наук

Дисертація містить результати власних досліджень. Використання ідей, результатів і текстів інших авторів мають посилання на відповідне джерело



Р. З. Буртієв

Науковий консультант: АЛКАЗ Василь Георгійович, доктор фізико-математичних наук,
заслужений діяч науки Республіки Молдова

Київ-2017

АНОТАЦІЯ

Буртієв Р. З. Методологія оцінки сейсмічної небезпеки на основі ймовірнісних моделей сейсмічності. – Кваліфікаційна наукова праця на правах рукопису.

Дисертація на здобуття наукового ступеня доктора фізико-математичних наук за спеціальністю 04.00.22 «геофізика» (103 – Науки про Землю). – Інститут геології і сейсмології Академії Наук Республіки Молдова, Кишинів, 2016; Інститут геофізики ім. С. І. Субботіна НАН України, Київ, 2017.

Дисертація присвячена розвитку методів оцінки сейсмічної небезпеки на основі ймовірнісних моделей сейсмічності. Розглянуто ряд питань виділення сейсмічних зон, розробки ймовірнісної моделі сейсмічного режиму сейсмічних зон, визначення місця виникнення і магнітуди майбутніх землетрусів, дослідження макросейсмічного поля землетрусів, оцінки сейсмічної небезпеки, розробки алгоритму та обчислювальної програми оцінки і картування сейсмічної небезпеки. При здійсненні досліджень застосований комплекс методів статистичного аналізу і побудови ймовірнісних моделей сейсмічності.

Виділення сейсмічних зон проводилося за допомогою кластерного аналізу. Щодо числа сейсмічних зон на території Румунії не існує єдиної думки, за деякими оцінками на даній території знаходяться 5 сейсмічних, а за іншими оцінками – 13. Домінуючим при оптимізації кластерного рішення (визначення числа кластерів у кластерних групах) виявилось 13 кластерів. Однак, кластерний аналіз, будучи формальним методом, не може враховувати геолого-геофізичні характеристики зони джерел землетрусів, але є цілком придатним для оперативного сейсмічного зонування. В якості пробної вибірки використовувався каталог землетрусів Південного Сходу Європи, що охоплює період 1900–1990 рр. У цьому каталозі містяться відомості про землетруси в 18 сейсмічних зонах. При порівнянні координат епіцентрів в сейсмічних зонах, виділених алгоритмом k -середніх, з координатами епіцентрів наведених в каталозі спостерігався 97% -ний збіг.

Дискримінантний аналіз застосовувався для поповнення каталогу землетрусів, що не ввійшли в нього відомостями про подальші землетруси.

Для виявлення статистичного зв'язку між різними типами магнітудних шкал використовувалися регресивні методи. У каталогах ISC, PDE, NEIC, USGS, ROMPLUS та ін., наведені значення різних типів магнітуд: *mb*, *MW*, *MS*, *ML*, *MD*, *MLH*, *MJ*, *ME*, *MPV*, *MLH* тощо. Для застосування закону загасання необхідно привести значення магнітуд до єдиного типу. У даній роботі перераховані типи магнітуд були перетворені в *mb* магнітуду.

Методи оцінки мінімальної магнітуди комплектності M_c ґрунтуються на припущенні про самоподібність сейсмічного процесу. В даному випадку, із багатьох методів обчислення M_c використовуються: метод максимальної кривизни; метод, заснований на стійкості M_c при оцінці коефіцієнта b - нахилу графіка повторюваності. Значення мінімальної магнітуди комплектності для каталогу ROMPLUS охоплює період 1978-2013рр., оцінене по обох методах, збігається і становить $M_c = 2.3$.

Між магнітудою і глибиною землетрусів виявлено слабкий статистичний зв'язок. Для землетрусів в межах земної кори зв'язок значно слабший, ніж для проміжних землетрусів. Поясненням цього факту є те, що в межах земної кори відбуваються слабкі землетруси, з магнітудою до $MW = 5.5$, а в проміжному шарі відбуваються землетруси з магнітудою до $MW = 7.7$ і область їх підготовки охоплює більший обсяг геофізичного середовища.

Виявлено слабкий зв'язок між річним числом землетрусів в межах земної кори і нижче границі Мохо. Статистична залежність між короною і підкороною сейсмічною активністю для землетрусів з магнітудою $M \geq 2.6$ практично відсутня. Між щомісячним числом корових і проміжних землетрусів також існує слабкий статистичний зв'язок. В рядах річного та щомісячного числа землетрусів в межах кори існує тренд, в проміжній сейсмічності припущення про відсутність тренду підтвердилося.

Рівняння макросейсмічної поля пов'язує з кожним землетрусом: координати епіцентру; глибина; магнітуда, інтенсивність сейсмічних впливів у пунктах спостережень. Отже, для передбачення інтенсивності сейсмічних впливів досить передбачити місце виникнення і магнітуду майбутніх землетрусів.

Сейсмічна зона Z^s , яка охоплює деякий простір геофізичного середовища $[\Phi \times \Psi \times H]$: Φ ; Ψ ; H – діапазони можливих значень широти, довготи і глибини землетрусів відповідно. Декартово множення даного простору на діапазон можливих значень магнітуди утворює чотиривимірний простір точки якого можуть виявитися можливими значеннями параметрів землетрусів. Тобто землетрус, що стався в момент часу t_n , на осі часу зображається точкою у чотиривимірному просторі, і значення параметрів землетрусів визначиться завданням чотиривимірного випадкового процесу $\Omega(t)$. Для розрахунку СН досить знати де відбудеться землетрус, і яка буде його магнітуда – визначити розподіл інтервальних значень параметрів, тобто обчислити вірогідність, того, що значення параметрів можливих землетрусів потраплять до деяких інтервали Φ_i, Ψ_j, H_k, M_m , які утворюються розбивкою діапазонів можливих значень параметрів землетрусів Φ, Ψ, H, M на R_ϕ, R_ψ, R_h, R_m непересічні інтервали, відповідно. Ймовірнісна модель послідовності землетрусів розробляється для визначення ймовірності того, що координати епіцентру, глибина і магнітуда землетрусів приймуть значення з інтервалів Φ_i, Ψ_j, H_k, M_m відповідно. Вирішення цього завдання передбачається здійснити на основі добре розробленої, витонченої теорії марковських процесів.

Неперервний Марківський ланцюг є адекватною моделлю потоку землетрусів. Дана модель підходить для опису траєкторії функціонування систем, які мають тенденцію залишатися в тому стані, в якому вже перебувають. Чотиривимірна модель Маркова є найбільш повною моделлю сейсмічного режиму і дозволяє передбачити місце виникнення і частотну структуру землетрусів.

Марковська модель сейсмічних зон може допомогти виявити міграцію осередків землетрусів уздовж сейсмічних поясів і наявності зв'язку між сейсмічними регіонами. Виявлено статистично значущий зв'язок між активністю деяких

сейсмічних регіонів Балкан. Динаміка сейсмічної активності балканського регіону - чергування вогнищ землетрусів в регіонах – відстежується елементами матриці ймовірностей переходу. Виявлено, що землетруси мають тенденцію відбуватися там, де воно вже відбулися. Наприклад, по матриці ймовірностей переходу видно, що після землетрусу в Румунії, наступний землетрус, швидше за все, трапиться теж в Румунії, тому що ймовірність переходу $\pi_{11}=0.224$ зі стану *ROM* в *ROM* найбільше. Другим за величиною є ймовірність $\pi_{13}=0.112$ переходу зі стану *ROM* в *BUL*, тобто, наступний після Румунії землетрус може статися, але з меншою ймовірністю в Болгарії. Ймовірність переходу $\pi_{14}=0.0$ зі стану *ROM* в *POL* дорівнює нулю. Це означає, що сейсмічні зони Румунія і Польща не пов'язані. З вектора граничних розподілів сейсмічних зон можна визначити зону, де і з якою ймовірністю може відбутися очікуваний землетрус. Критерії таблиці спряженості ознак також вказують на наявність середньої тісноти взаємозв'язку сейсмічної активності деяких зон. Марковська модель сейсмічних зон може допомогти виявити міграцію осередків землетрусів уздовж сейсмічних поясів.

У багатьох тимчасових послідовностях випадкових подій існує залежність від минулого. Модель, яка охоплює широке коло форми залежності між подіями, є реальним відображенням дійсності, однак це призводить до великих труднощів при вивченні їх ймовірнісної структури і статистичному аналізі. Тому приймається компромісне рішення – вибирається модель, яка враховує достатні для адекватності моделі залежності і піддається статистичному аналізу. Найкраще відповідність зазначеним вимогам проявляє клас марковських процесів, в яких моделюється залежність від результату спостереження в останній момент часу. Така модель є прийнятним компромісом між залежністю від всієї передісторії процесу і повною незалежністю випадкових подій в послідовності.

У ряді щомісячного числа землетрусів присутня довгострокова систематична складова. Припущення про стаціонарність часового ряду місячного числа землетрусів за середнім значенням і по дисперсії відкидається. Найкращими, за вибірковими значеннями статистичних критеріїв перевірки якості моделі, виявилися

проста сезонна, для прологарифмованого ряду, і адитивна модель Уінтерса. Найкращою моделлю для прогнозу часових рядів виявилася $ARIMA(0,1,2) \times (1,0,1)_s$ модель, підібрана конструктором моделей пакета SPSS.

У даній роботі проведено статистичний аналіз макросейсмічних даних землетрусів, що сталися 10.11.1940, 7.04.1977, 31.08.1986, 30.05.1990 та 31.05.1990 рр. Дослідження показують, що прийнятною згладжуючою лінією макросейсмічного поля є еліпс. За методом Фітцгіббона проведена апроксимація еліпсом багатьох точок однакової інтенсивності. Визначено параметри еліпсів, що апроксимують зони однакової бальності макросейсмічного поля. Коефіцієнт лінійної регресії між епіцентральною інтенсивністю I_0 та магнітудою для Румунських землетрусів, вийшов рівним $b=1.5$. Оптимальними значеннями коефіцієнтів рівняння макросейсмічного поля виявилися: $a=1.6$; $c=7.2$; найбільшого $b_{\max}=5.6$ і найменшого $b_{\min}=4.9$ значення осей еліпса загасання інтенсивності струсів; а кут між великою віссю еліпса загасання і позитивним напрямком осі абсцис рівним $\gamma_0=51^\circ$.

На основі отриманих значень побудована карта сейсмічної небезпеки території Молдови, Румунії та Болгарії. Для точок в ближній зоні обчислювалась відстань Джойнера-Бура, а в інших точках гіпоцентрально відстань. Карта сейсмічної небезпеки відображає інтенсивність сейсмічного ефекту в балах шкали *MSK-64* і складається з точок, в яких відповідна ймовірність позитивна, а карта, побудована за нормами "EUROCODE 8" складається з точок, в яких ймовірність дорівнює 0.1.

Запропонований метод дозволяє розрахувати сейсмічну небезпеку в населених пунктах, розташованих в радіусі впливу групи сейсмічних зон та побудувати карту сейсмічного районування великих територій, наприклад усієї Європи.

Основні положення дисертації викладені в 28 статтях у фахових журналах, в одній монографії та 4 тезах доповідей конференцій. Дисертація написана українською мовою, містить 310 сторінок, рисунки, таблиці і посилання з 342 найменування.

Ключові слова: імовірнісний аналіз сейсмічної небезпеки, макросейсміка, загасання інтенсивності, марківська модель, кластерний аналіз.

СПИСОК ПУБЛІКАЦІЙ

Монографія

1. Буртиев Рашид, Методология оценки сейсмической опасности, LAP LAMBERT Academic Publishing, 2017, 281.

Статті

2. Буртиев Р.З., Онофраш Н.И. О статистическом подходе к определению параметров сейсмичности. Вопросы инженерной сейсмологии. Задание сейсмических воздействий, Москва, 1993, 68–71.
3. Буртиев Р.З. Оценка максимальной магнитуды землетрясений очага Вранча по статистике экстремальных значений, Analele ATIC, Chisinau 2003, Vol. I(II), 110–117.
4. Буртиев Р.З. Оценка параметров марковских моделей сейсмического режима. Analele ATIC, Chisinau, 2003, I (IV), 117–128.
5. Буртиев Рашид, О применении статистики Гумбеля к оценке максимальной магнитуды, Buletinul Institutului de Geologie si Seismologie al ASM. 2007, Nr.2, 5–7.
6. Буртиев Р.З., Чобану Я.И., О применении оптимизационных методов в оценке параметров III распределения Гумбеля, Analele ATIC, Chisinau, 2007, Vol. I(II), 415–417.
7. Буртиев Рашид, Статистическая связь между сейсмическими зонами Балкан, Buletinul Institutului de Geologie și Seismologie al ASM. 2008, Nr.1. 31–34.
8. Буртиев Рашид, О статистическом подходе к картированию сейсмической опасности больших территорий. Buletinul Institutului de Geologie și Seismologie al ASM. 2008, Nr.2, 57–61.

9. Буртиев Рашид, Применение линейной регрессии для преобразования разных магнитуд в шкалу магнитуды mb. *Buletinul Institutului de Geologie și Seismologie al ASM*, 2009, Nr.1, 62–71.
10. Буртиев Рашид, О возможности применения многомерной линейной регрессии с коллинеарными факторами для преобразования магнитуд. *Buletinul Institutului de Geologie și Seismologie al ASM*, 2009, Nr.2, 35–44.
11. Буртиев Рашид, О применении регрессии в условиях гетероскедастичности сейсмических данных. *Analele ATIC. Chisinau*, 2010, I (IV), 177–180.
12. Буртиев Рашид, О статистической связи сейсмической активности Балканских сейсмических зон, *Buletinul Institutului de Geologie si Seismologie al ASM*, 2010, Nr.1, 5–15.
13. Буртиев Рашид, Применение кластерного анализа для сейсмического районирования Балканского региона. *Buletinul Institutului de Geologie si Seismologie al ASM*, 2010, Nr.2, 55–62.
14. Буртиев Рашид, Статистическая связь между коровой и подкоровой сейсмическими активностями очага Вранча. *Buletinul Institutului de Geologie si Seismologie al ASM*, 2011, Nr. 2, 133–148.
15. Telesca Luciano, Burtiev Rașid, Alcaz Vasile, Sandu Ilie, Time-clustering analysis of the 978-2008 sub-crustal seismicity of Vrancea region. *Journal Natural Hazards and Earth System Sciences*, 2011, 2335–2340.
16. Буртиев Р.З. Применение кластерного анализа для выделения сейсмических зон Румынии. *Analele ATIC, I (IV), Modelare matematică, Optimare și tehnologii informaționale*, Chisinau, 2012, 252–262.
17. Burtiev Rashid, Evaluation of seismic hazards from several seismic zones. *Environmental Engineering and Management Journal*, 2012, N12, 2141–2150.
18. Burtiev Rasid, Greenwell Fernanda, Kolivenko Vitalii, Statistical analysis time series of speed and temperature in Tiraspol, Moldova, *Environmental Engineering and Management Journal*, 2013, N1, 23-33.

19. Буртиев Рашид, Вероятностный метод оценки сейсмической опасности, *Buletinul INCERCOM*, 2013, Nr. 4, 41–47.
20. Буртиев Рашид, Павлов Петр, Исследование азимутальной неоднородности геологической среды по макросейсмическим данным. *Buletinul Institutului de Geologie si Seismologie al ASM*, 2013, Nr.2, 38–46.
21. Буртиев Рашид, Определение расстояния от точки наблюдения до элементарного источника землетрясений. *Buletinul Institutului de Geologie si Seismologie al ASM*, 2014, Nr.1, 23–29.
22. Burtiev Rashid, Fitting an Ellipse to the set of intensity data points of Vrancea earthquakes. *Journal of Engineering Studies and Research*, Baceu, Romania, 2014, Volume 20, No, 4, 15–24.
23. Burtiev Rashid, Time Series in the Study of Seismic Regime of Vrancea (Romania) Seismic Zone, *The Global Environmental Engineers*, Karachi, Pakistan, 2014, Vol.1, N2, 54–63.
24. Буртиев Рашид, Вероятностный анализ экстремальной сейсмической опасности в Молдове и Румынии. *Buletinul INCERCOM*, 2015, Nr. 4, 41–47.
25. Буртиев Рашид, Вероятностный анализ сейсмической опасности территории Молдовы и сопредельных районов, *Геофизический журнал*, Киев, 2015, № 6, Т. 37, 2, 124–130.
26. Буртиев Р.З., Павлов П.П. Вероятностная модель сейсмичности. *Buletinul Institutului de Geologie si Seismologie al ASM*, 2015, Nr.2, 97–104.
27. Burtiev Rashid, Alcaz Vasile, Cardanets Vladlen, Probabilistic Seismic Hazard Analysis on the Base of the Stochastic Models of Seismicity, *Indian Journal of Applied Research*, V.6, Issue 8, 2016, 454–466.
28. Буртиев Рашид, Методы расчета сейсмического риска и составления синтетического каталога землетрясений с использованием Монте-Карло моделирования, *Buletinul Institutului de Geologie si Seismologie al ASM*, 2016, Nr.2. 87–96.

29. Burtiev Rashid, Seismic Hazard Assessment Method Based on the Stochastic Models of Seismicity, Bulletin of the International Institute of Seismology and Earthquake Engineering, Tsukuba, Japan, 2017, Vol.5, 22–38.

Тези конференцій

30. Burtiev Rashid, Cluster Analysis to Select the Seismic Zones of Romania. First International Conference on Moldavian Risks—from Global to local scale. Book of Abstracts, 2012, Bacau, Romania. p. 25.
31. Burtiev Rashid, Probabilistic Seismic Hazard Analysis for the Territory of Moldova and Romania. Second International Conference on Moldavian Risks—from Global to local scale, Book of Abstracts, 2014, Bacau, Romania, p. 24.
32. Burtiev Rashid, A New Approach to Probabilistic Seismic Hazard Analysis, Georisks in the Mediterranean and Their Mitigation. University of Malta, 2015, p. 154.
33. Burtiev Rashid, Evaluation of Seismic Hazard on Base of Probabilistic Models of Vrancea Zones, The National Symposium “75 Years from November 10th 1940 Vrancea Earthquake”, Book of Abstracts, Bucarest, 2015, p. 14.

ABSTRACT

Burtiev R. Z. Methodology for seismic hazard assessment based on probabilistic models of seismicity. — Manuscript.

A thesis for the degree of Doctor of Physical and Mathematical Sciences, specialty 04.00.22 — geophysics. — Institute of Geology and Seismology of Academy of Science of Moldova; S. I. Subbotin Institute of Geophysics, National Academy of Sciences of Ukraine, Kyiv, 2017.

The thesis focuses on the development of the new methodology for seismic hazard assessment. The universal measure of seismic hazard is developed. It means that in a location of the Earth's surface during the time of t , ground shaking will occur n times and n_k of them with intensity of I^k . Such approach may be useful for earthquake resistant construction strategy, because every strong earthquake can affect the buildings' resistance.

To identify the area of earthquakes in the seismogenic zone cluster analysis is used. The Markov model of sequence all seismic events regarding for all seismogenic zones is elaborated, which allows to forecast the interval values of latitudes, longitudes, depth of the hypocenters and magnitudes of earthquakes. The coefficients of intensity attenuation equation for Moldova, elaborated on the base of macroseismic data set of the intermediate earthquakes occurred in Vrancea is defined. Macroseismic field of Vrancea intermediate earthquakes extended from the southwest on the northeast. A configuration, having the shape of an ellipse which major axis is turned on angle $\gamma_0=51^0$ concerning the positive direction of abscissa axis. The coefficients of the attenuation relationships between seismic intensity I and earthquake's characteristics: Magnitudes M_W , epicentral distance ρ , depth h in the Vrancea zone are represented by empirical azimuth-dependent attenuation equation. Their value was determined as follows: $a=1.6$; $c=7.2$, the biggest axis of the ellipse being $b_{max}=5.6$ and the smallest $b_{min}=4.9$. Probabilistic seismic hazard analysis has been performed and the map of seismic hazard of Moldova and adjacent territory is elaborated.

Key words: PSHA, macroseismic dates, attenuation law, Markov model, cluster analysis.

List of author's publications.

Monograph

1. Burtiev Rashid, Seismic Hazard Assessing Methodology, LAP LAMBERT Academic Publishing, 2017, 281.

Articles

2. Burtiev R.Z., Onofrash N.I., on the statistical approach to the definition of seismicity parameters. Engineering seismology questions, Moscow, 1993, 31-35.
3. Burtiev R.Z., Estimation of the maximum magnitude of earthquakes in the Vrancea seismogenic zones using extreme values statistics, Annals of ATIC, Chisinau, 2002, Vol. I(II), 110–117.
4. Burtiev R.Z., The assessment of the parameters of markov model of seismic regime, Annals of ATIC, Chisinau, 2003, I (IV), 117-128.

5. Burtiev Rashid, Application of Gumbel's statistics to the estimation of the maximum magnitude, Journal of the Institute of Geology and Seismology of the Academy of Sciences of Moldova, 2007, Nr. 2, 5–7.
6. Burtiev R.Z., Ciobanu I. I., Optimization methods in the estimation of the parameters of the iii distribution of Gumbel statistics, Annals of ATIC, Chisinau, 2007, Vol. I(II), 415–417.
7. Burtiev Rashid, Statistical relationship between seismic zones of the Balkans, Journal of the Institute of Geology and Seismology of the Academy of Sciences of Moldova, 2008, Nr.1. 31–34.
8. Burtiev Rashid, on the statistical approach to mapping the seismic hazard of large Areas, Journal of the Institute of Geology and Seismology of the Academy of Sciences of Moldova, 2008, Nr.2, 57–61.
9. Burtiev Rashid, The application of linear regression for the conversion of different magnitudes into a scale of magnitude mb, Journal of the Institute of Geology and Seismology of the Academy of Sciences of Moldova, 2009, Nr.1, 62–71.
10. Burtiev Rashid, On the possibility of applying multidimensional linear regression with collinear factors for the transformation of magnitudes. Journal of the Institute of Geology and Seismology of the Academy of Sciences of Moldova, 2009, Nr.2, 35–44.
11. Burtiev Rashid, On the application of regression under heteroscedasticity conditions of seismic data, Analele ATIC, Chisinau, 2010, I (IV), 177–180.
12. Burtiev Rashid, On the statistical relationship of seismic activity in the Balcan seismic zones, Journal of the Institute of Geology and Seismology of the Academy of Sciences of Moldova, 2010, Nr.1, 5–15
13. Burtiev Rashid, The Application of cluster analysis for seismic zoning of the balkan region, Journal of the Institute of Geology and Seismology of the Academy of Sciences of Moldova, 2010, Nr.2, 55–62.
14. Burtiev Rashid, Statistical relationship between the crustal and the subcrustal seismic activity of the vrancea focus, Journal of the Institute of Geology and Seismology of

- the Academy of Sciences of Moldova, 2011, Nr. 2, 133–148
15. Telesca Luciano, Burtiev Raşid, Alcaz Vasile, Sandu Ilie, Time-clustering analysis of the 978–2008 sub-crustal seismicity of vrancea region. *Journal Natural Hazards and Earth System Sciences*, 2011, 2335–2340.
 16. Burtiev R.Z., Application of cluster analysis for seismic zones identification in Romania, *Annals of ATIC*, Chisinau, 2012, I (IV), 252–261.
 17. Burtiev Rashid, Evaluation of seismic hazards from several seismic zones, *Environmental Engineering and Management Journal*, N12, 2012, 2141–2150.
 18. Burtiev Rasid, Greenwell Fernanda, Kolivenko Vitalii, Statistical analysis time series of speed and temperature in Tiraspol, Moldova, *Environmental Engineering and Management Journal*, 2013, N1, 23–33.
 19. Burtiev Rashid, Probabilistic method of seismic hazard assessment, *Bulletin of the INCERCOM*, Chisinau, 2013, N4, 41–47.
 20. Burtiev Rashid и Павлов Петр, Study of azimuthal inhomogeneity of geological environment on macroseismic data, *Journal of the Institute of Geology and Seismology of the Academy of Sciences of Moldova*, 2013, Nr.2, 38–46.
 21. Burtiev Rashid, Determination of the distance from the site to the elementary source of earthquakes, *Journal of the Institute of Geology and Seismology of the Academy of Sciences of Moldova*, 2014, Nr.1, 23–29.
 22. Burtiev Rashid, Fitting an ellipse to the set of intensity data points of vrancea earthquakes. *Journal of Engineering Studies and Research*, Baceu, Romania, 2014, Vol. 20, 2014, N4, 15–24.
 23. Burtiev Rashid, Time series in the study of seismic regime of vrancea (romania) seismic zone, *The Global Environmental Engineers*, 2014, Vol.1, N2, Karachi, Pakistan, 54–63.
 24. Burtiev Rashid, Probabilistic analysis of extreme seismic hazard in Moldova and Romania. *Bulletin of the INCERCOM*, Chisinau, 2015, N6, 162–177.
 25. Burtiev Rashid, Probabilistic seismic hazard analysis of Moldova and adjacent territories, *Geophysical Journal № 6*, T. 37, 2015, Киев, 124–130.

26. Burtiev R.Z., Pavlov P.P., Probabilistic model of seismicity, Journal of the Institute of Geology and Seismology of the Academy of Sciences of Moldova, Nr.2, 2015, 97–104.
27. Burtiev Rashid, Alcaz Vasile, Cardanets Vladlen, Probabilistic seismic hazard analysis on the base of the stochastic models of seismicity, Indian Journal of Applied Research, 2016, Vol.6, Issue 8, 2016, 454–466.
28. Burtiev Rashid, Methods for calculating of seismic risk and compiling synthetic catalog of earthquakes using of the Monte-Carlo modeling, Journal of the Institute of Geology and Seismology of the Academy of Sciences of Moldova, 2016, Nr.2. 87–96.
29. Burtiev Rashid, Seismic Hazard Assessment Method Based on the Stochastic Models of Seismicity, Bulletin of the International Institute of Seismology and Earthquake Engineering, Tsukuba, Japan, Vol.51, 2017, 22–38.

Abstracts of conferences

30. Burtiev Rashid, Cluster Analysis to Select the Seismic Zones of Romania. First International Conference on Moldavian Risks—from Global to local scale. Book of Abstracts, 2012, Bacau, Romania. p. 25.
31. Burtiev Rashid, Probabilistic Seismic Hazard Analysis for the Territory of Moldova and Romania. Second International Conference on Moldavian Risks—from Global to local scale, Book of Abstracts, 2014, Bacau, Romania, p. 24.
32. Burtiev Rashid, A New Approach to Probabilistic Seismic Hazard Analysis, Georisks in the Mediterranean and Their Mitigation. University of Malta, 2015, p. 154.
33. Burtiev Rashid, Evaluation of Seismic Hazard on Base of Probabilistic Models of Vrancea Zones, The National Symposium “75 Years from November 10th 1940 Vrancea Earthquake”, Book of Abstracts, Bucarest, 2015, p. 14.

ЗМІСТ		
ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ		21
ВСТУП		23
РОЗДІЛ 1. ЙМОВІРНІСНИЙ АНАЛІЗ СЕЙСМІЧНОЇ НЕБЕЗПЕКИ (СТАН ПРОБЛЕМИ)		32
1.1.	Огляд методів аналізу сейсмічної небезпеки	33
1.2.	Статистичні методи	41
1.3.	Монте-Карло моделювання	42
1.4.	Системний аналіз	45
1.5.	Сейсмічність республіки Молдова	47
Висновки до розділу 1		52
Список використаних джерел		53
РОЗДІЛ 2. ВИЯВЛЕННЯ СЕЙСМІЧНИХ ЗОН		60
2.1.	Кластерний аналіз	62
2.2.	Ієрархічний агломеративного метод кластерного аналізу	64
2.3.	Метод Варда (Ward-method)	68
2.4.	Алгоритм k -середніх	70
2.5.	Оптимальне кластерне рішення	72
	2.5.1. Мінімум сумарною дисперсії	73
	2.5.2. Метод Галінського-Харабаша	74
	2.5.3. Критерій мінімуму внутрішньо-кластерної дисперсії	75
	2.5.4. Статистика силуетів	76
	2.5.5. Метод Кржановського-Лая	77
	2.5.6. Непараметричний метод Сюгера-Джеймса	78
2.6.	Двокроковий кластерний аналіз	79
2.7.	Критерії однорідності кластерів	82

	2.7.1.	<i>F</i> -критерій	82
	2.7.2.	<i>t</i> -критерій	83
2.8.	Стійкість кластеризації		84
2.9	Дослідження унімодальності		87
Висновки до розділу 2			91
Список використаних джерел			92
РОЗДІЛ 3. СТАТИСТИЧНИЙ АНАЛІЗ КАТАЛОГІВ ЗЕМЛЕТРУСІВ			97
3.1.	Опис даних		98
3.2.	Опис використовуваних методів математичної статистики		99
3.3.	Однорідність даних		100
3.4.	Перетворення магнітудних шкал		101
3.5.	Оцінка параметрів і дослідження мультиколінеарності та гомоскедастичності лінійної багатовимірної регресії		106
	3.5.1.	Метод Фаррара – Глобера	112
	3.5.2.	Метод головних компонент	114
	3.5.3.	Дослідження гомоскедастичності в регресійному аналізі	115
	3.5.4.	Тест рангової кореляції Спірмена	116
	3.5.5.	Тест Глейзера	118
	3.5.6.	Дослідження однорідності сейсмічних даних	119
3.6.	Визначення мінімальної магнітуди комплектності		121
3.7.	Статистичний зв'язок між короною і підкороною сейсмічною активністю вогнища "Вранча"		123
3.8.	Дослідження статистичного зв'язку між магнітудою і глибиною землетрусів		129
3.9.	Взаємозв'язок між магнітудою і інтервалом часу між землетрусами		131
3.10.	Дослідження зв'язку між короною і підкороною сейсмічною активністю.		134

3.11.	Дослідження зв'язку між сейсмічністю вогнища Вранча і числом сонячних плям	136
Висновки до розділу 3		137
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ		139
РОЗДІЛ 4. МОДЕЛІ СЕЙСМІЧНОГО РЕЖИМУ		146
4.1.	Ймовірнісні моделі сейсмічного режиму	146
4.2.	Марковська модель потоку землетрусів	147
4.3.	Дослідження вкладеного марковського ланцюга	156
4.3.1.	Оцінка параметрів неперервного марковського ланцюга	163
4.3.2.	Статистика марковських ланцюгів	167
4.3.3.	Спільний розподіл кінцевої сукупності параметрів землетрусів	170
4.4.	Ймовірнісна модель балканського сейсмічного регіону	172
4.4.1.	Дослідження статистичного зв'язку між сейсмічною активністю Балканських сейсмічних зон	174
4.4.2.	Таблиці спряженості ознак	176
4.4.3.	Вибіркове кореляційне відношення	178
4.5.	Марковська модель сейсмічних зон	180
4.5.1.	Марковська модель румунських сейсмічних зон	181
4.5.2.	Марковська модель балканських сейсмічних зон	184
Висновки до розділу 4		186
Список використаних джерел		188
РОЗДІЛ 5. СТАТИСТИЧНИЙ АНАЛІЗ МАКРОСЕЙСМІЧНОГО ПОЛЯ		192
5.1.	Рівняння загасання	192
5.2.	Математичний опис еліпса	195
5.3.	Визначення епіцентральної відстані	206
5.4.	Обчислення значень коефіцієнтів рівняння макросейсмічного поля	209

5.5.	Дослідження неоднорідності геофізичної середовища за макросейсмічними даними		214
Висновки до розділу 5			221
Список використаних джерел			222
РОЗДІЛ 6. ЧАСОВІ РЯДИ У ДОСЛІДЖЕННІ СЕЙСМІЧНОГО РЕЖИМУ			228
6.1.	Часовий ряд числа місячних землетрусів		229
6.2.	Аналіз даних		230
6.3.	Дослідження структури ряду		232
6.4.	Автокореляційні функції рядів		237
6.5.	Аналіз тренда в часових рядах		241
	6.5.1.	Вибір оптимальних моделей часових рядів	243
	6.5.2.	Метод експоненціального згладжування і прогнозування часових рядів	245
6.6.	Розкладання рядів на компоненти		249
6.7.	Спектральний аналіз часового ряду		250
6.8.	Лінійні моделі часових рядів		255
Висновки до розділу 6			260
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ			261
РОЗДІЛ 7. ЙМОВІРНІСНИЙ АНАЛІЗ СЕЙСМІЧНОЇ НЕБЕЗПЕКИ			265
7.1.	Оцінка сейсмічної небезпеки		267
7.2.	Розподіл інтенсивності сейсмічних впливів		271
7.3.	Метод побудови карти сейсмічної небезпеки		274
7.4.	Обчислення сейсмічного ризику		284
7.5.	Оцінка максимальної магнітуди землетрусів за статистикою Гумбеля		286
	7.5.1.	Перший граничний розподіл	289
	7.5.2.	Третій граничний розподіл	291

7.6.	Складання стохастичного каталогу землетрусів	295
	Висновки до розділу 7	300
	Список використаних джерел	301
	Загальні висновки	308
	Додаток 1- Список публікацій здобувача за темою дисертації	314

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ

- ЙАСН – Ймовірнісний аналіз сейсмічної небезпеки
- ВВЗ – Виникнення вогнищ землетрусів
- ДАСН – Детерміністський аналіз сейсмічної небезпеки
- ЗСР 97 – Загальне сейсмічне районування 1997 р.
- ЗСР 2012 – Загальне сейсмічне районування 2012 р.
- СН – сейсмічна небезпека
- СР – сейсмічне районування
- СР-78 – сейсмічне районування 1978 р.
- ЕДЗ – елементарне джерело землетрусів
- EUROCODE 8 (EC8) – будівельні норми Європейського Союзу ЄС
- IDP – Intensity Data Points (точки даних бальності)
- ISC – International Seismological Centre (<http://www.isc.ac.uk>)
- MGR – магнітуда Гутенберга–Ріхтера
- mb – магнітуда за телесеїсмічною короткоперіодною об'ємною хвилею
- mB – магнітуда за середньо періодною об'ємною хвилею (M_{pv})
- MS – магнітуда за поверхневою хвилею
- MSK-64 – шкала інтенсивності Медведєва–Шпонхойера–Карника 1964 р.
- MW – моментна магнітуда
- ML – локальна магнітуда
- MD – магнітуда по тривалості запису землетрусу
- MJ (JMA) – магнітуда для регіональних землетрусів, використовується в Японії
- NEIC – National Earthquake Information Center (<http://earthquake.usgs.gov/regional/neic/>)
- PDE – The NEIC global earthquake bulletin called the Preliminary Determination of epicenters.
- PGA – пікові прискорення

ROMPLUS – каталог румунських землетрусів, database catalogue:

<http://www.infp.ro/>)

USGS – United States Geological Survey (<http://www.usgs.gov>)

Сейсмічні регіони:

ROM – Румунія

CRO – Хорватія

BUL – Болгарія

BH – Боснія

ALB – Албанія

CZH – Чехія

D – Німеччина

SLK – Словачія

MON – Чорногорія

HUN – Угорщина

SLO – Словенія

SR – Сербія

UKR – Україна

MAC – Македонія

A – Австрія

GR – Греція

TUR – Туреччина

POL – Польща

GR – Греція

TUR – Туреччина

POL – Польща

ЧР – часовий ряд

ВСТУП

Сейсмічну небезпеку неможливо зменшити, можна зменшити прийнятний сейсмічний ризик, забезпечуючи інженерів–будівельників, проєктувальників та інших фахівців, які приймають рішення, достатньою для цього інформацією. Прийнятний сейсмічний ризик – це певний компроміс між рівнем сейсмічної безпеки і спроможністю суспільства забезпечити її в даний момент часу. Актуальність розвитку методів оцінки сейсмічної небезпеки зростає у зв'язку із зростанням народонаселення, будівництвом висотних і надвисотних будівель, збільшенням кількості об'єктів стратегічної важливості і розвитком міст.

Фундаментальні дослідження сейсмічності і прогнозу сейсмічної небезпеки носять неперервний характер, оскільки удосконалюються системи спостережень, надходить якісно нова інформація, розробляються нові математичні та обчислювальні моделі сейсмічності і проєктування сейсмостійких споруд. Стратегія сейсмостійкого будівництва і зменшення сейсмічного ризику потребує постійного вдосконалення методів розрахунку і складання все більш інформативних нормативних прогнозних карт сейсмічної небезпеки.

У 80-ті рр. в Інституті геофізики і геології АН Молдови під керівництвом Н.І. Онофраша були розпочаті дослідження з розвитку кількісних методів оцінки сейсмічної небезпеки. Передбачалося описати вогнище землетрусів у вигляді лінійної інтерференційної системи елементарних випромінювачів, що дозволило б описати механізм вогнища землетрусів, обчислити функцію розподілу параметрів такого випромінювача і висловити сейсмічну небезпеку в інженерних параметрах впливу сейсмічних хвиль, що і потрібно для проєктування сейсмостійких об'єктів. В ході виконання роботи початкові плани зазнали деякі змін, і придбали форму, що викладена в даній роботі.

Як відомо, сейсмологія – це розділ геофізики, яка займається вивченням землетрусів і пов'язаних з ними явищ. Основне завдання сучасної сейсмології – навчитися передбачати силу, час і місце виникнення землетрусів. Стратегія

сейсмостійкого будівництва також базується на інформації про потенційну сейсмічну небезпеку.

У визначенні параметрів сейсмічності випадковість відіграє велику роль. Тому детерміноване прогнозування сейсмічної небезпеки має істотні обмеження. Параметри сейсмічності не є невідомими, їх значення, як випадкові величини, можна задавати конкретною функцією розподілу, список яких широкий. Завдяки цьому, багато емпіричних розподілів вдається апроксимувати адекватними теоретичними функціями розподілів.

Перші дослідження сейсмічної небезпеки були виконані ще в кінці XIX століття в Англії [Musson, Sargeant; 2007]. Метод базувався на спостереженнях, простим способом, визначалися території, які в різній мірі постраждали від землетрусів. Більш ніж за вікову історію методи аналізу сейсмічної небезпеки зазнали великих змін. Інтенсивне зростання міст і висотне будівництво на початку XX століття ставили великі вимоги до якості визначення потенційної сейсмічної небезпеки. На зміну споглядальних методів прийшли детерміністські (ДАСН) і ймовірнісні (ЙАСН) методи аналізу сейсмічної небезпеки. Поява електронно–обчислювальних машин сприяла розвитку кількісних методів оцінки сейсмічної небезпеки, які розвивалися з урахуванням всієї передісторії методів аналізу сейсмічної небезпеки, і залишаються фактично детерміністськими. Завдання ймовірності перевищення інтенсивністю сейсмічних впливів деякого порогу в задані інтервали часу несуттєво змінює ситуацію, оскільки параметри оцінюються з детерміністських позицій.

Оцінка сейсмічної небезпеки, яка носить прогнозний характер, залежить від декількох параметрів вогнища землетрусу. Це географічні координати епіцентру, глибина джерела, магнітуда та функція загасання інтенсивності сейсмічного ефекту з відстанню. В ймовірнісному підході до аналізу сейсмічної небезпеки всі параметри землетрусу вважаються випадковими величинами, для визначення ймовірнісної структури яких використовуються відомості з каталогів землетрусів. У деяких випадках інтервал часу, охоплюваний каталогом, є коротшим від періоду повторення землетрусу з найбільшою силою. Функція загасання визначається на

основі макросейсмічних даних – містять похибки спостережень, пов'язані з неврахованими факторами і невизначеностями. Все сказане є переконливим аргументом на користь розвитку методів оцінки сейсмічної небезпеки на принципово новій основі:

Припустимо, що будівельний майданчик Q , з координатами (φ, ψ) , знаходиться в радіусі дії групи з S сейсмічних зон. Проектувальник прогнозує якої інтенсивності сейсмічний вплив може викликати майбутній землетрус, і яка може бути сейсмічна обстановка в найближчі, наприклад, 475 років? Сейсмічна обстановка визначається рядом природних умов: землетрус деякої магнітуди відбувається в одній з сейсмічних зон; вогнище розташовується в деякій точці геофізичного простору, який займає сейсмічна зона; землетрус викликає сейсмічний поштовх певної інтенсивності на будівельному майданчику, яка визначається функцією загасання по мірі віддалення від вогнища; за фіксований період часу в сейсмічних зонах відбувається якесь число землетрусів, які викликають у точці $Q(\varphi, \psi)$ деяке число сейсмічних поштовхів певної інтенсивності. Виникнення сейсмічних подій відбувається за деякою закономірністю, незвіданою до кінця на сучасному рівні розвитку наук про Землю. Тому, рішення охопити адекватною ймовірнісною моделлю ланцюга сейсмічних подій для прогнозування сейсмічної обстановки є виправданим. Методологія продемонстрована на прикладі сейсмічного районування території Румунії і сусідніх країн. Внесок деякої сейсмічної зони Z^s , $s=1, \dots, S$ в сумарну сейсмічну небезпеку оцінюється ймовірністю γ_s – того, що очікуваний землетрус відбудеться в деякій зоні Z^s . Дана подія позначається тим же ідентифікатором. Оцінкою максимальної правдоподібності ймовірності γ_s є відносна частота:

$$\gamma_s = \frac{N_s}{N} \quad (1)$$

де N_s – спостережена кількість сейсмічних подій в зоні Z^s , а N – загальна кількість землетрусів у всіх сейсмічних зонах.

Діапазон можливих значень інтенсивності сейсмічних впливів розбивається на K_{\max} інтервалів. Землетрус, що стався в зоні Z^s , з деякою ймовірністю може

викликати в пункті (φ, ψ) сейсмічний поштовх зі значенням інтенсивності з інтервалу $(I^k-0.5, I^k+0.5)$ в балах шкали *MSK-64*, або ж $(a_k-\Delta a, a_k+\Delta a)$ у значеннях параметрів руху ґрунту. Наприклад, розподіл інтенсивності I^k сумарних сейсмічних впливів у пункті (φ, ψ) , викликаний землетрусами з усіх зон, визначається формулою:

$$\beta^k = \sum_{s=1}^S \gamma_s \beta_s^k \quad (2)$$

Ймовірності $\beta_k, k=1, \dots, K_{\max}$, складають вектор розподілу інтенсивності сейсмічних струсів:

$$\vec{\beta} = (\beta_1, \dots, \beta_{K_{\max}}) \quad (3)$$

Цей вектор і ймовірнісна модель потоку сейсмічних подій, дозволяють здійснювати в найбільш повному, – на даному етапі розвитку сейсмології обсязі, ймовірнісний аналіз сейсмічної небезпеки.

Нехай у пункті спостереження (φ, ψ) за час t з ймовірністю $p(t, n)$ відбуваються n сейсмічних поштовхів. Тоді умовна ймовірність появи m_1 поштовхів з інтенсивністю I^1 , m_2 рази з інтенсивністю I^2, \dots , і m_{12} випадків з інтенсивністю I^{12} , $(m_1 + \dots + m_{12} = n)$ відповідає поліноміальному співвідношенню:

$$p_n(m_1, m_2, \dots, m_{12}) = \frac{n!}{m_1! \dots m_{12}!} \beta_1^{m_1} \dots \beta_{12}^{m_{12}} \quad (4)$$

руху

$$p_n(m_k) = \frac{n!}{m_k! (n - m_k)!} \beta_k^{m_k} (1 - \beta_k)^{n - m_k} \quad (5)$$

Середня кількість перевищень рівня інтенсивності I^k або ж параметра руху a_k за час t визначиться за формулою:

$$\alpha_k = \lambda t \beta_k \quad (6)$$

де λ – інтенсивність виникнення землетрусів в групі сейсмічних зон.

Нехай, виникнення сейсмічних впливів в деякій точці з інтенсивністю I , або ж значенням параметра руху ґрунту a , що перевищують поріг інтенсивності I^k , або ж зміщення a_k ґрунту відповідно, являє собою пуассонівський процес:

$$p(t, n) = e^{-\alpha_k t} \frac{(\alpha_k t)^n}{n!} \quad (7)$$

В такому випадку, оцінкою міри сейсмічної небезпеки служить ймовірність:

$$P_{\varphi,\psi}(t,n,m,I^k) = p(t,n) \frac{n!}{m!(n-m)!} \beta_k^m (1-\beta_k)^{n-m} \quad (8)$$

Ця ймовірність пропонується в якості превентивної сейсмічної небезпеки. Карта сейсмічної небезпеки видається масивом точок з координатами (φ_i, ψ_j) , $i=1, \dots, L_\varphi$; $j=1, \dots, L_\psi$, генерованих з деяким кроком дискретності $\Delta\varphi$ і $\Delta\psi$ за широтою та довготою, відповідно. Тобто територія, яка піддається впливу землетрусів покривається географічної мережею, і в вузлах мережі розраховується СН.

Всі існуючі міри сейсмічної небезпеки і сейсмічного ризику є окремими випадками цієї міри. Наприклад, пов'язуючи з кожним поштовхом деякої інтенсивності I^k економічні збитки, отримаємо вірогідну оцінку економічного ризику. Ймовірність того, що з t років в деякій сейсмічній зоні хоча б один раз трапиться землетрус, який в деякій точці (φ, ψ) викличе вплив, що перевищує поріг інтенсивності I^k балів, або ж рівень a_k – параметра зсувів ґрунту, виражається формулою:

$$1 - P(50,0) = 1 - e^{-\alpha_k t} \quad (9)$$

Вважаючи $t=50$ років, отримаємо оцінку сейсмічної небезпеки в термінах престандарта *ЕС8*. Можна визначити найбільше значення інтенсивності поштовхів, характеристик зсувів ґрунту – тобто те значення, ймовірність, перевершити яке дорівнює нулю. Таким чином, для застосування методу необхідно знати місце розташування вогнища землетрусів, його магнітуду, закон загасання інтенсивності сейсмічного впливу і частоту виникнення землетрусів. Рішення перерахованих завдань є головною ідеєю, яка спонукала на здійснення даних досліджень.

Запропонована методологія ґрунтується на припущенні про стаціонарність, тобто незалежності від часу сейсмічного процесу.

Зв'язок роботи з науковими програмами, планами і темами.

Робота виконана в 2007–2015 рр. в лабораторії сейсмології Інституту геології і сейсмології Академії Наук Молдови в рамках науково-дослідних тем: характеристика сейсмічності території Республіки Молдова; ймовірнісний метод

оцінки сейсмічної небезпеки; дослідження процесів, що визначають сейсмічність зони Вранча і факторів сейсмічного ризику на території РМ (2001–2005); геодинамічні процеси в зоні Вранча і розвиток методів розрахунку сейсмічної небезпеки і ризику (2006–2010), ІГС-АН РМ; розробка методів кількісної оцінки сейсмічної небезпеки і ризику відповідно до будівельних норм *EC8*, ІГС-АН РМ; характеристика сейсмічності території Республіки Молдова; ймовірнісний аналіз сейсмічної небезпеки територій Республіки Молдова в рамках міжнародного проекту Італія–Молдова 11.820.08.03/itf "Characterization of seismicity of Moldova Republic territory. Contribution to seismic hazard assessment "; сейсмічне районування території Туреччини.

Мета і завдання дослідження: Розробка методології оцінки сейсмічної небезпеки на основі ймовірнісних моделей сейсмічності; розробка алгоритму і обчислювальної програми оцінки і картування сейсмічної небезпеки. Для досягнення поставленої мети необхідно було вирішити такі основні завдання в чотири етапи:

На першому етапі, як і в ДАСН, визначаються сейсмічні зони, і передбачається, що вогнища землетрусів розподілені з певною ймовірністю в геофізичному просторі, який займає зона вогнищ землетрусів. У цьому полягає основна відмінність від ДАСН, де фіксується найближчий до об'єкту активний розлом, тобто розлом, в якому з ймовірністю, що дорівнює одиниці може виникнути осередок землетрусу.

На другому етапі будується марковська модель сейсмічного режиму. На основі ергодичного неперервного марковського ланцюга можна передбачити магнітуду і їх середнє число за фіксований період часу. У ДАСН цьому відповідає закон повторюваності Гуттенберга-Ріхтера.

Третім кроком є визначення параметрів загасання сейсмічного ефекту. В даному випадку, аналіз сейсмічної небезпеки проводиться на основі бальності макросейсмічної шкали *МСК-64* і в прискореннях зміщення ґрунту по рекомендованим будівельним нормам *EC8*.

На четвертому етапі для конкретної території, яка знаходиться в радіусі дії групи сейсмічних зон, обчислюється сумарна міра сейсмічної небезпеки

Об'єкт дослідження: сейсмічність Румунії, сейсмічна зона Вранча і макросейсмічні дані підкорових вранчських землетрусів.

Предметом дослідження є розроблена нова методологія оцінки сейсмічної небезпеки і алгоритму сейсмічного районування на прикладі, території Молдови, Румунії та Болгарії.

Методи дослідження:

- Кластерний аналіз каталогів землетрусів для виявлення сейсмічних зон
- Визначення мінімальної магнітуди комплектності
- Регресійний аналіз для перетворення магнітуд
- Метод найменших квадратів для обчислення коефіцієнтів рівняння макросейсмічного поля і метод Фітцгіббона для апроксимації поля еліпсом
- Оптимізаційні методи для обчислення оптимальних значень коефіцієнтів рівняння загасання
- Теорія марковських процесів для визначення розподілу параметрів землетрусу
- Теорія часових рядів для дослідження стаціонарності потоку сейсмічних подій
- Лемма Неймана-Пірсона для перевірки гіпотези про відповідність стандартного нормального розподілу відхилень реальних значень в балах макросейсмічною інтенсивності в населених пунктах від їх теоретичних значень .

Наукова новизна отриманих результатів:

- Розроблена принципово нова методологія розрахунку сейсмічної небезпеки, яка дозволяє розрахувати сейсмічну небезпеку в населених пунктах, розташованих в радіусі впливу групи сейсмічних зон і побудувати карту сейсмічного районування великих територій
- Застосовано кластерний аналіз для виявлення сейсмічних зон
- Розроблена марковська модель сейсмічності, яка дозволяє передбачити місце виникнення і магнітуду можливих землетрусів в сейсмічних зонах
- Досліджено статистичний зв'язок між балканськими сейсмічними зонами

- Застосовані математичні методи для апроксимації макросейсмічного поля кривими другого порядку і виявлення еліптичності макросейсмічного поля на прикладі підкорових вранчських землетрусів.
- Здійснено статистичний аналіз макросейсмічного поля, складена модель загасання інтенсивності струшувань та визначені регіональні значення коефіцієнтів рівняння загасання на прикладі вранчських землетрусів
- Розроблено алгоритм розрахунку і картування сейсмічної небезпеки на основі найбільш інформативної міри потенційної сейсмічної загрози, як ймовірності

$$P_{(\varphi, \psi)}(t, n, m, I^k) \quad (10)$$

того, що в географічній точці $Q(\varphi, \psi)$ за час t трапиться n сейсмічних поштовхів, m з інтенсивністю I^k .

- Здійснено ймовірнісний аналіз сейсмічної небезпеки і побудована загальна карта сейсмічної небезпеки територій Молдови, Румунії, Болгарії та території Туреччини.

Практичне значення отриманих результатів:

Розроблено методологію, яка дозволяє розрахувати сейсмічну небезпеку в населених пунктах, розташованих в радіусі впливу групи сейсмічних зон, і побудувати карту сейсмічного районування великих територій, наприклад, для всієї території Європи.

Даний підхід дозволяє застосувати обчислювальні засоби на всіх етапах розрахунку сейсмічної небезпеки: статистичний аналіз макросейсмічних даних, виділення сейсмічних зон, побудова ймовірнісних моделей сейсмічного режиму, визначення функції загасання інтенсивності сейсмічних впливів, розрахунок і картування сейсмічної небезпеки.

На прикладі вранчських землетрусів здійснений статистичний аналіз макросейсмічних даних, визначені регіональні закономірності загасання сейсмічного ефекту, розроблено алгоритм і обчислювальна програма для оцінки сейсмічної небезпеки і побудовані карти сейсмічного районування території Республіки Молдова, Туреччини та загальної території Молдови, Румунії та Болгарії. Обчислювальна програма «A Fortran code for the seismic hazard assessment

and mapping based on a Markovian model of seismic regime» зареєстрована в фонді програм Інституту методів екологічного аналізу в місті Тіто-Скала, Італія (Istituto di metodologie di analisi ambientale. Tito Scalo, Italy).

Особистий внесок автора

В основу дисертації покладені результати роботи здобувача в Інституті геофізики і геології (геології і сейсмології, з 2005р.) АН Молдови. Наведені результати отримані автором самостійно.

Апробація роботи. Основні результати роботи доповідалися: на Генеральній Асамблеї Європейської сейсмологічної комісії (ЕСК XIX, Москва, 1984), Всесоюзних нарадах (Агверан, Кишинів, Чорноголовка, період 1983–1988), міжнародних конференціях: ECBR WORKSHOP– 2011, The National Symposium "75 Years from November 10th 1940 Vrancea Earthquake"-2015 (Бухарест); ESNET (Стамбул, 2012; Варна, 2013, Констанца, 2014 року); International Conference on Moldavian Risks-from Global to local scale (Вакеу 2012, 2014 року); INCERCOM - 2014 року, Conferinta stiintifica consacrata aniversarii de 110 ani de la nasterea geologului basarabean I.M. Suhov - 2015 (Кишинів), Georisks in the mediterranean and their mitigation (Валлетта, 2015).

Публікації:

Результати роботи викладені в 28 статтях у фахових журналах, в одній монографії та 4 тезах доповідей конференцій.

Структура і обсяг роботи. Дисертація, загальним обсягом 317 сторінок, складається з вступу, семи розділів, висновків, списку використаних джерел, що включає 342 найменування. Робота проілюстрована рисунками і таблицями.

Автор зберігає вдячну пам'ять про перших наставників Г.А. Алібекова, Н.І. Онофраша і викладачів механіко-математичного факультету Київського Національного Університету ім. Т.Г. Шевченко. Висловлює щирю подяку колегам по роботі: В.Г. Алказу, І.Е. Вербицькому, В.Ю. Карданцю, Н.Л. Тропіну, В.І. Чорному, В.І. Шумілі.

РОЗДІЛ 1 ЙМОВІРНІСНИЙ АНАЛІЗ СЕЙСМІЧНОЇ НЕБЕЗПЕКИ (СТАН ПРОБЛЕМИ)

Перші дослідження сейсмічної небезпеки були виконані ще в XIX столітті [Musson, Sargeant, 2007]. Метою цих робіт було бажання ранжувати простим способом території в різній степені постраждали від впливу землетрусів. На цьому етапі не ставилося завдання кількісних оцінок очікуваних сейсмічних впливів, досить було показати сейсмічний ефект історичних землетрусів. Наступним кроком став перехід від наглядних методів до детерміністських (ДАСН) і ймовірнісних (ЙАСН) методів аналізу сейсмічної небезпеки. Сейсмічний процес пов'язаний зі складними геологічними, фізико-хімічними процесами, що відбуваються в надрах Землі, з переплетенням фізичних полів Землі і гравітаційної сили небесних тіл, і їх впливом на глобальну тектоніку. Ці процеси характеризуються великою кількістю ознак, як і сейсмічні процеси. Зв'язок між сейсмічністю і перерахованими вище причинними ознаками має чотири форми [Malachow, 1981]:

- Закону, якщо залежність між ознаками встановилася, невідворотна, однозначна і поява землетрусу повторюється кожного разу, коли спостерігаються причинні процеси
- Статистичного закону, коли певна сейсмічна подія може трапитися, може, й ні, хоча частина або всі причинні ознаки спостерігаються
- Теорії або принципу, коли немає певної закономірності між перерахованими вище процесами і сейсмічністю
- Гіпотези, коли немає достатньо підстав для теорії, то для пояснення феномена виникнення землетрусів при спостереженні процесів, здатних викликати сейсмічну подію, висувається гіпотеза.

У даній роботі, на всіх етапах ймовірнісного аналізу сейсмічної небезпеки, елемент випадковості має місце: розташування вогнища землетрусу; час виникнення;

магнітуда; загасання інтенсивності сейсмічних впливів з віддаленням від вогнища і т.д. У дослідженні сейсмічності беруть участь всі чотири форми залежності.

1.1. Огляд методів аналізу сейсмічної небезпеки

Оцінка сейсмічної небезпеки – перший крок до сейсмічного районування та оцінки сейсмічного ризику. У загальних словах, сейсмічна небезпека (СН) визначається як максимальні сейсмічні впливи, очікувані з певною ймовірністю на заданій території в заданому інтервалі часу і зумовлені повторюваністю землетрусів. Сейсмічна небезпека доповнюється також вторинними явищами, що супроводжують землетрус (поверхневі розриви, зсуви, обвали, цунамі і т.п.) і впливають на нормальну життєдіяльність.

Сейсмічне районування в США і в деяких інших країнах, почало проводитися інженерами - будівельниками, за участю сейсмологів і геологів, з метою забезпечення сейсмостійкого будівництва. Картування сейсмічної небезпеки проводилося в параметрах руху ґрунту [Musson, Sargeant, 2007].

Перша в світі офіційна карта сейсмічного районування (СР) була складена Г.П. Горшковим в 1937 для всієї території СРСР. Сейсмічна небезпека полягала в максимальних проявах сейсмічних впливів. Потім відбувся перехід від принципу, що землетрус відбудеться там, де вони відбувалися, на 2-х стадійну. Виділялися реальні і потенційні вогнища, потім розраховувався очікуваний струсу на земній поверхні. Карти СР регулярно оновлювалися приблизно кожні 10 років, з урахуванням відомостей про ті, що відбулися землетруси і в зв'язку з удосконаленням методів розрахунку сейсмічної небезпеки. Необхідність поновлення карти сейсмічного районування була пов'язана з тим, що в зонах, які вважалися раніше асейсмічними або зовсім несейсмічними, траплялися сильні землетруси. Так як ці землетруси відбувалися в районах, де сейсмостійкість в будівництві не передбачалася, руйнування часто виявлялися катастрофічними [Уломов, 1993; Уломов, Богданов, 2013]. Оцінка сейсмічної інтенсивності зводилася до реєстрації макросейсмічного ефекту, вираженого в цілочисельних балах, що

охоплюють при сильних землетрусах великий діапазон інтенсивності сейсмічних впливів. Це нерідко призводило до невиправданого завищення або заниження оцінок прогнозованого сейсмічного ефекту, що позначалося на якості і вартості сейсмостійкого будівництва. На відміну від карт СР зразка 1937, 1957, 1968 і 1978 рр., для ЗСР-97 складений комплект карт, де враховані недоліки попередніх карт сейсмічного районування.

Детерміністські методи аналізу сейсмічної небезпеки (ДАСН) ґрунтуються на геологічних даних, при цьому головним є визначення найближчого активного розлому і обчислення максимального можливого для нього землетрусу. Потім обчислюється найбільша можлива інтенсивність сейсмічних поштовхів на поверхні Землі. Розрахунки проводяться на основі теорії поширення сейсмічних хвиль з урахуванням їх загасання по мірі віддалення від вогнища землетрусів і впливу місцевих ґрунтів, які зазвичай подаються у вигляді горизонтально шаруватого середовища. При цьому розглядаються плоскі фронти сейсмічних хвиль, що поширюються вертикально вгору і викликають сейсмічні впливи для вільної земної поверхні, не навантаженої будь-яким спорудою [Алказ, 2007; Гусєв, Шумилина, 1995; Уломов, Богданов, 2013].

Детерміністські методи використовують наявні дані про сейсмічності для аналізу сейсмічної небезпеки. Головною ідеєю ДАСН є визначення найгіршого сценарію розвитку сейсмічної ситуації на деякій території при виникненні землетрусу.

Метод передбачає виконання таких кроків [Golbs, 2008]:

1. Виявлення та характеристика всіх потенційних вогнищ землетрусів, здатних викликати значні зсуви ґрунту в заданій точці. Характеристика джерел землетрусів означає визначення геометрії вогнища і обчислення максимальної можливої магнітуди землетрусу.
2. Обчислення найкоротшого відстані між точкою і вогнищем землетрусів

3. Вибір землетрусу, здатного викликати найбільші зміщення ґрунту на досліджуваній території (домінуючий землетрус). Вибір відбувається порівнянням значень зсувів ґрунту, спричинених усіма землетрусами
4. Визначення за законом загасання значення руху ґрунту на епіцентральної відстані R , викликаного землетрусом з магнітудою M .

Сейсмічна небезпека для досліджуваної території визначається найбільшими зсувами ґрунту. Значення ДАСН і його місця в сучасних підходах до оцінки сейсмічної небезпеки викладені в роботах [Abrahamson, Bommer, 2005; Baker, 2008; Golbs, 2008; Mcguire, 1999; Grünthal et al., 1996; Krinitzsky, 2003].

Перевагами даного методу є:

- Відносна простота і швидкість отримання оцінок
- Висока точність результатів.
- Придатність як тестового критерію адекватності моделі сейсмічності.

Недоліки методу:

- Не завжди вдається однозначно виявити найближчий сейсмоактивний розлом і визначити його геометрію. Не виявлені, приховані розломи не розглядатимуться як сейсмоактивні, і не будуть враховані в оцінці небезпеки
- Кількісно вираз найбільшої можливої сейсмічної події пов'язано з великими невизначеностями.

В СРСР практично всі карти сейсмічного районування, починаючи з першої карти 1937 р до карти 1978 року, були детерміністськими. Вони не враховували основні характеристики сейсмічного режиму сейсмоактивних територій. На відміну від карт СР зразка 1937, 1957, 1968 і 1978 рр., Карти ЗСР-97 складаються з трьох карт, де в вузлах географічної сітки, яка покриває досліджувану територію з деяким кроком дискретності, розрахована ймовірність неперевищення значення відповідної бальності [Уломов, 1993; Уломов, Богданов, 2013].

Метод ЗСР-97 і його вдосконалений варіант ЗСР-2012 використовують в повній мірі геометрію вогнища, особливості геофізичної середовища і сейсмічних процесів в ній, уявлення про граничну величину максимальної можливої магнітуди

землетрусів, і є значним просуванням на шляху розповсюдження інформації збагачення міри сейсмічної небезпеки. ЗСР-2012 передбачає збільшення періодів повторюваності і введення дробових балів, проте дані методи, по суті, залишаються гібридними, так як присутність фактора ймовірності неперевикнення деякого порогу не надає їм характеру ймовірнісного методу і характеризується тими ж, властивими ДАСН, недоліками. Концепції ДАСН і ЙАСН не повинні протиставлятися, їх потенціал повинен бути використаний для вирішення загального завдання: удосконалення методів розрахунку сейсмічної небезпеки і підвищення якості карт СН. Адаптація висотних і надвисотних конструкцій, де застосовуються нові технології сейсмостійкого будівництва, пред'являє високі вимоги до якості та інформаційної наповненості мір сейсмічної небезпеки. Недоліки в обліку даних та невизначеності, які супроводжують чинники, що визначають СН, обмежують застосування детерміністських методів. ДАСН використовується тоді, коли непрямі збитки, в разі похибкової оцінки сейсмічної небезпеки, не будуть високими. Детерміністські методи, які засновані на минулих сейсмічних спостереженнях, виявляють велику частку випадковості в результатах аналізу сейсмічної небезпеки. Детерміністський підхід корисний при оцінці небезпеки територій окремих міст і стратегічно важливих об'єктів. Похибки в оцінці параметрів землетрусів, невизначеності при створенні математичних моделей геофізичних процесів, відсутність обліку економічної складової забезпечення сейсмостійкого будівництва і оцінки сейсмічного ризику обмежують застосування детерміністського підходу до сейсмічного районування.

Детерміністські карти ЗСР не враховують особливостей сейсмічного режиму досліджуваних територій. Недоліки і невизначеності, що існують в ДАСН, а також похибки у визначенні параметрів сейсмічності сприяло розвитку методології оцінки сейсмічної небезпеки на ймовірнісній основі [Онофрасс, 1981; Ризниченко, 1965, 1966; Algermissen, Perkins, 1976; Baker, 2008; Cornell, 1968; Golbs, 2008; Karnik et al, 1981; Mcguire, 2004; Reiter, 1990; Solomos et al., 2008; Ismail-Zadeh, et al., 2007, 2012; Sokolov, Ismail-Zadeh, 2015, 2016]. Було б неможливим прямими спостереженнями

над сейсмічністю змодельовати розподіл інтенсивності майбутніх сейсмічних поштовхів, тому що період повторюваності великих землетрусів іноді становить десятки і сотні років, тоді як важливий продукт сейсмології, – каталоги землетрусів, охоплюють надто короткий, для збору достатньої кількості спостережень над сейсмічністю, інтервал часу. Ймовірнісний підхід до оцінки сейсмічної небезпеки, в умовах існуючих невизначеностей в параметрах сейсмічності і похибок в їх оцінках, використовує добре розроблені методи теорії ймовірностей і математичної статистики. У своїй основі ЙАСН є кількісним методом оцінки сейсмічної небезпеки, який дозволяє об'єднати всі доступні відомості про сейсмічності: розташування вогнища, частота виникнення сейсмічних подій, магнітуда землетрусів, а також модель поширення сейсмічних коливань і т.д. Ймовірнісний підхід до оцінки сейсмічної небезпеки, що з'явився в середині 60-х років минулого століття, швидко набув широкого поширення в більшості країн світу. В СРСР технологія ЙАСН була розпочата під керівництвом Ю.В.Різніченко [Різниченко, 1965, 1966], і отримала подальший розвиток в роботах [Онофраш, 1981; Шумила, 1985]. У загальному випадку, метод, який з виникненням деякого сейсмічної події пов'язує вірогідну міру, називається ймовірнісним. Ймовірнісний аналіз сейсмічної небезпеки ґрунтується на випадковий характер самих сейсмічних подій, що виражається в розсіюванні їх параметрів. На відміну від детерміністських методів ймовірнісний підхід дозволяє досліджувати великий обсяг даних.

Запропонований Корнеллом в 1968р. ймовірнісний метод аналізу сейсмічної небезпеки на базі сейсмостатистики становить основу ЙАСН. Публікація Корнелла [Cornell, 1968] привела до побудови карт сейсмічного районування в термінах ймовірності перевищення інтенсивністю сейсмічних впливів деякого порогу в заданий інтервали часу. Ця ймовірність була більш інформативною мірою для обґрунтування проектних і конструктивних рішень. Як характеристики сейсмічних впливів розглядається пікове прискорення ґрунту (PGA) a_g .

Загальноприйнятого стандарту представлення результатів оцінки сейсмічної небезпеки не існує. Багато країн прийняли за основу американські стандарти,

пристосувавши їх до регіональних особливостей сеймотектонічних умов і місцевих норм будівництва. Ряд сейсмологів [Baker, 2008; Burton, 1979; Garcia-Fernandez et al., 1989; Ismail-Zadeh, et al., 2007, 2012; Mcguire, 1999, 2004; Reiter, 1990; Rosenhauer, 1999; Sokolov, Ismail-Zadeh, 2015, 2016] вдосконалили запропонований Корнеллом метод ймовірнісного аналізу сейсмічної небезпеки на основі моделей сейсмічності. Активні розлами (лінійні і точкові джерела) і поверхні розриву розглядаються як потенційні джерела землетрусів. На даний момент часу алгоритм ЙАСН складається з 5 основних етапів [Baker, 2008]:

1. На відміну від ДАСН, де розглядається тільки найбільша сейсмічна подія, враховуються всі джерела землетрусів, які викликають небезпечні для цілісності інфраструктури зміщення ґрунту на досліджуваній території. Це можуть бути розломи, які зазвичай мають форму плоскої поверхні, виявленої із спостережень над розташуванням історичних землетрусів і з геологічних даних. Не розпізнаний розлом, як джерело землетрусів представляється в формі лінії або кола в будь-якій точці яких може статися землетрус. Джерело в формі кола, часто використовується для обліку фонові сейсмічності, або ж для землетрусів, які не пов'язані ні з яким розломом. Лінійне джерело підходить для моделювання розломів, які існують на кордоні двох тектонічних плит. Територія, що не охоплена каталогом, але на якій, за історичними відомостями, спостерігалися землетруси розбивається на потенційно небезпечні географічні сейсмічні зони за допомогою діаграми Вороного за принципом рівної щільності епіцентрів, які не пов'язані з реальними сейсмічними зонами.

В авторському варіанті перевага віддається реальним сейсмічним зонам, а в разі їх відсутності в каталозі землетрусів проводиться розпізнавання сейсмічних зон за допомогою кластерного аналізу.

2. На тектонічних розломах можуть відбуватися землетруси різної магнітуди. Розподіл магнітуд визначається на основі закону повторюваності Гутенберга-Ріхтера.

Автор пропонує застосувати статистику екстремальних значень для визначення максимальної магнітуди.

3. Визначається розподіл відстані між географічною точкою і всіма джерелами землетрусів. Передбачається, що землетруси з однаковою ймовірністю можуть відбутися в будь-якій точці сейсмічної зони.

Автор пропонує розбити чотиривимірний простір, який є декартовим простором діапазонів можливих значень широти, довготи, глибини і магнітуди (параметри землетрусів) на елементарні джерела землетрусів (ЕДЗ). Спостережені в ЕДЗ землетруси, утворюють потік сейсмічних подій, який є реалізацією ергодичного неперервного 4-вимірного марковського ланцюга. На основі марковської моделі стає можливим визначення спільного кінечновимірного розподілу землетрусів по ЕДЗ. В залежності від віддалення географічної точки від ЕДЗ обчислюється епіцентрально відстань або Джойнер-Бура.

4. Наступним кроком є побудова моделі зсувів ґрунту, яка допомагає визначити розподіл ймовірностей інтенсивності коливань ґрунту в залежності від магнітуди, відстані, механізму землетрусу, геологічних особливостей в приповерхневих шарах, наявність ефекту спрямованості і т.д. Оскільки число факторів велике, розглядається спрощена модель загасання сейсмічного ефекту в залежності від магнітуди і відстані.

Проведений автором статистичний аналіз макросейсмічних даних проміжних вранчських землетрусів показує, що макросейсмічне поле апроксимується еліпсом, і відхилення спостережених значень інтенсивності від теоретичних відповідає нормальному розподілу.

5. На останньому етапі обчислюється 10%-на ймовірність перевищення інтенсивністю сейсмічних впливів заданого порогу a_k , і знаходиться річна частота цих перевищень для землетрусів з деякого джерела. Якщо ж джерел кілька, то сумарна частота складається з суми частот усіх джерел. Далі обчислюється ймовірність перевищення деякого порогу інтенсивності сейсмічних впливів за 50 років. Якщо ж джерел землетрусів кілька, то підсумовуються ймовірності перевищень для кожного землетрусу.

Для розрахунку СН, за стандартами *EC8*, обчислюється річна інтенсивність перевищень $w=w(a_g)$ порогу зсувів ґрунту a_g в точці $Q(\varphi, \psi)$. Період середньої повторюваності T_R перевищень даного порогу зміщення ґрунту в даній точці є величиною зворотною до інтенсивності w перевищення порогу a_g [Golbs, 2008; Solomos et al., 2008]:

$$T_R = \frac{1}{w} \quad (1.1)$$

Якщо інтенсивність перевищень за рік дорівнює $w=1/T_R$, то інтенсивність перевищень в T_L років дорівнює: $wT_L=T_L/T_R$. Передбачається, що число перевищень є випадковою величиною, підпорядкованою закону розподілу Пуассона. Згідно пуассонівської моделі можна обчислити ймовірності [Musson, 2007]:

$P[n$ подій в T_L років, поріг a_{gR}]:

$$P(n) = \frac{e^{-wT_L} (wT_L)^n}{n!} \quad (1.2)$$

$P[n=0$ подій в T_L років, поріг a_{gR}]:

$$P(n=0) = e^{-wT_L} \quad (1.3)$$

P [хоча б одне перевищення порога a_{gR} в T_L років]:

$$P_R = 1 - P(n=0) = 1 - e^{-wT_L} = 1 - e^{-\frac{T_L}{T_R}} \quad (1.4)$$

Зв'язок між ймовірністю перевищення P_R в T_L років деякого рівня сейсмічних зсувів a_{gR} і середнім періодом повторення даного рівня сейсмічних рухів задається виразом:

$$T_R = -\frac{T_L}{\ln(1 - P_R)} \quad (1.5)$$

В Європейському стандарті СН *EC8* [Solomos et al., 2008] використовується поріг максимальних прискорень, який буде перевищено з ймовірністю $q=0.1$, хоча б один раз в 50 років. Відповідний період повторення перевищення в пункті $Q(\varphi, \psi)$ становить 475 років.

1.2. Статистичні методи

У загальному випадку статистика ґрунтується на спостережених даних і моделях, а математична статистика пов'язує множина даних з множиною ймовірнісних моделей. У загальному випадку статистика ґрунтується на фіксованих даних і моделях, а математична статистика пов'язує множина даних з множиною ймовірнісних моделей. Основне завдання математичної статистики полягає в отриманні висновків про масові явища за даними спостережень над ними і експериментів. Статистичні висновки являють собою твердження про загальні характеристики цих явищ (ймовірності, законах розподілу і їх параметрах: математичних очікуваннях, дисперсіях і т.д.). Мета математичної статистики – оцінити характеристики генеральної сукупності за вибірковими даними. Властивості генеральної сукупності, як правило, невідомі, і всі висновки про генеральну сукупність здійснюються за вибіркою.

Підхід, запропонований деякими сейсмологами [Писаренко и др., 2014; Родкин, 2005; Kijko, Sellevoll, 1989; Knopoff, Kagan, 1977; Grieser et al., 2011], заснований на статистичному аналізі даних про виникнення землетрусів для кількісного вираження, за статистикою Гумбеля, ймовірного рівня інтенсивності майбутніх сейсмічних поштовхів. Основна ідея полягає у тому, що виникнення екстремальних подій відповідає певному статистичному розподілу. Застосувавши цей розподіл до великих історичних землетрусів, можна оцінити сейсмічну небезпеку. Багато сейсмологів скептично ставляться до ідеї про можливість отримання хороших результатів, коли аналіз починається з відкидання близько 90% фіксованих сейсмічних даних. Наприклад, при оцінці максимальної магнітуди землетрусів зони Вранча за даними каталогу землетрусів, що охоплює 106-річний період часу, відбиралися найбільші значення магнітуд землетрусів, що трапилися за рік, в результаті чого складена вибірка обсягом 106 елементів [Буртиев, 2003]. Оцінки, отримані методами екстремальних значень, мають тенденцію зростання, коли період повторення наближається до тривалості тимчасового інтервалу, охопленого каталогом землетрусів [Raschke et al., 2011]. Часто при дослідженні статистичних

залежностей нехтується можливістю появи великих, але рідкісних подій. Під розподілом з важкими «хвостами», мається на увазі розподіл, хвіст яких не можна відрізати, тобто в даному випадку, не можна нехтувати великими, але рідкісними землетрусами. Великі землетруси являють собою сейсмічні події, збиток від яких може перевищити сумарний збиток від всіх інших землетрусів. При попередньому аналізі вибірки слід застосовувати методи виявлення хвоста і визначення ступеня тяжкості хвоста, наприклад, за оцінкою Хілла [Шепель, Акимов, 2014]. Некоректне застосування статистики екстремальних значень може дати спотворений результат.

На практиці не вдається отримати точні оцінки деяких параметрів сейсмічності. Наприклад, ще недостатньо точно встановлений закон загасання енергії землетрусів, а також геометрія джерела землетрусів. Слід додати, що при застосуванні статистичних методів складно знайти спосіб обліку початкових невизначеностей, особливо випадкового характеру зсувів, властивий ґрунтам.

Значення вказаних невизначеностей розглядаються як випадкові величини з певним типом функції розподілу і з відповідними вибірковими оцінками параметрів розподілу. Період повторення великих землетрусів становить десятки, а іноді сотні років, тоді як каталоги землетрусів охоплюють більш короткий період спостережень. Тому неможливо отримати надійну кількісну оцінку з одних тільки наглядових даних. Метод Монте-Карло застосовується для імітації сейсмічності і створення стохастичних каталогів землетрусів в умовах недостатньої сейсмостатистики.

1.3. Монте–Карло моделювання

Існуючі в параметрах сейсмічності невизначеності можуть бути враховані і досліджені з застосуванням методу Монте-Карло [Беврані, Анічкін, 2005; Hardebeck et al., 2008; Ismail–Zadeh et al., 1999; Kagan, 1991; Lapajne et al., 1997; Metropolis, Milutinovic, 2013; Robinson et al., 2009; Ulam, 1949; Perelman, 1963; Shapira, 1983; Siebert, Sommer, 2006; Woo, 1996]. Причиною популярності ЙАСН методів є можливість врахування великого обсягу даних при складанні моделей. Такі моделі

використовуються в оцінці сейсмічної небезпеки за методом Монте-Карло, який дозволяє чисельними методами вирішити математичні (аналітичні) завдання:

- вирішувати багатовимірні задачі
- використовуючи генератор випадкових чисел, отримувати приблизні результати
- реальні процеси моделювати на основі випадкових чисел
- застосовувати в задачах, де враховуються похибки вимірювань.

Так як модель сейсмічності є описом процесу виникнення (генерації) землетрусів в деякому регіоні, то виникає питання, чи можливо використовувати дану модель для складання штучного каталогу землетрусів. Кожен каталог представляє історію сейсмічності минулого, і передбачається, що її характеристики в регіоні зберігаються без змін у наступні сотні років. На основі статистичного аналізу сейсмічного ефекту великої кількості змодельованих землетрусів обчислюються ймовірнісні оцінки параметрів. Цей метод дозволяє враховувати невизначеності у вхідних параметрах завданням їх функції розподілу з вибіркоким середнім і стандартним відхиленням. Значення різних параметрів можуть бути отримані з розподілу вибіркоких значень кожної моделі сейсмічності.

Метод Монте-Карло, як один з підходів до оцінки сейсмічної небезпеки, набув широкого поширення [Cramer et al., 1996; Cox, Sibert, 2006; Musson, 2000]. Термін "Монте-Карло" використаний в аналізі сейсмічної небезпеки з урахуванням припущень [Golbs, 2008]:

- Монте-Карло моделювання використовується для здійснення вибірки гілок великого логічного дерева, що представляє різні параметри, які виражають існуючі в сейсмічності невизначеності.
- статистичний розподіл історичних землетрусів із застосуванням методу Монте-Карло здійснюється в припущенні, що значення магнітуди і координати епіцентру землетрусів є випадковими величинами.
- для дослідження сейсмічності застосовуються звичайні ймовірнісні методи, а для оцінки сейсмічної небезпеки генерується стохастичний каталог землетрусів, і аналізуються результати: зміщення, швидкості, прискорення.

На відміну від деяких інших підходів, цей метод сумісний зі звичайними ймовірнісними дослідженнями сейсмічної небезпеки [Cramer et al., 1996; Ebel, Kafka, 1999; Lapaune et al., 2000]. Ідея методу така, сейсмічна ситуація моделюється 10000 разів і тільки раз інтенсивність сейсмічних впливів перевищує деяке порогове значення параметрів руху ґрунту. Метод заснований на сейсмічних спостереженнях і оцінює сейсмічну небезпеку в термінах сейсмічних зсувів, прискорень. Монте-Карло моделювання застосовується також для поповнення бази сейсмічних даних, при нестачі сейсмостатистики.

Перевагами методу є те, що він простий і не вимагає використання складного математичного апарату. Процес моделювання можна зупинити на будь-якому етапі для перевірки адекватності моделі. Це означає, що сейсмічна небезпека може бути оцінена різними способами, можна показати які землетруси викликають найбільші рухи ґрунту і, переглядаючи результати, можна побачити, як відбуваються рухи згідно прийнятої моделі. Багато методів оцінки сейсмічної небезпеки подібні на зразок чорного ящика, а Монте-Карло моделювання є відкритим і контрольованим.

Загалом, модель – це просто множина, яка забезпечена правилом пов'язування елементів, що містяться в ній. Сейсмологія є, здебільшого, емпіричною наукою, так як твердження, як і моделі ґрунтуються на фактах і спостереженнях. Метою емпіричних наук є охоплення абстрактною моделлю результатів спостережень реального світу. У практичних застосуваннях кращою вважається модель, яка є найбільш загальним і спрощеним описом спостережень і має високий потенціал прогнозу. Так як неможливо охопити моделлю всю множину спостережень, то використовуються методи теорії ймовірностей для того, щоб висновки про загальні реальні взаємозв'язки у всій сукупності спостережень були зроблені з деякою (найбільшою) ймовірністю.

Метою стохастичного моделювання є розпізнавання істинної структури системи з спостережених даних. Подання картини майбутньої поведінки і визначення майбутньої форми моделі є головним завданням емпіричних наук. Основне завдання моделювання полягає в розробці і виборі математичної моделі,

яка зможе максимально точно описати процес. Залежно від початкових умов, можемо мати кілька «найкращих» моделей. Теорія ймовірностей у своєму розпорядженні оперує потужними статистичними критеріями вибору оптимальної моделі, прикладами яких є інформаційні критерії Акаїке і Байеса. Підхід до вибору ймовірнісної моделі, в залежності від цілей, можна розбити на 2 класи: структурний підхід, де визначається структура даних, і теоретичний підхід, який визначає прогнозний потенціал моделі. У методі, який застосовується для обґрунтування теоретичної моделі, може міститися інформація і про структуру моделі. Зворотне не завжди має місце, тобто структурну модель не завжди можна застосувати для прогнозу [Buhlmann, 2000]. Образно кажучи, модель - це «окуляри», через які проглядається система.

1.4. Системний аналіз

Система – це множина взаємопов'язаних об'єктів, але вона є чимось більшим, ніж об'єднання окремих складових елементів [Волкова, Денисов, 2010; Спіцнадель, 2000; Шамис и др., 2013; Gruber, 2012; Imboden, Koch, 2003; Ulrich, Probst, 1991]. Тобто система має властивості, не завжди притаманні кожному елементу окремо [Gruber, 2012]. У тих випадках, коли системи ускладнюються, стає важче визначити залежність між елементами і параметри цієї залежності. Для вирішення цієї проблеми застосовується системний аналіз. Під поняттям «системний аналіз» мається на увазі множина методів і уявлень про взаємозв'язок елементів. При застосуванні системного аналізу визначаються обмеження на систему і елементи, розташовані всередині і поза системою. Далі складається математична модель системи, реакція системних параметрів на відносини у зовнішніх змінних в статичної і в динамічної моделей. Встановлюються внутрішні і зовнішні взаємозв'язки елементів системи.

Внутрішні взаємозв'язки: взаємодія між елементами в межах системи.

Зовнішні взаємозв'язки: вплив елементів, розташованих поза системою не внутрішні елементи системи.

Математичні моделі системи: утворення «чорного ящика», визначення обмежень системи, визначення системних змінних (число параметрів моделі), встановлення внутрішніх і зовнішніх співвідношень елементів системи. Характеристика моделей:

- 1) у дискретної за часом моделі значення змінних системи скачкоподібно змінюються в дискретні моменти часу
- 2) неперервна в просторі модель – значення змінних змінюються в просторі і в часі неперервно. Наприклад, сейсмічний процес
- 3) стохастична модель – зміна значень системної змінної визначеної детерміністськими математичними рівняннями, а відбувається випадково при різних невизначеностях в елементах і в визначеннях співвідношень.

Прикладом системи в сейсмостійкому будівництві є будівля. Тут внутрішніми змінними є: фундамент споруди, параметри несучих конструкцій, співвідношення висоти поверхів до товщини стін; конструктивні міри щодо забезпечення сейсмостійкості. В якості зовнішніх змінних розглядаються: літологічний склад, фізико-механічні властивості ґрунтів, гідрогеологічні умови, рельєф, фізико-геологічні процеси, наявність тектонічних розривів, ампліудно-частотний склад сейсмічних і інших коливань і т.д. В системі ґрунт-споруда, споруда являє собою «чорний ящик».

Сейсмічна зона також являє собою систему. Внутрішніми змінними системи є характеристики землетрусів: координати гіпоцентру; час виникнення і магнітуда. Взаємозалежність (для вогнищевих зон, де вона є) між магнітудою і глибиною вогнища; залежність магнітуди від геометричних розмірів вогнища; інтервали часу між послідовними землетрусами і т.д. є статистичними зв'язками внутрішніх змінних системи. Інтенсивність тектонічного процесу, розміри об'єму земної кори, фізико-хімічні характеристики середовища на різних глибинах, хвилі приливних деформацій, швидкість обертання Землі і т.д. є зовнішніми для системи змінними. Ясно, що зовнішні змінні впливають на внутрішні змінні. Наприклад, виявлено

статистично значущий зв'язок середньої сили між щомісячним числом землетрусів в Румунії і числом сонячних плям [Буртиєв, 2011].

1.5. Сейсмічність Республіки Молдова

Сейсмічність – схильність Землі або окремих територій до землетрусів. Землетрус є відбиток геологічних процесів, що відбуваються в земних надрах, і відбуваються в різних за структурою і характером розвитку регіонах Землі. Сейсмічність залежить від інтенсивності тектонічного процесу; від розмірів блоку земної кори, в якому можуть накопичуватись напруження; від фізико-хімічних умов на різних глибинах; від межі міцності і межі плинності гірських порід на цих глибинах і особливостей геологічної будови території.

Епіцентри землетрусів мають тенденцію групуватися в просторі. Це явище використовується для географічного розподілу територій на сейсмічні зони. Карти сейсмічного районування (СР) служать для здійснення сейсмостійкого будівництва, забезпечення безпеки населення, охорони навколишнього середовища та інших мір, спрямованих на зменшення шкоди при сильних землетрусах. Сейсмічне районування, в залежності від завдань і масштабу досліджень, ділиться на три категорії:

- загальне сейсмічне районування (ЗСР), що представляє оцінку сейсмічної небезпеки на території всієї країни для планування соціально-економічного розвитку та інфраструктури великих регіонів; масштаб карт ЗСР 1: 2. 500.000 - 1: 8. 000.000
- детальне сейсмічне районування (ДСР), служить для визначення можливих сейсмічних впливів на діючі та заплановані споруди, території населених пунктів і окремих районів; масштаб карт ДСР 1: 500 000 і крупніше
- сейсмічне мікрорайонування (СМР) оцінює вплив ґрунтових умов на сейсмічні коливання в межах будівельних площ і на території населених пунктів; масштаб карт СМР 1:50 000 і крупніше.

Сейсмічна небезпека для території республіки Молдова головним чином визначається сейсмічною зоною "Вранча". Перша карта сейсмічної небезпеки Молдови, заснована на сейсмологічних і геологічних даних, була складена Інститутом Фізики Землі Академії наук СРСР в 1957. Південно-західна частина Молдови була віднесена до зони сейсмічної інтенсивності VIII балів, решта території знаходилася в VII бальній зоні. Згідно з картою СР-69 (Сейсмічне районування території СРСР, 1969), були зроблені поправки в попередній карті сейсмічного районування для Північної і Північно-східної території Молдови, де рівень потенційних сейсмічних впливів був знижений до 6 балів. Карта СР-78 сейсмічного районування території республіки Молдови, яка включає ізолінії сейсмічної інтенсивності 6, 7 і 8 ступенів відповідно до шкали *MSK-64* [Алказ, 2007; Alcaz та ін., 2005], все ще є нормативним документом в Республіці Молдова.

Ймовірнісні карти сейсмічної небезпеки виявилися корисні при сейсмостійкому будівництві [Borman, 2011; Burtiev, 1986, 2012; Gavrilesco, Manta, 2011; Ismail-Zadeh, et al., 2007; Petrova, 2011; Rotaru, Kolev, 2010; Sokolov, Ismail-Zadeh, 2015, 2016]. Вирішенню цієї проблеми присвячено багато робіт і були запропоновані різні оцінки мір сейсмічної небезпеки: сейсмічне струшування; економічний збиток від сейсмічних впливів; ймовірність перевищення порогу інтенсивності деякого значення; період повторення максимальних сейсмічних подій і т.д. Методи оцінки сейсмічної небезпеки постійно вдосконалювалися, оновлювалися карти СР, які ставали все якіснішими. Однак, існуючі оцінки сейсмічної небезпеки і карти СР все ж недостатньо інформативні для оптимального планування і будівництва в сейсмонебезпечних регіонах. Оцінка сейсмічної інтенсивності проводиться за макросейсмічними даними, вираженими в цілочисельних балах шкали сейсмічної інтенсивності, що охоплюють широкий діапазон сейсмічних впливів при сильних землетрусах. Такий підхід призводить до зниження якості карт сейсмічного районування, оскільки передбачені значення сейсмічного ефекту виявляються завищеними або заниженими. Наприклад, в картах СР-2012 замість однієї карти пропонується створити комплект динамічно змінюваних в часі карт, в яких

проявилися деякі технічні ситуації, що ускладнили будівельникам підхід до забезпечення сейсмостійкості будівель і споруд.

Недоліком методу, який обговорюється фахівцями [Алказ, 2007; Гусев, Шумилина, 1995; Уломов, 1993; Уломов, Богданов, 2013], є те, що за широкими інтервалами в градації шкали бальності, ділянки одних і тих же територій часто мали одне і те ж значення бала на різних картах комплекту ЗСР-97. Професор Уломов зазначив: «При створенні комплекту карт ЗСР-2012, зберігається методологія ЗСР-97 для ідентифікації ЛДФ-моделі зон ВВЗ і ймовірнісний підхід до оцінки сейсмічної небезпеки» [Уломов, Богданов, 2013]. У новому комплекті карт загального сейсмічного районування ступінь сейсмічної небезпеки по картах ЗСР-2012 відповідає можливостям 39% (карта *A*), 10% (карта *B*), 5% (карта *C*), 2% (карта *D*), 1% (карта *E*) і 0,5% (карта *F*) можливого перевищення розрахункових максимальних значень інтенсивності, що вказані на відповідних картах, протягом 50-річних інтервалів часу. Для визначення порогу максимальних значень характеристик зсувів ґрунту, який буде перевищено з ймовірністю $q = 0.1$, потрібно знати вірогідну структуру параметрів зсувів, далі необхідно обчислити ймовірність перевищення цього порогу хоча б один раз в 50 років, що також передбачає знання розподілу перевищень в часі.

Загасання інтенсивності сейсмічних поштовхів, як функції епіцентральної відстані, є важливим фактором при оцінці сейсмічної небезпеки. Значення коефіцієнтів не завжди задовольняють реальній картині загасання сейсмічних коливань, так як, дуже часто фактична інтенсивність виявляється істотно вище розрахункової. Отже, розрахунок інтенсивності струсів з використанням вищевказаних коефіцієнтів може дати занижені результати, що в свою чергу позначиться на оцінці сейсмічної небезпеки. У зв'язку з цим була зроблена спроба обчислити нові значення коефіцієнтів, що відповідають реальним макросейсмічним проявам і заново розрахувати значення інтенсивності струшувань.

В якості запобіжної сейсмічної небезпеки пропонується ймовірність $P_{(\varphi, \psi)}(t, n, m, I^k)$ того, що в деякій точці земної поверхні, з координатами (φ, ψ) ,

протягом часу t можуть виникати n сейсмічних поштовхів, і m з них з інтенсивністю рівної I^k . Ймовірність $P_{(\varphi, \psi)}(t, n, m, I^k)$ є найбільш загальною мірою сейсмічної небезпеки, так як всі існуючі оцінки розрахункової сейсмічної небезпеки можна дістати з запропонованої міри [Burtiev, 1986, 2012].

Для здійснення дослідження використовуються макросейсмічні дані землетрусів, що сталися 10.11.1940, 7.04.1977, 31.08.1986, 30.05.1990 та 31.05.1990 рр. Досить великий обсяг макросейсмічних спостережень по всьому азимуту із землетрусів 1977 і 1986 рр., 4088 і 2119 точок відліку, дозволяє провести детальний аналіз макросейсмічної поля.

Територія Республіки Молдова знаходиться в радіусі дії землетрусів Карпатської зони, домінуючим джерелом землетрусів якої є вогнище Вранча. Зона Вранча – унікальна сейсмоактивна зона на Європейському континенті, розташована на стику Південних і Східних Карпат, Румунія [Ismail-Zadeh, et al., 2007, 2012]. Вогнища землетрусів концентровані на території 60×80 км² в межах кори, а також у верхній мантії на глибинах 80 – 200 км. Найбільшу небезпеку становлять землетруси, що виникають на великих глибинах, які здатні викликати на території Молдови сейсмічні струшування до 8 балів (за шкалою *MSK-64*), і горизонтальне прискорення, що перевищує або близько $0,2g$ [Алказ, 2007]. Масив Вранча – плита Східноєвропейської платформи, яка відкололася, можливо, в результаті тиску, що чиниться в північно-західному напрямку Чорноморською плитою. Гідростатичні сили сприяють зануренню плити, а в'язкість і сили тертя цьому перешкоджають. На проміжних глибинах виникають пружні напруги геофізичного середовища і енергія, накопичена внаслідок деформації, раптово вивільняючись, викликає землетруси [Balan, 1982].

Земна кора на території Молдови має складну будову, серією тектонічних розломів вона роздроблена на різні за розмірами геологічні структури (блоки). Багато блоків під впливом ендегенних процесів відчувають підняття і опускання. Ці рухи земної кори іноді провокують окремі слабкі землетруси на території Молдови. Тектонічні рухи в Карпатах ще тривають, про що свідчать геодезичні спостереження

– гірські Карпати продовжують підніматися і зсуваються в північно-східному напрямку зі швидкістю декількох сантиметрів на рік [Balan, 1982]. Про активні процеси у верхній мантії Землі під Карпатами свідчать останні землетруси, що сталися в 1977, 1986, 1990 роках в горах Вранча.

Функція загасання інтенсивності сейсмічних впливів є одним з факторів, що визначають оцінку сейсмічної небезпеки. При аналізі сейсмічної небезпеки і при побудові карт сейсмічного районування Республіки Молдова були випробувані різні модифікації функції загасання. Однак більшість варіантів функції загасання дають завищені значення розрахункової інтенсивності в балах, відповідно нереальними є засновані на них карти сейсмічної небезпеки. Існуюча неадекватність в картах сейсмічної небезпеки, обумовлена відмінностями застосованих функцій загасання, спонукала автора до проведення власного аналізу загасання інтенсивності сейсмічних впливів.

Загасання сейсмічних коливань характеризується: геометрією вогнища, фізичними властивостями середовища, типами ґрунтів, особливостями геологічної будови середовища і т.д. Особливість сейсмічного впливу кожного землетрусу визначається такими його характеристиками, як тектоніка, глибина вогнища, механізм, напрямок і хід процесу розриву гірських порід. Ці деталі стають відомі після того, як стався землетрус, за результатами аналізу отриманої сейсмологами інформації. Охоплення математичною моделлю всіх перерахованих факторів не представляється можливим. Крім того, рівняння, яке пов'язує інтенсивність впливів з параметрами вогнища землетрусів і з характеристиками середовища не є алгебраїчним. Отже, для дослідження макросейсмічного поля і побудови рівняння загасання інтенсивності сейсмічних впливів слід застосувати методи обчислювальної математики.

Перша карта сейсмічної небезпеки Молдови, заснована на сейсмологічних і геологічних даних, була складена Інститутом Фізики Землі Академії наук СРСР в 1957. Південно-західна частина Молдови була віднесена до зони сейсмічної інтенсивності VIII балів, решта території знаходилася в VII бальній зоні. Згідно з

картою СР-69 (сейсмічне районування території СРСР, 1969), були зроблені поправки в попередній карті сейсмічного районування для північної і північно-східної частини території Молдови, де рівень потенційних сейсмічних впливів був знижений до 6 балів. Оновлена (2009р.) Карта СР-78 сейсмічного районування території Республіки Молдова, яка включає ізолінії сейсмічної інтенсивності 6, 7 і 8 балів відповідно до шкали *MSK-64* [Алказ, 2007], з періодом повторення 100 років, все ще є нормативним документом в Республіці Молдова. Карта СР-78 є детерміністською і не враховує основні характеристики сейсмічного режиму сейсмоактивних територій. Самі карти недостатньо інформативні для проектування і проведення сейсмостійкого будівництва. Очевидно, що проектувальники при прийнятті рішень повинні володіти найповнішою інформацією про потенційну сейсмічну небезпеку, а карти у формі ізоліній інтенсивності різної бальності, з періодом повторення 100 і більше років, малоінформативні. Більш диференційовану інформацію надають карти ЗСР-97, які базуються на моделях вогнища і на ймовірності неперевищення певного порогу із заданою вірогідністю протягом певного інтервалу часу.

Висновки до розділу 1

Проаналізовано основні результати робіт в області оцінки сейсмічної небезпеки, обгрунтовано актуальність розвитку в даній роботі напрямків. Зроблено огляд літературних джерел, присвячених теорії виникнення землетрусів і їх зв'язку з аномаліями фізичних полів Землі. Пропонується формальний підхід до вирішення проблем геофізики для пошуку методів прогнозу землетрусів. Передбачається, що аномалії фізичних полів Землі, провісники і сейсмічні події є елементами деякої сукупності чи множини. Якщо в даній множині додати ще взаємозв'язок між елементами, то виходить модель (система). Для передбачення часу виникнення землетрусів необхідно поповнити модель новими елементами і новими зв'язками між ними.

Оцінка сейсмічної небезпеки, яка носить прогнозований характер, залежить від декількох параметрів вогнища землетрусу. Це географічні координати епіцентру, глибина вогнища, магнітуда і функція загасання інтенсивності сейсмічного ефекту з відстанню. У ймовірнісному підході до аналізу сейсмічної небезпеки всі параметри землетрусу представляються випадковими величинами, для визначення ймовірнісної структури яких використовуються відомості з каталогів землетрусів. Концепції ДАСН і ЙАСН не повинні протиставлятися, їх потенціал повинен бути використаний для вирішення загального завдання: удосконалення методів розрахунку сейсмічної небезпеки і підвищення якості карт СР. Адаптація висотних і надвисотних конструкцій, де застосовуються нові технології сейсмостійкого будівництва, пред'являє високі вимоги до якості та інформаційної наповненості міри сейсмічної небезпеки.

Список використаних джерел

- Алказ В.Г. Основы прогноза сейсмической опасности и сейсмического риска территории Республики Молдова. Кишинев, 2007, 229.
- Беврани Х., Аничкин К. МКО. 2005, ч. 2, 493 – 495.
- Буртиев Р.З. Оценка максимальной магнитуды землетрясений очага Вранча по статистике экстремальных значений. *Analele ATIC. Vol. I(II). Chisinau. 2003, 110–117.*
- Буртиев Р.З. Статистическая связь между коровой и подкоровой сейсмическими активностями очага “Вранча”. *Buletinul Institutului de Geologie si seismologie al ASM. Nr.2. Chisinau, 2011, 133–148.*
- Burtiev R.Z. Seismic Hazard Mapping Based on Multi-dimensional Markov Model of a Seismic Regime // *Gerlands Beitrage zur Geophysik, Leipzig, 95, 1986, 499– 508.*
- Burtiev Rashid. Evaluation of seismic hazards from several seismic zones. *Environmental Engineering and Management Journal. 12, 2012. 32–42.*
- Волкова В.Н., Денисов А.А. Теория систем и системный анализ. Москва, Юрайт, 2010, 241.

- Гусев А.А., Шумилина Л.С. Некоторые вопросы методики общего сейсмического районирования // Сейсмичность и сейсмическое районирование Северной Евразии. Отв. ред. В.И. Уломов, Москва, ОИФЗ РАН, 1995. Вып. 2–3, 289–300.
- Онофреш Н.И. Статистические методы картирования сейсмической опасности. Кишинев, Штиинца, 1981, 109.
- Писаренко В. Ф., Родкин М. В., Рукавишникова Т. А., Оценка вероятности сильнейших сейсмических катастроф на основе теории экстремальных значений. Физика Земли. Наука, 2014, 3–17.
- Родкин М. В., Теория экстремальных значений, распределения вероятностей для хвоста распределения для основных типов распределений, вопросы применения. Международный институт теории прогноза землетрясений и математической геофизики РАН, 2005, 27.
- Ризниченко Ю.В. От активности очагов землетрясений к сотрясаемости земной поверхности. Изв. АН СССР, Физика Земли, 1965, № II, 1–12.
- Ризниченко Ю.В. Расчет сотрясаемости точек земной поверхности от землетрясений в окружающей области. Изв. АН СССР, Физика Земли, 1966, №5, 16–32
- Спицнадель В.Н. Основы системного анализа: Учеб. пособие. Москва, Бизнес–пресса, 2000, 137.
- Уломов В.И., Сейсмичность и сейсмическое районирование Северной Евразии. Отв. Ред. Вып. 1, Москва, ИФЗ РАН, 1993, 303.
- Уломов В.И., Богданов М.И. Новый комплект карт общего сейсмического районирования территории Российской Федерации (ОСР-2012). Инженерные изыскания, 8, 2013. 8-17.
- Шамис Е.Е., Присяжнюк М.И., Цуркану Н.Г., В.Д. Иванов. Системный анализ проблем как практический метод решения производственных задач промстройматериалов. Buletinul INCERCOM, Nr. 4, 2013, 63–70.
- Шепель В.Н., Акимов С.С. Использование оценки хилла для различения законов распределения вероятности вестник ОГУ №1. 2014, 75–78.

- Шумила В.И. Статистические модели интерпретации поля сейсмической интенсивности. Кишинев. Штиинца. 1985. 128.
- Abrahamson, N A, and Bommer, J J. Probability and uncertainty in seismic hazard analysis. *Earthquake Spectra*, Vol. 21, 2005, 603–608.
- Algermissen S. and D. Perkins: A probabilistic estimate of maximum acceleration in rock in the contiguous United Staates, U.S. Geol. Surv. Open–File Rep. 76–416, Reston, Va., 1976, 45.
- Balan Ștefan, Cristescu Valeriu, Cornea Ion, Cutremurul de pământ din Romania de la 4 martie 977. Editura Academiei Republicii Socialiste România. București, 1982, 513.
- Baker W. Jack. *An Introduction to Probabilistic Seismic Hazard Analysis (PSHA)*, 2008, 1-72.
- Borman P. From earthquake prediction, research to time–variable seismic hazard assessment applications, *Pure and Applied Geophysics*, 168, 2011, 329–366.
- Buhlmann Peter, *Von Daten zu Stochastischen Modellen. Elemente der Mathematik*. Birkhauser Verlag, Basel, 2000. 18.
- Burtiev R.Z. Seismic Hazard Mapping Based on Multi–dimensional Markov Model of a Seismic Regime // *Gerlands Beitrage zur Geophysik*, Leipzig, 95, 1986, 499– 508.
- Burtiev Rashid. Evaluation of seismic hazards from several seismic zones. *Environmental Engineering and Management Journal*. 12, 2012. 32–42.
- Burton, P. W. Seismic risk in Southern Europe through India examined using Gumbel’s third distribution of extreme values, *Geophysical Journal of the Royal Astronomical Society*. Vol. 59, 1979, 249–280.
- Cornell, C.A., *Engineering Seismic Risk Analysis*, *Bull. Seism. Soc. Am.*, Vol. 58, N. 5, 1968, 1583–1606.
- Cox, M.G; Siebert, B. R.: The use of a Monte Carlo method for evaluating uncertainty and expanded uncertainty. *Metrologia*, Vol. 43, 2006, 178–188
- Cramer, C.H., Petersen, M. D, and Reichle, M. S., *A Monte Carlo Approach in Estimating Uncertainty for a Seismic Hazard Assessment of Los Angeles, Ventura, and Orange*

- Counties, California, Bulletin of the Seismological Society of America, Vol. 86, 1996, 1681–1691.
- Ebel, J. E. and Kafka, A. L. A Monte Carlo approach to seismic hazard analysis, Bulletin of the Seismological Society of America, Vol. 89, 1999, 854–866.
- Garcia–Fernandez M., Jimenez M. J., Kijko A., Seismic hazard parameters estimation in Spain from historical and instrumental catalogues. Tectonophys, Warszawa, Inst. Electrical & Electronics Engineers, vol. 167, No. 9, 1989, 245–251.
- Gavrilesco M., Manta V. Advances in the visualization of three–dimensional seismic volume data, Environmental Engineering and Management Journal, 10, 2011, 567–578.
- Golbs Christian, Probabilistische seismische Gefährdungsanalysen auf der Grundlage von Epizentrendichten und ihre ingenieurpraktischen Anwendungsgebiete, Dissertation zur Erlangung des akademischen Grades, Weimar, Deutschland, 2008, 193.
- Grieser Jürgen, Staeger Tim, Schönwiese Christian, Zur Schätzung der Wiederkehrzeiten von Starkniederschlägen. Institut für Atmosphäre und Umwelt der J. W. Goethe – Universität Frankfurt/M. 2009, 35.
- Gruber Nicolas. Systemanalyse. Environmental Physics, 2012. 43.
- Grünthal G., Bosse G., Musson C., Gariel R., Crook J. C., Verbeiren T., Camelbeeck R., Imboden D.M., Koch S., Systemanalyse. Springer–Verlag Heidelberg GmbH&Co.KG, 2003, 255.
- Ismail-Zadeh, A., Sokolov, V. and Bonjer, K., Tectonic stress, seismicity, and seismic hazard in the southeastern Carpathians, Nat. Hazards, Springer, 2007, 1-22.
- Ismail-Zadeh, A., Matenco, L., Radulian, M., Cloetingh, S., and Panza, G., Geodynamic and intermediate-depth seismicity in Vrancea (the south-eastern Carpathians): Current State-of-the-Art, Tectonophysics, 50-79, 2012 (http://www.mitp.ru/~aismail/papers/Vrancea_Tectonophysics.pdf)
- Karník, V., Schenkova, Z. and V. Schenk: Seismic hazard estimates for low seismicity region – example of Bohemia, PAGEOPH 119, 1981, 1077–1092.

- Kijko, A, and Sellevoll, M A. Estimation of earthquake hazard parameters from incomplete data files : Part 1 – Utilisation of extreme and complete catalogues with different threshold magnitudes, *Bulletin of the Seismological Society of America*, Vol. 79, 1989, 645–654.
- Knopoff, L, and Kagan, Y Y. Analysis of the theory of extreme values as applied to earthquake problems, *Journal of Geophysical Research*, Vol. 82, 1977, 5647–5657.
- Krinitzsky, E L. How to combine deterministic and probabilistic methods for assessing earthquake hazards, *Engineering Geology*, Vol. 70, 2003, 157–163.
- Lapajne, J K, Sket Motnikar, B, Zabukovec, B and Zupancic, P. Spatially– smoothed seismicity modeling of seismic hazard in Slovenia, *Journal of Seismology*, Vol. 1, 1997, 73–85.
- Lapajne, J. Some features of the spatially smoothed seismicity approach, *Seismicity modelling in seismic hazard mapping*. Lapajne, J, and Idrih, R (editors). (Ljubljana: Slovenian Geophysical Survey.), 2000, 27–33.
- Malachow A., *Geheimnisse des Erdinneren*, Leibzig, Fachbuchverlag, 1981, 196.
- Mayer–Rosa T. and Lenhardt D., Joint seismic hazard assessment for the central and western part of GSHAP Region 3 (Central and Northwest Europe). *Seismology in Europe*. B. Thorkelsson, (editor) (Reykjavik: Icelandic Met. Office.) W., 1996, 339–342.
- McGuire K. Robin, *Deterministic vs. Probabilistic Earthquake Hazards and risks*. Risk Engineering, Colorado, USA, 1999, 7.
- McGuire, R. K.: *Seismic Hazard and Risk Analysis*, Earthquake Engineering Research Institute, EERI Publication No. MNO–10, Oakland CA, 2004, 26–34.
- Metropolis, N.; Ulam, S.: The Monte Carlo method. *Journal of the American Statistical Association*, Vol. 44, 1949, 335–341
- Musson R.M.W., *Intensity–based seismic risk assesment*. *Soil Dynamics and Earthquake Engineering*, Elsevier, 2000, 353–360.
- Musson R. M. W., Generalised seismic hazard maps for the Pannonian Basin using probabilistic methods, *Pure Appl. Geophys.*, 157, 2000, 147–169.

- Musson RMW, Sargeant S.L., British geological survey, seismology and geomagnetism programme, technical report cr/07/125. Eurocode-8 seismic hazard zoning maps for the UK. Keyworth, Nottingham British Geological Survey, 2007, 62.
- Perelman, J.I.: Unterhaltsame Geometrie. Eine Sammlung allgemeinverständlicher geometrischer Aufgaben zur Unterhaltung und Übung. Verlag Volk und Wissen, Berlin, 1963, 118.
- Petrova E. Critical infrastructure in Russia: geographical analysis of accidents triggered by natural hazards, *Environmental Engineering and Management Journal*, 10, 2011, 53–58.
- Raschke, M., Kröger, W., Bilis, V. (2011): Erdbeben und Wintersturmrisko des Schweizer Stromübertragungsnetzes. *SEV/VSE-Bulletin*, 2011, 5.
- Reiter, L. *Earthquake Hazard Analysis*, Columbia University Press, New York, 1990, 252.
- Rosenhauer W.: Benutzungs–Anleitung für das Programm PSSAEL zur probabilistischen seismischen Standort Analyse, Bericht im Auftrag der Technischen Vereinigung der Grosskraftwerksbetreiber e.V. (VGB), 1999, 126.
- Rotaru A., Kolev C. Addressing issues of geoenvironmental risks in Dobruja, Romania/Bulgaria, *Environmental Engineering and Management Journal*, 9, 2010, 961–969.
- Shapira, A. Potential earthquake risk estimations by application of a simulation process. *Tectonophysics* 95, 1983, 75–89.
- Siebert, B.R.L.; Sommer, K.–D.: Grundlagen der Monte–Carlo–Methode für die Unsicherheitsberechnung. *VDI/VDE–GMA – Workshop ”Neue Entwicklungen in der Messunsicherheitsermittlung“*, Ilmenau, 2006, 98.
- Solomos G., Pinto A., Dimova S., A Review of The Seismic Hazard Zonation in National Building Codes in the context of Eurocode 8, European Commission Joint Research Centre, Luxembourg, 2008, 72.
- Sokolov, V., and Ismail-Zadeh, A., Seismic hazard from instrumentally recorded, historical and simulated earthquakes: Application to the Tibet-Himalayan region,

Tectonophysics, 657, 187-204, 2015 (http://www.mitp.ru/~aismail/papers/VS_AIZ_Tecto_2015.pdf).

Sokolov, V., and Ismail-Zadeh, A. On the Use of Multiple-Site Estimations in Probabilistic Seismic Hazard Assessment, Bull. Seismol. Soc. Am., 106(5), 2233–2243, 2016 (http://www.mitp.ru/~aismail/papers/2016_BSSA_Multiple-Site.pdf

Ulrich H., Probst G., Anleitung zum ganzheitlichen Denken und Handeln. Haupt, 1991. 98.

WOO, G. Kernel estimation methods for seismic hazard area source modelling. Bulletin of the Seismological Society of America, Vol. 86, 1996, 353-362.

РОЗДІЛ 2

ВИЯВЛЕННЯ СЕЙСМІЧНИХ ЗОН

Сейсмічність – схильність Землі або окремих територій впливу землетрусів. Істотний крок в дослідженні сейсмічності країн або територій – це розбиття на окремі сейсмічні зони. Спостереження показують, що вогнища землетрусів розподілені нерівномірно, мають тенденцію групуватися, утворюючи сейсмічні зони. Сейсмічні зони є відображенням географічного поділу областей з різними геолого-структурними, тектонічними і сейсмічними особливостями. Детальність і точність аналізу сейсмічної небезпеки визначається багатьма факторами, в тому числі, і сейсмічними зонами. Вони повинні відображати геолого-геофізичні особливості регіону і не повинні бути великими, інакше розподіл сейсмічності не буде однорідним. Розмір сейсмічних зон визначається ще й достатнім для аналізу сейсмічної небезпеки обсягом даних про сейсмічні події. При невеликих розмірах сейсмічних зон розподіл сейсмічності стає більш однорідним, але зменшується необхідний для якісного аналізу сейсмічної небезпеки обсяг даних. При сейсмічному зонуванні повинні бути узагальнені як геотектонічні особливості, так і регіональні характеристики сейсмічності. Сейсмічне зонування в значній мірі ґрунтується на суб'єктивних критеріях, з чим і пов'язаний фактор неоднозначності при виділенні сейсмічних зон.

Сейсмічна зона, в якій виникають землетруси певної потужності, характеризується: геолого-геофізичними ознаками, особливостями геологічного розвитку регіону (інтенсивність рухів і будова тектонічних блоків), геодезичними даними про сучасні рухи земної кори, геофізичними даними, що описують глибинну будову регіону і розподілом фізичних властивостей гірських порід на різних глибинах. Для оцінки сейсмічної небезпеки необхідно зібрати вихідну інформацію про історичні сейсмічні події, і за період інструментальних спостережень в формі каталогів землетрусів. Таким чином, при вирішенні завдання сейсмічного районування, повинні бути виділені зони вогнищ землетрусів і складений

відповідний каталог сейсмічних подій. Каталоги землетрусів є важливою продукцією сейсмології, і їх складання пов'язано з обробкою великого обсягу геолого-геофізичних і сейсмічних даних.

Просторовий розподіл землетрусів не є рівномірним, отже, на земній кулі можна виділити зони з різними особливостями сейсмічного режиму: сейсмічно активні, менш активні і асейсмічні зони. У каталогах землетрусів містяться необхідні для аналізу сейсмічної небезпеки відомості, і часто вказується ідентифікатор сейсмічної зони. Наприклад, каталог землетрусів Центральної і південно-східної Європи містить відомості по сейсмічних зонах [SHEBALIN et al. 1998], каталог Національної сейсмічної служби США NEIC містить графу номерів сейсмічних регіонів по Флінну-Енгдалу. Однак, у багатьох випадках, каталоги складені за територіальною ознакою без вказівки сейсмічних зон або ж зовсім не вказані. В такому випадку, для оперативного виділення сейсмічних зон можна застосувати формальні методи: діаграма Вороного, кластерний аналіз. Часто на картах епіцентрів землетрусів виявляються області рівної щільності епіцентрів, для виділення яких застосовується діаграма Вороного. Та обставина, що вогнища землетрусів групуються в просторі, наштовхнуло на думку про застосування кластерного аналізу для виділення компактних груп осередків землетрусів. Кластерний аналіз знаходив застосування в сейсмології, наприклад, для розподілу вогнищ землетрусів на структуровані і розсіяні використовувалися методи кластерного аналізу на основі «індексу Морішіті», або методу найближчого сусіда [Ареф'єв, 2003; Шебалин, 1968]. Для географічної ідентифікації сейсмічних зон часто використовується метод Вороного, наприклад, виявлялися географічні прямокутники, що містять епіцентри землетрусів, і далі з прямокутників з однаковим числом епіцентрів склалися сейсмічні зони [Golbs, 2008]. Кластерний аналіз знайшов застосування і в розпізнаванні місць виникнення сильних землетрусів [Gvishiani et al., 2013].

2.1. Кластерний аналіз

Перше застосування знайшов в соціології. Назва кластерний аналіз походить від англійського слова cluster – гроно, скупчення. Вперше в 1939 англійським математиком Тріона був визначений предмет кластерного аналізу і зроблено його опис [Hudec, 2003]. Головне призначення кластерного аналізу - розбивка множини досліджуваних об'єктів на однорідні у відповідному розумінні (за деякою метрикою) групи (кластери). Це означає, що вирішується завдання класифікації даних і виявлення відповідної структури в цій сукупності. Методи кластерного аналізу застосовуються в самих різних областях. Кластерний аналіз дозволяє скорочувати розмірність даних, виробляючи їх класифікацію. Тільки змінні повинні вимірюватися в порівнюваних шкалах. У кластерному аналізі об'єкти (землетрусу) деякого сукупності розбиваються на кластери за принципом схожості. Об'єкти, що належать одному кластеру, повинні бути максимально подібними, а об'єкти, що належать різним кластерам різнорідними. В даному випадку об'єктами є епіцентри землетрусів. Кластерний аналіз включає в себе більше 100 різних алгоритмів, дозволяє аналізувати показники різного типу даних, скорочувати розмірність простору даних, робити їх структуру наочною. Єдине обмеження: шкали, в яких вимірюються дані, повинні бути порівнювані. При практичних застосуваннях, при такій кількості методів кластерного аналізу, виникає питання, який же метод вибрати? Існують рекомендації фахівців щодо вибору методу [Бюль, Цефель, 2001; Кім et al. 1989; Шалымов, Граничин, 2010; Timm, 2002; Milligan, 1985; Greutert, 2004; Stein et al., 2007]:

1. генетичні алгоритми і штучні нейронні мережі добре розпаралелювати
2. генетичні алгоритми і метод гарту здійснюють глобальний пошук, але метод гарту сходиться повільно
3. генетичні алгоритми добре працюють тільки для одно і (дво)-мірних об'єктів, зате не потрібна неперервність координат
4. *k*-means швидко працює і простий в реалізації, але створює тільки кластери, схожі на гіперсфери

5. ієрархічні алгоритми дають оптимальне розбиття на кластери, але їх трудоємкість квадратична обсягу вибіркового даних
6. на практиці найкраще зарекомендували себе гібридні підходи, де кластеризація виконується методом *k-means*, а початкове розбиття - одним з сильніших методів, наприклад, ієрархічним алгоритмом.

При виборі алгоритму кластеризації потрібно виходити від обсягу даних, враховувати апіорну інформацію про структуру кластерів.

Подібність двох об'єктів визначається мірою схожості. Чим більше міра, тим більш схожими є об'єкти. Мірою несхожості об'єктів служить міра дистанції. Дистанція між об'єктами, в даній задачі, визначається відстанню між епіцентрами, які задаються географічними координатами. Залежно від способу визначення відстані можуть виходити різні групи кластерів.

Кластерний аналіз проводився автором за схемою:

1. вибір змінних і підготовка даних
2. визначення міри схожості між об'єктами
3. вибір методу кластеризації
4. оптимізація кластерного рішення і перевірка його якості.

З існуючого сімейства методів кластерного аналізу використовувалися ієрархічні агломеративні, неієрархічні ітеративні і двокрокові методи. Кластерний аналіз реалізований в пакеті SPSS (Statistical Package for social science).

В цьому розділі вирішувалися наступні завдання:

- кластеризації епіцентрів землетрусів. Перевірка якості кластерного рішення на прикладі каталогу землетрусів Балкан і Туреччини
- доповнення даного каталогу розподілами за сейсмічними зонами землетрусів, що сталися після 1990 року
- застосування дискримінантного аналізу для класифікації епіцентрів спостережених і подальших землетрусів за сейсмічними зонами.

Існує велика кількість алгоритмів кластерного аналізу, які використовують різні міри відстані і критерії об'єднання об'єктів і кластерів, передбачені в SPSS. З

урахуванням особливостей розв'язуваної задачі застосовувалися такі методи для кластеризації: обчислення середньої відстані між кластерами і метод Варда. Так як координати епіцентрів є неперервними величинами на площині, то в якості міри подібності об'єктів можна використовувати відстань Евкліда.

2.2. Ієрархічний агломеративного метод кластерного аналізу

Сейсмічність Румунії зосереджена в декількох сейсмічних зонах, які являють собою територію, де зосереджена сейсмічність, де сейсмічна активність і орієнтація поля напружень вважається відносно однорідною [Ardeleanu et al., 2005; Oncescu et al., 1999; Radulian et al., 2000. 2009 ; Ghica, 2009]. Визначення довгострокових характеристик землетрусів в межах кожної сейсмічної зони має велике значення при проведенні аналізу сейсмічної небезпеки (Pora 2009; Popescu, 2009). Число виявлених на території Румунії румунськими сейсмологами сейсмогенних зон коливається від 4 до 13. Однак, в каталозі румунських землетрусів не містяться відомості про сейсмічні зони. Визначення оптимального числа сейсмічних зон, і виявлення приналежності землетрусу до тієї, чи іншої зони можна здійснити за допомогою кластерного аналізу. Кластерний аналіз може бути корисним, як «експрес метод» для оперативного виділення сейсмічних зон.

В якості пробної вибірки використовуються дані каталогу землетрусів на Балканах за період 1900-1990pp. [Shebalin et al., 1998]. В каталозі містяться відомості про 18 сейсмічних зон, виділених експертами. Порівнюючи виділені експертами і кластерним аналізом сейсмічні зони можна скласти уявлення про потенціал кластерного аналізу для вирішення завдання сейсмічного районування.

Перед початком кластеризації всі епіцентри вважаються окремими кластерами, які об'єднуються на наступних кроках. На першому кроці пара точок, які є найближчими по деякій мірі схожості, об'єднується в один кластер. Тобто, після першого кроку маємо 1 кластер, який містить 2 точки, і $(n-2)$ кластера містять по одному елементу. На кожному кроці кількість кластерів зменшується на одиницю. Процедура повторюється, поки не залишиться один кластер (рис. 2.1). На першому

кроці дві точки z_i, z_j , що знаходяться у внутрішньому замкнутому контурі, складають перший кластер. Кластеру присвоюється номер, рівний меншому з індексів i, j . Далі знову утворений кластер об'єднується з точкою z_k , але вже за методом кластеризації.

Як приклад, кластерному аналізу піддається масив епіцентрів 3003 землетрусів з каталогу, що спостерігались за період 1900-1990 рр. на Балканах. Об'єктами є епіцентри, які задаються географічними широтою і довготою вогнищ землетрусів. Відстань між цими об'єктами визначається обраною мірою схожості.

Міри схожості залежать від типу змінної і шкали, до якої вона відноситься. Для даних, що відносяться до інтервальної шкали, існує множина мір подібності об'єктів. Найпоширенішою мірою визначення відстані між двома точками на площині $Z_1(x_1, y_1), Z_2(x_2, y_2)$ (рис. 2.2), утвореної координатними осями, є міра Евкліда ($r = 2$):

$$d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} \quad (2.1)$$

При застосуванні метрики Евкліда кластери матимуть вигляд кола [Батуркин и др., 2010 року; Корпроф, Корр 2009; Tibshirani, 2001].

Таким чином, для застосування кластерного аналізу необхідно визначитися з мірою подібності епіцентрів землетрусів і подібності кластерів.

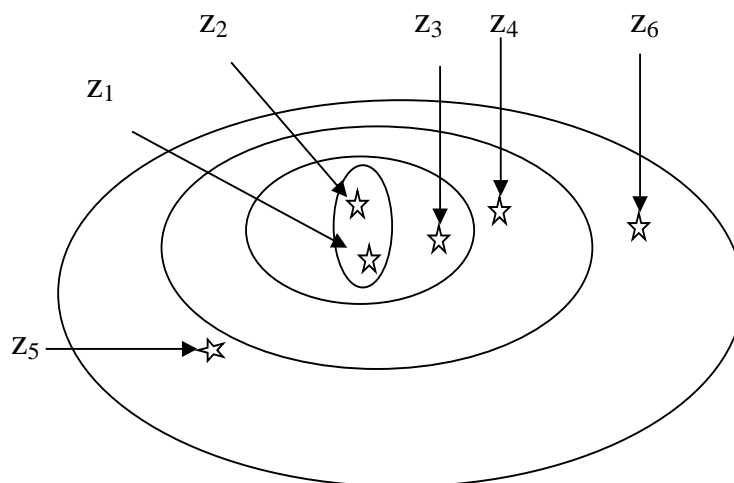


Рис. 2.1. Схема об'єднання кластерів в агломеративному методі.

Далі необхідно визначити відстань між парою кластерів, для чого застосовується метод, заснований на обчисленні середньої відстані між кластерами – середня відстань між групами (Between-groups linkage) (рис. 2.3). Відстань між двома кластерами визначається як середня відстань між усіма точками (об'єктами), що містяться в кластерах:

$$d(C_r, C_h) = \frac{1}{n_r n_h} \sum_{x_r \in C_r} \sum_{x_h \in C_h} d(x_r, x_h) = \frac{1}{n_r n_h} \sum_{i=1}^{n_r} \sum_{j=1}^{n_h} \sqrt{(x_{ri} - x_{hj})^2 + (y_{ri} - y_{hj})^2} \quad (2.2)$$

де n_r, n_h - число точок в кластерах C_r, C_h , відповідно.

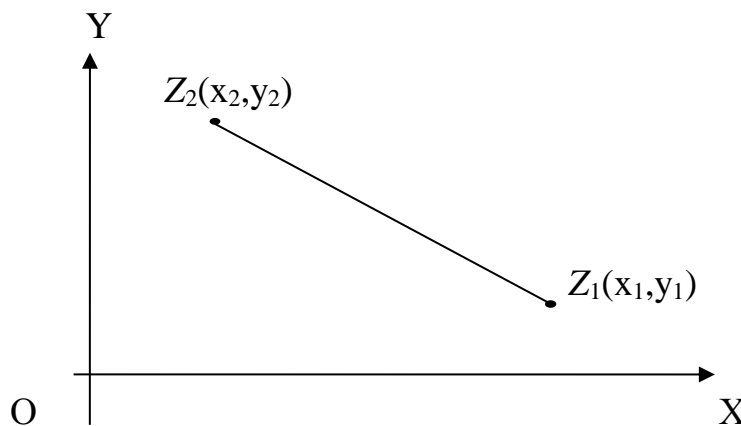


Рис. 2.2. Відстань Евкліда

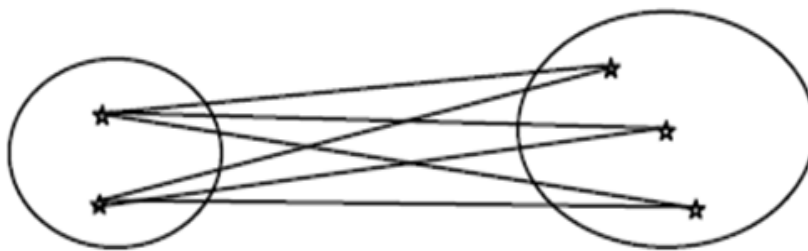


Рис. 2.3. Середні зв'язку між групами

В таблиці “Порядок агломерации” (табл. 2.1) пакета SPSS можна простежити за процедурою кластеризації вогнищ землетрусів. Так, в колонці “Об'єднання в кластери” можна бачити послідовність об'єднання 2-х епіцентрів в кластери (табл. 2.1). Наприклад, якщо на першому кроці епіцентри с номерами 1926 и 2974 утворюють кластер під номером 1926, то кластер 2974 в таблиці більше не

появляється. На слідуючому кроці проводиться об'єднання кластерів 2899 и 2900, далі 2836 и 2879 и т.д.

Цей підхід визначає оптимальне розбиття на кластери, проте визначення числа кластерів суб'єктивно. Скільки кластерів залишити?

Таблиця 2.1

Схема об'єднання кластерів в агломеративному методі.

Порядок агломерації - середня відстань між кластерами						
	Об'єднання в кластери		Коефіцієнти	Крок, на якому кластер з'являється вперше		
	кластер 1	кластер 2		кластер 1	кластер 2	
1	1926	2974	0.000	0	0	1344
2	2899	2900	0.000	0	0	1719
3	2836	2879	0.000	0	0	10
...

Продовження таблиці

крок	кластер 1	кластер 2	коефіцієнти	кластер 1	кластер 2	Наступний крок
2981	517	2553	2.006	2960	2851	2989
2982	19	21	2.028	2969	2971	2993
2983	17	32	2.112	2964	2967	2992
2984	569	690	2.245	2931	0	2999
2985	11	1698	2.311	2976	2979	2990
2986	18	57	2.465	2972	2978	2995
2987	88	175	2.499	2952	2959	2994
2988	2	1074	2.528	2920	2946	2999
2989	34	517	2.690	2927	2981	2995
2990	11	15	2.891	2985	2980	2997
2991	60	562	2.939	2974	2970	3000
2992	1	17	3.003	2975	2983	2998
2993	19	35	3.180	2982	2958	2996
2994	88	423	3.440	2987	2977	2997

2995	18	34	3.784	2986	2989	2996
2996	18	19	4.309	2995	2993	2998
2997	11	88	4.479	2990	2994	3000
2998	1	18	5.155	2992	2996	3001
2999	2	569	5.239	2988	2984	3002
3000	11	60	5.715	2997	2991	3001
3001	1	11	7.053	2998	3000	3002
3002	1	2	10.758	3001	2999	0

Там, де відстань між кластерами (колонка "Коефіцієнти"), визначена на підставі обраної міри схожості збільшується скачкоподібно, процес необхідно зупинити, тому що в протилежному випадку будуть об'єднані кластери, що знаходяться дуже далеко один від одного. У колонці "Коефіцієнти" вказані відстані між двома відповідними кластерами, визначені на підставі обраної міри подібності. Число кластерів $K=3002+1-2985=18$ визначалося різницею числа спостережень 3002 плюс одиниця і кроку 2985, на якому відбувається скачок значення відстані (коефіцієнта) (табл. 2.1) від значення 2.311 до 2.465 (табл. 2.1).

2.3. Метод Варда (Ward-Method)

При використанні даного методу рекомендується застосовувати метод «Зв'язок між групами» (Between-groups linkage). У методі Варда, як відстані між кластерами розглядається приріст суми квадратів відстаней епіцентрів до центрів кластерів, в результаті їх об'єднання. Відстань між двома кластерами визначається як середня відстань між усіма епіцентрами, що містяться в кластерах:

$$d(C_r, C_h) = \frac{n_r n_h}{n_r + n_h} d^2(\bar{Z}_r, \bar{Z}_h) = \left\| \bar{Z}_r - \bar{Z}_h \right\|_2 \quad (2.3)$$

$$\left\| \bar{Z}_r - \bar{Z}_h \right\|_2 = \sqrt{(\bar{x}_r - \bar{x}_h)^2 + (\bar{y}_r - \bar{y}_h)^2} \quad (2.4)$$

де n_r, n_h – кількість епіцентрів в кластерах C_r і C_h , відповідно. Центри кластерів визначаються за формулою:

$$\bar{Z}_r = \frac{1}{n_r} \sum_{z \in C_r} z; \quad \bar{Z}_h = \frac{1}{n_h} \sum_{z \in C_h} z \quad (2.5)$$

Для оцінки відстаней між кластерами використовуються методи дисперсійного аналізу. На кожному кроці алгоритму об'єднуються такі два кластери, які призводять до мінімального збільшення цільової функції, тобто внутрішньо групової суми квадратів. Процедура кластеризації епіцентрів землетрусів така ж, як в агломеративного методі (табл. 2.2). Число кластерів $K=3002+1-2986=17$ визначалося різницею числа спостережень 3002 плюс одиниця і кроку 2986, на якому відбувається скачок значення відстані – коефіцієнта – від значення 1444.512 до 8238.520 (табл. 2.2).

Таблиця 2.2

Сума квадратів похибок (метод Варда)

Крок	Об'єднання в кластери	коефіцієнти	Крок, на якому кластер з'являється вперше	Наступний крок	Крок	Об'єднання в кластери
1	1891	2964	0.000	0	0	1527
2	2905	2906	0.000	0	0	1981
3	2599	2886	0.000	0	0	942
...

2981	34	220	1208.346	2943	2970	2989
2982	19	32	1246.265	2968	2960	2992
2983	35	38	1288.220	2969	2963	2995
2984	18	57	1336.663	2951	2975	2989
2985	15	175	1388.987	2977	2948	2987
2986	17	39	1444.512	2957	2974	2996
2987	15	423	1505.201	2985	2964	2997
2988	21	23	1567.479	2962	2978	2992
2989	18	34	1635.476	2984	2981	2995
2990	86	88	1720.542	2980	2967	2996
2991	275	341	1806.602	2976	2979	2994
2992	19	21	1901.989	2982	2988	3000

2993	11	60	2042.227	2972	2956	2997
2994	1	275	2187.528	2973	2991	2999
2995	18	35	2455.518	2989	2983	2998
2996	17	86	2734.799	2986	2990	2999
2997	11	15	3065.992	2993	2987	3001
2998	2	18	3468.335	2952	2995	3000
2999	1	17	3924.612	2994	2996	3001
3000	2	19	4517.923	2998	2992	3002
3001	1	11	5767.875	2999	2997	3002
3002	1	2	8238.520	3001	3000	0

Число кластерів, визначене за агломеративного методу об'єднання кластерів, майже збігається з числом, знайденим по методу Варда. Даний спосіб визначення числа кластерів є суб'єктивним, він використовується для прийняття оперативного рішення про кількість кластерів. Кластерний аналіз був застосований також для виділення сейсмічних зон Туреччини. Каталог землетрусів Туреччини містить відомості про 9200 землетрусів, що відбулися в період 1900–2010 рр.

2.4. Алгоритм k -середніх

При великій кількості спостережень ієрархічні методи кластерного аналізу не придатні, оскільки розрахунковий час зростає пропорційно третьому степеню числа спостережень. У таких випадках використовуються неієрархічні методи, засновані на поділі даних, які представляють собою ітеративні методи дроблення вихідної сукупності. У процесі поділу нові кластери формуються до тих пір, поки не буде виконано правило зупинки процесу кластеризації (об'єднання даних). Неієрархічні методи засновані на оптимізації деякої цільової функції, яка визначає, за деяким критерієм, розбиття множини епіцентрів (спостережень) на кластери. Алгоритм k -середніх в якості цільової функції використовує суму квадратів відхилень координат епіцентрів від центрів відповідних кластерів. Алгоритм чутливий до вибору початкових наближень (центрів кластерів).

Алгоритм k -середніх буде k кластерів, розташованих на максимально

можливих віддалях один від одного. Основний тип задач, які вирішує алгоритм k -середніх, які дозволяють зробити припущення щодо числа кластерів, при цьому вони повинні бути різні настільки, наскільки це можливо. Вибір числа кластерів є складним питанням. Якщо немає припущень щодо цього числа, рекомендується створити 2 кластера, потім 3, 4, 5 і т.д., і знайти оптимальне кластерне рішення, порівнюючи значення цільових функцій.

Випадковий вибір центрів на першому кроці може призводити до похибкової кластеризації. На практиці для формування початкового наближення виділяються k найбільш віддалених точок вибірки. Перші дві точки визначаються по максимуму всіх парних відстаней, кожна наступна точка вибирається так, щоб відстань від неї до найближчої, вже виділеної точки, було максимальною.

Загальна ідея алгоритму:

Вибирається число k , і на першому кроці ці k епіцентрів вважаються "центрами" кластерів. Кожному кластеру відповідає один центр. Вибір початкових центроїдів може здійснюватися в такий спосіб:

- вибір k - епіцентрів для максималізації початкової відстані
- випадковий вибір k - епіцентрів
- вибір перших k - епіцентрів.

В результаті кожен епіцентр буде призначений певному кластеру. Задане число кластерів k зіставляється кластерам так, що б середні значення в кластері (для всіх змінних) якомога більше відрізнялися один від одного.

Обчислюються центри кластерів, якими потім і далі вважаються по координатно (широта і довгота) середні, що містяться в кластерах епіцентрів. Далі, ітеративно, епіцентри перерозподіляються в кластери, поки не буде досягнуто оптимальне значення функції мети, що визначає якість кластеризації. Процес обчислення центрів і перерозподілу епіцентрів триває до тих пір, поки не буде виконано одну з умов:

- кластерні центри стабілізувалися, тобто всі спостереження належать кластеру, до якого належали до поточної ітерації

- число ітерацій дорівнює максимальному числу ітерацій.

Після отримання результатів кластерного аналізу, методом k -середніх, слід перевірити якість кластеризації (тобто оцінити, наскільки кластери відрізняються один від одного). Для цього розраховуються покоординатно середні значення епіцентрів для кожного кластера. При оптимальній кластеризації повинні бути отримані дані що значно відрізняються середні для всіх вимірювань або хоча б більшої їх частини. Тому для групування епіцентрів застосовується алгоритм k -середніх, який точніше визначає положення центра сейсмічних зон.

2.5. Оптимальне кластерне рішення

Існують різні способи розбиття заданої сукупності спостережень на кластери – складання кластерних груп, які розрізняються числом кластерів в них (сейсмічні зони). Представляє практичний інтерес завдання порівняльного аналізу якості кластерних рішень, які визначаються на множині всіх можливих розбитків простежуваної вибірки епіцентрів землетрусів на групи кластерів. Тобто, в процесі кластеризації виходять групи, що складаються з різного числа кластерів, і визначається група з оптимальним числом кластерів. Під найкращим розбиттям на кластери мається на увазі таке розбиття, при якому досягається екстремум обраного функціоналу якості. Для пошуку найкращого кластерного рішення проводиться кластеризація при різних значеннях числа кластерів G групах і вибирається те рішення G_0 , при якому досягається поліпшення функціоналу якості – ступеня схожості об'єктів всередині кластерів. В якості початкового числа кластерів використовується число кластерів, визначене за агломеративного методу.

При виборі числа кластерів потрібно керуватися компромісом між однорідністю кластерів і бажанням скоротити обсяг даних. У даній роботі використовувалися деякі критерії оптимізації кластерного рішення з багатьох існуючих методів.

2.5.1. Мінімум сумарної дисперсії

Одним з критеріїв якості кластерного рішення є мінімум сумарної дисперсії [Батуркин и др., 2010; Greutert, 2004; Jäger et al., 2009; Kursten, 2006; Milligan, 1985]:

$$Q(X, C) = \sum_{r=1}^K \sum_{x \in C_r} \|x - \bar{c}_r\|^2; \quad \bar{c}_r = \frac{1}{n_r} \sum_{x \in C_r} x \quad (2.6)$$

де K – число кластерів, C_r – кластери, $r=1, \dots, K$.

За величиною дисперсії визначається ступінь однорідності кластерів - чим менше дисперсія, тим вище схожість об'єктів всередині даного кластера. Критерій (2.6) має важливу властивість для алгоритмічного визначення оптимального розбиття на кластери. Квадрат відстані Евкліда кожного елемента кластера до його середини не перевищує квадрат відстані цієї точки до середини інших кластерів.

Однозначного оптимального кластерного рішення за критерієм Елбоу для землетрусів Туреччині не існує. На графіку критерію Елбоу видно кілька рівнозначних точок скачкоподібних значень відстаней між кластерами: 16; 24; 32; 38 і 50, які визначають оптимальне число кластерів.

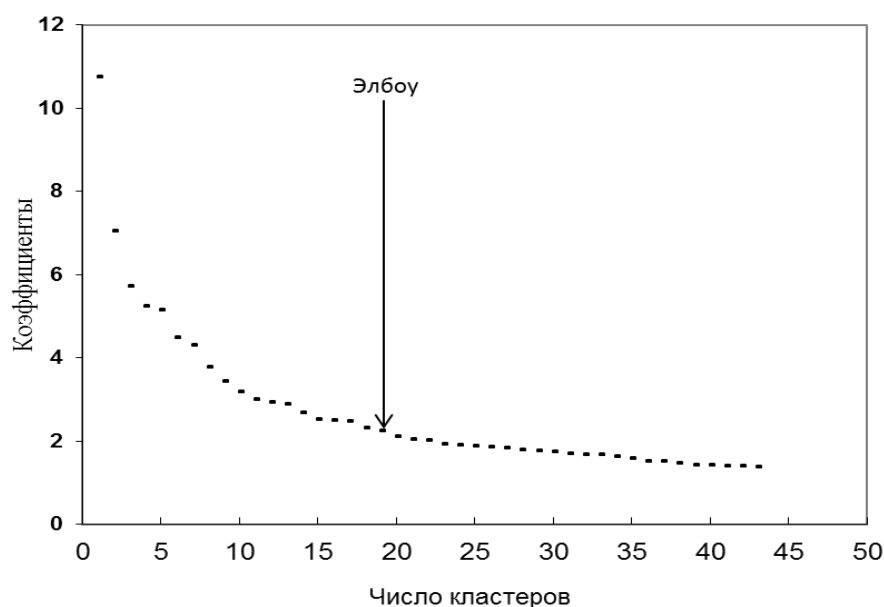


Рис. 2.4. Визначення числа кластерів за критерієм Елбоу в ієрархічному агломеративного методі

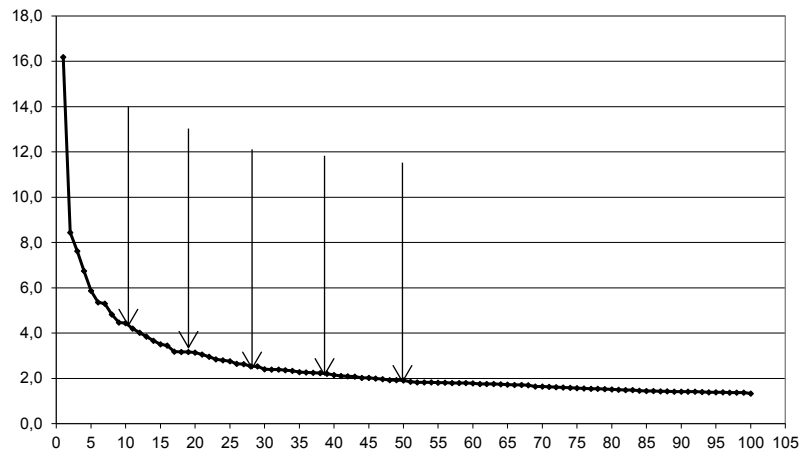


Рис. 2.5. Визначення числа кластерів території Туреччини за критерієм Елбоу в ієрархічному агломеративного методі.

Для землетрусів з каталогу південного сходу Європи [Shebalin et al., 1998], мінімум цільової функції (2.6) досягається при числі кластерів $K=18$. Число кластерів визначається номером кроку, на якому відбувається скачок графіка суми квадратів відхилень. На цьому заснований метод Елбоу (рис. 2.4; 2.5). Стрілкою позначено точка, на якій відбувається скачок значень відстані між кластерами (коефіцієнти). Оптимальним кластерним рішенням для сейсмічного регіону Балкани є 18 сейсмічних зон, так як, в цій точці відбувається скачок значень відстані.

2.5.2. Метод Галінського-Харабаша

У цьому методі оптимальне кластерне рішення визначається з максимуму індексу Галінського-Харабаша [Krieg, 2007; Mingjin, 2005; Sugar, James, 2003]:

$$Ch_k = \frac{\text{trace}(B_k)(n-k)}{\text{trace}(W_k)(k-1)}, \quad (2.7)$$

де B_k , W_k – матриця між кластерних і внутрішньо кластерних сум квадратів відстаней (дисперсії); n , k – число спостережень і кластерів; trace – слід матриці [http://books.google.de/books, 2014; Mingjin, 2005]:

$$B_k = \sum_{g=1}^G \sum_{x \notin C_g} (x - \bar{c}_g)(x - \bar{c}_g)^T; \quad W_g = \sum_{g=1}^G \sum_{x \in C_g} (x - \bar{c}_g)(x - \bar{c}_g)^T; \quad \bar{c}_l = \frac{1}{n_l} \sum_{x \in C_l} x \quad (2.8)$$

Індекс Ch_k визначається тільки для значень k більших ніж 1, тому що значення B_k не визначене, коли $k=1$. Для каталогу [Shebalin et al., 1998], оптимальним є рішення в $K=22$ кластера, статистика критерію (2.6) при цьому значенні максимальна. Максимальне значення індексу (2.7), для більшості землетрусів з каталогу ROMPLUS [http://www.infp.ro], $Ch_k = 1.2.105$ досягається при $K=13$ кластерах. В Італії виділено 35 сейсмічних зон, але кластерний аналіз, на основі каталогу землетрусів, що містить дані про більш ніж 60000 землетрусів, спостережених за період 2005–2010рр., визначив 33 кластера. При такому числі кластерів значення індексу Галінського-Харабаша є найбільшим. Найкращим рішенням для сейсмічності Туреччини, за величиною індексу Ch_k (2.7), є 48 кластерів.

2.5.3. Критерій мінімуму внутрішньо-кластерної дисперсії

Оптимальним кластерним рішенням для сукупності румунських землетрусів є $K=13$ кластерів, при яких внутрішньо кластерна сума квадратів відстаней W_g [Milligan, 1985]:

$$W_g = \sum_{g=1}^G \sum_{x \in C_g} (x - c_g)(x - c_g)^T \quad (2.9)$$

мінімальна $W_{13} = 1.11 \times 10^3$. Для землетрусів південного сходу Європи [Shebalin et al., 1998] рішення в $K=18$ кластерів виявилось оптимальним. За цим критерієм на території Італії виділяються 33 кластера, а в Туреччині 48 кластерів.

2.5.4. Статистика силуетів

Статистика силуетів (silhouette statistic) є мірою щільності розташування даних усередині кластера, і визначається з виразу [Sugar, James, 2003; Tibshirani et al., 2001]:

$$s(i) = \begin{cases} 0, & \text{если } j\text{-е наблюдение находится в единичном кластере} \\ \frac{b(i) - a(i)}{\max(a(i), b(i))}, & \text{иначе} \end{cases} \quad (2.10)$$

$$a(i) = \sum_{j \in \{C_g \setminus i\}} \frac{d_{ij}}{n_g - 1}, \quad b(i) = \min_{\eta \neq g} \sum_{j \in C_\eta} \frac{d_{ij}}{n_\eta} \quad (2.11)$$

1. $a(i)$: середня дистанція між i -м епіцентром і епіцентрами всередині кластера C_g .
2. $b(i)$: найменша середня відстань між епіцентром і епіцентрами з відмінним від C_g кластерів
3. $d_{i,j}$: міра схожості – відстань Евкліда між i -м і j -м епіцентрами. Вибирається число кластерів \hat{G} , при якому середня широта силуету [Sugar, James, 2003]:

$$\bar{S}_g = \frac{1}{n} \sum_{i=1}^n s(i), \quad g \in (2, \dots, G) \quad (2.12)$$

максимальна. Оцінкою числа кластерів служить функціонал [Tibshirani et al., 2001]:

$$\hat{G} = \arg \max_{2 \leq g \leq G} \bar{S}_g \quad (2.13)$$

Значення статистики силуетів можуть змінюватись в межах $-1 \leq s(i) \leq 1$.

Негативні значення небажані, тому що вони вказують на те, що середня відстань $a(i)$ до точок кластера, більше мінімальної середньої відстані $b(i)$ до точок інших кластерів. Якщо значення статистики силуетів знаходиться в інтервалі $(0,1)$, де $a(i) < b(i)$, то кластерне рішення є правильним. Чим ближче значення $a(i)$ до нуля, тим вище якість кластерного рішення. У цьому випадку значення $s(i)$ близько до 1. Великі значення $s(i)$ говорять про хорошу якість кластеризації.

Для даних з каталогу [Shebalin et al., 1998] значення статистики силуетів максимально при 19 кластерах. Найкращим рішенням для спостережених землетрусів Румунії з ROMPLUS виявилось 5 кластерів. За методом силуетів, для Італії, оптимальним виявилось рішення в 33 і 34 кластера. Оптимальне рішення для

епіцентрів землетрусів Туреччини визначається з максимуму $S_{49}=S_{51}=0.585$ (2.13), тобто адекватними є рішення в 49 і 51 кластери.

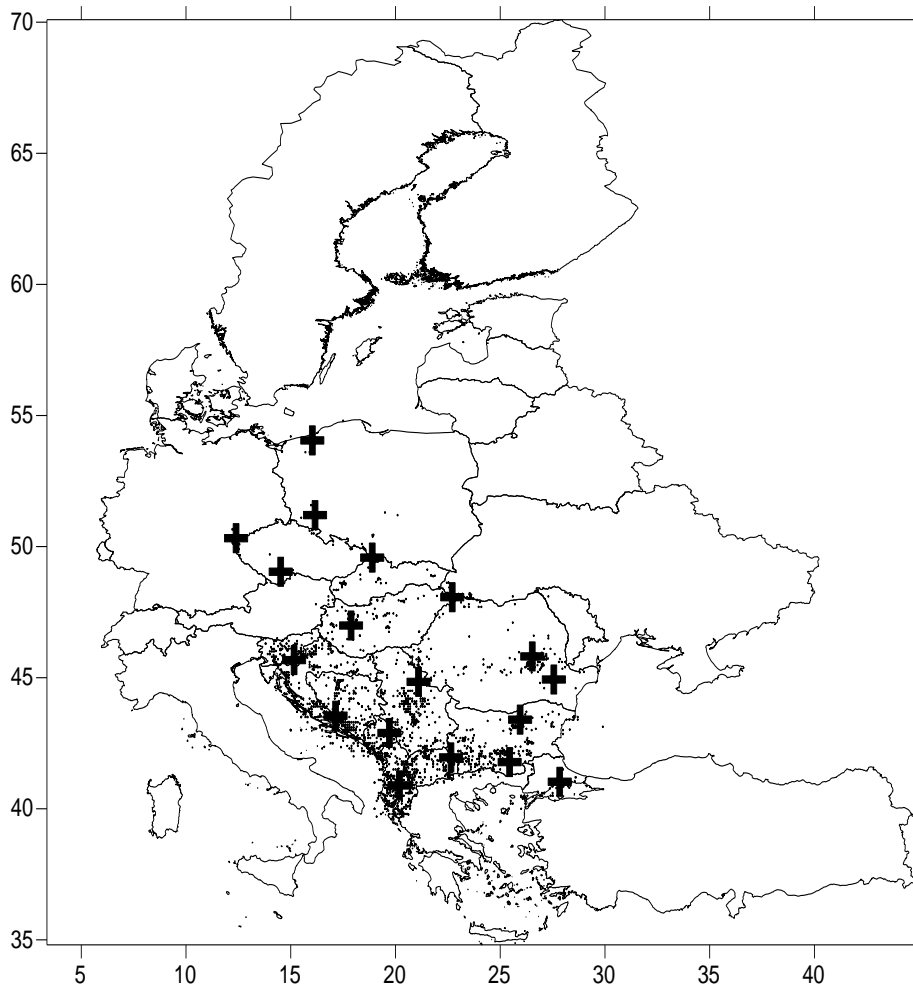


Рис. 2.7. Географічні середини 18 кластерів (сейсмічних зон)

2.5.5. Метод Кржановського-Лая

Метод максимілізує функцію [Шалымов и Граничин, 2010; Greutert, 2004; Milligan, 1985]:

$$KL(K) = \left| \frac{DIFF(K)}{DIFF(K+1)} \right| \quad (2.14)$$

$$DIFF(K) = (K-1)^{2/p} W(K-1) - K^{2/p} W(K) \quad (2.15)$$

Оптимальне, по методу Кржановського-Лая, рішення для території південного сходу Європи, Румунії та Італії складає 21, 13, 33 кластери відповідно. Найбільше значення індексу $KL(K)=7.12$, для території Туреччини досягає при 49 кластерах.

2.5.6. Непараметричний метод Сюгера-Джеймса

Метод заснований на використанні середніх всередині кластерних відстаней, званих "спотвореннями" [Sugar, James, 2003; Tibshirani et al., 2001; Mingjin, 2005].

1. На множині даних з урахуванням поточної кількості кластерів запускається будь-який алгоритм кластеризації. Як правило, це k -means. Множина розбивається на окремі кластери, для різних значень G , підраховується внутрішня дисперсія:

$$d_g = \sum_{x \in C_g} (x - c_g)(x - c_g)^T \quad (2.16)$$

Із значень дисперсій d_g , $g=1, \dots, G$, вибирається найменше, яке приймається за мінімальне "спотворення" і стає значенням функціоналу при даній кількості кластерів.

2. Вибирається степінь трансформації $\lambda > 0$. Зазвичай приймається $\lambda = p/2$, розмірність простору об'єктів. У нашому випадку $p=2$.
3. Обчислюються скачки по формулі [Mingjin, 2005]:

$$\widehat{d}_G = \min d_g; J_G = \widehat{d}_G^{-\lambda} - \widehat{d}_{G-1}^{-\lambda}; g \in (1, G) \quad (2.17)$$

4. За підсумкове значення кількості кластерів вибирається те значення, яке відповідає найбільшому скачку:

$$G^* = \arg \max_G J_G \quad (2.18)$$

Найкращим рішенням для землетрусів південного сходу Європи [Shebalin et al., 1998], Румунії та Італії складає 22, 11 і 13, 33 кластери, відповідно. Оптимальне рішення для Туреччини визначено по максимуму значення стрибка $J_k=0.0142$, яке спостережено при 45 кластерах.

На рис. 2.7 знаками (+), (*) позначені середньоарифметичні значення 18 кластерів і координати осередків балканських землетрусів.

2.6. Двокроковий кластерний аналіз

Стійкість кластеризації є характеристикою, яка б показала, наскільки різними виходять підсумкові розбиття на групи після багаторазового застосування алгоритмів кластеризації для одних і тих же даних. Невелика розбіжність результатів вказує на високу стійкість. Кількість кластерів, яке максимілізує кластерну стійкість, може стати найкращою оцінкою їх реальної кількості. На цьому принципі діє процедура двокрокового кластерного аналізу, в якому число кластерів може бути визначено автоматично. Масив даних спочатку аналізується і розбивається на підкластери невеликого розміру, який визначається структурою дерева кластерних властивостей [Vacher et al., 2006]. Кількість підкластерів залежить від заданої глибини (рівнів) кластерного дерева, від заданого числа гілок (вузлів) цього дерева. У процедурі кластерного аналізу пакету SPSS розмір підкластеру визначається ще й обсягом виділеної для обчислень пам'яті. В ході аналізу це дерево піддається багаторазовим змінам, проте база даних залишається незмінною. На цьому етапі можна провести аналіз викидів в даних.

Відповідно до інформаційних критеріїв Акайке (AIC) і Байеса (BIC) найкращою, серед альтернативних моделей, є модель з найменшим значенням AIC або BIC . Тому, попередня оцінка числа кластерів на першому кроці здійснюється, за значеннями інформаційних критеріїв AIC , або BIC :

$$AIC_k = 2l_k + 2r_k, \quad BIC_k = 2l_k + r_k \log n \quad (2.19)$$

де r_k число незалежних параметрів, d_k – відстань Евкліда між кластерами C_i і C_j :

$$d_k(i, j) = \sqrt{\sum_{m=1}^2 \bar{x}_{im} - \bar{x}_{jm}}, \quad \bar{x}_{im} = \frac{1}{n_i} \sum_{x \in C_i} x, \quad \bar{x}_{jm} = \frac{1}{n_j} \sum_{x \in C_j} x \quad (2.20)$$

n_i, n_j число елементів в кластерах C_i і C_j , відповідно, m - розмірність простору ознак (епіцентрів). Ця відстань, в якості запобіжної подібності кластерів, використовується для неперервних змінних, якими є координати епіцентрів. Тому

що, зазвичай, при зростаючій кількості кластерів значення AIC і BIC спочатку зменшується, а потім зростає, то потрібно шукати значення числа кластерів, при якому AIC або BIC мінімальне. Мінімальне значення інформаційного критерію Акайке досягається при 22 кластерах $AIC_{22}=333.29$ (табл. 2.3).

Таблиця 2.3

Двокроковий кластерний аналіз: автоматичне визначення числа кластерів

Число кластерів	Інформаційний критерій Акайке (AIC)	Варіація (AIC)	Відношення варіації AIC	Відношення відстані
1	4170.042			
2	2604.288	-1565.754	1.000	1.921
3	1793.147	-811.141	0.518	2.587
4	1484.568	-308.579	0.197	1.065
5	1195.238	-289.330	0.185	1.024
6	912.829	-282.409	0.180	2.025
7	777.429	-135.400	0.086	1.713
8	701.701	-75.728	0.048	1.062
9	630.828	-70.873	0.045	1.070
10	565.081	-65.747	0.042	1.140
11	508.390	-56.691	0.036	1.052
12	454.913	-53.477	0.034	1.625
13	425.084	-29.829	0.019	1.183
14	401.111	-23.972	0.015	1.266
15	383.859	-17.252	0.011	1.113
16	369.164	-14.695	0.009	1.116
17	356.819	-12.345	0.008	1.330
18	349.526	-7.294	0.005	1.117
19	343.830	-5.696	0.004	1.123
20	339.636	-4.193	0.003	1.018
21	335.659	-3.978	0.003	1.155
22	333.290	-2.368	0.002	1.493
23	334.346	1.056	0.000	1.016
24	335.515	1.168	0.000	1.100
25	337.304	1.789	-0.001	1.047
26	339.373	2.070	-0.001	1.128
27	342.117	2.744	-0.002	1.013

Таким чином, попередня оцінка числа сейсмічних зон південного сходу Європи, виділених двукроковим кластерним аналізом, збігається з кількістю приведених в каталозі зон. Якщо ж, $AIC_1 < AIC_2$ або $BIC_1 < BIC_2$, тоді кількість кластерів вважається рівним $K=1$, і робота алгоритму завершується.

На другому етапі аналізу, як і в ієрархічному агломеративного методі, на кожному наступному кроці об'єднуються найближчі два підкластери, і утворюється новий кластер. При цьому підкластери обробляються як окремі спостереження. Коли, кількість кластерів досягає певного на першому кроці порогового числа кластерів, то алгоритм переходить до визначення оптимального кластерного рішення:

$$R(k) = \frac{d_{k-1}}{d_k} \quad (2.21)$$

де $d_k - 1$ є відстанню, яка виходить після приєднання k кластерів до $(k-1)$ кластеру. Відстань d_k визначається з наступних виразів:

$$d_k = l_{k-1} - l_k, \quad l_v = (2r_v - AIC_v) / 2 \quad (2.22)$$

$v=k, k-1$. Найкращим кластерним рішенням є певна кількість з ряду $k=1, \dots, k_{max}$ (k_{max} - певне в першому кроці максимальне число кластерів), на якому відбувається скачок зміни відношення:

$$R(k_1) / R(k_2) \quad (2.23)$$

де $R(k_1), R(k_2)$ найбільші значення $R(k)$. Якщо зміна ставлення (2.23) перевершує початковий поріг зміни відстані c_2 , яке визначається специфікою розв'язуваної задачі, число кластерів вважається рівним значенню k_1 , інакше воно дорівнює більшому із значень k_1, k_2 . З цього випливає, що на другому кроці алгоритм двокрокового кластерного аналізу визначив рішення в 6 кластерів для землетрусів південного сходу Європи. Далі обчислюється відношення варіації $R(k)$ в відстанях для k кластерів [Bacher et al., 2006].

2.7. Критерії однорідності кластерів

Вище було сказано, що існує множина мір схожості об'єктів, алгоритмів обробки кластерів і способів визначення оптимального кластерного рішення. Однак не існує загального критерію визначення числа кластерів. Для порівняння альтернативних кластерних рішень можна використовувати однорідність виділених кластерів. Найкращим вважається кластерне рішення, де однорідність вища. Поряд з однорідністю кластерів, ще потрібно враховувати відповідність кластерних рішень структурі даних і перевірити адекватність виділених кластерів. При прийнятті рішень потрібно звернути увагу на стійкість і доцільність кластерних рішень. Стійкість кластеризації визначає, наскільки різними є групи кластерів, виділені при багаторазовому застосуванні кластерного аналізу на одній і тій же множині епіцентрів. Оптимальне кластерне рішення визначається кількістю кластерів в тій групі, в якій для всіх кластерів досягається найкраща стійкість.

2.7.1. F -критерій

Перевірка однорідності кластерів здійснюється за допомогою F -критерію [Бюль, Цефель, 2001; Корпроф, Корр 2009; Mingjin, 2005]:

$$F = \frac{V(J, G)}{V(J)}, \quad (2.24)$$

де $V(J, G)$, $V(J)$ – дисперсія змінної J в кластері G і в загальній сукупності. Чим менше F , тим незначніше розсіювання змінної в деякому кластері в порівнянні розсіювання цієї змінної в основній вибірці. Величина F не повинна бути більше одиниці, інакше спостереження в кластері мають більше розсіювання, ніж в основний вибірці. Слід зазначити, що F -критерій повинен застосовуватися до всіх змінних у всіх виділених кластерах в деякій групі кластерів. Для сукупності землетрусів з каталогу ROMPLUS [www.infp.ro/arhiva] рішення в 5 кластерів виявилось найкращим. При такій кількості кластерів значення F -критерію для широти і довготи епіцентрів не

перевищують значення $F_{5,X}=0.44$ і $F_{5,Y}=0.21$, відповідно. За значенням F -критерію оптимальним рішенням у сукупності епіцентрів землетрусів Туреччини є 50 кластерів. При цьому числі кластерів спостерігається найменша розбіжність значень F -критерію в групі кластерів, при їх кількості від 10 до 55 у відповідних групах. Вибіркові значення критерію (2.24) знаходяться в допустимих межах: $0.04 < F_{50,X} < 0.88$; $0.1 < F_{50,Y} < 0.63$. Проведений аналіз показує, що для всіх груп кластерів значення F -критерію знаходиться в допустимих межах (0,1), тобто всі групи кластерного розбиття однорідні. Незначна розбіжність значень критерію (2.24) для всіх груп кластерів вказує на високу стійкість процедури кластерного аналізу при виділенні сейсмічних зон.

2.7.2. t -критерій

$$t = \frac{\bar{X}(J, G) - \bar{X}(J)}{S(J)} \quad (2.25)$$

$\bar{X}(J, G)$ – середнє значення змінної J в кластері G

$\bar{X}(J)$ – середнє значення змінної J в основній вибірці

$S(J)$ – стандартне відхилення змінної J в основній вибірці.

Негативні значення t -критерію означають, що змінна недостатньо представлена в кластері, в порівнянні з основною вибіркою. Позитивні значення навпаки, змінна надлишково представлена в кластері, ніж в основній вибірці [Greutert, 2004; Sugar, James, 2003]. Для румунських землетрусів мінімальне розходження значень t -критерію спостерігається при вирішенні в 13 кластерів. Для сукупності епіцентрів землетрусів Туреччини найкраща кластерна стійкість досягається для групи з 50 кластерів (сейсмічних зон). Вибіркові значення критерію (2.25) для широти і довготи епіцентрів для всіх кластерів в групі з 50 сейсмічних зон (кластерів) знаходяться в діапазоні незначних розбіжностей: $0.4 < t_{50,X} < 0.79$; $0.03 < t_{50,Y} < 0.69$. Номер групи визначається кластерів, що містяться в ній. Наприклад, $t_{k,X}$ означає значення t -критерію для ознаки X в групі, що складається з k кластерів.

2.8. Стійкість кластеризації

У кластерному аналізі не існує однозначно найкращого критерію якості кластерного рішення. Існує множина евристичних алгоритмів кластеризації, метрики дистанції, мір близькості кластерів, вибір яких суб'єктивний [Батуркин и др., 2010; Kürsten, 2006; Pfeffer, 2004]. Теорія математичної статистики не має стандартних критеріїв оптимізації кластерних рішень, хоча метод надає досить адекватну картину подібності об'єктів, що належать одному кластеру і відмінності об'єктів, що належать різним кластерам. Різні методи, використані на одній і тій же вибірці епіцентрів землетрусів, можуть визначити різні кластерні рішення, тобто спостерігається нестійкість методу. Наприклад, при виділенні сейсмічних зон в Румунії за даними з каталогу ROMPLUS, за деякими методами оптимізації процесу кластеризації, найкращим є рішення в 5 кластерів, а за іншими в 13 кластерів. З географічного розподілу епіцентрів румунських землетрусів (рис. 2.8) видно, що між областями скупченості епіцентрів немає чіткої межі [Lungu et al., 2008], тому результат кластерного аналізу буде залежати від методів кластеризації.

Перші дослідження з визначення сейсмічних зон на території Румунії проводилися на основі тектонічних і сейсмічних даних, які корелювали з характеристиками основних тектонічних одиниць, активних розломів і розподілом епіцентрів землетрусів [Constantinescu, Marza, 1980; Radu et al., 1980]. При розбитті території Румунії на сейсмічні зони, враховувалися простий географічний розподіл сейсмічної активності і межі адміністративних округів. На той період питання про зв'язок сейсмічних зон з особливостями регіональної тектоніки не розглядалися. Сейсмічна активність в Румунії зосереджена на контактних поверхнях між основними тектонічними одиницями, велика частина якої представлена землетрусами проміжної глибини в зоні Вранча, яка займає вузьку область площею близько 3000 км² в південно-східній частині Карпатської дуги. Сейсмічна зона Вранча є складним геологічним регіоном континентальної конвергенції, де сходяться тектонічні одиниці: Східно-Європейська плита, Внутрішньоальпійська

плита, край Мизийської плити і складчасті споруди Північної Добруджі [Constantinescu et al., 1976, Grasu, 1997].

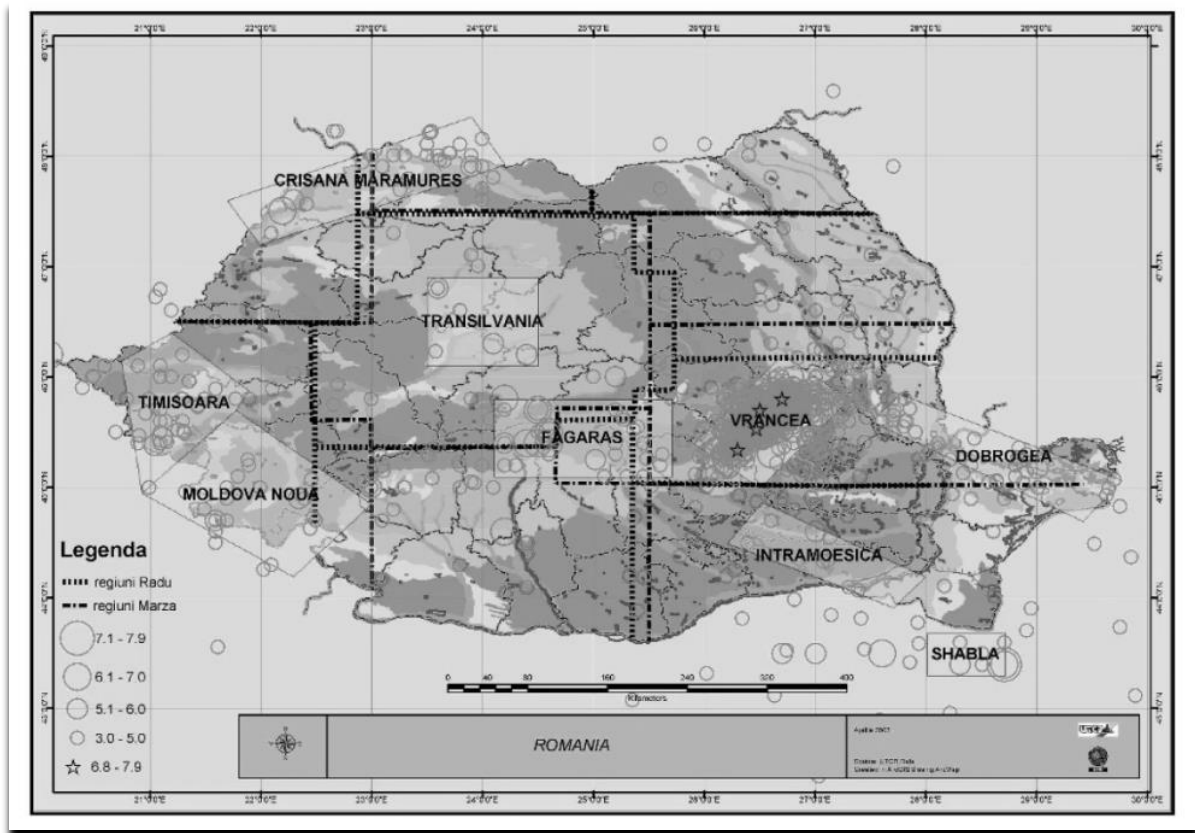


Рис. 2.8. Карта епіцентрів регіональних землетрусів ROMPLUS.

Сейсмічність зони землетрусів проміжної глибини Вранча є предметом багатьох досліджень [Enescu, 1980; Oncescu, Trifu, 1987; Enescu, Zugravescu, 1990; Ismail-Zadeh et al., 1999, 2008, 2012; Radulian et al., 2000; Гинтов и др., 2014; А. В. Кендзера и др., 2015].

Сейсмічність земної кори, пов'язана з зоною вогнищ землетрусів «Східна Вранча» частково перекриває зону землетрусів проміжної глибини Вранча, але з розширенням на схід і південний схід, до Мизийської платформи, між внутрішнім Мизийським розломом і розломом Тротуш. У районі гір Вранча кластерний аналіз виділив дві сейсмічні зони: Вранча і Східна Вранча, перший і шостий кластери на малюнку (рис. 2.10), відповідно.

Механізм вогнища землетрусу є одним з головних факторів, що характеризують сейсмічну подію, що визначають конфігурацію макросейсмічного поля і відображають геолого-геофізичні умови формування вогнища, просторову орієнтацію осей головних напружень, можливих площинних розломів і зрушень в епіцентрі землетрусу.

Існує ряд робіт, присвячених дослідженню механізмів вогнищ землетрусів зони Вранча [Enescu, 1980; Enescu і Zugravescu, 1990; Oncescu і Trifu, 1987; Ismail-Zadeh et al. 1999; Radulian et al., 2000; Степаненко и др., 2013 Гинтов и др., 2014; Кендзера и др., 2015; Степаненко, Карданец, 2016]. Дослідження показали, що для орієнтації площини розлому існує три типових рішення:

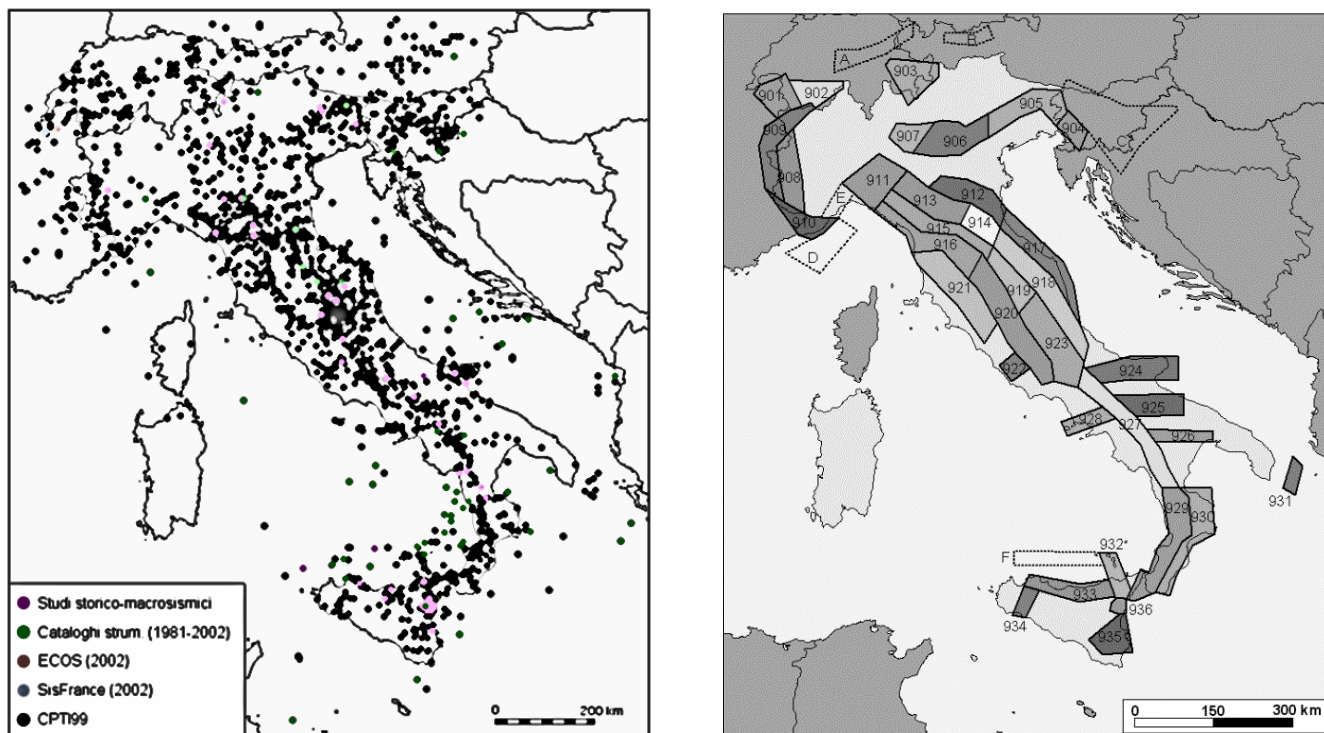
- центральний тип – площина розлому орієнтована, в основному, в напрямку північний схід - південний захід (*NE-SW*), і вісь стиснення *P*, перпендикулярна до гірської дуги
- північний тип – площина розлому орієнтована головним чином в напрямку північний захід - південний схід (*NW-SE*), і вісь стиснення *P* перпендикулярна Східним Карпатам
- південний тип – площина розлому орієнтована головним чином в напрямку захід-схід (*W-E*), і вісь стиснення *P* перпендикулярна Південним Карпатам.

У всіх трьох випадках вісь растягування (*T*) близька до вертикальної, яка стискає (*P*) – близька до горизонтальної.

Одним з основних факторів, що визначають якість аналізу сейсмічної небезпеки, є закон загасання, який не враховує всі доступні відомості про сейсмічності. Причиною є відсутність цих відомостей в каталогах землетрусів і необхідність деякого компромісу між адекватністю моделі загасання і її ускладненням, при включенні додаткових параметрів в модель. Тим не менше, хотілося б в майбутньому мати каталоги землетрусів, де будуть міститися і відомості про механізми вогнища землетрусів.

Епіцентри землетрусів Італії (рис.2.9, а) за оцінками сейсмологів розподілені по 35 сейсмічних зонах (рис. 2.9.b) [Bosi, Galli, 2004; Del Gaudio et al., 2009; Gruppo di

Lavogo, 2004]. Кластерний аналіз визначив 33 сейсмічні зони, число яких не дуже відрізняється від експертної оцінки. При визначенні сейсмічних зон Туреччини, за відомостями з каталогу турецьких землетрусів [Kalafat et al., 2011], методи кластерного аналізу та оптимізації кластерного рішення дають різні оцінки кількості кластерів. Домінуючим рішенням виявилася 50 сейсмічних зон (кластерів).



a)

b)

Рис. 2.9. Сейсмічна карта Італії: а) епіцентри землетрусів; б) сейсмічні зони.

2.9. Дослідження унімодальності

Кластерні методи передбачають унімодальне дослідження розподілу відстані між епіцентрами в деякому кластері C_r і його серединою Z_r (2.5), тобто кожен кластер описується однією середньою точкою: модою відповідного розподілу. Унімодальність розподілу даних в кластері є показником якості кластеризації і вказує на оптимальний розподіл кластерів. Наочне уявлення про мультимодальність дає гістограма щільності розподілу відстаней між епіцентрами в 1-м кластері (зона Вранча) і його центром. У представленій гістограмі для сейсмічної зони Вранча

(рис. 2.10) (перший кластер на малюнку) є тільки один пік, і це означає, що набір даних має унімодальний розподіл. Число інтервалів відстані k визначалося за формулою: $k=(2n)^{0,333}$, n – число епіцентрів в кластері. Негативна кореляція $r=-0.28772$ між магнітудою землетрусів і віддаленням епіцентрів від середньої точки першого кластера є значущим на рівні значущості $\alpha=0.01$. Тобто більші сейсмічні події стягуються до центру сейсмічної зони.

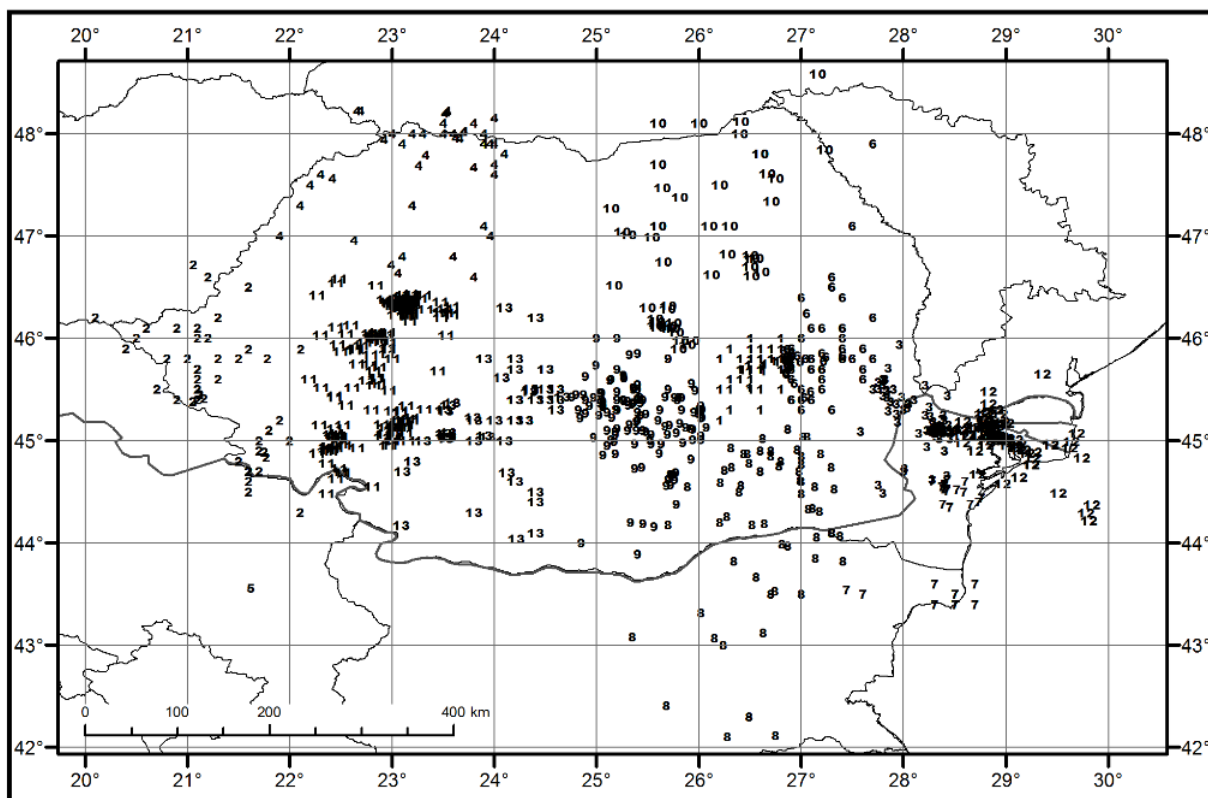


Рис. 2.10. Карта кластерів (каталог ROMPLUS): цифрами позначені епіцентри в кластерах (сейсмічних зонах).

Формальним методом аналізу структури кластера на унімодальність є dip-test критерій Хартігана, побудований на основі тестування кумулятивного емпіричного розподілу набору даних [Hartigan J. A., Hartigan P. M., 1985]:

$$H_n = \sqrt{n} \sup_x |F_n(x) - U(x)| \quad (2.26)$$

де n – число епіцентрів в кластері, $F_n(x)$ - емпірична функція розподілу відстаней і $U(x)$: функція рівномірного розподілу. Нехай нульова гіпотеза H_0 означає, що епіцентри в кластерах має унімодальний розподіл при альтернативній гіпотезі H_1 , що складається в тому, що розподіл є мультимодальний.

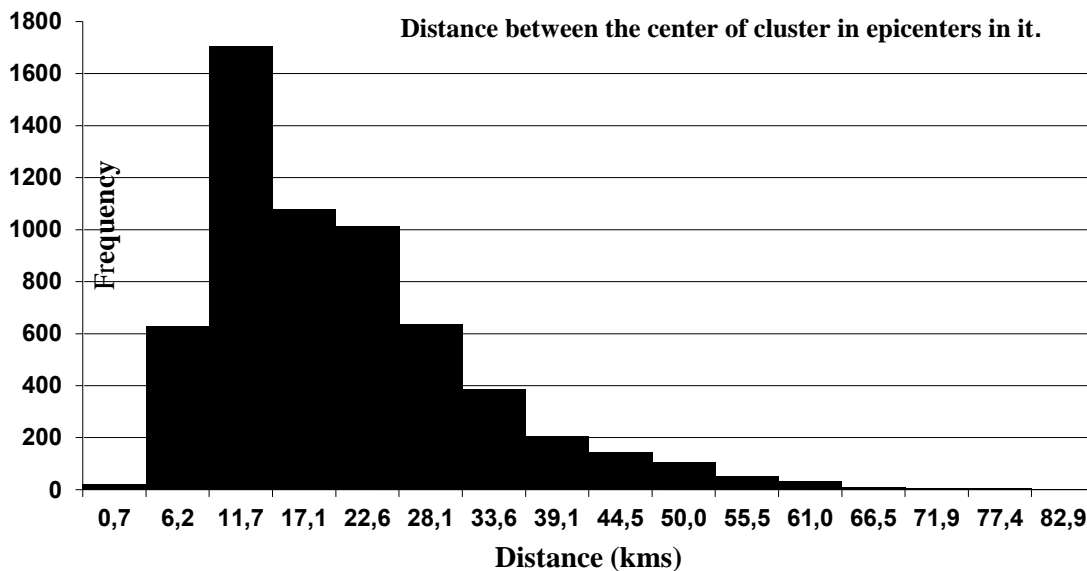


Рис. 2.11. Гістограма розподілу відстані між епіцентрами в 1-ому кластері (зона Вранча) і його центром. На осі абсцис наведені інтервали відстаней, висота прямокутників над ними пропорційна частоті.

Виявлення унімодальності є природним способом визначити, що розподіл епіцентрів всередині кластера є унімодальним, не містить нерозділених кластерів, і немає необхідності перерозподілу даних. Якщо ж гіпотеза H_0 відхиляється на користь альтернативної гіпотези H_1 , то даний кластер розбивається на кластери, число яких визначається кількістю мод. Для перевірки гіпотези H_2 , що складається в тому, що розподіл епіцентрів є бімодальною при альтернативі H_0 , застосовується коефіцієнт бімодальності (BC) [Кларп, 2007]:

$$BC = \frac{m_3^2 + 1}{m_4 + \frac{3(n-1)^2}{(n-2)(n-3)}} \quad (2.27)$$

де m_3 – асиметрія, і m_4 – ексцес. Значення BC більше або рівні 0.555 (значення для рівномірно розподіленої генеральної сукупності даних) вказує на бімодальність розподілу епіцентрів. Знак (+) означає, що гіпотеза H_0 повинна бути прийнята, тому що ймовірність значущості p перевершує рівень значущості $\alpha=0.01$ dip-test критерію, а значення BC менше величини 0.555. Знак (-) означає, що гіпотеза H_0 повинна бути відкинута на користь конкуруючої гіпотези H_1 (табл. 2.4).

Таблиця 2.4

Величина ймовірності значущості p dip-test статистики та коефіцієнта бімодальних попарних відстаней між епіцентрами в кластерах.

Номер кластера	1	2	3	4	5	6	7	8	9	10	11	12	13
P-значення Dip-test статистики	0	0.037	0	0	0	0.5	0.5	0.99	0	0	0	0.004	0
Правдоподібність гіпотези	ні	так	ні	Ні	ні	так	так	так	ні	ні	ні	ні	Ні
Коефіцієнт бімодальності	0.522	0.482	0.404	0.541	0.694	0.533	0.508	0.633	0.44	0.49	0.512	0.736	0.482
Унімодальне	ні	так	так	Так	так	ні	так	так	ні	так	ні	так	Так

Таблиця 2.5

Величина ймовірності значущості p dip-test статистики та коефіцієнта бімодальності віддалення епіцентрів від центру кластерів.

Номер кластера	1	2	3	4	5	6	7	8	9	10	11	12	13
Ймовірність значущості p	0.003	0.665	0.708	0.043	0.368	0.001	0.843	0.366	0.002	0.81	0.0	0.663	0.452
Правдоподібність гіпотези	ні	так	так	Так	так	ні	так	так	ні	так	ні	так	Так
Коефіцієнт бімодальне	0.532	0.504	0.397	0.403	0.426	0.588	0.609	0.677	0.440	0.543	0.38	0.69	0.539
Правдоподібність гіпотези	так	так	так	так	так	ні	ні	ні	так	так	так	ні	Так

Для проведення ЙАСН необхідно виділити групи епіцентрів, щоб побудувати вірогідну модель послідовності спостережених в них землетрусів. Крім кластерного аналізу для виділення сейсмічних зон можна застосувати і метод групування

епіцентрів землетрусів. Кластерний аналіз і метод групування об'єктів вирішують різні завдання, хоча використовуються загальні алгоритми. При кластеризації об'єкти, що належать одній групі подібні, а що належать різним групам різні, тобто самі групи розрізняються, будучи ізольованими між груповими межами. Групування епіцентрів теж створює групи подібних елементів, але самі групи не зобов'язані бути різними, тобто об'єкти в сусідніх групах можуть бути схожими. Метод групування не впливатиме на адекватність ймовірнісної моделі сейсмічного режиму сейсмічних зон.

Висновки до розділу 2

Існує множина евристичних алгоритмів кластеризації. Однак, в кластерному аналізі не існує однозначно найкращого критерію якості кластерного рішення. Теорія математичної статистики не має стандартних критеріїв оптимізації кластерних рішень, хоча методи кластерного аналізу надають досить адекватну картину подібності об'єктів, що належать одному кластеру, і відмінності об'єктів, що належать різним кластерам.

В якості пробної вибірки використовувався каталог землетрусів південно-східної частини Європи, що охоплює період 1900-1990 рр., в якому містяться відомості про землетруси в 18 сейсмічних зонах. Каталог використовувався в якості навчального та при порівнянні координат епіцентрів в сейсмічних зонах, виділених алгоритмом *k*-середніх, з координатами епіцентрів, наведеними в каталозі південно-східної частини Європи спостерігався 97% - ний збіг. Висока частка збігів є показником ефективності методу. Однак, кластерний аналіз, будучи формальним методом, не може враховувати геолого-геофізичні характеристики зони вогнищ землетрусів, але цілком придатний для оперативного сейсмічного зонування (групування вогнищ землетрусів). Дискримінантний аналіз застосовувався для поповнення каталогу землетрусів, що не ввійшли в нього відомостями про подальші землетруси.

Крім кластерного аналізу для виділення сейсмічних зон можна застосувати і метод групування епіцентрів землетрусів. Кластерний аналіз і метод групування об'єктів вирішують різні завдання, хоча використовуються загальні алгоритми. При кластеризації, об'єкти, що належать одній групі подібні, а що належать різним групам різні, тобто самі групи розрізняються, будучи ізольованими між груповими межами. Групування епіцентрів теж створює групи подібних елементів, але самі групи не зобов'язані бути різними, тобто об'єкти в сусідніх групах можуть бути схожими. Метод групування не впливатиме на адекватність ймовірнісної моделі сейсмічного режиму сейсмічних зон.

Кластерний аналіз є корисним інструментом для оперативного виділення сейсмічних зон, хоча заснований на формальній ознаці подібності епіцентрів і не враховує фізичну природу осередків.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

- Арефьев С.С. Эпицентральные сейсмологические исследования. М.: ИКЦ «Академкнига», 2003. 375 с.
- Батуркин С.А., Батуркина Е.Ю., Зименко В.А., Сигинов И. В., Статистические алгоритмы кластеризации данных в адаптивных обучающих системах. Вестник РГРТУ. № 1 (выпуск 31). Рязань, 2010, 34– 46.
- Бюль А., Цефель П. SPSS: искусство обработки информации. – СПб.: Диа Софт, 2001, 186.
- О. Б. Гинтов, А.В. Муровская, Т.П. Егорова, Ю.М. Вольфман, Т.А. Цветкова и др., Геофизический журнал, №3, Т. 37, 2015, 22-49.
- Левин Б.В., Ким Ч.У., Соловьев В.Н., Оценка сейсмической опасности и результаты детального сейсмического районирования для городов о. Сахалин, Тихоокеанская Геология, том. 31, №5, 2012, 93–103.
- А. В. Кендзера, А. Н. Островной, Т. А. Цветкова, Землетрясения в зоне Вранча и мантийные сейсмические границы, Доповіді Національної Академії наук України, 2015, № 9, 74-78.

- Шалымов Д. С., Граничин О. Н., “Рандомизированный алгоритм нахождения количества кластеров”, СПбГУ, 2010, 47.
- Шебалин Н.В. Методы использования инженерно–сейсмологических данных в сейсмическом районировании. В кн.: Сейсмическое районирование СССР, Ч.1, гл. 6. Москва. Наука. 1968, 54–69.
- Bacher Johann, Wenzig Knut, Vogler Melanie, SPSS TWOSTEP CLUSTER – A FIRST EVALUATION. Universität Erlangen–Nürnberg, 2006, 20.
- Bosi V. and Galli P., Incorporating paleoseismological data in PSHA: The case of Calabria (southern Italy), *Bollettino di Geofisica Teorica ed Applicata* vol. 45, N. 4, 2004, 255-270.
- Constantinescu, L. and Marza, V., A Computer-compiled and Computer-oriented Catalogue of Romania’s Earthquakes During a Millennium (AD 984 –1979), *Rev. Roum. Geol., Geophys., Geogr., Ser Geophys.* 24, 1980, 171–191.
- Enescu, D., Contributions to the Knowledge of the Focal Mechanism of the Vrancea Strong Earthquake of March 4, 1977, *Rev. Roum. Geol., Geophys., Geogr., Ser. Geophys*, 24, 1980, 3–18
- Enescu D. and Zugravescu D., Geodynamic Considerations Regarding the Eastern Carpathians Arc Bend, Based on Studies on Vrancea Earthquakes, *Rev. Roum. Geophysique* 34, 1990, 17–34.
- Grasu Constantin, *Geologie structurala*, Editura Technica, Bucuresti, 1996, 244.
- Greutert Andreas. Methoden zur Schätzung der Clusteranzahl. Eidgenössische Technische Hochschule Zürich. 2004, 117.
- Gruppo di Lavoro. Redazione della mappa di pericolosità sismica prevista dall’Ordinanza PCM 3274 del 20 marzo 2003. Rapporto conclusivo per il Dipartimento della Protezione Civile, INGV–Milano Roma, 2004, 65.
- Gvishiani Alexei, Dobrovolsky Mikhail, Agayan Sergei, Dzeboev Boris. Fuzzy-based clustering of epicenters and strong earthquake – prone areas, *Environmental Engineering and Management Journal*, Vol. 12, 1, 2013. 1–10.

- V. Del Gaudio, Pierri P. and Calcagnile G., Seismogenic zonation and seismic hazard estimates in a Southern Italy area (Northern Apulia) characterised by moderate seismicity rates. *Nat. Hazards Earth Syst. Sci.*, 9, 161–174, 2009
- Hartigan J. A., Hartigan P.M., The Dip Test of unimodality, *Annals of Statistics*, Volume 13, Issue 1, 1985, 70-84.
- Hudec Marcus. Einführung in die Clusteranalyse. Universität Wien. 2003.
<http://books.google.de/books?id=mg>
<http://www.infp.ro>
- Ismail-Zadeh A.T., Keilis-Borok V.I., Soloviev A.A., Numerical modelling of earthquake flow in the southeastern Carpathians (Vrancea): effect of a sinking slab. *Physics of the Earth and Planetary Interiors*. Volume 111, Issues 3–4, 1999, 267–274, http://www.mitp.ru/~aismail/papers/vran_slab.pdf
- Ismail-Zadeh, A., Sokolov, V. and Bonjer, K., Tectonic stress, seismicity, and seismic hazard in the southeastern Carpathians, *Nat. Hazards*, Springer, 2007, 1-22.
- Ismail-Zadeh, A., Matenco, L., Radulian, M., Cloetingh, S., and Panza, G., Geodynamic and intermediate-depth seismicity in Vrancea (the south-eastern Carpathians): Current State-of-the-Art, *Tectonophysics*, 50-79, 2012 (http://www.mitp.ru/~aismail/papers/Vrancea_Tectonophysics.pdf).
- Kalafat D, Güneş Y, Kekovalı K, Kara M, Deniz P, Yılmaz M. (2011). Bütünleştirilmiş Homojen Türkiye Deprem Kataloğu (1900-2010; $M \geq 4.0$). İstanbul, Turkey: Boğaziçi Üniversitesi, Kandilli Rasathanesi ve Deprem Araştırma Enstitüsü, İstanbul, 2011.
- Kopprof Johannes, Kopp Johannes, Lois Daniel. Clusteranalyse. Institut für Soziologie, TU Chemnitz. 2009, 71.
- Knapp R. Thomas, Bimodality Revisted. *Journal of Modern Applied Statistics Methods*, Vol. 6, No. 1, 2007, 8-20.
- Kürsten Jens. Systematisierung und Evaluierung von Clustering-Verfahren im Information Retrieval. Technische Universität Chemnitz. Chemnitz, 2006, 148.

- Krieg Verena. Lineare Klassifikationsmethoden. Universität Ulm. Fakultät für Mathematik und Wirtschaftswissenschaften, 2007, 48.
- Jäger B., Wodny M., Biebler K. E., Rudolph P.E., Matthies K., Diskriminanzanalyse mit binären Daten. E.–M.–Arndt–Universität. Greifswald. 2009, 8.
- Milligan G.W. and Cooper M.C. An examination of procedures for determining the number of clusters in a data set, *Psychometrika*, 1985, 50, 159–179.
- Mingjin Y., *Methods of Determining the Number of Clusters in a Data Set and a New Clustering Criterion*, Blacksburg, Virginia, 2005, 48.
- Lungu D., Aldea A., Arion C., Romania's Seismicity and Seismic Hazard: From Historical Records to Design Codes, *Harmonization of Seismic Hazard in Vrancea Zone*, Springer Science + Business Media B.V. 2008, 1–16.
- Oncescu M. C. and Trifu C. I., Depth Variation of the Moment Tensor Principal Axes in Vrancea (Romania) Seismic Region, *Ann. Geophysicae* 5B, 1987, 149–154.
- Pfeffer Fabian, *Einführung in die Clusteranalyse*, Arbeitsgemeinschaft Statistische Methoden der Sozialwissenschaften
- Radu, C., Apopei I., and Utale, A., Contributions to the study of the seismicity of Romania (in Romanian). In *Progrese in Fizica Symposium*, Cluj-Napoca, 1980.
- Radulian M., Vaccari F., Mandrescu N., Panza G. F., Moldoveanu C., Seismic hazard of Romania: A deterministic approach, *Pure Appl. Geophys.*, 157, 2000, 221-247.
- Shebalin N.V., Leydecker G., Mokrushina N.G., Tatevossian R. E. and Vassiliev V. YU.: *Earthquake Catalogue for Central and Southeastern Europe 342 BC–1990 AD*. European Commission, Report No. ETNU CT 93 – 0087, Brussels, 1998.
- Sugar C., James G., “Finding the number of clusters in a data set : An information theoretic approach”, *Journal of the American Statistical Association*. N98. 2003, 750–763.
- Stein Petra, Monika Pavetic, Marcel Noack. *Multivariate Analyseverfahren*. Duisburg, 2007, 102.
- Tibshirani Robert, Guenther Walther, Trevor Hastie. Estimating the number of clusters in a data set via the gap statistic. *J.R. Statist. Soc. B.*, 63, Part2, 2001, 411–423.

Timm Heiko. Fuzzy–Clusteranalyse: Methoden zur Exploration von Daten mit fehlenden Werten sowie klassifizierten Daten. Dissertation zur Erlangung des akademischen Grades. Magdeburg, 2002, 97.

РОЗДІЛ 3

СТАТИСТИЧНИЙ АНАЛІЗ КАТАЛОГІВ ЗЕМЛЕТРУСІВ

Одним із значущих чинників, що визначають якість аналізу сейсмічної небезпеки, є модель сейсмічності зони вогнищ землетрусів. Сейсмічність: схильність Землі або окремих територій впливу землетрусів. Сейсмічність характеризується:

- Частотою виникнення землетрусів
- Статистичним розподілом сили поштовхів (магнітуди)
- Просторовим розподілом осередків
- макросейсмічними спостереженнями сильних сейсмічних подій (інтенсивність сейсмічних впливів, картина завданих збитків).

Визначення перерахованих характеристик ґрунтується на сейсмостатистиці, у вигляді каталогів землетрусів сейсмічних зон. Каталог землетрусів є важливою продукцією сейсмології, що містить необхідну і достатню для аналізу сейсмічної небезпеки інформацію про сейсмічні події. Отже, другим кроком при аналізі сейсмічної небезпеки, після виділення сейсмічних зон, є створення бази даних у формі каталогів землетрусів, спостережених у сейсмічних зонах. У геофізичних процесах присутній елемент невизначеності і складності, властиві випадковим подіям. Для дослідження таких явищ застосовуються методи теорії ймовірностей, математичної статистики та теорії випадкових процесів.

Математична статистика, головне завдання якої полягає у пов'язуванні реального світу даних зі світом теоретичних моделей, розвивається на основі теорії ймовірностей і використовується для визначення параметрів моделі за вибірковими даними, прогнозу, перевірки гіпотез, і в якійсь мірі вирішує зворотні завдання. Математична статистика виявляє ступінь правдоподібності якихось гіпотез щодо відповідних явищ. Правдоподібність: не є доказом, а вказує лише на те, що немає достатньо для відхилення як неправдоподібної нульової гіпотези. Методи математичної статистики можуть допомогти досліднику в отриманні висновків про

масові явища за спостереженнями над ними, виявити найбільш значущі чинники і напрямки досліджень.

3.1. Опис даних

Палеосейсмологічні дослідження показують, що тектонічні процеси, що призводять до виникнення вогнищ землетрусів, тривають мільйони років [Солоненко, 1973; Wenzel et al., 1997], тоді як період інструментальних спостережень почався тільки в кінці XIX століття.

Сейсмологія є емпіричною наукою, тому що твердження, як і моделі, ґрунтуються на фактах і спостереженнях. Сейсмічні процеси є незапланованими експериментами, які ставить природа. Отже, в розпорядженні дослідників є одна єдина вибірка: частина генеральної сукупності, – по ній доведеться судити про властивості генеральної сукупності. На відміну від звичайної вибірки порядок послідовності (моменти часу настання землетрусів t_1, t_2, \dots, t_N) має тут істотне значення.

Результати спостережень містять випадкові і систематичні похибки. Похибки, невраховані фактори і невизначеності в геофізиці, існуючі на даному етапі розвитку наук про Землю, впливають на якість геофізичних моделей, і немає можливості повторити спостереження, вибрати моменти його виникнення. Отже, створення адекватних моделей висуває вимоги до якості вибірових даних і способів їх підготовки. Сейсмічні процеси відбуваються і розвиваються в часі і просторі під дією внутрішнього детермінізму глобальної тектоніки. Невизначеності, пов'язані з переплетенням внутрішніх фізичних полів Землі і гравітаційної сили небесних тіл, їх впливом на глобальну тектоніку, вносять елемент випадковості в моделі сейсмічності. Тривалість періоду інструментальних спостережень є лише миттю в геологічній історії зони вогнищ землетрусів, і тому на відносно короткому відрізку часу сейсмічність зони не може зазнавати значних змін, тобто повинна залишатися стаціонарною.

Сейсмічний режим є потоком сейсмічних подій, які реєструються в порядку їх виникнення і утворюють часовий ряд (ЧР). Тимчасовим називається послідовність значень одного або значної кількості показників певних явищ, упорядкованих в порядку виникнення. Теоретичною базою для аналізу часових рядів є теорія випадкових процесів.

Землетрус можна охарактеризувати багатьма параметрами, однак для аналізу сейсмічної небезпеки досить знати: час виникнення; координати епіцентрів; глибину вогнища і магнітуду землетрусу. Отже, маємо багатовимірний часовий ряд, компонентами якого є ряди, упорядковані у часі значень параметрів землетрусів, тобто вибірки: на мові математичної статистики. Для дослідження сейсмічного режиму і аналізу сейсмічної небезпеки будуть достатні послідовності значень параметрів. Основні припущення, що лежать в основі аналізу часових рядів, полягають в стаціонарності процесу і можливості його адекватного опису за допомогою нижчих моментів розподілу ймовірностей. Крім стаціонарності при аналізі часових рядів зазвичай, використовуються припущення щодо ергодичності до ймовірнісних характеристик, тобто, коли характеристики, обчислені по реалізації і за часом, збігаються. Наприклад, ергодичність до математичного сподівання означає, що середнє по множині реалізацій дорівнює середньому по часу.

3.2. Опис використовуваних методів математичної статистики

Методи математичної статистики діляться на описову статистику і на статистичні висновки [Кендалл, 1973; Fisz, 1966; Soong, 2004]. Основні методи описової статистики – обробка емпіричних даних, їх систематизація, наочне уявлення у формі графіків і таблиць, а також їх кількісний опис за допомогою основних статистичних показників. А завдання поширення ймовірнісних характеристик вибіркового даних на всю генеральну сукупність вирішується в рамках статистичних висновків. Описова статистика використовує три основні методи агрегування даних [Аргучинцева, 2007]: табличне представлення; графічне зображення і розрахунок статистичних показників.

Агрегування даних означає процес збору, обробки та подання інформації в остаточному вигляді. Основні статистичні характеристики поділяються на дві групи: міри середнього рівня і міри розсіювання [Аргучинцева, 2007; Дембицький, 2013]. Міри середнього рівня обчислюють вибіркoву характеристику об'єктів за певною ознакою: середнє значення, стандартна похибка, стандартне відхилення, ексцес, асиметрія, інтервал, мінімум, максимум, рахунок, медіана, мода, квантиль, математичне очікування, рівень надійності 95%, міра розсіювання, дисперсія випадкової величини, середньоквадратичне відхилення, розмах варіації, міжквартильний розмах, середнє абсолютне відхилення. Міри розсіювання показують, наскільки добре дані значення представляють дану вибірку.

Пакет SPSS є універсальною статистичною системою програм, необхідних для аналізу даних і призначеній для реалізації повної послідовності кроків аналізу даних: від перегляду даних, створення таблиць і обчислення описових статистик до складного статистичного аналізу. Графічні засоби, вбудовані в статистичні процедури, полегшують розуміння даних і інтерпретацію результатів аналізу, - вони корисні для представлення результатів аналізу.

3.3. Однорідність даних

Проводиться статистичний аналіз вибірки, складений з 1015 значень магнітуд типу mb і MS за даними міжнародного сейсмологічного центру ISC. Попереднім етапом статистичної обробки повинен бути етап перевірки однорідності вибірки в сенсі незмінності ймовірнісних характеристик. Мірою однорідності статистичної сукупності є коефіцієнт варіації [Welkowitz et al., 1988]:

$$V = \frac{\sigma}{\bar{a}} 100\% \quad (3.1)$$

де σ – середньоквадратичне відхилення, \bar{a} – середнє арифметичне вибірки. Сукупність вважається однорідною, якщо коефіцієнт варіації не перевищує 33% [Смирнов, Дунин-Барковский, 1965].

Однією з основних задач математичної статистики є оцінювання тісноти лінійного зв'язку між двома випадковими величинами. Для встановлення зв'язку між

залежними і незалежними змінними, найбільш часто, застосовуються регресійний і дисперсійний аналіз. Регресійний аналіз застосовується тоді, коли всі фактори є кількісними, в яких залежна змінна є кількісною, проте залежна змінна в регресійному аналізі є кількісною, а в дисперсійному аналізі якісною. При змішаній моделі, коли одні фактори є кількісними, а інші якісними застосовується коваріаційний аналіз [Toutenburg, Neumann, 2008]. У сейсмології знаходять застосування всі три методи аналізу статистичного зв'язку. Для виявлення статистичного зв'язку між різними типами магнітуд використовувалися регресивні методи.

3.4. Перетворення магнітудних шкал

Рівняння макросейсмічною поля:

$$I = am - b \log \sqrt{\Delta^2 + h^2} + c \quad (3.2)$$

пов'язує з кожним землетрусом магнітуди m , глибини вогнища h , інтенсивність сейсмічних струсів в точках земної поверхні, розташованих на епіцентральному відстані Δ від вогнища, a , b , c – коефіцієнти. У каталогах ISC, PDE, NEIC, USGS, ROMPLUS і т.д., наведені значення різних типів магнітуд: mb , MW , MS , ML , MD , MLH , MJ , ME , MPV , MLH і т.д. Для застосування закону загасання (3.2) необхідно привести значення магнітуд до єдиного типу, та визначати значення коефіцієнтів рівняння (3.2) для кожного типу магнітуди. У даній роботі зроблена спроба перетворення різних типів магнітуд в магнітуду шкали mb . Найпростішою моделлю регресії є одновимірний лінійний модель:

$$y = \alpha + \beta x + \varepsilon_i \quad (3.3)$$

де x – регресорів, y – відгук. В даному випадку змінна x також є випадковою величиною. У цьому рівнянні параметр α – вільний член; він представляє відрізок ординати в системі прямокутних координат. З точки зору аналітичної геометрії параметр β – коефіцієнт регресії, визначає нахил лінії регресії по відношенню до осей координат [Лешінській та ін., 2003; Себер, 1980; Schmidt, 2006]. Наносячи спостережені значення (x_i, y_i) , $(i=1, \dots, n)$ двовимірної дискретної випадкової

величини (X, Y) на координатну площину, можна попередньо визначити форму зв'язку. Надалі, у означатиме магнітуду mb , а x – значення одного з типів магнітуд: $mb, MW, MS, ML, MD, MLH, MJ, ME, MPV, MLH$.

В основу сейсмології закладені сейсмічні вимірювання, які схильні до систематичних і випадкових похибок. Статистичними методами вдається зменшити внесок випадкових похибок в експеримент. Типові джерела випадкових похибок в сейсмології – це невеликі похибки реєстрації, перешкоди і неточності самої сейсмічної апаратури. Спостережені значення магнітуди землетрусів визначаються з деякою похибкою $m = M + \varepsilon$, де M – справжнє значення магнітуди, ε – похибка. На значення похибки впливають багато незалежних факторів, і можна застосувати центральну граничну теорему [Гнеденко, 1961; Феллер, 1967; Крамер, 1975]. Передбачається, що похибки є випадковими величинами з математичним очікуванням рівним нулю і не залежать від самих випадкових змінних і їх розподілів. Для аналізу моделей, де є випадкові: відгук і регресори, застосовується метод умовної регресії залежної змінної на регресори. Для цього використовуються звичайні оцінки найменших квадратів, працюючи в термінах математичних очікувань [Thorsten, 2006]. Серед сукупності спостережень випадкової величини (X, Y) можуть виявитися повторювані точки, що відповідають експерименту з повторними спостереженнями.

Регресія магнітуди mb на MS .

Коефіцієнти регресії оцінювалися за формулами [Лешінській та ін., 2003]:

$$\beta = \frac{n \sum_{i=1}^n x_i y_i - \sum_{i=1}^n x_i \sum_{i=1}^n y_i}{n \sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2}, \quad \alpha = \bar{Y} - \beta \bar{X} \quad (3.4)$$

де \bar{X} і \bar{Y} – середньоарифметичні значення регресорів і відгуку відповідно. Простий мірою відповідності моделі лінійної регресії вибірковими даними є коефіцієнт кореляції:

$$r = \frac{\sum (Y_i - \bar{Y})(X_i - \bar{X})}{n \sigma_x \sigma_y} \quad (3.5)$$

σ_x, σ_y – середньоквадратичне відхилення відгуку і регресорів. Для перевірки гіпотези про рівність нулю коефіцієнта кореляції: $H_0: r=0; H_1: r \neq 0$, застосовується t -розподіл з $n-2$ ступенями свободи. Для перевірки гіпотези про значущість регресії $H: \beta=0$ використовувався також t -розподіл Стьюдента [Крамер, 1975]:

$$T = \frac{r}{\sqrt{\frac{(1-r^2)}{n-2}}} \quad (3.6)$$

Перевірка гіпотез щодо параметрів і значущості регресії ґрунтується на величині залишкової суми квадратів. Наприклад, гіпотеза H_0 може полягати в тому, що регресія незначуща: $\beta=0$, в разі лінійної регресії $y=\alpha+\beta x$. Вибіркові коефіцієнти регресії і кореляції складають: $\alpha=2.266, \beta=0.5508, r=0.8825$ відповідно. Гіпотеза про значимість регресії і кореляції виявилася правдоподібною. Вибіркові значення t -розподілів перевершують порогове значення $t_q=1.96$, відповідні 5% рівнем значущості с числом ступенів свободи більше 120.

Вибіркові оцінки статистичних параметрів є випадковими величинами з деяким апріорним розподілом ймовірностей. У практичних додатках для значень параметрів потрібно привести ймовірнісні оцінки. У загальному випадку дуже складно визначити точний розподіл вибірових характеристик. Часто, як і в даному випадку, визначенню підлягає просто невідома постійна. Для практичних застосувань досить вказати довірчий інтервал, який з деякою довірчою ймовірністю перекрив би невідомі значення параметрів. В одновимірному випадку, довірчі інтервали для параметрів регресії α, β виражаються формулами [Себер, 1980]:

$$\hat{\alpha} \pm u_{2,n-2,p}^\alpha S \left\{ \frac{\sum x_i^2}{n \sum (x_i - \bar{x})^2} \right\}^{\frac{1}{2}}, \quad \hat{\beta} \pm u_{2,n-2,p}^\alpha S \left\{ \frac{1}{n \sum (x_i - \bar{x})^2} \right\}^{\frac{1}{2}} \quad (3.7)$$

$$S^2 = \frac{1}{n-2} \left\{ \sum_{i=1}^n (Y_i - \bar{Y})^2 - \beta^2 \sum_{i=1}^n (x_i - \bar{x})^2 \right\} \quad (3.8)$$

де $u_{2,n-2,p}^\alpha$ визначається зі співвідношення $P\{U \geq u_{2,n-2,p}^\alpha\} = \alpha$, де U -максимум

абсолютних величин k випадкових змінних мають t -розподіл Стьюдента з $n-2$ ступенями свободи, кореляція між якими дорівнює ρ [Fisz, 1966]. Обчислені 95% -ні довірчі інтервали для коефіцієнтів регресії α , β вийшли рівними: [2.179; 2.358] і [0.530; 0.5718] відповідно. Величина r^2 – коефіцієнт детермінації – характеризує ступінь відповідності апроксимуючої лінії вибірковими даними. Отже, передбачені значення магнітуди mb пояснює 77.89% коливань оцінок магнітуди MS .

Середня похибка оцінки коефіцієнта кореляції $E(r)$ обчислюється за формулою [Себер, 1980]:

$$E(r) = \frac{1-r^2}{\sqrt{n-p-1}} \quad (3.9)$$

де n – обсяг вибірки, p – число параметрів регресії і становить: $E(R)=0.0070$. Якість регресії можна оцінити за залишковим середньоквадратичним відхиленням, обчисленим за формулою [Елисеєва, Юзбашев, 2004]:

$$\sigma_{y/x} = \sigma_y \sqrt{(1-r^2)} \quad (3.10)$$

где σ_y – залишкове середньоквадратичне відхилення залежної змінної, r – коефіцієнт кореляції. Ця оцінка абсолютна і, отже, не можуть бути порівняні різні лінії регресії одна з одною. Тому оцінюється відносна похибка рівнянь, які виражаються в процентах і служать для точності передбачення результатів одного показника по заздалегідь відомим значенням іншого. Відносна похибка лінійної регресії дорівнює $\delta=10.21\%$ визначалася за формулою:

$$\delta_{y/x} = \frac{\sigma_y \sqrt{(1-r^2)}}{\bar{y}} 100\% \quad (3.11)$$

Залишкове середньоквадратичне відхилення характеризує коливання Y відносно лінії регресії по X , отже, чим менше відносна похибка рівняння регресії, тим точніше прогноз значень залежної змінної за відомими значеннями незалежної змінної. Відносна похибка $\delta=10.21\%$ є незначною. Отже, регресія значень магнітуд mb на MS має вигляд: $mb=2.266+0.5508MS$.

Регресія магнітуди mb на MW

Коефіцієнти: регресії $\alpha=0.660$, $\beta=0.8339$, кореляції $r=0.9006$. Відносна похибка виявилася незначною $\delta=6.98\%$. Передбачені значення магнітуди mb пояснює 77.89% коливань оцінок магнітуди MW . 95% -ні довірчі інтервали для коефіцієнтів регресії α , β рівні: $[0.287; 1.0326]$ і $[0.760; 0.9075]$, відповідно. За критерієм Стюдента виявлено значимість регресії і кореляції між магнітудами mb і MW . Регресія значень магнітуд mb на MW має вигляд: $mb=0.660+0.8339MW$.

Регресія магнітуди mb на MD

Коефіцієнти: регресії $\alpha=0.599$, $\beta=0.8585$, кореляції $r=0.7640$. Відносна похибка виявилася незначною $\delta=8.21\%$. Передбачені значення магнітуди mb пояснює 58.37% коливань оцінок магнітуди MD . 95% -ні довірчі інтервали для коефіцієнтів регресії α , β рівні: $[0.017; 1.1814]$ і $[0.711; 1.0063]$ відповідно. За критерієм Стюдента виявлено значимість регресії і кореляції між магнітудами mb і MD . Регресія має вигляд:

$$mb=0.599+0.8585MD.$$

Регресія магнітуди mb на ML

Коефіцієнти: регресії $\alpha=0.665$, $\beta=0.8788$, кореляції $r=0.8767$. Відносна похибка виявилася незначною $\delta=7.53\%$. Передбачені значення магнітуди mb пояснює 76.87% коливань оцінок магнітуди MD . 95% -ні довірчі інтервали для коефіцієнтів регресії α , β рівні: $[0.354; 0.9759]$ і $[0.808; 0.9496]$, відповідно. За критерієм Стюдента виявлено значимість регресії і кореляції між магнітудами mb і ML . Вибіркове значення t -статистики одне $T=28.1796$. Рівняння регресії має вигляд: $mb=0.665+0.8788ML$.

Довірчим інтервалом Уоркінга-Хотеллінга для передбачених значень ϵ :

$$y = \bar{Y} + \bar{\beta}(x - \bar{x}) \pm \lambda S \left\{ \frac{1}{n} + \frac{(x - \bar{x})^2}{\sum_{i=1}^n (x_i - \bar{x})^2} \right\}^{1/2} \quad \lambda = \sqrt{2F_{2,n-2}^\alpha} \quad (3.12)$$

Для застосування цієї формули інтервал (a, b) перетворюється в стандартний варіант [Себер, 1980]:

$$c = \frac{(b-a)s_x}{\left\{s_x^2 + (a-\bar{x})^2\right\} \left\{s_x^2 + (b-\bar{x})^2 + s_x^2 + (a-\bar{x})(b-\bar{x})\right\}}, \quad (3.13)$$

далі за таблицями знаходиться відповідне йому значення λ . При побудові довірчого інтервалу для передбачених значень магнітуди mb по магнітудами MS величина стандартного варіанту вийшла рівною: $c=4.9404$. Передбачалося, що діапазон можливих значень незалежної змінної дорівнює $[a,b]=[0;9.3]$.

3.5. Оцінка параметрів і дослідження мультиколінеарності та гомоскедастичності в лінійній багатовимірній регресії

Багатовимірна лінійна модель регресії має вигляд:

$$y = \alpha + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_i x_i + \dots + \varepsilon \quad (3.14)$$

де x_i , ($i = 1, \dots, p$) – регресорів. Як регресорів можна розглядати магнітуду, глибину, механізм вогнища або деякі функції значень магнітуд. Якщо є вибірка обсягу n , то отримаємо n рівнянь

$$\begin{aligned} y_1 &= \alpha + \beta_1 x_{11} + \dots + \beta_{p1} + \varepsilon_1 \\ y_2 &= \alpha + \beta_1 x_{12} + \dots + \beta_{p2} + \varepsilon_2 \\ &\dots \\ &\dots \\ y_n &= \alpha + \beta_1 x_{1n} + \dots + \beta_{pn} + \varepsilon_n \end{aligned} \quad (3.15)$$

У векторній формі: $y = X\beta + \varepsilon$

$$\begin{pmatrix} y_1 \\ \cdot \\ \cdot \\ \cdot \\ y_n \end{pmatrix} = \begin{pmatrix} 1 & x_{11} & \cdot & x_{p1} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ 1 & x_{1n} & \cdot & x_{pn} \end{pmatrix} \begin{pmatrix} \beta_1 \\ \cdot \\ \cdot \\ \beta_n \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \cdot \\ \cdot \\ \cdot \\ \varepsilon_n \end{pmatrix} \quad (3.16)$$

Оцінка найменших квадратів вектора параметрів в матричній формі має вигляд:

$$\bar{\beta} = (X^T X)^{-1} X^T Y \quad (3.17)$$

Залишкова сума квадратів обчислюється з виразу:

$$\varepsilon^T \varepsilon = Y^T Y - \beta^T \bar{X}^T X \bar{\beta} \quad (3.18)$$

де символ T -позначає операцію транспонування векторів і матриць.

У науковій літературі зустрічаються роботи, присвячені застосуванню регресійного аналізу в сейсмології, наприклад, [Grünthal, Wahlstrom; 2003; Vakov, 1996; Stromejer et al., 2004]. Однак, в них не розглядається питання виконання умов мультиколінеарності та гомоскедастичності, необхідних для отримання незміщених, ефективних і достатніх оцінок параметрів регресії. У разі гетероскедастичності, тобто відсутності сталості дисперсії залишків регресії, оцінки параметрів регресії є лінійними і незміщеними, але вони не будуть ефективними, і дисперсії оцінок параметрів будуть мати зміщення.

Мультиколінеарність означає високу взаємну кореляцію змінних регресії. Відсутність колінеарності регресорів є однією з умов застосування методу найменших квадратів для оцінки параметрів багатовимірної лінійної регресії. У практичних додатках, при побудові рівняння регресії, може спостерігатися висока кореляція змінних. Колінеарність може з'являтися і при агрегуванні даних. Вважається, що два регресори x і z колінеарні, якщо парний коефіцієнт кореляції $r_{xz} > 0.7$ [Габріелян, 2006]. Чим сильніше мультиколінеарність, тим менш надійна оцінка параметрів регресії (3.17). Коефіцієнти регресії (3.17) існують тоді і тільки тоді, коли визначник матриці $(X^T X)$ відмінний від нуля, тобто стовпці матриці X незалежні. Кількісних критеріїв для виявлення мультиколінеарності не існує. Ознаками наявності мультиколінеарності є: близькість до нуля визначника матриці парних коефіцієнтів кореляції; мінімального власного значення λ_{min} матриці $(X^T X)$; велике значення відношення максимального власного значення λ_{max} до мінімального власного значення λ_{min} матриці $(X^T X)$; високе значення коефіцієнта детермінації:

$$R^2 = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (3.19)$$

де n – число спостережень; y_i – спостережені значення, \hat{y}_i – модельні значення і \bar{y} – середнє арифметичне значення залежної змінної.

Мультиколінеарність може проявлятися в функціональній та стохастичній формах. При функціональній формі мультиколінеарності, по крайній мірі, один з парних зав'язків між регресорами є функціональним. В цьому випадку матриця $(X^T X)$ особлива, тому що містить лінійно залежні вектори: стовпці, і її визначник дорівнює нулю, тобто порушується передумова регресійного аналізу, що робить неможливим вирішення відповідної системи нормальних рівнянь, і, застосування методу найменших квадратів для оцінки параметрів регресійної моделі.

В якості початкової математичної моделі розглядається регресійні рівняння:

$$mb = \beta_0 + \beta_1 MS + \beta_2 h + \varepsilon \quad (3.20)$$

де MS – значення магнітуди, h – глибина; ε – похибка; $\beta_0, \beta_1, \beta_2$ – параметри регресії. Далі в рівняння додавалися нові регресори – ступеня значень MS . Відповідно до методу найменших квадратів, параметри регресії оцінюються за формулою [Себер, 1980; Welkowitz et al., 1988; Лещинський та ін., 2003]:

$$\hat{\beta} = (X^T X)^{-1} X^T Y \quad (3.21)$$

де X – матриця значень регресорів. Аналіз матриці парних коефіцієнтів кореляції між усіма факторами з таблиць 1–2 показує, що між магнітудою mb і магнітудою MS існує тісний кореляційний зв'язок. Кореляційний зв'язок між глибиною вогнища і магнітудами mb і MS слабка.

Шукані регресивні рівняння мають вигляд

$$mb = 1,047MS + 0,007h, \quad mb = 2,134MS + 0,001h - 0,358MS^2 + 0,026MS^3 \quad (3.22)$$

Обчислення теоретичних значень залежної змінної можна виконувати в матричній формі:

$$\hat{Y} = X\beta \quad (3.23)$$

де X – регресійна матриця; β – вектор коефіцієнтів регресії.

Перевірка адекватності регресійної моделі проводиться за схемою [Лещинський та ін., 2003; Габріелян, 2006]:

1. відносна похибка обчислених значень і середнє значення відносних похибок визначаються за формулами:

$$\delta_i = \frac{\hat{u}_i}{y_i}, \quad \bar{\delta} = \frac{\sum_{i=1}^n \delta_i}{n} \quad (3.24)$$

де $\hat{u}_i = y_i - \hat{y}_i$.

Таблиця 3. 1.

Матриця парних кореляцій

<i>MS</i>	<i>H</i>	<i>MS</i> ²	<i>MS</i> ³	<i>Mb</i>
1.0	0.007	0.991	0.964	0.881
0.007	1.0	-0.008	-0.021	0.039
0.991	-0.008	1.0	0.991	0.878
0.964	-0.021	0.991	1.0	0.858
0.881	0.039	0.878	0.858	1.0

середнє значення відносних похибок $\bar{\delta} = 0.03391$ є незначним. Дисперсія залишків регресії (3.22) оцінювалася за формулою:

$$\sigma_u^2 = \frac{\sum (y_i - \hat{y}_i)^2}{n - m - 1} \quad (3.25)$$

Значення середньоквадратичної похибки дисперсії залишків $\sigma_u = 0.587$ виявилася незначною, що вказує на адекватність регресії.

2. значення вибіркової статистики розподілу Стьюдента $t_q = 28.790$ і $t_q = 78.391$ більше табличного значення $t_{0,025,1012} = 1.96$, що вказує на значущість коефіцієнта множинної кореляції між магнітудою *mb* і регресорами.
3. для перевірки значущості коефіцієнтів регресії обчислювалася вибіркова статистика *t*-розподілу:

$$t_j = \frac{\beta_j}{\sqrt{\sigma_u^2 c_{jj}}} \quad (3.26)$$

де σ_u^2 – дисперсія залишків (3.23); c_{jj} – діагональні елементи матриці $(X^T X)^{-1}$.
Вибіркові статистики: $t_1 = 59.26$; $t_2 = 2.18$; $t_1 = 56.384$; $t_2 = 2.249$; $t_3 = -22.245$; $t_4 = 15.712$, для рівнянь (3.22) відповідно, більше табличного значення $t_{0,025, 1012} = 1,96$. Отже, параметри моделі є значущими.

4. значення коефіцієнта детермінації (3.19) вийшли рівними: $R^2=0.450$, $R^2=0.763$ – для рівнянь (3.22), відповідно. Внесок нового додаткового регресора визначається корегованим коефіцієнтом детермінації Тейла R_T^2 і Амеція R_A^2 [Лещинский та ін., 2003]:

$$R_T^2 = 1 - (1 - R^2) \frac{n-1}{n-m-1}; \quad R_A^2 = 1 - (1 - R^2) \frac{n-1}{n-m-1} \quad (3.27)$$

Скоригований коефіцієнт детермінації в деякій степені збільшується в міру збільшення числа регресорів, в порівнянні зі звичайним коефіцієнтом детермінації, і продовжує рости при введенні нових регресорів, поки t-статистика для цих регресорів більше одиниці [Лещинський та ін., 2003]. При великому обсязі вибірки і невеликому числі регресорів ефект, пов'язаний з включенням в регресію нової змінної, повинен бути незначним, що і підтвердилося. Відповідні коефіцієнти кореляції мають значення: $R_1=0.671$; $R_2=0.874$. Високі значення коефіцієнтів детермінації і кореляції при регресійному рівнянні (3.22) вказують на те, що 76.33% зміни значень магнітуди m_b пояснюються регресією. При $m=4$ і $n=1003$, значення скоригованих коефіцієнтів детермінації досить великі, які складають відповідно:

$R_T^2 = 0.762$, $R_A^2 = 0.761$, і трохи відрізняються від звичайного коефіцієнта детермінації, що також є показником якості моделі. Знайдемо граничні значення вкладу j -го регресора, в відповідний коефіцієнт регресії:

$$\Delta R_j^2 = \frac{(1 - R^2) t_j^2}{n - m - 1} \quad (3.28)$$

$\Delta R_1^2 = 0.754$, $\Delta R_2^2 = 0.332$, $\Delta R_3^2 = 0.747$, $\Delta R_4^2 = 0.748$. Серед рівнянь регресії, з різним числом змінних, при однакових значеннях інших показників якості регресії, перевага віддається рівнянням з великим значенням скоригованого коефіцієнта детермінації. Тобто, для порівняння регресії, з різним числом змінних використовується скоригований коефіцієнт детермінації, тому що нескоригований коефіцієнт збільшується з додаванням нових регресорів, хоча поліпшення якості в регресійній моделі може не відбутися. Скоригований коефіцієнт детермінації може

зменшуватися при введенні в модель нових регресорів, які не роблять значущий вплив на залежну змінну.

Вибіркові статистики F -розподілу $F_q=414.420$ і $F_q=1563.63$ перевершують табличні значення $F_{0.05,2,1012}=3.00$ і $F_{0.05,4,1012}=2.37$. Отже, регресія в обох випадках є значущою.

5. коваріаційна матриця оцінок параметрів регресії визначається з формули:

$$\text{var}(\hat{\beta}) = \sigma_u^2 (X^T X)^{-1} \quad (3.29)$$

де σ_u^2 – дисперсія залишків (3.26). Діагональні елементи коваріаційної матриці: 0.0378; 0.00486; 0.0161; 0.00165 є вибірковими дисперсіями оцінок параметрів регресії (3.22).

6. якість прогнозу можна оцінити за середньою похибкою апроксимації за формулою:

$$Q = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| 100\% \quad (3.30)$$

Значення середньої похибки (3.25) для рівнянь (3.22) і (3.23), що становлять 10,55%; і 5,21%, відповідно, вказують на адекватність моделі. Якщо величина Q менше 10%, то регресійна модель вважається статистично значущою.

7. ще однією оцінкою середньої похибки апроксимації магнітуди mb є вираз [Себер, 1980; Лещинский та ін., 2003]:

$$A = \frac{1}{n} \sum_{i=1}^n \frac{y_i - \hat{y}_i}{y_i} 100\% \quad (3.31)$$

Незначна середня похибка для обох регресій: 3.391% і -0.384%, також свідчить про значущість регресій.

Якщо мірою відповідності моделі лінійної регресії вибірковими даними є множинний коефіцієнт детермінації R^2 , то мірою значущості кожного з регресорів може бути його коефіцієнт кореляції з залежною змінною. Якщо сукупність регресорів впорядкувати по модулю коефіцієнта кореляції:

$$|R(y, x_{i_1})| \geq |R(y, x_{i_2})| \geq \dots \geq |R(y, x_{i_p})| \quad (3.32)$$

то з цього ряду можна відібрати деяке число значущих регресорів. У рівняння лінійної багатовимірної регресії додавалися степені значень магнітуди і глибини вогнища землетрусів. Коефіцієнт парної кореляції між степенями значень магнітуди MS більше 0.7 – порогової величини, при перевищенні значенням коефіцієнта кореляції між регресорами, якої мультиколінеарність вважається встановленою.

Високі значення парних коефіцієнтів кореляції, незначимість коефіцієнтів регресії, виявлена по t -критерію, близькість до нуля детермінанта кореляційної матриці регресорів є достатніми для наявності мультиколінеарності умовами. Високе значення відношення $\lambda_{\max}/\lambda_{\min}$ максимального і мінімального власних значень матриці $(X^T X)$, так само вказує на наявність мультиколінеарності.

Точних кількісних критеріїв у визначенні мультиколінеарності не існує. Проте, є деякі евристичні підходи по її виявленню. Найбільш повним методом дослідження мультиколінеарності є алгоритм Фаррара-Глобера.

3.5.1. Метод Фаррара - Глобера

Досліджується три види мультиколінеарності:

- всіх регресорів
- кожного регресора з іншими
- кожної пари регресорів.

Для дослідження мультиколінеарності нормуються значення регресорів $X_j, j=1, \dots, m$ [Лещинский та ін., 2003]:

$$X_{ij}^H = \frac{x_{ij} - \bar{x}_j}{\sqrt{n\delta_j^2}}, \quad (3.33)$$

де n – число спостережень; $m=4$ число регресорів, \bar{x}_j – вибіркове середнє арифметичне і δ_j – середнє квадратичне відхилення j -й змінної в рівнянні (3.22).

а) визначається кореляційна матриця:

$$R = (X^H)^T X^H \quad (3.34)$$

б) вибіркове значення $\chi_q=114,563$ критерію:

$$\chi^2 = -(n-1 - \frac{1}{6}(2m+5)) \ln|R| \quad (3.35)$$

перевершує табличне значення $\chi_{tab}=12.592$, що відповідає рівню значущості $\alpha=0.05$ з числом ступенів свободи $df=0.5m(m-1)=6$. Отже, в масиві регресорів існує мультиколінеарність.

с) мультиколінеарності кожного регресора з іншими визначаються за критерієм Фішера [Лещинський та ін., 2003; Наконечний та ін., 2004]. Для цього визначається зворотна матриця $D=R^{-1}$, де R кореляційна матриця. Вибіркові значення F-критерію Фішера визначаються з виразу:

$$F_k = (d_{kk}-1) \frac{n-m}{m-1} \quad (3.36)$$

де d_{kk} – діагональні елементи матриці D .

Вибіркові значення $F_1 = 441.973$; $F_2=4.51$; $F_3 = 1794.78$; $F_4=482.248$ порівнюються з табличним значенням $Fq=8,53$, що відповідає рівню значущості $\alpha=0.05$ F- критерію зі ступенями свободи: $df_1=1012$, $df_2=3$.

Порівняння показує, що регресорів MS , MS^2 , MS^3 колінеарні з іншими, а глибина вогнища землетрусів не корелює з регресорами: MS ; MS^2 ; MS^3 .

д) для виявлення мультиколінеарності кожної пари змінних обчислюються парні коефіцієнти кореляції:

$$r_{ij} = \frac{-d_{ij}}{\sqrt{d_{ii}d_{jj}}} \quad (3.37)$$

Для перевірки гіпотези про значущість кореляції між кожною парою регресорів використовується t - критерій:

$$t_{ij} = \frac{|r_{ij}| \sqrt{n-m}}{\sqrt{1-r_{ij}^2}} \quad (3.38)$$

Вибірковими значеннями парних кореляції $r_{12}=0.034$, $r_{13}=0.994$, $r_{14}=-0.979$, $r_{23}=-0.023$, $r_{24}=0.012$, $r_{34}=0.993$, $r_{12}=0.034$ відповідають значення t -критерію $t_{12}=1.1$; $t_{13}=301.2$; $t_{14}=153.8$; $t_{23}=0.7$; $t_{24}=0.4$; $t_{34}=286.1$. З порівняння значень t -статистики з табличним значенням $t_{0,05,1010}=1.96$ [Welkowitz et al., 1988], що відповідає рівню значущості $\alpha = 0.05$ с числом ступенів свободи $df=1010$, знаходимо, що t_{12} , t_{23} , t_{24} менше табличного значення, отже, глибина не корелює зі ступенями магнітуд, і ясно, що ступені магнітуди колінеарні між собою.

е) Коефіцієнти детермінації для кожного з регресорів обчислюються за формулою:

$$R_k^2 = 1 - \frac{1}{d_{kk}} \quad (3.39)$$

Якщо значення коефіцієнта детермінації наближається до одиниці, то, відповідна змінна є мультиколінеарною з іншими змінними. Серед значень коефіцієнтів детермінації: $R_1^2 = 0.999$; $R_2^2 = 0.013$; $R_3^2 = 1.0$; $R_4^2 = 0.999$ – тільки значення, відповідне глибини є незначними. Отже, змінні MS , MS^2 , MS^3 мультиколінеарні з іншими. Виключенням з регресії колінеарних регресорів зменшується число параметрів моделі.

Для оцінки параметрів моделі, в яку входять мультиколінеарні змінні, застосовується метод головних компонент.

3.5.2. Метод головних компонент

Алгоритм виконується за наступний схемою:

1. Нормалізація значень регресорів виконується за формулою:

$$x_{ij}^H = \frac{x_{ij} - \bar{x}_j}{\delta_j} \quad (3.40)$$

де x_j – середнє арифметичне; δ_j – середнє квадратичне відхилення j -ої змінної регресії.

2. Кореляційна матриця нормалізованих регресорів обчислюється множенням матриць:

$$R = \frac{1}{n} (X^H)^T X^H \quad (3.41)$$

З рішення характеристичного рівняння $|R - \lambda E| = 0$ визначаються власні числа кореляційної матриці [Аллен, 1963; Райнш, 1976; Райс, 1984].

3. Власні числа упорядковуються в порядку зменшення: $\lambda_1=2.961$; $\lambda_2=0.999$; $\lambda_3=0.035$; $\lambda_4=0.00036$.

Таблиця 3.3

Кореляційна матриця нормалізованих регресорів

MS	H	MS^2	MS^3
0.999	0.0074	0.9895	0.9633
0.0074	0.999	-0.0077	-0.0215
0.9895	-0.0077	0.999	0.9904
0.9633	-0.0215	0.9904	0.999

4. Обчислюються відповідні власні вектори кореляційної матриці:

$$R = \begin{pmatrix} 0.5755 & 0.0189 & 0.7123 & -0.4014 \\ -0.0064 & 0.9998 & -0.0212 & 0.0003 \\ 0.5808 & 0.0033 & -0.0107 & 0.8140 \\ 0.5757 & -0.0111 & -0.7015 & -0.4199 \end{pmatrix} \quad (3.42)$$

5. Визначаються головні компоненти – вектори $z_k = X^H \cdot c_k$, где c_k – власний вектор, що відповідає власному значенню λ_k , $k=1, \dots, m$. Вектори матриці головних компонент, повинні задовольняти умовам [Наконечный, 2004]:

$$\sum_{j=1}^n z_{ij} = 0; \quad i=1, \dots, m; \quad \frac{1}{n} Z_k^T Z_k = \lambda_k; \quad Z_i^T Z_k = 0; \quad i = \overline{1, m}, i \neq k. \quad k = \overline{1, m} \quad (3.43)$$

6. Параметри моделі визначаються з рішення матричного рівняння:

$$\widehat{Y} = Z\widehat{b}, \quad \widehat{b} = Z^{-1}Y \quad (3.44)$$

7. Параметри багатовимірної регресії, після усунення ефекту мультиколінеарності, визначаються з рівняння:

$$mb = \beta_1 MS + \beta_2 h + \beta_3 MS^2 + \beta_4 MS^3, \quad \widehat{\beta} = A\widehat{b}, \quad \widehat{\beta} = (\beta_1, \beta_2, \beta_3, \beta_4) \quad (3.45)$$

Шукане регресійне рівняння має вигляд:

$$mb = 2.289MS + 0.001h - 0.423MS^2 + 0.032MS^3 \quad (3.46)$$

8. Далі основна вибірка обсягом в 1015 спостережень розбивалася на дві підвибірki зі значеннями магнітуди більше і менше: 5; 5,5; 6-відповідно. Загальна регресійні рівняння складається з двох ліній регресії:

$$mb < 5 : mb = 2.182MS + 0.002h - 0.383MS^2 + 0.029MS^3 \quad (3.47)$$

$$mb \geq 5 : mb = 1.771MS + 0.003h - 0.13MS^2 + 0.0004MS^3 \quad (3.48)$$

$$mb < 5.5 : mb = 2.182MS + 0.002h - 0.383MS^2 + 0.029MS^3 \quad (3.49)$$

$$mb \geq 5.5 : mb = 1.771MS + 0.003h - 0.13MS^2 + 0.0004MS^3 \quad (3.50)$$

$$mb < 6 : mb = 2.182MS + 0.002h - 0.383MS^2 + 0.029MS^3 \quad (3.51)$$

$$mb \geq 6 : mb = 1.771MS + 0.003h - 0.13MS^2 + 0.0004MS^3 \quad (3.52)$$

Максимальне відхилення вибірових значень магнітуди mb від їх теоретичних значень, обчислених по регресійній моделі (3.47–3.52) становить $\delta_{\max}=0.822$. Максимальне відхилення вибірових значень магнітуди mb від їх теоретичних значень, обчислених по регресійній моделі (3.47–3.52) дорівнювало $\delta_{\max}=0.822$. Якість регресії, яка виражається середньою похибкою апроксимації (3.30) для рівнянь (3.47–3.52), не перевищує 10%-ні значення, відповідає значущому зв'язку між залежною і незалежною змінними. При $mb=6$ спостерігається краще значення якості регресії Q (3.30) для обох підвбірок основної вибірки: 5.18%; 1.46%. Таким чином, параметри рівняння регресії магнітуди mb на MS залежать від діапазону значень магнітуди.

Адекватність моделей (3.46–3.52) досліджувалася за схемою, викладеної в параграфі 3.5. Всі моделі регресії виявилися статистично значущими. Після усунення ефекту мультиколінеарності збільшився коефіцієнт множинної кореляції, і покращилася якість регресії Q (3.30).

3.5.3. Дослідження гомоскедастичності в регресійному аналізі

Гомоскедастичність, це сталість дисперсії залишків регресії. При відсутності сталості дисперсії залишків, тобто при гетероскедастичності, оцінки параметрів регресії будуть лінійними і незміщеними, але вони не будуть ефективними і дисперсії оцінок параметрів можуть мати зміщення.

В якості початкової моделі розглядається регресійне рівняння:

$$mb = \beta_0 + \beta_1MS + \beta_2h + \varepsilon \quad (3.53)$$

де MS – магнітуда; h – глибина; $\beta_0, \beta_1, \beta_2$ – параметри регресії. Далі в рівняння додавалися нові регресори - ступеня значень MS .

Застосування методу найменших квадратів в багатовимірній лінійній регресійній моделі передбачає виконання таких умов [Мхітарян та ін., 2008]:

1. Математичне сподівання випадкової складової дорівнює нулю: $M(\varepsilon)=0$
2. Компоненти вектора залишків не корельовані і мають постійну дисперсію:

$$M(\varepsilon^T \varepsilon) = \sigma^2 E$$

3. Регресори не корельовані із залишками
4. Регресори неколінеарні між собою

Наслідками гетероскедастичності (відсутність гомоскедастичності):

- 1) неспроможність оцінок коефіцієнтів регресії;
- 2) заниження оцінок дисперсії
- 3) зміщення оцінок коефіцієнтів
- 4) зміщення оцінок дисперсії коефіцієнтів.

Параметри регресії оцінювалися за формулою [Себер, 1980]: Параметри регресії оцінювалися за формулою [Себер, 1980]:

$$\hat{\beta} = (X^T X)^{-1} X^T Y \quad (3.54)$$

де X – матриця значень регресорів, Y – залежна змінна.

Як приклад розглядається регресія mb на магнітуду MS і глибину вогнища землетрусу h :

$$mb = 2.289MS + 0.001h - 0.423MS^2 + 0.032MS^3. \quad (3.55)$$

Дослідження гетероскедастичності проводилося за допомогою тесту рангової кореляції Спірмена і методу Глейзера [Габриелян, 2008; Наконечный та ін., 2004].

3.5.4. Тест рангової кореляції Спірмена

Значення залишків регресії обчислюються як різниці між тими, що спостерігаються і передбаченими регресією значеннями магнітуд:

$$\varepsilon_j = y_j - \hat{y}_j \quad (3.56)$$

де y_j – вибіркоче значення залежної змінної, \hat{y}_j – відповідне теоретичне значення. Передбачається, що абсолютні значення залишків (3.56) і значення регресорів некорельовані. Якщо в моделі кілька регресорів, то перевірка гіпотези про

гомоскедастичність проводиться окремо для кожного з них. Для застосування методу Спірмена значення залишків та регресорів упорядковуються по зростанню. Потім обчислюється коефіцієнт рангової кореляції:

$$R(x^k, \varepsilon) = 1 - 6 \frac{\sum_{i=1}^n \delta_i^2}{n(n^2 - 1)} \quad (3.57)$$

де δ_i – різниця між рангами значень залежної змінної або деякого з регресорів x^k , ($k=1, \dots, p$) і значеннями вектора залишків ε . В даному прикладі обчислюється рангова кореляція між значеннями залишків і значеннями магнітуди MS і між значеннями залишків глибини h землетрусів. Для оцінки значущості коефіцієнта рангової кореляції $r=0.036$ використовується t -критерій [Мхитарян и др., 2008; Габриелян, 2008]. Якщо коефіцієнт кореляції дорівнює нулю, то статистика:

$$t = \frac{R(x, \varepsilon) \sqrt{n-2}}{\sqrt{1 - R(x, \varepsilon)^2}} \quad (3.58)$$

має розподіл Стюдента з числом ступенів свободи $\mu=n-2$. Статистика t -критерію $t_q=1.149$, менше табличного значення $t_{0,05,1013}=1.96$, t -розподілу, що відповідає рівню значущості $\alpha=0.05$, з числом ступенів свободи $\mu=1013$. Отже, між значеннями відхилень і значеннями магнітуди MS не існує кореляційного зв'язку. Гіпотеза $H_0: R(h, \varepsilon)=0$, про те, що коефіцієнт рангової кореляції між значеннями глибини вогнищ землетрусів і відхиленнями дорівнює нулю, також виявилася правдоподібною, тому що відповідне значення критерію Стюдента: $t_q=1.495$ менше табличного значення.

3.5.5. Тест Глейзера

У методі Глейзера визначається регресійна залежність абсолютних значень відхилень $|\varepsilon_i|$ від регресорів [Наконечний та ін., 2004]. Розглядається регресійна модель:

$$|\varepsilon_i| = \alpha + \beta x_i + v_i \quad (3.59)$$

Статистична значимість коефіцієнта регресії β означає наявність гетероскедастичності. У регресії залишків $|\varepsilon_i|$ на значення магнітуди MS коефіцієнти вийшли рівними: $\alpha=0.19$; $\beta=0.00740$. Статистика t-критерію $t_q=1.254$, менше табличного значення $t_{0,05,1013}=1.96$, t-розподілу, що відповідає рівню значущості $\alpha=0,05$, з числом ступенів свободи $\mu=1013$, свідчить про незначущість коефіцієнта регресії β , і на відсутність регресії відхилень ε_i на значення магнітуди MS. Гіпотеза про статистичну значущість регресійної залежності між значеннями відхилень і глибини вогнищ землетрусів, також не підтвердилася. Отже, умова сталості дисперсії відхилень виконується, і лінійна багатовимірна регресія може ефективно застосовуватися для перетворення магнітуд.

3.5.6. Дослідження однорідності сейсмічних даних

Найпростіших мірою однорідності статистичної сукупності є коефіцієнт варіації (3.1), який становить 39.4% для значень магнітуд усіх землетрусів з каталогу ROMPLUS. Для магнітуд усіх землетрусів зони Вранча коефіцієнт варіації становить 32.7%, а 24.4% для магнітуд проміжних і 44.5% для землетрусів в межах земної кори. Сукупність вважається однорідною, якщо коефіцієнт варіації не перевищує 33%, тобто тільки сукупність значень магнітуд проміжних вранчських землетрусів є однорідною. Однорідною, за значенням коефіцієнта варіації, який становить 14.8%, виявилася сукупність з 9200 значень магнітуд з каталогу турецьких землетрусів, де наводяться значення: *mb*, *MS*, *MD*, *ML*, *MW* магнітуд для всіх 9200 сейсмічних подій. Однак, за допомогою статистики Dip-test критерію Хартігана, гіпотеза про унімодальний розподіл значень магнітуд всіх вибірок відкидається при рівні значущості 0.05. Аналіз форми гістограми, в якому є кілька піків, також дозволяє робити висновки про мультимодальність вибірки. Відсутність унімодального розподілу вказує на те, що дані вибірки складаються з суміші двох і більше генеральних сукупностей, що є само собою зрозумілим, враховуючи принципи складання каталогів.

Землетруси є відбиток геологічних процесів, що відбуваються в земних надрах, і виникають в абсолютно різних за структурою і характером геологічного розвитку районах Землі. Сейсмічність залежить від інтенсивності тектонічного процесу; від розмірів блоку земної кори, в якому можуть накопичуватися пружні напруги; від фізико-хімічних умов на різних глибинах надр Землі; від межі міцності і межі плинності гірських порід на цих глибинах і від особливостей геологічної будови території. Каталоги землетрусів, складені за географічною ознакою, якими є ROMPLUS і каталог турецьких землетрусів, містять значення параметрів землетрусів всіх сейсмічних зон на даній території. У різних сейсмічних зонах, геолого-геофізичні процеси, що призводять до землетрусів, відбуваються по-різному. Тому, є некоректним застосування методів математичної статистики, до масиву даних складених за відомостями про параметри сейсмічності різних сейсмічних зон.

Методами теорії ймовірностей можна вивчати ті випадкові події, які можуть бути відтворені в одних і тих же умовах. Тобто, при здійсненні фізичного експерименту, пов'язаного з вимірами значень параметрів, умови, в яких проводиться експеримент, не повинні змінюватися. Особливо істотно з'ясування умов незмінності для експериментів, які проводяться протягом тривалого часу. У геофізичних дослідженнях, експерименти ставить природа, і дослідник приймає пасивну участь в цих дослідженнях, яке обмежується: підготовкою та обробкою даних; адекватною постановкою завдання і підбором відповідного математичного апарату. Повертаючись до перетворення магнітуд, після виявлення наявності в основній вибірці суміші двох і більше генеральних сукупностей потрібно визначити межі відповідних у підвибірках розподілів. Одним з критеріїв визначення меж є:

$$T = \max_{t_1, t_2, \dots, t_k} SC(t_1, t_2, \dots, t_k) \quad (3.60)$$

$$SC(t_1, t_2, \dots, t_k) = 1 - \frac{Var(0, t_1) + Var(t_1, t_2) + \dots + Var(t_{k-1}, t_k)}{Var(0, t_k)} \quad (3.61)$$

де t_1, t_2, \dots, t_k – шукані значення границі розподілів.

3.6. Визначення мінімальної магнітуди комплектності

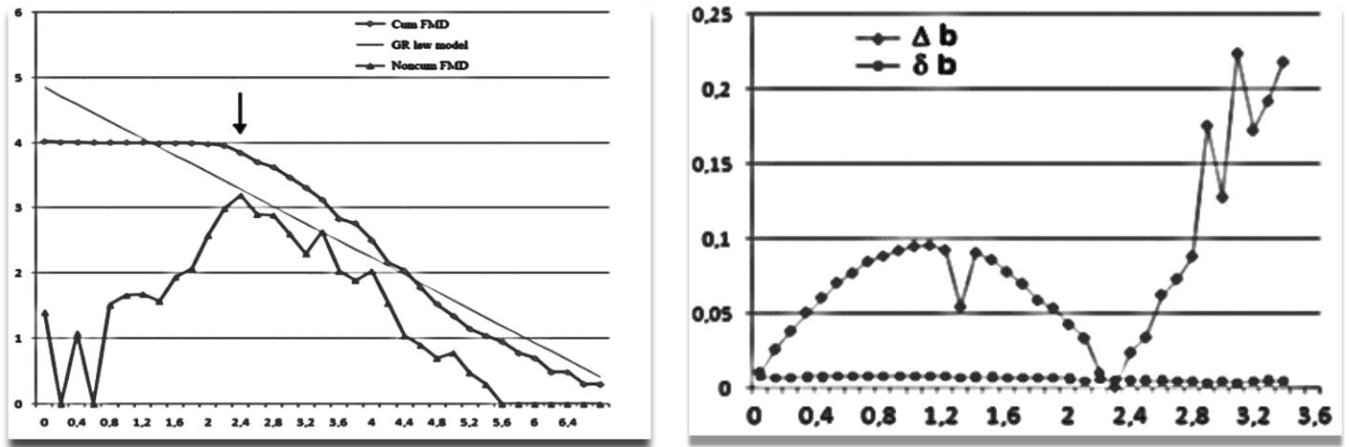
Сейсмічність Карпатського регіону складається з корової і подкорової сейсмічності. Головним генератором землетрусів даного регіону є зона "Вранча", сейсмічний режим якого детально досліджений [Алказ, 2007; Онофреш, 1981; Balan, 1982; Ismail-Zadeh, et al., 1999, 2006, 2008, 2012; Pora 2009]. Щодо середньої глибини кордону M існує думка, що вона становить 60 км [Pora 2009]. Каталог землетрусів зони "Вранча" містить дані про землетруси, починаючи з кінця X століття [www.infp.ro/arhiva].

Мінімальна магнітуда комплектності каталогу є важливим фактором у дослідженні сейсмічності. За відомостями з каталогів землетрусів будуються моделі сейсмічного режиму зони вогнищ землетрусів, і проводиться аналіз сейсмічної небезпеки. Якість результатів залежить від наглядності каталогу, тобто від мінімальної магнітуди комплектності M_c [Ameer et al., 2005; Cao A., Cao S., 2002; Marzocchi, Sandri, 2003; Mignan et al., 2011; Mignan, Woessner, 2010; Schorlemmer, Woessner, 2008; Woessner, Wiemer, 2005; Wiemer, Wyss, 2003; Wiemer, 2001]. Магнітуда M_c – це нижній поріг значень магнітуди, починаючи з якого всі землетруси реєструються без пропусків. Методи оцінки M_c ґрунтуються на припущенні про взаємоподібність сейсмічних процесів. В даному випадку, з множини методів обчислення M_c використовується:

- 1) метод максимальної кривизни (МАХС) [Woessner, Wiemer, 2005]
- 2) метод, заснований на стійкості M_c при оцінці коефіцієнта b - нахилу графіка повторюваності (MBS) [Cao A., Cao S., 2002; Schorlemmer, Woessner, 2008].

Метод МАХС визначає точку найбільшої кривизни, як мінімальну магнітуду комплектності, при обчисленні максимальної величини першої похідної магнітудно-частотної кривої. Метод МАХС - швидкий і достовірний спосіб оцінки M_c , полягає в знаходженні інтервалу магнітуди з найбільшим значенням щільності магнітудно-частотного розподілу [Woessner, Wiemer, 2005; Дулепов, Лєскова 2009]. Досліджено

сукупність румунських землетрусів за період 1978-2013рр., що міститься в каталозі ROMPLUS.



a)

b)

Рис. 3.1. Визначення значення мінімальної магнітуди комплектності:

а) кумулятивний і некумулятивний магнітудно-частотний розподіл (FMD),

б) стандартне відхилення δb і різниця $\Delta b = |b - b_{ave}|$

c)

Величина b нахилу графіка повторюваності, обчислюється для значень магнітуд, що перевищують M_c . Кумулятивний і некумулятивний магнітудно-частотний розподіли (FMD) показані на (рис. 3.1).

MBS метод: M_c визначається як перше значення нижнього порогу магнітуди, для якої виконується умова $\Delta b = |b_{ave} - b| \leq db$ (db – похибка). Похибка обчислення нахилу графіка повторюваності db визначається за формулою:

$$\delta b = 2.3b^2 \sqrt{\frac{\sum_{i=1}^N (m_i - \bar{m})^2}{N(N-1)}} \quad (3.62)$$

Значення мінімальної магнітуди комплектності, оцінене для обох методів, збігається і становить $M_c = 2.3$. Середнє b_{ave} , рівне арифметичному середньому значень b , відповідних послідовним значенням нижнього порогу магнітуди, рівних половині довжини інтервалу магнітуд $dM = 0.4$:

$$b_{ave} = \frac{\sum_{M_{co}^i}^{N_c} b(M_{co}^i)}{N_c} \quad (3.63)$$

N_c – число значень нижнього порогу магнітуди, $b(M_{co})$ – b значення, що відповідають значенням порогових магнітуд:

$$M_c^1, M_c^2, M_c^3, \dots, M_c^{N_c} \quad (3.64)$$

Перший раз, при значенні $M_c=2.3$ нижнього порогу магнітуди, виконується умова $\Delta b \leq db$ ($\Delta b=0.0002$, $db=0.0053$) (рис. 3.2).

3.7. Статистичний зв'язок між коровою і підкоровою сейсмічною активністю вогнища "Вранча"

Мовою статистики, досліджувані властивості об'єктів (подій) називаються ознаками. Для аналізу ступеня можливого взаємозв'язку будь-яких двох ознак складається кореляційна таблиця спостережених частот цих ознак: таблиця спряженості ознак. Число рядків (r), і стовпців (s) дорівнює числу градацій першої і другої ознак, відповідно. У кожному осередку на перетині i -го рядка і j -го стовпця вказано число об'єктів n_{ij} (подій), що відносяться до i -ої градації за першою ознакою і до k -ої градації – по другому [Thorsten, 2006; Durner, 1999]. Перш ніж досліджувати можливу залежність ознак, повинна бути перевірена нульова гіпотеза – припущення попарної незалежності самих ознак. Нехай нульова гіпотеза H_0 стверджує, що x і y є взаємно незалежними випадковими величинами. Для перевірки гіпотези H_0 обчислюється статистика критерію [Крамер, 1975; Сиденко и др., 2011]:

$$\chi^2 = N \sum_{i=1}^r \sum_{j=1}^s \frac{(n_{ij} - \frac{n_i \cdot n_j}{N})^2}{n_i \cdot n_j}; \quad n_i = \sum_{j=1}^s n_{ij}; \quad n_j = \sum_{i=1}^r n_{ij} \quad (3.65)$$

χ^2 – розподілена випадкова величина з $(r-1)(s-1)$ ступенями свободи.

За величиною статистики критерій χ^2 приймається рішення про наявність зв'язку між двома змінними за допомогою таблиць спряженості. Якщо гіпотеза H_0 відкидається, це означає, що існує статистично значущий зв'язок між двома

змінними. Однак статистично значущий зв'язок не визначає тісноту зв'язку. В якості запобіжної взаємозалежності застосовуються коефіцієнти: спряженості, Крамера, кореляції, кореляційного відношення. Для оцінки тісноти зв'язку в таблицях будь-якого розміру використовується коефіцієнт спряженості ознак [Крамер, 1975]:

$$C = \sqrt{\frac{\chi^2}{\chi^2 + N}} \quad (3.66)$$

Для оцінки тісноти нелінійного кореляційної зв'язку використовується кореляційне відношення η_{yx} . Вибірковим кореляційним відношенням називається відношення міжгрупового середнього відхилення до загального середньоквадратичного відхилення:

$$\eta_{y/x} = \frac{\sigma(\bar{y}_x)}{\sigma_y}; \quad \sigma^2(\bar{y}_x) = \frac{1}{N} \sum n_x \bar{y}_x^2 - \bar{y}^2 \quad (3.67)$$

де σ_y – середньоквадратичне відхилення змінної y .

Кореляційне відношення є спряженості тісноти будь-якого зв'язку. У цьому його перевага перед коефіцієнтом кореляції, який оцінює ступінь тісноти тільки лінійного зв'язку. Для оцінки достовірності кореляційного відношення необхідно визначити похибку квадрата цього показника з виразу [Себер, 1980]:

$$m_{\eta^2} = (1 - \eta^2) \frac{k - 1}{n - k} \quad (3.68)$$

де k – число інтервалів порядкової статистики; n – обсяг вибірки. Далі вибіркове значення критерію Фішера [Birkner, 2009]:

$$F_{k-1, n-k} = \frac{\eta_{y/x}^2 / (k-1)}{(1 - \eta_{y/x}^2) / (n-k)} \quad (3.69)$$

порівнюється зі стандартним значенням F -розподілу, відповідним значущості ймовірності $\alpha=0.05$, зі ступенями свободи $\nu_1=k-1$ і $\nu_2=n-k$. Величина:

$$\varepsilon = \eta_{y/x}^2 - r^2 \quad (3.70)$$

застосовується в якості міри відхилення від лінійної залежності. Якщо для значення критерію нелінійності справедлива нерівність:

$$F = \frac{(\eta^2 - r^2)(n - k)}{(1 - \eta^2)(k - 2)} > F_{\nu_1, \nu_2} \quad (3.71)$$

то зв'язок між випадковими величинами нелінійна [Сиденко и др., 2011]. В якості запобіжного зв'язку виступає також коефіцієнт рангової кореляції Спірмена [Елисеєва, Юзбашев, 2004]:

$$R(x, \varepsilon) = 1 - 6 \frac{\sum_{i=1}^n \delta_i^2}{n(n^2 - 1)} \quad (3.72)$$

де δ_i – різниця між рангами значень змінних. Значимість коефіцієнта рангової кореляції оцінюється по t-критерію. Якщо коефіцієнт кореляції дорівнює нулю, то статистика [Сиденко и др., 2011]:

$$t = \frac{R(x, \varepsilon)\sqrt{n-2}}{\sqrt{1-R(x, \varepsilon)^2}} \quad (3.73)$$

має розподіл Стьюдента з числом ступенів свободи $\mu=n-2$.

Осередок Вранча характеризується відносно стабільним сейсмічним режимом, тобто землетруси в заданих інтервалах магнітуд відбуваються досить ритмічно [Уломов и др., 2007]. Це дозволяє здійснювати їх довгостроковий прогноз і визначати інтервали часу, в межах яких можуть статися землетруси.

Нехай N_T – випадкова величина, що означає число землетрусів за час T . Далі, припущення H_0 про те, що, що N_T має пуассонівський закон розподілу з параметром λ , є нульовою гіпотезою. При цьому припущенні, достатньою статистикою для параметра λ , при фіксованому значенні T , є спостережуване число землетрусів N . Як точкової оцінки величини λ використовується співвідношення [Гайский, 1970; Кокс, Льюис, 1969]:

$$\lambda = \frac{N}{T} \quad (3.74)$$

Довірчими межами для середнього значення $\eta=\lambda T$ пуассонівського процесу є [Сиденко и др., 2011]:

$$N + \frac{1}{2}c_{\frac{\alpha}{2}}^2 \pm c_{\frac{\alpha}{2}}\sqrt{N} \quad (3.75)$$

де c_α – верхній α -квантиль одиничного нормального розподілу. Для перевірки гіпотези H_0 можна скористатися критерієм згоди для порівняння:

а) розподілу частот числа подій в інтервалах часу постійної довжини h з пуассонівським розподілом

б) розподілу частот довжин інтервалів між наступними землетрусами з показовим розподілом.

Нехай

$$0 \leq t_1 \leq t_2 \leq \dots \leq t_n \leq T \quad (3.76)$$

послідовність моментів настання землетрусів, а $\tau_1, \tau_2, \tau_3, \dots, \tau_{n-1}$ – послідовність інтервалів часу між послідовними землетрусами, щодо яких висуваються три гіпотези:

- 1) землетрус, який стався в проміжку (t_1, t_2) , стимулює появу землетрусу в наступному проміжку (t_2, t_3)
- 2) те, що відбувся землетрус не робить ніякого впливу на виникнення наступного землетрусу
- 3) те, що відбувся землетрус перешкоджає виникненню землетрусу в наступному проміжку.

При справедливості другої гіпотези інтервали між землетрусами повинні відповідати показниковій функції розподілу [Гайский, 1970]:

$$F(t) = P\{\tau < t\} = 1 - e^{-\lambda t} \quad (3.77)$$

де λ – параметр показникового розподілу. Тоді величини:

$$f_i = 1 - e^{-\frac{\lambda}{T} \tau_i} \quad (3.78)$$

повинні мати рівномірну функцію розподілу:

$$P\{f_i < u\} = u \quad (3.79)$$

на відріжку $0 \leq u \leq 1$. Для перевірки гіпотези про рівномірний розподіл величин (3.79) застосовується критерій Колмогорова-Смирнова. Критерій Колмогорова-Смирнова доречно застосовувати в тих випадках, коли потрібно перевірити–підпорядковується спостережена випадкова величина деякому закону розподілу,

відомому з точністю до параметрів. Нехай X_n – вибірка незалежних однаково розподілених випадкових величин, $F_n(x)$ – емпірична функція розподілу, $\Phi(x)$ – функція розподілу генеральної сукупності, з якої вилучена вибірка. Тоді статистика критерію для перевірки гіпотези H_0 про те, що вибірка підпорядковується нормальному розподілу $\Phi(x)$ визначається наступним чином:

$$D_n = \sup_x |F_n(x) - \Phi(x)| \quad (3.80)$$

Гіпотеза H_0 відкидається, якщо статистика $n^{0.5}D_n$ перевищує табличне значення k_α , що відповідає рівню значущості α . Далі будуюмо теоретичну і емпіричну функції розподілу:

$$y = u, \quad y = f(u) = \frac{s(u)}{N} \quad (3.81)$$

де $s(u)$ – число величин $f_i \leq u$. При справедливості другої гіпотези обидві криві повинні йти близько одна від одної. Залежно від характеру розбіжності між зазначеними кривими, можна скласти уявлення про характер відхилення закону розподілу $F(t)$ від показникового закону. Якщо для малих значень u графік функції $f(u)$ знаходиться вище прямої $y=u$, і далі залишається вище цієї прямої, тоді правдоподібна перша гіпотеза. Якщо крива $f(u)$ опускається нижче цієї прямої, то це говорить на користь третьої гіпотези. Для перевірки гіпотези про показниковий розподіл інтервалів між землетрусами можна скористатися критерієм, якій базуються на рівномірному розподілі нормованих моментів настання землетрусів [Буртиев, 1983; Капур, Ламберсон, 1980]:

$$u_i = \frac{t_i}{T}, i \in (1, n) \quad (3.82)$$

Цей критерій є найбільш потужним при конкуруючій гіпотезі про наявність тренда виду:

$$\lambda(t) = e^{\alpha + \beta t} \quad (3.83)$$

Процедура перевірки гіпотези $\beta=0$: полягає в обчисленні виразу:

$$u = \frac{\sum_{i=1}^n t_i / n - T/2}{T \sqrt{1/12n}} \quad (3.84)$$

тобто в порівнянні середнього арифметичного спостережених моментів виникнення землетрусів (3.82) з середньою точкою періоду спостережень.

Розподіл величини (3.84) швидко сходиться до стандартного нормального розподілу при справедливості гіпотези $\beta=0$ [Кофман, Крюон, 1965]. Перевірка цієї гіпотези про наявність тренда виду:

$$\lambda_i = \alpha + \beta z_i \quad (3.85)$$

де z_i – деяка незалежна від τ_i випадкова величина – зводиться до визначення рівня значущості величини [Капур, Ламберсон, 1980]:

$$v = \frac{\sum_{i=1}^n (z_i - \bar{z})(\tau_i - \bar{\tau})}{\left(\sum_{i=1}^n \tau_i \right) \sqrt{\sum_{i=1}^n (z_i - \bar{z})^2 / n(n+1)}} \quad (3.86)$$

за таблицями стандартного нормального розподілу.

Нехай послідовність випадкових величин x_1, x_3, \dots, x_n означає число землетрусів на i -му інтервалі часу. Для перевірки правдоподібності гіпотези про відсутність систематичного зсуву в послідовності (математичне очікування постійно $M(x_i)=const$), застосовується критерій Аббе [Гайский, 1970]. Критерій складається в порівнянні величин:

$$q^2 = \frac{1}{2(n-1)} \sum_{i=1}^{n-1} (x_{i+1} - x_i)^2; \quad r = \frac{q^2}{s^2} \quad (3.87)$$

Величина r має асимптотично нормальний розподіл з параметрами:

$$M(r) = 1; \quad D(r) = \frac{1}{n+1} \left(1 - \frac{1}{n+1} \right) \quad (3.88)$$

Якщо виявиться, що для даних x_1, x_3, \dots, x_n виконується умова $r < r_\alpha$, це означає, що в ряді спостережень є систематичний зсув. Значна регресія величин x_i на номер інтервалу i також вказує на наявність тренда в частоті виникнення землетрусів з часом [Буртиев, 1983].

З послідовності інтервалів між послідовними землетрусами

$$\tau_1, \tau_2, \dots, \tau_{n-1} \quad (3.89)$$

утворюються дві послідовності:

$$\tau'_1, \tau'_2, \dots, \tau'_n \quad (3.90)$$

$$\tau''_1, \tau''_2, \dots, \tau''_m \quad (3.91)$$

де (3.90) – послідовність інтервалів часу між віддаленими більше, ніж на Δ доби послідовними землетрусами; (3.91) – послідовність інтервалів часу між іншими землетрусами.

3.8. Дослідження статистичного зв'язку між магнітудою і глибиною землетрусів

За період 1978–2011 рр. відбулося 5031 землетрусів з магнітудою, що перевищує мінімальну магнітуду комплектності $M_c=2.3$ [www.infp.ro/arhiva]. Для перевірки гіпотези H_0 , про статистичну незалежність магнітуди і глибини землетрусів, обчислюється статистика критерію (3.65). Вибіркова статистика критерію $\chi_q=3940$ більше табличного значення $\chi_{0.05;357}=422.094$, χ^2 – розподіленої випадкової величини, яка відповідає рівню значущості $\alpha=0.05$ с 357 ступенями свободи, свідчить про те, що гіпотеза H_0 повинна бути відхилена, на користь альтернативної гіпотези про статистичний значущий зв'язок між магнітудою і глибиною. Для цієї вибірки отримані наступні значення зв'язку: кореляція Спірмена $R_s=0.535$; $\mu=0.56$; $C=0.56$. Значимість кореляції Спірмена підтверджується тестовим критерієм (3.72).

Таблиця 3.1

Шкала класифікації тісноти зв'язку Чеддока

μ	0÷0.1	0.11÷0.3	0.31÷0.5	0.51÷0.7	0.71÷0.9	0.91÷0.99	0.991÷1
Характер зв'язку	відсутня	слабка	помірна	помітна	тісний	сильна	Функціональна

Чисельні значення емпіричних коефіцієнтів тісноти зв'язку можна класифікувати за шкалою Чеддока (табл. 3.1) [Сиденко, 2011]. Для землетрусів в межах земної кори значення: $C=0.32$; кореляція Спірмена $R_s=0.12$; $\mu=0.13$ –

вказують на слабку залежність між магнітудою і глибиною. Для 4441 проміжного землетрусу статистичні характеристики взаємозв'язку, рівні $R_s=0.426$; $C=0.47$; $\mu=0.42$; $\chi_q=1273$ яки свідчать про більш значущий зв'язок між магнітудою і глибиною проміжних землетрусів, ніж для корових землетрусів.

Таким чином, проведений статистичний аналіз показав, що зв'язок між магнітудою і глибиною землетрусів вогнища Вранча в межах земної кори практично відсутній, а для проміжних землетрусів зв'язок має середню силу.

Загальна кількість наведених в каталозі землетрусів вогнища Вранча, за період 1978–2011 рр., становить 6043. Величина критерію $\chi_q=3703$ перевершує значення $\chi_{0.05;357}=93,21$ для χ^2 - розподіленої випадкової величини, яка відповідає рівню значущості $\alpha=0.05$ зі 64 ступенями свободи. Отже, гіпотеза про незалежність магнітуди і глибини землетрусів є неправдоподібною. Значення коефіцієнтів: пов'язаності $C=0.61$; рангової кореляції Спірмена критерію $R_s=0.64$; кореляційного відношення $\mu=0.66$ також вказує на статистичний зв'язок середньої сили між магнітудою і глибиною землетрусів. Значимість коефіцієнта рангової кореляції Спірмена $R_s=0.64$ підтверджується статисткою t-критерію (3.73). Такий же результат виходить для 8149 румунських землетрусів, спостережених за період 1978–2011 рр.: кореляція Спірмена $R_s=0.61$; $C=0,60$; $\mu=0.65$; $\chi_q = 4541$ вказують на зв'язок середньої сили. У таблиці (табл. 3.2) наведені параметри взаємозв'язку між магнітудою і глибиною землетрусів, з якої видно, що зв'язок магнітуди з глибиною слабшає з ростом порога магнітуди.

Таблиця 3.2

Регресія магнітуди на глибину, і тіснота залежності

Нижній поріг магнітуди	Вільний член (a)	Коефіцієнт регресії (b)	Коефіцієнт кореляції	Кореляційне відношення
$M_w \geq 2.3$	-0.01	0.02926	0.417	0.55
$M_w \geq 2.9$	-0.21	0.031	0.23	0.27
$M_w \geq 3.5$	-0.09	0.03157	0.121	0.16
$M_w \geq 4.1$	-0.04	0.03295	0.069	0,11

Якщо відкинути землетруси, з магнітудою меншою деякого порогового значення M_{co} , то ймовірна структура вибірки змінюється, і з ростом порогової магнітуди ступінь залежності зменшується. Так, для вибірки з 1899 землетрусів з магнітудою більшою $M_{co}=4.1$, правдоподібність гіпотези H_0 – відсутність статистичного зв'язку між магнітудою і глибиною землетрусів – підтверджується за всіма наведеними вище критеріями:

$$M_{co}=3.2: R_s=0.187; C=0.21; \mu=0.18; \chi_q=88.1$$

$$M_{co}=3.5: R_s=0.089; C=0.21; \mu=0.17; \chi_q=51.1$$

$$M_{co}=3.8: R_s=0.126; C=0.23; \mu=0.15; \chi_q=30.9$$

$$M_{co}=4.1: R_s=0.021; C=0.24; \mu=0.11; \chi_q=15.2.$$

У сейсмологічній літературі зустрічаються роботи, присвячені дослідженню зв'язку між параметрами землетрусів [Ardeleanu, 2002; Буртиев, 2009, 2010; Oncescu, 1987; Oncescu, Bonjer, 2001; Radu, Oancea, 1988].

3.9. Взаємозв'язок між магнітудою і інтервалом часу між землетрусами.

Рангова кореляція Спірмена, для 5031 землетрусів вогнища Вранча, виявилася незначною $R_s=0.05$. Значення коефіцієнта спряженості також виявилось незначним: $C=0.13$. Гіпотеза про відсутність значущого нелінійного статистичного зв'язку ($\mu=0.06$) підтверджується значенням F критерію (3.69). Вибіркове значення критерію $F_q=1.7$ менше теоретичного значення $F_t=1.94$ критерію (3.69), що відповідає рівню значущості $\alpha=0.05$. Отже, гіпотеза H_0 , про статистичну незалежність магнітуди і інтервалів часу між землетрусами, правдоподібна.

Для землетрусів в межах земної кори рангова кореляція Спірмена становить: $R_s=-0.029$. Значення коефіцієнтів спряженості рівні: $\chi_q=40.1$; $C=0.24$. Величина критерію $\chi_q=40.1$ менше табличного значення $\chi_{0,05;36}=58.62$ для χ^2 -розподіленої випадкової величини, яка відповідає рівню значущості $\alpha=0.05$ с 36 ступенями свободи. Припущення про слабку нелінійність статистичної зв'язку ($\mu=0.09$) відкидається за критерієм (3.70). Тобто магнітуда і інтервали часу між землетрусами в межах земної кори незалежні. Для землетрусів нижче кордону Мохо статистичні

характеристики взаємозв'язку, що мають значення $R_s=0.04$; $C=0.08$; $\mu=0.04$ свідчать на користь гіпотези про незалежність магнітуди і інтервалів часу. Вибіркова статистика критерію $\chi_q=27.0$ менше табличного значення $\chi_{0.05;36}=58.62$. Вибіркові статистики критеріїв перевірки незалежності магнітуди від інтервалів часу для румунських землетрусів з каталогу "ROMPLUS" виявилися рівними: $R_s=0.15$; $C=0.17$; $\mu=0.09$; $X_q=107.9$ – що вказує на правдоподібність H_0 .

Залежно від характеру розбіжності між зазначеними кривими можна скласти уявлення про характер відхилення закону розподілу $F(t)$ від показникового закону. З порівняння наведених на рисунку (рис.3.3) графіків функцій розподілу $f(u)$ і $y(u)$ послідовності інтервалів часу можна зробити висновок, що між землетрусами вогнища Вранча переважають короткі інтервали. Графік функції $f(u)$ знаходиться вище прямої лінії $y=i$, отже, землетрус, що стався в проміжку (t_1, t_2) стимулює появу землетрусу на сусідньому відрізку (t_2, t_3) .

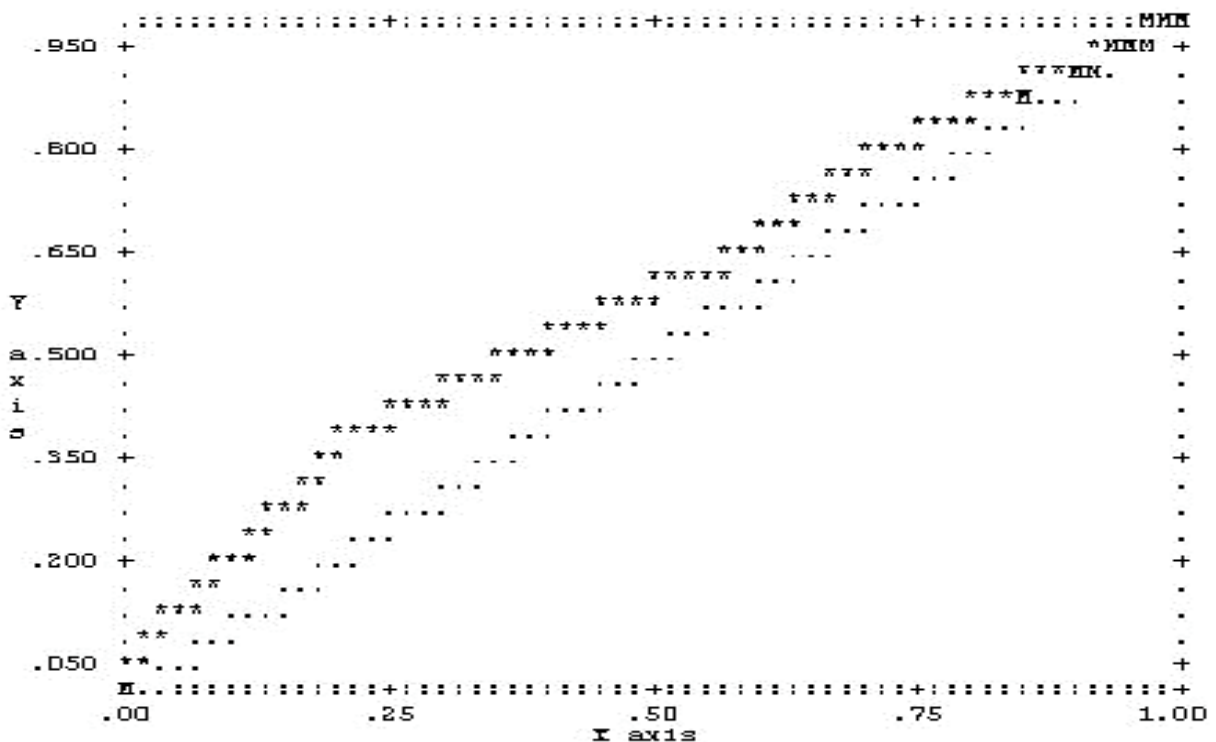


Рис. 3.3. Розподіл моментів настання землетрусів і функція рівномірного розподілу.

За критерієм Колмогорова, гіпотеза про показниковий розподіл інтервалів між послідовними землетрусами, і про рівномірний розподіл величин (3.82) повинна бути відкинута. На це також вказує відмінний від одиниці коефіцієнт варіації $C=0.55$. Гіпотеза про відсутність експоненціального і лінійного трендів (3.83), (3.85) в послідовності (3.76) відкидається, так як вибіркові значення 11.23 і 11.5 більше 95% -ного квантиля розподілу $N(0,1)$. Далі складалися спостережені вибірки (3.90), (3.91), де величина Δ приймала значення $\Delta=3, 6, 9$ і 12 діб. При значенні $\Delta=3$, для вибірок (3.90) і (3.91), гіпотеза про відсутність експоненціального і лінійного трендів підтверджується.

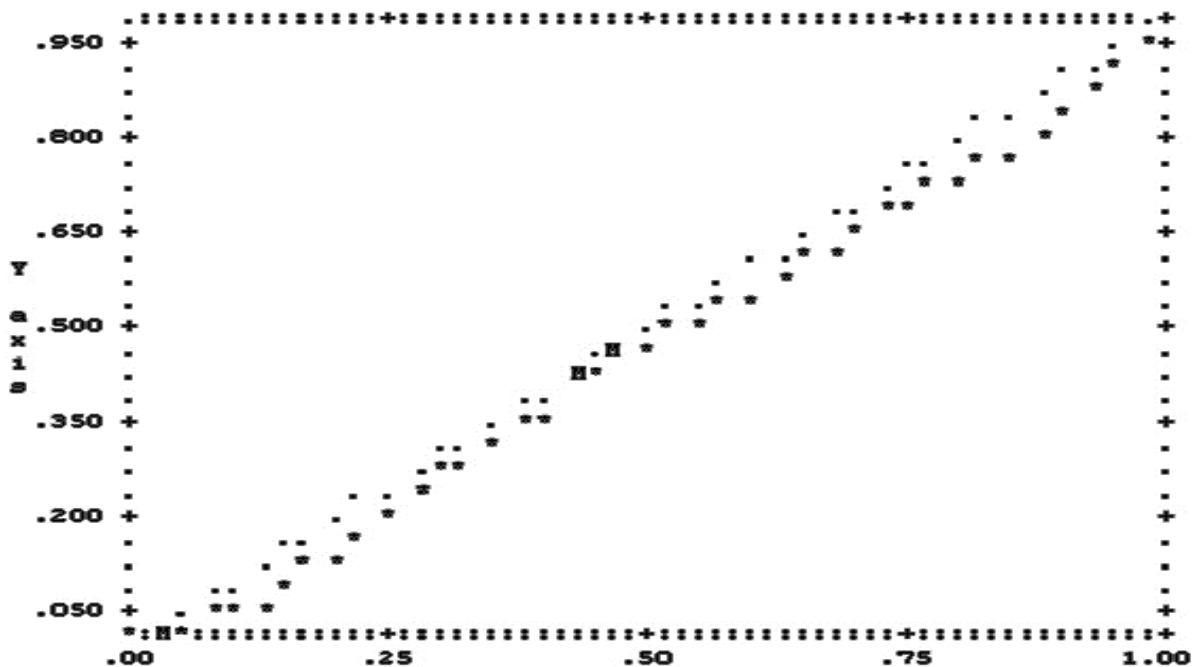


Рис. 3.4. Розподіл нормованих моментів настання землетрусів і функція рівномірного розподілу. Збігаються значення позначені точкою М.

У сукупності землетрусів, з магнітудою більшою значення $m_0=3.2$, статистичний зв'язок зменшується: $R_s=0.16$; $C=0.12$; $\mu=0.04$; $\chi_q=32.5$. Розподіл інтервалів часу між землетрусами наближається до показникового закону розподілу. Припущення про те, що спостережені інтервали становлять вибірку з генеральної сукупності, що має показниковий закон розподілу, підтверджується для землетрусів

з магнітудою більше $M_{co}=3.8$. Вибіркова статистика критерію χ^2 , $\chi_q=8.58$ менше критичного значення $\chi_{0.05;10}=18.3$, що відповідає рівню значущості $\alpha=0.05$, з десятима ступенями свободи.

3.10. Дослідження зв'язку між коровою і підкоровою сейсмічною активністю

З спостережених річних і місячних землетрусів з магнітудою більше мінімальної магнітуди комплектності M_c , склалися вибірки:

$$n_1, n_2, n_3, \dots, n_N \quad (3.92)$$

$$m_1, m_2, m_3, \dots, m_M \quad (3.93)$$

Для спостереженої послідовності (3.92) землетрусів Вранчі, припущення про відсутність тренда правдоподібно, величина статистики критерію Аббе (3.87), що дорівнює $r=0.769$, більше значення $r_p=0.722$, відповідного 5%-ному рівню табличного значення критерію Аббе. Хоча кореляційне відношення 0.37 і рангова кореляція Спірмена 0.323 між річним числом землетрусів і роком вказують на наявність слабого тренда по роках. Для пошуку зв'язку між коровою і підкоровою сейсмічною активністю складаються послідовності числа землетрусів за роками в межах земної кори і нижче межі М:

$$n_1', n_2', n_3', \dots, n_N' \quad (3.94)$$

$$n_1'', n_2'', n_3'', \dots, n_N'' \quad (3.95)$$

Значення статистики критерію Аббе складають 0.55 і 0.83 відповідно, 5%-не порогове значення дорівнює $r_p=0.72$. Це говорить про те, що в рядах річного числа корових землетрусів існує тренд, а в сукупності підкорових землетрусів Вранча тренд відсутній. Рангова кореляція Спірмена між послідовностями (3.94) і (3.95) виявилася рівною $R_s=0.315$, а кореляційне відношення дорівнює: $\mu=0.24$. Значення $R_s=0.315$ менше порогової величини ($R_t=0.334$), яка відповідає рівню значущості $q=0.05$ t -статистики з $df=33$ ступенями свободи. Вибіркове значення F -критерію (3.69) ($F_q=0.61$) менше порогового значення ($F_t=2.91$), відповідного ймовірності значущості $\alpha=0.05$ с $v_1=3$, $v_2=31$ ступенями свободи, отже, має місце нелінійна

залежність. Значення коефіцієнта спряженості рівне $C=0.44$ і вибіркова величина $\chi_q=8.46$ критерію (3.65) вказують на слабкий зв'язок між річним числом землетрусів в межах земної кори і нижче границі Мохо:

$$M_c=2.6: R_s=-0.04; C=0.33; \mu=0.13; \chi_q=4.31$$

$$M_c=2.9: R_s=-0.29; C=0.29; \mu=0.22; \chi_q=3.38.$$

Далі з щомісячного числа вранчських землетрусів складається вибірка в 408 спостережень. Коефіцієнт варіації $V=0.69$ вказує на відхилення від пуассонівського закону розподілу щомісячного числа землетрусів. Вибіркова статистика критерію Аббе $r=0.48$ менше табличного значення $r_p=0.916$, відповідного 5%-не рівня значущості. Кореляція Спірмена між щомісячним числом землетрусів і номером місяця дорівнює $R_s=0.235$. Кореляційне відношення $\mu=0.17$; значення коефіцієнта спряженості $C=0.22$ між числом землетрусів по місяцях і номером місяця вказують на наявність слабого тренда в спостережуваній послідовності (3.95). Для спостережених послідовностей щомісячного числа корових і підкорових землетрусів (3.97), (3.98) значення критерію Аббе (0.77; 0.48) менше табличного значення 0.92, що відповідає рівню значущості $\alpha=0.05$. За значеннями показників: коефіцієнта спряженості $C=0.22$; кореляції Спірмена $R_s=0.24$; значення критерію (3.68) $\chi_q=74.6$; коефіцієнта лінійної кореляції $r=0.17$; кореляційного відношення $\mu=0.17$; регресії послідовності (3.97) на (3.98) – можна сказати, що між коровою і підкоровою сейсмічною активностями існує слабкий статистичний зв'язок:

$$m' = 2.92 + 0.0576m'' \quad (3.96)$$

$$m'_1, m'_2, m'_3, \dots, m'_N \quad (3.97)$$

$$m''_1, m''_2, m''_3, \dots, m''_N \quad (3.98)$$

і зі збільшенням магнітуди зв'язок слабшає. Так, для землетрусів з магнітудою більше $M_{co}=2.5$ значення статистичних критеріїв складають: кореляція Спірмена $R_s=0.161$; коефіцієнти спряженості: $C=0.124$; $\chi_q=6.42$; кореляційне відношення $\mu=0.11$; коефіцієнт лінійної кореляції $r=0.11$. Статистика $F_q=0.255$ критерію (3.69), перевірки відхилення форми зв'язку між щомісячним числом корових і підкорових землетрусів від лінійної форми, менше теоретичного значення F -розподілу ($F_t=3.02$),

відповідають ймовірності значущості $\alpha=0.05$ зі ступенями свободи $\nu_1=2$, $\nu_2=405$. Дійсно, в цьому випадку, коефіцієнт лінійної кореляції $r=0.11$ дорівнює кореляційному відношенню $\mu=0.11$.

У послідовності числа румунських землетрусів за роками існує тренд. Величина критерію Аббе $r=0.543$ менше теоретичного значення $r_p=0.71$, відповідає ймовірності значущості $\alpha=0.05$. Значення критеріїв оцінки зв'язку між річним числом землетрусів і роком рівні: $R_s=0.379$; $C=0.47$; $\mu=0.40$.

Статистики критерію Аббе для послідовності корових і проміжних землетрусів Румунії рівні: 0.48; 0.832 – відповідно. З порівняння з табличним значенням $r_p=0.722$, відповідним 5%-ним рівнем значущості, слідує, що в коровій сейсмічності присутній тренд, а в сукупності проміжних землетрусів відсутній. Значення критеріїв оцінки зв'язку між річним числом корових і підкорових землетрусів рівні: $R_s=0.643$; $C=0.51$; $\chi_q=123.8$; $\mu=0.58$. Таким чином, між річним числом землетрусів в межах земної кори і в проміжному шарі існує статистично значущий зв'язок. Для сукупності щомісячного числа землетрусів спостерігається така ж статистична картина: $R_s=0.426$; $C=0.58$; $\mu=0.46$. На наявність тренда вказує статистично незначна на 5%-ном рівні значущості величина 0.3 статистики критерію Аббе.

У сукупності корових і підкорових землетрусів присутній тренд, значення критерію Аббе рівні: 0.293; 0.47.

Статистики критеріїв залежності вказують на статистично значимий зв'язок середньої сили: $R_s=0.479$; $C=0.48$; $\chi_q=125.8$; $\mu=0.43$. Регресія числа корових землетрусів на щомісячне число проміжних румунських землетрусів має вигляд:

$$n' = 9.86 + 0.2086n \quad (3.99)$$

3.11. Дослідження зв'язку між сейсмічністю вогнища Вранча і числом сонячних плям

Негативні значення коефіцієнтів лінійної кореляції $r=-0.278$ та кореляції Спірмана $R_s=-0.235$ вказують на те, що між числом сонячних плям [<http://www.sidc.be/silso/datafiles>] і місячним числом землетрусів в Румунії існує

слабка негативна кореляційна залежність. Статистики критеріїв залежності вказують на статистично значимий зв'язок середньої сили: $R_s = -0.235$; $C = 0.405$; $\chi_q = 80.2$; $\mu = 0.29$; коефіцієнт лінійної кореляції $r = -0.278$. Абсолютне значення кореляції Спірмана $R_s = 0.235$ більше порогової величини ($R_t = 0.097$), яка відповідає рівню значущості $q = 0.05$ t -статистики з $df = 406$ ступенями свободи. Значимо і кореляційне відношення $\mu = 0.29$ – вибіркове значення F -критерію (3.69) $F_q = 7.26$ більше порогового значення $F_t = 2.24$, відповідного значення ймовірності значущості $\alpha = 0.05$, з числом ступенів свободи $\nu_1 = 5$, $\nu_2 = 402$. Вибіркова статистика критерію χ^2 : ($\chi_q = 80.2$) більше критичного значення $\chi_{0.05;25} = 44.31$ – відповідає рівню значущості $\alpha = 0.05$ с числом ступенів $df = 25$, тобто по таблиці спряженості ознак виявлена значуща залежність сейсмічності Румунії від числа сонячних плям. Лінійна регресія числа землетрусів на число сонячних плям визначається формулою:

$$N = 23.88 - 0.06043S \quad (3.100)$$

Коефіцієнт детермінації дорівнює: $R^2 = 0.055$. Вибіркова статистика $F_q = 23.79$, F -розподілу (3.69) перевершує табличне значення $F_{0.05,1,406} = 3.86$. Отже, регресія є значущою..

Якщо вірити статистиці, то гіпотеза про залежність сейсмічності Румунії від числа сонячних плям правдоподібна. Негативна рангова кореляція Спірмана і коефіцієнт лінійної кореляції вказують на слабке зменшення сейсмічної активності Румунії при збільшенні числа сонячних плям.

Висновки до розділу 3

Виконання умов мультиколінеарності та гомоскедастичності необхідно для отримання незміщених, ефективних і надійних оцінок параметрів регресії. При гетероскедастичності оцінки регресії є лінійними і незміщеними, але вони не будуть ефективними, і дисперсії оцінок параметрів будуть мати зміщення. Здійснений в даному розділі статистичний аналіз показав, що для лінійної багатовимірної регресії виконуються умови теореми Гаусса-Маркова. Найбільш повний алгоритм дослідження мультиколінеарності – алгоритм Фаррара-Глобера – використовується

для виявлення колінеарності між змінними регресії. Для оцінки параметрів моделі, де має місце мультиколінеарність регресорів, застосовується метод головних компонент – це метод, який дозволяє зменшити число змінних у регресійній моделі і розмірність вибірових даних. Таке завдання є актуальним в сейсмології, наприклад, в дослідженні технічного стану житлового фонду, яке часто є трудомістким і дорогим. Пошкоджуваність споруд залежить від багатьох факторів: природних і антропогенних, число яких визначається ресурсами, кваліфікацією фахівців, які приймають рішення про технічний стан споруд, наявністю і рівнем розвитку відповідних технологій. Скорочення числа вводиться в аналіз факторів шляхом виявлення їх колінеарності, виділення головних з них, і є важливим завданням при складанні математичних моделей стійкості споруд.

Для дослідження умови гомоскедастичності застосовувалися: тест рангової кореляції Спірмена і метод Глейзера. При виборі найкращої, серед регресійних моделей вирішальними були: мінімум стандартної похибки регресії і максимум коефіцієнта детермінації R^2 .

Гіпотеза про наявність кореляційного зв'язку між магнітудою і глибиною землетрусів не підтвердилася. Оцінка мінімальної магнітуди комплектності M_c , виконана за методами *MAXC* і *MBS*, збігається і дорівнює $M_c=2.3$. У рівняння взаємозалежності магнітудних шкал включалися додаткові регресори, проте це не привело до істотного поліпшення якості регресії.

Між магнітудою і глибиною землетрусів відзначений слабкий статистичний зв'язок, причому, для землетрусів у межах земної кори зв'язок значно слабше, ніж для проміжних землетрусів. Область підготовки більш сильних землетрусів охоплює більший обсяг геофізичного середовища в межах земної кори відбуваються землетруси, з магнітудою до $MW=5.5$, а в проміжному шарі з магнітудою до $MW=7.7$. Значення магнітуд і інтервалів часу між землетрусами зони Вранча, в межах земної кори, пов'язані слабкою лінійною залежністю, тоді, як для землетрусів нижче границі Мохо, статистичні характеристики послідовності вранчських сейсмічних подій взаємозв'язку свідчать про незалежність магнітуди від інтервалів

часу. Виявлено слабкий зв'язок між річним числом землетрусів в межах земної кори і нижче границі Мохо. Статистична залежність між короною і подкороною сейсмічною активністю для землетрусів з магнітудою $M \geq 2.6$ практично відсутня.

Також, відзначений слабкий статистичний зв'язок між щомісячним числом корових і проміжних землетрусів. В рядах річного та щомісячного числа землетрусів в межах кори присутній тренд. У проміжній сейсмічності підтвердилося припущення про відсутність тренда.

У завданнях сейсмології по виявленню взаємозалежностей між параметрами сейсмічності, висновки робляться по їх спостережених значеннях, які, зазвичай, задаються в формі таблиць (x_i, y_i) , $i=1, \dots, n+1$. Як впливає з основної теореми алгебри, через $n+1$ точку, можна провести криву, описувану многочленом ступеня n [Курош, 1965]. Однак, при великому числі точок, використовуються багаточлени невисоких степенів для визначення кривої, що проходить через всі точки, зокрема, так званий гладкий кубічний сплайн.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

- Аргучинцева А.А., Методы статистической обработки и анализа гидрометеорологических наблюдений. Учебное пособие. Иркутск, 2007, 105.
- Буртиев Р.З., Исследование стационарности потока Карпатских землетрясений. Известия АН МССР. Сер. физ.–тех. и мат. наук. Кишинев, 1983, №2, 32–36.
- Буртиев Р.З. О возможности применения многомерной линейной регрессии с коллинеарными факторами для преобразования магнитуд. Buletinul Institutului de Geologie și Seismologie al Academiei de Științe a Moldovei. 2009.No. 2, pp. 35-44
- Буртиев Р.З. О статистической связи сейсмической активности Балканских сейсмических зон. Buletinul Institutului de Geologie si Seismologie al ASM. 2010. Nr.1. 1–5.
- Дембицкий С. Виды статистических методов: описательная статистика [Электронный ресурс].– Режим доступа: <http://www.soc-research.infoquantitat>

/ive6.html—See more at: <http://www.soc-research.info/quantitative/6.html#sthash.ZGAEuNBv.dpuf>, 2013.

Дулупов В.И., Лескова О.А., Майоров И.С. Системная экология, 2009, http://abc.vvsu.ru/Books/sistemnaja_ekologija_up/page0001.asp.

Елисеева И. И., Юзбашев М. М., Общая теория статистики. М.: Финансы и статистика, 2004, 656.

Габриелян Р. Г., Эконометрика, Видное, 2008, 85.

Гайский В.Н., Статистические исследования сейсмического режима Наука, Москва, 1970, 121.

Гнеденко Б.В., Курс теории вероятностей. Изд. Физмат литературы. Москва, 1961, 406.

Лешинский О. Л., Рязанцева В. В., Юнькова О. О., Економетрія, Київ, 2003, 208.

Капур К., Л. Ламберсон. Надежность и проектирование систем. Мир, Москва, 1980, 696.

Кендалл М. Дж., Стьюарт А., Статистические выводы и связи. Наука, Москва, 1973, 898.

Кокс Д., Льюис П. Статистический анализ последовательности событий, Москва, Мир, 1969. 310 с.

Кофман А., Крюон Р., Массовое обслуживание, Мир, Москва, 1965, 302.

Крамер Гаральд. Математические методы статистики. Москва. Мир. 1975. 648.

Курош А. Г., Курс высшей алгебры. Наука, Москва, 1965, 431.

Мхитарян В.С., Архипова М. Ю., Сиротин В.П., Економетрика, Москва, 2008, 142.

Наконечный С. И., Терешенко Т. О., Романюк Т.П., Економетрія. КНЕУ, КИЇВ, 2004, 520.

Себер Дж. Линейный регрессионный анализ. Мир, Москва, 1980, 456.

Сиденко А.В., Вишняков В.В., Исаев С.М., Теория статистики. Учебник. М.: МАКС-Пресс, 2011, 376.

Смирнов, И.В. Дунин-Барковский Н.В., Курс теории вероятностей и математической статистики. Наука, Москва, 1965, 511

- Солоненко В.П. Палеосейсмогеология //Изв. АН СССР. Физика Земли, 1973, №9, 3–16.
- Уломов В.И., Севостьянов В.В., Миндель И.Г., Трифонов Б.А. Оценка сейсмической опасности для высотных зданий в г. Москве // Современное высотное строительство. Монография (464с.). М.: ГУП "ИТЦ Москомархитектуры", 2007. 94–100
- Феллер В., Введению в теорию вероятностей и ее приложения. Мир, Москва, 1967, 498.
- Ameer A. S., M. L. Sharma, H. R. Wason, S. A. Alsinawi. Probabilistic Seismic Hazard Assessment for Iraq Using Complete Earthquake Catalogue Files. Pure applied geophysics. Birkhäuser Verlag, Basel, 2005, 951–966.
- Luminita Ardeleanu. Space and time patterns of the seismicity of the Vrancea, Romania. Geophysique, Editura Academiei Romane, Tome 46, 2002, 39–46.
- Birkner Matthias. Angewandte Stochastik. Felix Hofmann, Universität Bonn, 2009, 72.
- Cao, A. M., and S. S. Gao. Temporal variation of seismic b-values beneath northeastern Japan island arc, Geophys. Res. Lett. 29, No. 9, 2002.
- Durner Wolfgang. Einführung in die Statistik. Skript zur Vorlesung an der Universität Bayreuth SS, 1999, 73.
- Fisz M., Wahrscheinlichkeitsrechnung und Mathematische Statistik. VEB Deutscher Verlag der Wissenschaften, Berlin, 1966, 551.
- Grunthal Gottfried, Wahlstrom Rutger. An MW based earthquake catalogue for central, northern and northwestern Europe using a hierarchy of magnitude conversions, Journal of Seismology Kluwer Academic Publishers, Netherlands, 2003, 507–531.
- <http://www.sidc.be/silso/datafiles>
- Ismail-Zadeh A.T. Keilis-Borok V.I., Soloviev A.A., Numerical modelling of earthquake flow in the southeastern Carpathians (Vrancea): effect of a sinking slab. <http://www.sciencedirect.com/science/article/pii/S0031920198001666> #CORR1

- Marzocchi Warner and Sandri Laura. A review and new insights on the estimation of the b-value and its uncertainty. *ANNALS OF GEOPHYSICS*, VOL. 46, N. 6, December 2003, 12.
- Mignan, A., J. Woessner, Completeness magnitude in earthquake catalogs, Community Online Resource for Statistical Seismicity Analysis, doi:10.5078/corssa-00180805, 2010, Available at <http://www.corssa.org>.
- Mignan A., Werner J., Wiemer S., Chen C., and Wu Y.-M., Bayesian Estimation of the Spatially Varying Completeness Magnitude of Earthquake Catalogs. *Bulletin of the Seismological Society of America*, Vol. 101, No.3, 2011, 1371–1385.
- Oncescu M. C. and Trifu C. I., Depth Variation of the Moment Tensor Principal Axes in Vrancea (Romania) Seismic Region, *Ann. Geophysicae* 5B, 1987, 149–154.
- Oncescu M. C. and Bonjer, R.P. A note on the depth recurrence and strain release of large intermediate Vrancea earthquakes. *Tectonophysics*, 2001, 291–302.
- Popa Mihaela. Monitorizarea seismicității produse pe teritoriul României în vederea evaluării hazardului seismic. Cercetâr privind managementul dezastrelor generate de cutremurele românești, coordonator Marmureanu, Tehnopress. Iași, 2009, 68–93.
- Radu C., Oancea V., Recurrence models for intermediate earthquakes of the Vrancea region. In proceedings XXI General Assembly of ESK, Sofia, 1988, 94–100.
- Schmidt Thorsten, *Mathematische Statistik I*. Universität Leipzig, 2006, 80.
- Soong T.T., *Fundamentals of probability and statistics for engineers*. State University of New York at Buffalo. New York, USA, 2004, 298.
- Schorlemmer D. and Woessner J., Probability of Detecting an Earthquake. *Bulletin of the Seismological Society of America*, Vol. 98, No. 5, 2008, 2103–2177.
- Stromeyer Dietrich, Grünthal Gottfried, Wahlström Rutger. Chi-square regression for seismic strength parameter relations, and their uncertainties, with applications to an MW based earthquake catalogue for central, northern and northwestern Europe. *Journal of Seismology*. Kluwer Academic Publishers, Netherlands, 2004, 143–153.

- Toutenburg Helge, Heumann Christian, Deskriptive Statistik. Eine Einführung in Methoden und Anwendungen mit R und SPSS. Springer–Verlag Berlin–Heidelberg. 2008. 127.
- Vakov V. Alexander. Relationships between earthquake magnitude, source geometry and slip mechanism. *Tectonophysics* 26, 1996, 97–113.
- Welkowitz Jean, Robert B. Ewen, Cohen Jacob, Introductory Statistics, Harcourt Brace Jovanovich, Publishers. Printed in the USA, 1988, 375.
- Wenzel F., Lorenz F.P., Sperner B. and Oncescu M.C., Seismotectonics of the Romanian Vrancea Area. Vrancea earthquakes: Tectonics, Hazard and Risk Mitigation, 1997, Kluwer Academic Publishers, 15–25.
- Wiemer Stefan and Wyss Max, Minimum Magnitude of Completeness in Earthquake Catalogs: *Bulletin of the Seismological Society of America*, Vol. 93, No. 4, 2003, 859–869.
- Wiemer Stefan. A software package to analyze seismicity: ZMAP, *Seismological Research Letters, Electronic Seismologist*, 2001, 10.
- Woessner Jochen, Wiemer Stefan. Assessing the Quality of Earthquake Catalogues: Estimating the Magnitude of Completeness and its Uncertainty. *Bulletin of the Seismological Society of America*, Vol. 95, No. 2, 2005, 684–698.

www.infp.ro/arhiva

РОЗДІЛ 4

МОДЕЛІ СЕЙСМІЧНОГО РЕЖИМУ

Для аналізу сейсмічної небезпеки, після виділення сейсмічних зон і створення каталогів їх землетрусів, необхідно визначити розподіл параметрів землетрусів. Це завдання вирішується на основі ймовірнісних моделей сейсмічного режиму зони вогнищ землетрусів – частотою виникнення в часі вогнищ землетрусів різних значень магнітуд в сейсмічній зоні. При дослідженні сейсмічності ми маємо справу з змінним у часі сейсмічним процесом. Багато задач сейсмології пов'язані з умінням обчислювати ймовірності певних подій, пов'язаних з процесом і визначенням його ймовірнісної структури. В основі вирішення цих та багатьох інших завдань сейсмології лежить теорія випадкових процесів.

Найбільш часто зустрічаються в літературі ймовірнісні моделі в часовій послідовності землетрусів засновані на припущенні про те, що землетруси відбуваються в часі згідно пуассонівського процесу, і що магнітуди рівномірно розподілені. При цьому передбачається, що вогнища землетрусів також розподілені рівномірно в фізичному просторі (x, y, h) . Для моделювання потоку землетрусів багато дослідників намагалися застосувати методи аналізу часових рядів. Розраховувався пуассонівський показник розсіювання, будувалися статистичні розподіли інтервалів часу між землетрусами і числа сейсмічних подій в непересічних інтервалах, автокореляційні функції, функції ризику і т.д [Богданов та ін., 2010; Буртієв, Онофраш, 1993; Естева, 1968; Кнопов, Маркушевич, 1984; Эрисоглу и др., 2011; Шумила, 1985; Bottari et al., 1992; Delic, Radojicic, 2005; Kagan, Knopoff, 1976; Lombardi, 2007; Patwardhan et al., 1980; Vere-Jones, 2010].

Ухвалення пуассонівської моделі означає, що розподіл часу очікування до наступного землетрусу не зміниться від того, що відомо час, що минув з моменту виникнення попереднього землетрусу. Тоді як фізичні моделі, які передбачають поступове накопичення енергії пружних деформацій, і їх миттєва розрядка вимагає створення більш широкої моделі

регенеруючого процесу. Ця модель, на відміну від пуассонівської моделі повинна враховувати, що очікувана тривалість часу до наступної сейсмічної події зменшується з часом і залежить від магнітуди землетрусу [Гайський, 1960]. Для малих обсягів сейсмогенних зон, часто, спостережені послідовності землетрусів суперечать пуассонівській моделі через групування сейсмічних подій у часі. При дослідженні загальних ймовірнісних закономірностей розподілу вогнищ в просторі (φ, ψ, h) , за часом t , і магнітуди m спостерігається нерівномірність розподілу координат φ, ψ, h, t, m п'ятимірного простору (φ, ψ, h, t, m) . Розподіл землетрусів в фізичному просторі (φ, ψ, h) та за магнітудою m відрізняється приналежністю осередків до вузьких сейсмічних поясів і до області малих магнітуд.

Спостерігаються також певні відхилення тимчасового ходу сейсмічності від пуассонівського процесу. Виявлено позитивний вплив події в інтервалі (t_1, t_2) землетрусу на його виникнення в інтервалі (t_2, t_3) [Гайський, 1960; Буртиев, 1983]. Показано, що ймовірнісній структурі такої послідовності подій більше відповідає розподіл Пойя [Буртиев, Онофреш, 1993]:

$$P(N) = \frac{m(m+d)(m+2d)+\dots+(m+(N-1)d)}{N!(1+d)^{m+N}} \quad (4.1)$$

де m -математичне сподівання випадкової величини N , а дисперсія визначається за формулою:

$$d = \frac{\sigma^2}{m} - 1 \quad (4.2)$$

Існують моделі альтернативні до пуассонівських моделей – це моделі тригерного типу [Кнопов, Маркушевич, 1984]. У них процес виникнення землетрусів розглядається як накладення деякого числа часових рядів, що виходять з різних точок. Моменти виникнення землетрусів утворюють пуассонівський процес, який є окремим випадком неперервного марковського ланцюга. У напівмарковські й моделі магнітуди утворюють дискретний марковський ланцюг, а інтервали часу між землетрусами мають розподіл $F(t)$ відмінний від показникового закону. Така

модель була застосована для опису потоку сильних тихоокеанських землетрусів [Patwardhan et al., 1980].

Для обліку ефекту групування землетрусів застосовувався складний пуассонівський процес, який складався з основного (основні сейсмічні події) і допоміжного (рої землетрусів) потоків сейсмічних подій. В основному, ймовірнісний підхід до аналізу каталогів застосовується для визначення просторово-часового розподілу параметрів землетрусів, тобто для опису сейсмічного режиму.

4.1. Ймовірнісні моделі сейсмічного режиму

Як відомо, під випадковою величиною розуміють змінну величину, значення якої залежить від випадку і для якої визначена функція розподілу ймовірностей:

$$F(x) = P\{\zeta < x\} \quad (4.3)$$

Випадковий процес - це множина випадкових величин - функція двох аргументів:

$$\zeta(t) = \psi(\varepsilon, t), \quad t \in T, \quad \varepsilon \in \Omega \quad (4.4)$$

де Ω – множина елементарних подій. Для кожного значення t функція $\zeta(t)$ є функцією елементарної події ε , може приймати значення з деякої кінцевої множини E , тобто є випадковою величиною [Huisinga, Meerbach, 2005; Winkler, 2000]. Для певної елементарної події функція $\zeta(t)$ залежить тільки від дійсної змінної. Кожна така функція називається реалізацією випадкового процесу $\zeta(t)$. Випадковий процес являє собою або множина випадкових величин, або множина реалізацій [Гнеденко, 1961].

Для аналізу сейсмічної небезпеки, міра якої заснована на обчисленні ймовірності:

$$P_{(\varphi, \lambda)}(t, n, m, I^k) \quad (4.5)$$

необхідно знати місце виникнення і магнітуду майбутніх землетрусів. Вирішення цього завдання передбачається здійснити на основі добре розробленої, витонченої теорії марковських процесів.

4.2. Марковська модель потоку землетрусів

Рівняння макросейсмічної поля пов'язує з кожним землетрусом параметри (φ, ψ, h, m) : φ, ψ — координати епіцентру; h — глибина; m — магнітуда, інтенсивність сейсмічних впливів у пунктах спостережень (φ_i, ψ_i) . Отже, для передбачення інтенсивності сейсмічних впливів досить передбачити місце виникнення і магнітуду майбутніх землетрусів.

Сейсмічність визначається: інтенсивністю тектонічного процесу; розмірами тектонічних блоків, в якому можуть накопичуватися пружні напруги; фізико-хімічними умовами на різних глибинах надр Землі; межею міцності і плинності гірських порід на цих глибинах і особливостями геологічної будови території. Вона залежить також і від багатьох невизначених і невідомих, — на даному етапі розвитку наук про Землю — факторів. Однак, залежність сейсмічності від перерахованих характеристик не є однозначною, має форму статистичного закону, тобто при спостереженні процесів, що впливають на виникнення сейсмічних подій, землетрус може статися, а може, й ні. Однак, при відносно тривалому періоді спостереження над сейсмічністю, в появі землетрусів існує стійкість параметрів, що характеризують сейсмічність: частота виникнення землетрусів; статистичний розподіл магнітуди; просторовий розподіл вогнищ.

Перераховані фактори, поряд з існуючою неточністю, неоднозначністю і невизначеністю в описі комплексу процесів, що призводять до виникнення сейсмічних подій, припускають застосування методів теорії ймовірностей у вивченні сейсмічності.

Розглянемо сейсмічну зону Z^s , яка охоплює деякий простір геофізичного середовища $[\Phi \times \Psi \times H]$: Φ ; Ψ ; H — діапазони можливих значень широти, довготи і глибини землетрусів відповідно. Декартово множення даного простору на діапазон M — можливих значень магнітуди утворює $\Xi^s = (\Phi \times \Psi \times H \times M)$ чотиривимірний простір точки якого можуть виявитися можливими значеннями параметрів землетрусів. Тобто землетрус, що стався в момент часу t_n , на осі часу зображається точкою у чотиривимірному просторі Ξ^s , і значення параметрів землетрусів

визначиться завданням чотиривимірного випадкового процесу $\Omega(t)$. Для розрахунку СН досить знати де відбудеться землетрус, і яка буде його магнітуда. Ймовірнісна модель послідовності землетрусів розробляється для визначення ймовірності того, що координати епіцентру, глибина і магнітуда землетрусів приймуть значення з інтервалів Φ_i, Ψ_j, H_l, M_n відповідно.

У каталогах землетрусів значення параметрів наводяться з кроком дискретності від 0.001 до 0.1, відповідно параметри можуть мати настільки велике число можливих значень, що їх можна подати у вигляді неперервних випадкових величин. Їх слід групувати і визначити розподіл інтервальних значень параметрів, тобто обчислити вірогідність, того, що значення параметрів можливих землетрусів потраплять до деяких інтервали Φ_i, Ψ_j, H_l, M_n . Отже ймовірність того, що в деякій точці, на глибині h_0 , з координатами (φ_0, ψ_0) , відбудеться землетрус з магнітудою m_0 , буде незначною величиною. Групування проводиться при великому обсязі вибіркового даних, для спрощення їх обробки і форми подання. Його слід здійснити тоді, коли вибіркові дані є наближенням реальних значень параметрів і спостереження над неперервними величинами обробляються методами статистичного аналізу вибіркового значень дискретних випадкових величин. Крім того, при групуванні збільшується представництво вибіркового даних.

Якщо простір Ξ^s розбити на непересічні чотиривимірні інтервали, то параметри кожного землетрусу будуть належати одному і тільки одному інтервалу. Інтервали Φ_i, Ψ_j, H_l, M_n утворюються розбивкою діапазонів можливих значень параметрів землетрусів на $R_\varphi, R_\psi, R_h, R_m$ частин, відповідно. Простір можливих значень параметрів землетрусів буде складатися з чотиривимірних інтервалів $\Xi^s = [\Phi_i \times \Psi_j \times H_l \times M_n]$, $i=1, \dots, R_\varphi; j=1, \dots, R_\psi; l=1, \dots, R_h; n=1, \dots, R_m$, яких назвемо елементарними джерелами землетрусів (ЕДЗ). З огляду на лексикографічний порядок індексів, при якому зростає останній індекс, потім передостанній і так далі, можна перетворити 4-вимірної індекси $ijkl$ в зручні для практичного застосування одномірні індекси r [Соколов, 1972]:

$$r = (i-1)R_\psi R_h R_m + (j-1)R_h R_m + (l-1)R_m + n \quad (4.6)$$

У ЕДЗ тенденція сейсмічних подій групування за часом і в просторі може ослабнути, або зовсім зникнути. Тут відбувається випадкове "просіювання" вихідного потоку землетрусів, подібно до утворення потоків Ерланга. Тільки, для потоку Ерланга k -го порядку в вихідному потоці збереглося б кожне k -е спостереження. В даному випадку відбір спостереження (землетрусу) для j -го потоку з вихідного, відбувається з певною ймовірністю π_j , яка визначиться за стохастичною моделлю сейсмічного режиму. Зона вогнищ землетрусів розглядається як фізична система, яка змінює свої стан в випадкові моменти часу [Burtiev, 1986, 2012]. На інтервалі часу між зміною станів траєкторія функціонування системи вважається постійною і рівною значенням параметрів землетрусу, що стався. У такому випадку положення системи в будь-який момент часу визначиться завданням чотиривимірною випадкового процесу:

$$\Omega(t) = \{\xi_\varphi(t), \xi_\psi(t), \xi_h(t), \xi_m(t)\} \quad (4.7)$$

з простором станів Ξ^s , де $\xi_\varphi(t)$, $\xi_\psi(t)$, $\xi_h(t)$, $\xi_m(t)$ – одномірні скачкоподібні випадкові процеси, реалізацією яких є послідовності значень широти, довготи, глибини і магнітуди спостережених землетрусів.

Для зручності інтервал $(\Phi_{ix}\Psi_{jx}H_{ix}M_n)$ позначається через Σ_r , число яких складе: $R_s = R_\varphi R_\psi R_h R_m$. В такому випадку, простором станів випадкового процесу(4.7) є об'єднання всіх: $\Xi^s = \{\Sigma_r\}$, $r=1, \dots, R_s$, де R_s – число станів (4-х мірні інтервали). Якщо в момент часу t_n , відбувається землетрус в елементарній зоні землетрусів Σ_i , то це буде означати, що спостережена випадкове сейсмічне подія Σ_i в момент часу t_n .

Послідовність спостережених в сейсмічній зоні Z^s сейсмічних подій, яка охоплюється каталогом землетрусів:

$$\Sigma_{r_0}(t_0), \Sigma_{r_1}(t_1), \Sigma_{r_2}(t_2), \dots, \Sigma_{r_N}(t_N) \quad (4.8)$$

є однією з реалізацій випадкового процесу (4.7).

На сучасному рівні розвитку сейсмології, моменти часу виникнення землетрусів і значення їх параметрів розглядаються як випадкові величини і спостереження послідовності землетрусів: $\{\varphi_n, \psi_n, h_n, m_n\}$, $n=1, \dots, N$ є реалізацією

випадкового процесу (4.7). Вичерпним способом визначення випадкового процесу є задання його спільної скінченномірної функції розподілу, визначення якої, за винятком деяких тривіальних випадків, є завданням практично нездійсненним.

Однак існує клас марковських процесів, де спільні скінченномірні функції розподілу виражаються через одне і двовимірні розподіли. В цьому випадку, стає можливим визначення розподілу сейсмічності — спільний розподіл параметрів кінцевої сукупності землетрусів, тобто розподіл визначення усіх місць виникнення і магнітуди землетрусів. Отже, якщо випадковий процес (4.7), задовольняє умові марковості, то стане можливим визначення спільного скінченномірного розподілу ймовірності сейсмічних подій (4.8), тобто значень параметрів землетрусів.

У багатьох тимчасових послідовностях випадкових подій існує залежність від минулого. Модель, яка охоплює широке коло форми залежності між подіями, є реальним відображенням дійсності, однак це призводить до великих труднощів при вивченні їх ймовірнісної структури і статистичному аналізі. Тому приймається компромісне рішення, вибирається модель, яка враховує достатні для адекватності моделі залежності і піддається статистичному аналізу. Найкраще відповідність зазначеним вимогам проявляє клас марковських процесів, в яких моделюється залежність від результату спостереження в останній момент часу. Така модель є прийнятним компромісом між залежністю від всієї передісторії процесу і повною незалежністю випадкових подій в послідовності. Марковський процес підходить для опису явищ, зміни яких відбуваються в будь-який момент часу [Валтер, 1976].

Марковський процес – це процес, що володіє тією властивістю, що якщо відомо значення випадкової величини ξ_t , то розподіл ймовірності величини ξ_z , $z > t$ не залежить від ξ_u , $u < t$. Іншими словами, ймовірність будь-якої події, пов'язаної з майбутньою поведінкою процесу за умови, що відомий стан в даний момент t , не зміниться від знання поведінки процесу до моменту часу t . Формально процес є марковським, якщо [Гнеденко, 1961; Романовский, 1948; Langrock, Jahn, 1971]:

$$P \left\{ a \leq \xi_{t_{j+1}} \leq b \middle/ x_1 < \xi_{t_1} \leq x_1 + \Delta x_1, \dots, x_j < \xi_{t_j} \leq x_j + \Delta x_j \right\} = P \left\{ a \leq \xi_{t_{j+1}} \leq b \middle/ x_j < \xi_{t_j} \leq x_j + \Delta x_j \right\} \quad (4.9)$$

Станом системи є ЕДЗ Σ_r , $r=1, \dots, R_s$, тобто значення, прийняті марковським процесом пронумеровані, і в подальшому використовуються номери станів. Якщо в момент t_1 , стався землетрус в ЕДЗ Σ_{r_1} , а момент t_{1+1} спостережено землетрус в ЕДЗ Σ_{r_2} , то це означає, що за час $\tau_1=t_{1+1}-t_1$ відбувся перехід системи r_1 -го стану в r_2 -го стану. Спостережені числа переходів зі стану в стан утворюють матрицю частот переходу [Валтер, 1976; Карлин, 1971]:

$$U = (u_{r_1 r_2}); \quad r_1, r_2 \in (1, R_s) \quad (4.10)$$

Щодо послідовності (4.8) висувається гіпотеза H_0 про те, що вона є реалізацією однорідного чотиривимірною неперервного марковського ланцюга з простором станів: $\Xi^s = \{\Sigma_r\}$, $r=1, \dots, R_s$.

Коли говорять, що функціонування деякої фізичної системи описується неперервним і марковським ланцюгом, це означає наступне:

У початковий момент часу $t=t_0$ система знаходиться в одному з R_s можливих станів Σ_i , перебуваючи в цьому стані випадково час τ_i , показово розподілене з параметром $\alpha_i > 0$, з певною ймовірністю π_{ij} миттєво переходить в деякий стан Σ_j (рис. 4.1). У цьому стані система перебуває випадкове час τ_j , також показово розподілене з параметром $\alpha_j > 0$ і т.д.

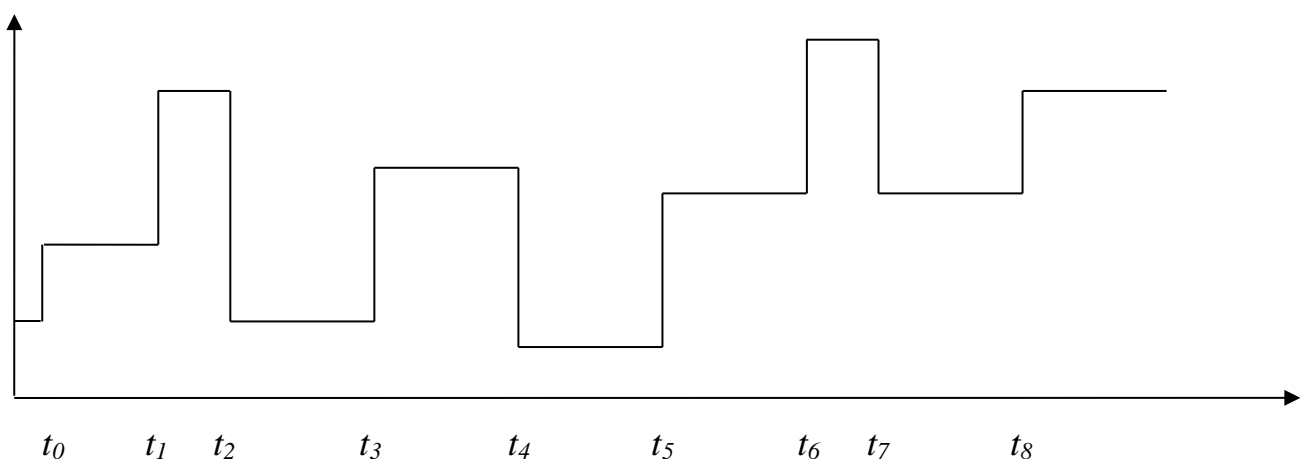


Рис. 4.1. Приклад реалізації неперервного марковського ланцюга

Складовими неперервного марковського ланцюга є випадковий процес, званий вкладеним марковським ланцюгом, при якому різноманітні переходи в просторі станів Ξ^s задовольняють умови марковості, і випадковий потік подій, під впливом якого відбуваються самі різноманітні переходи.

Щодо послідовності спостережених сейсмічних подій: $\{\Sigma_i(t_l), i=1, \dots, R_s; l=1, \dots, L\}$, без урахування моментів часу їх виникнення, висувається гіпотеза про те, що дана послідовність є реалізацією чотиривимірного вкладеного однорідного дискретного марковського ланцюга. Тенденцію переходів зі стану в стан, тобто динамічний режим сейсмічної зони описується матрицею перехідних ймовірностей, вкладеного марковського ланцюга.

Умовна ймовірність:

$$\pi_{ij}^n(k) = P\{\Sigma_j^{n+k} / \Sigma_i^n\} \quad (4.11)$$

того, що після землетрусу на n -му кроці, з параметрами зі стану Σ_i , на $(n+k)$ -му кроці відбувається землетрус з параметрами зі стану Σ_j , називається ймовірністю переходу дискретної марковської ланцюзи з i -го Σ_i в j -е стан Σ_j за k кроків.

Матриця, з елементами (4.11) називається матрицею ймовірностей переходу за k кроків. При значенні $k=1$ отримуємо матрицю однокрокових ймовірностей переходу:

$$\Pi = \begin{pmatrix} \pi_{11} & \pi_{12} & \pi_{13} \dots & \pi_{1R_s} \\ \pi_{21} & \pi_{22} & \pi_{23} \dots & \pi_{2R_s} \\ \dots & \dots & \dots & \dots \\ \pi_{R_s 1} & \pi_{R_s 2} & \pi_{R_s 3} \dots & \pi_{R_s R_s} \end{pmatrix} \quad (4.12)$$

Оцінкою максимальної правдоподібності перехідних ймовірностей служить [Langrock, Jahn, 1979; Hillston, 2009; Валтер, 1971]:

$$\pi_{ij} = \frac{n_{ij}}{\sum_{j=1}^{R_s} n_{ij}}; \quad i, j \in (1, R_s) \quad (4.13)$$

де n_{ij} – число переходів марковського ланцюга зі стану Σ_i в стан Σ_j за період спостережень. Ймовірнісна картина всіляких змін, пов'язаних з переходами зі стану в стан, задається матрицею перехідних ймовірностей (4.13).

Еволюція однорідного дискретного марковського ланцюга повністю визначається вектором початкових розподілів π_0 , матрицею однокрокових ймовірностей переходу (4.12) і простором станів Ξ^s [Джадж и др., 1977; Groß, 2009; Gunkel, 2008; Winkler, 2000].

Оцінки (4.13) мають асимптотичний нормальний розподіл з математичним очікуванням E_{ij} дисперсією D_{ij} :

$$E_{ij} = E(\bar{\pi}_{ij}) = \pi_{ij}; \quad D_{ij} = D(\bar{\pi}_{ij}) = \bar{\pi}_{ij} \frac{(1 - \bar{\pi}_{ij})}{n_i} \quad (4.14)$$

де n_i – спостереження частоты стану Σ_i . Інтервал:

$$\left(\bar{\pi}_{ij} - t_q \frac{\sigma_{ij}}{\sqrt{n_i}}, \bar{\pi}_{ij} + t_q \frac{\sigma_{ij}}{\sqrt{n_i}} \right) \quad (4.15)$$

є відповідним ймовірності $(1-q)$ довірчим інтервалом для невідомих ймовірностей переходу π_{ij} , при відомому середньоквадратичному відхиленні. З значень нормованої функції Лапласа [Смирнов, Дунин-Барковский, 1965]:

$$\Phi(t) = \frac{1}{\sqrt{2\pi}} \int_0^t e^{-\frac{z^2}{2}} dz \quad (4.16)$$

знаходимо значення t_q , відповідне ймовірності $(1-q)$.

У разі невідомого середнього відхилення, при великій кількості спостережень n_i , значення σ_{ij} можна замінити його заможної достатньою оцінкою s_{ij} і вважати інтервал:

$$\left(\bar{\pi}_{ij} - t_q \frac{s_{ij}}{\sqrt{n_i}}, \bar{\pi}_{ij} + t_q \frac{s_{ij}}{\sqrt{n_i}} \right) \quad (4.17)$$

довірчим інтервалом для ймовірностей π_{ij} . У розглянутому прикладі, оцінки ймовірностей переходу знаходяться в межах довірчих інтервалів відповідні ймовірності $1-q=0.95$.

Вектор початкових розподілів визначиться за формулою:

$$\pi_0 = (\pi_0^j), \quad j \in (1, R_v); \quad \pi_0^j = \frac{N^j}{N} \quad (4.18)$$

де N – загальне число землетрусів, N^j число землетрусів з параметрами зі стану Σ_j . Ланцюг Маркова, для якої ймовірність переходу не залежить від кроку n , на якому він здійснюється, – тобто виконується умова:

$$\pi_{ij}^n(k) = \pi_{ij}(k) \quad (4.19)$$

називається однорідним дискретним ланцюгом Маркова. Якщо ж, виконується умова:

$$\lim_{k \rightarrow \infty} \pi_{ij}^n(k) = \pi_j \quad (4.20)$$

то ланцюг Маркова є ергодичним.

Ергодичні марковські ланцюги володіють чудовою для практичних застосувань властивістю – стаціонарний граничний розподіл такого ланцюга, з будь-якою мірою точності, можна визначити з однієї досить довгої її реалізації:

[Langrock, Jahn, 1979]:

$$P \left\{ \lim_{N \rightarrow \infty} \frac{N^j}{N} = \pi_j^+ \right\} = 1 \quad (4.21)$$

Це означає, що стане можливим опис розподілу сейсмічного режиму і визначення розподілу параметрів землетрусів. Наявність властивості ергодичності ланцюга визначається спектральними властивостями матриці перехідних ймовірностей (4.11) [Гантмахер, 1967; Курош, 1965; Ланкастер, 1978; Романовский, 1948].

Землетруси відбуваються у випадкові моменти часу, і проміжки часу між ними теж випадкові. Спостереження показують, розподіл вогнищ землетрусів в просторі, і моментів їх виникнення на осі часу не є однорідними, мають тенденцію групуватися в просторі і в часі. Однорідність за часом означає стаціонарність випадкового потоку сейсмічних подій. Зазвичай на практиці зустрічаються потоки стаціонарні на обмеженому інтервалі часу, тобто інтенсивність подій залишається постійною для всього інтервалу. Багато фізичних процесів, які вважаються стаціонарними, є такими лише на вузькому інтервалі часу, і поширення стаціонарності до нескінченності є зручним для практичного застосування

спрощенням. На відміну від багатьох фізичних процесів, відсутність стаціонарності в сейсмічних процесах проявляється на коротких інтервалах і при збільшенні інтервалу часу статистичні характеристики сейсмічного режиму набувають стійкості. Якщо моменти виникнення землетрусів нанести на вісь часу отримаємо випадковий потік однорідних сейсмічних подій, що відбуваються в випадкові моменти часу. Найважливішою характеристикою потоку є закон розподілу $F(t)$ інтервалів часу між двома сусідніми сейсмічними подіями в потоці.

Стаціонарні потоки, де інтервали часу між сусідніми сейсмічними подіями є незалежними і однаково розподіленими випадковими величинами є потоками Пальма. Потоки, що володіють властивістю стаціонарності, ординарності і відсутністю післядії називаються пуассонівськими, і є окремим випадком потоків Пальма: потоків з обмеженою післядією [Вентцель, Овчаров, 2000]. В геологічному масштабі, за одиницю часу можна взяти інтервал довжиною в роки і десятки років. Пуассонівські потоки є найпростішими потоками, так як для дослідження систем, що знаходяться під їх впливом застосовуються прості методи.

Вище було показано, що зону Вранча можна представити у вигляді об'єднання $\Xi^s = \{\Sigma_r\}$, $r=1, \dots, R_s$, где R_s – число елементарних зон. На основі марковської моделі сейсмічного режиму були визначені розподіли ймовірності виникнення сейсмічних подій в ЕДЗ $\{\Sigma_v, v=1, \dots, R_s\}$. Моменти виникнення землетрусів сейсмічної зони Ξ^s , нанесені на вісь часу утворюють випадковий потік подій Π , який назвемо вихідним потоком землетрусів. Вихідний потік складається з R потоків, які утворюються з вихідного потоку, шляхом випадкового відбору подій з ймовірністю p_v . Тобто, моменти виникнення сейсмічних подій з деякою Σ_v утворює потік з інтенсивністю λ_v . Даний потік назвемо вкладеним потоком першого порядку, і позначимо через $\Pi_{1,v}$. Кількість вкладених потоків дорівнює числу станів марковського ланцюга, тобто числу елементарних зон R_s . Динамічним режимом сейсмічної зони, виникненням землетрусів в ЕДЗ управляє випадковий потік подій. В термінах марковських моделей це означає перехід системи зі стану Σ_η в стан Σ_μ . Назвемо цей потік вкладеним потоком другого порядку $\Pi_{2,\eta\mu}$, який утворюється з деякого вкладеного

потоків першого порядку $\Pi_{1,\mu}$, також шляхом випадкового відбору подій з ймовірністю $\pi_{\eta\mu}$. Таким чином, відбувається випадкове розрідження потоку Пальма: з вихідного потоку Π складаються R_s вкладених потоків першого порядку Π_1 , і $(R_s)^2$ потоків другого порядку Π_2 . Потоки Пальма, розріджені випадковим чином прагнуть до пуассонівського потоку [Вентцель, Овчаров, 2000]. Для випадкових потоків має місце теорема, аналогічна центральній граничній теоремі: якщо підсумувати достатньо велике число потоків різної структури, то сумарний потік буде близький до пуассонівського потоку. Головною умовою тут є можливість співставлення інтенсивностей потоків. Ця теорема дійсна для ординарних стаціонарних і нестаціонарних потоків [Вентцель, Овчаров, 2000].

Велике число станів марківського процесу, наприклад, кілька сотень, може привести до виникнення проблеми, відомої під назвою «прокляті розмірності» в додатку теорії марковських процесів. Велике число станів, відповідно великі розмірності матриць, що виникають при дослідженні марковських процесів, є головною перешкодою на шляху їх застосування. Групування значень параметрів землетрусів призводить до укрупнення станів ланцюга Маркова, внаслідок чого вдається обійти труднощі, пов'язані з розмірністю простору станів. Для аналізу СН досить передбачити інтервали, що містять значення параметрів. При цьому повинен існувати деякий компроміс між розмірами інтервалів, об'ємом вибірок і втратою інформації про особливості параметрів сейсмічності. Наприклад, при групуванні спостережених значень в деякому інтервалі можуть виявитися значення параметрів корових і проміжних землетрусів.

4.3. Дослідження вкладеного ланцюга Маркова

Алгоритм побудови марковської моделі демонструється на прикладі 705 землетрусів, з магнітудою понад 4, зони Вранча з каталогу ROMPLUS.

Перш ніж приступити до дослідження марковського ланцюга наведемо кілька необхідних з теорії марковських процесів визначення: стан Σ_j називається досяжним зі стану Σ_i , якщо $\pi_{ij}(k) > 0$ для деякого числа $k > 0$, тобто ймовірність того, що,

вирушаючи зі стану Σ_i , ланцюг за кінцеве число кроків потрапить в стан Σ_j , позитивна.

Досяжні один з одного стану називаються сполученими. Марковський ланцюг називається незвідним – якщо всі її стани є сполученими. Інакше множина станів розбивається на класи еквівалентності – підпростору, сполучених один з одного станів. Для послідовності параметрів землетрусів це означає, що можливе виникнення землетрусу з параметрами з будь-якого стану. Простір станів Ξ^s складається з об'єднання $R=27$ ЕДЗ (станів) (таб. 4.1).

Періодом $d(j)$ стану Σ_j називається найбільший спільний дільник всіх цілих чисел $k \geq 1$, для яких $\pi_{jj}(k) > 0$. Марковський ланцюг, кожний стан якого має період, що дорівнює одиниці, називається неперіодичном. Стан називається поворотним, якщо [Langrock, Jahn, 1979]:

$$\sum_{k=0}^{\infty} \pi_{jj}^k = 1 \quad (4.22)$$

Таблиця 4.1

Стани чотиривимірного марковського ланцюга.

Номер стану	Інтервал магнітуди	Інтервал глибини	Інтервал широти	Інтервал довготи
1	(4.00–4.53)	(9.9 – 79.4)	(45.20– 46.00)	26.00– 26.41
2	(4.00–4.53)	(9.9 – 79.4)	(45.20– 46.00)	26.41– 26.82
3	(4.00–4.53)	(79.4 –148.9	(45.20– 46.00)	(26.00– 26.4
4	(4.00–4.53)	(79.4 –148.9	(45.20– 46.00)	(26.41– 26.8
5	(4.00–4.53)	(148.9 –218.4)	(45.20– 46.00)	26.00– 26.41
6	(4.00–4.53)	(148.9 –218.4)	(45.20– 46.00)	26.41– 26.82
7	(4.53–5.06)	(9.9 – 79.4)	(45.20– 46.00)	26.41– 26.82
8	(4.53–5.06)	(79.4 –148.9)	(45.20– 46.00)	26.00– 26.41
9	(4.53–5.06)	(79.4 –148.9)	(45.20– 46.00)	26.41– 26.82
10	(4.53–5.06)	(148.9 –218.4)	(45.20– 46.00)	26.00– 26.41
11	(4.53–5.06)	(148.9 –218.4)	(45.20– 46.00)	26.41– 26.82
12	(5.06–5.59)	(9.9 – 79.4)	(45.20– 46.00)	26.41– 26.82
13	(5.06–5.59)	(79.4 –148.9)	(45.20– 46.00)	26.00– 26.41

14	(5.06–5.59)	(79.4 –148.9)	(45.20– 46.00)	26.41– 26.82
15	(5.06–5.59)	(148.9 –218.4)	(45.20– 46.00)	26.00– 26.41
16	(5.06–5.59)	(148.9 –218.4)	(45.20– 46.00)	26.41– 26.82
17	(5.59–6.12)	(9.9 – 79.4)	(45.20– 46.00)	26.41– 26.82
18	(5.59–6.12)	(79.4 –148.9)	(45.20– 46.00)	26.00– 26.41
19	(5.59–6.12)	(79.4 –148.9)	(45.20– 46.00)	26.41– 26.82
20	(5.59–6.12)	(148.9 –218.4)	(45.20– 46.00)	26.00– 26.41
21	(5.59–6.12)	(148.9 –218.4)	(45.20– 46.00)	26.41– 26.82
22	(6.12–6.64)	(9.9 – 79.4)	(45.20– 46.00)	26.41– 26.82)
23	(6.12–6.64)	(79.4 –148.9)	(45.20– 46.00)	26.00– 26.41
24	(6.12–6.64)	(79.4 –148.9)	(45.20– 46.00)	26.41– 26.82
25	(6.64–7.17)	(79.4 –148.9)	(45.20– 46.00)	26.41– 26.82
26	(7.17–7.70)	(79.4 –148.9)	(45.20– 46.00)	26.41– 26.82
27	(7.17–7.70)	(148.9 –218.4)	(45.20– 46.00)	26.41– 26.82

Для потоку землетрусів зворотність означає, що можна визначити ймовірність між повторенням двох землетрусів з параметрами з Σ_j , виникнення землетрусу з параметрами з іншого стану Σ_i . Нехай $\pi_{ij}(k)$ ймовірність події полягає в тому, що k -е, після землетрусу з параметрами зі стану Σ_i , землетрус буде з параметрами зі стану Σ_j . Ймовірності $\pi_{ij}(k)$ складають матрицю k -крокових ймовірностей переходу $\Pi(k)$.

Для однорідних ланцюгів $\Pi(k)=\Pi^k$, де Π – матриця однокрокових ймовірностей переходу: розподіл ймовірностей параметрів наступної сейсмічної події [Huisinga, 2005; Langrock, Jahn, 1979]. У практичних додатках обчислення степеней матриці здійснюється за формулою:

$$\Pi(k)=\Lambda H^k \Lambda^{-1} \quad (4.23)$$

де Λ – повна система правих власних векторів, відповідних власним значенням: $\lambda_1, \dots, \lambda_R$, Λ^{-1} – матриця обернена до матриці Λ . Матриця обернена до матриці Λ – повна система лівих власних векторів, H – діагональна матриця:

$$H = \begin{pmatrix} \lambda_1 & 0 & 0 & 0 & 0 \\ \cdot & \lambda_2 & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & 0 & \lambda_{R_v} \end{pmatrix} \quad (4.24)$$

Власними значеннями матриці Π є рішення характеристичного рівняння:

$$|\Pi - \lambda E| = 0 \quad (4.25)$$

Серед власних значень:

$\lambda_1=(1.000, 0.000)$; $\lambda_2=(0.515, 0.000)$; $\lambda_3=(0.430, 0.000)$; $\lambda_4=(0.250, 0.065)$; $\lambda_5=(0.250, -0.065)$; $\lambda_6=(-0.210, 0.000)$; $\lambda_7=(0.113, 0.125)$; $\lambda_8=(0.113, -0.125)$; $\lambda_9=(-0.039, 0.143)$; $\lambda_{10}=(-0.039, -0.143)$; $\lambda_{11}=(-0.144, 0.000)$; $\lambda_{12}=(0.018, 0.126)$; $\lambda_{13}=(0.018, -0.126)$; $\lambda_{14}=(0.094, 0.051)$; $\lambda_{15}=(0.094, -0.051)$; $\lambda_{16}=(-0.076, 0.065)$; $\lambda_{17}=(-0.076, -0.065)$; $\lambda_{18}=(-0.078, 0.000)$; $\lambda_{19}=(0.076, 0.000)$; $\lambda_{20}=(0.007, 0.062)$; $\lambda_{21}=(0.007, -0.062)$; $\lambda_{22}=(-0.058, 0.000)$; $\lambda_{23}=(0.000, 0.000)$; $\lambda_{24}=(0.000, 0.000)$; $\lambda_{25}=(0.000, 0.000)$; $\lambda_{26}=(0.000, 0.000)$; $\lambda_{27}=(0.000, 0.000)$, одиниця (λ_1) виявилася найбільшим власним значенням і з кратністю один.

Отже, марковський ланцюг, реалізацією якого є спостережений порядок опцій землетрусів вогнища Вранча – нескоро чуваний, неперіодичний, і позитивно поворотний [Langrock, Jahn, 1979]. Всі степені Π^k , починаючи з $k > 5$, виявилися однаковими і всі рядки матриці рівними:

$$\pi_{ij}(k) = \pi_j; k \geq 5 \quad (4.26)$$

Для динамічних систем, функціонування яких описується марковським ланцюгом, умова (4.22) означає здатність системи "забувати своє минуле". У послідовності незалежних випадкових величин забування минулого відбувається на наступному кроці, в ергодичних ланцюгах через кінцеве число кроків. Кожний ергодичний марковський ланцюг має, незалежним від початкового розподілу граничний розподіл π^+ . Цей розподіл є стаціонарним щодо матриці перехідних ймовірностей Π :

$$\sum_{i=1}^{R_v} \pi_i^+ \pi_{ij} = \pi_j^+ \quad (4.27)$$

Граничні ймовірності вкладеного марковського ланцюга характеризують середнє число станів протягом досить тривалого періоду спостереження за системою. Тобто визначає середнє число сейсмічних подій в ЕДЗ.

Для досліджуваної послідовності землетрусів умова (4.26) виконується. Отже, сейсмічний процес знаходиться в стаціонарному режимі – ймовірність виникнення землетрусу з параметрами з деякого стану Σ_j не залежить від часу його виникнення. Компоненти вектора граничних розподілів станів обчислюються за формулою [Langrock, Jahn, 1979]:

$$\pi_j^+ = \frac{D_{jj}}{\sum_{i=1}^{R_v} D_{ii}} \quad (4.28)$$

де D_{ii} – головний мінор визначника $|E-P|$, отриманий викреслюванням i -го рядка i , i -го стовпця. Вектор граничних розподілів станів: $\pi^+ = (0.0028; 0.0099; 0.0685; 0.3942; 0.0344; 0.0671; 0.0028; 0.0185; 0.1861; 0.0128; 0.0270; 0.0015; 0.0056; 0.0567; 0.0071; 0.0142; 0.0028; 0.0126; 0.0474; 0.0014; 0.0097; 0.0027; 0.0014; 0.0057; 0.0043; 0.0014; 0.0014)$ збігається з рядками степеней матриці перехідних ймовірностей і незначній мірі відрізняється від початкового розподілу станів вкладеного марковського ланцюга. Оцінка ймовірностей (4.28) є головною метою побудови марковської моделі потоку землетрусів.

Нехай випадкова величина $\eta^j(N)$ означає число землетрусів з параметрами з деякого стану Σ_j , при загальній кількості N землетрусів. Математичне сподівання величини $\eta_j(N)$ визначається за формулою [Сираждинов, 1955]:

$$M\eta^j(N) = N\pi_j^+ + O(1) \quad (4.29)$$

Теоретичне число землетрусів в ЕДЗ, обчислені за формулою (4.29): (6.0; 36.0; 45.0; 532.2; 823.1; 927.1; 12.0; 6.0; 12.0; 11.0; 208.8; 152.1; 137.0; 6.0; 3.0; 3.0; 29.8; 28.8; 83.5; 2.0; 3.0; 9.0; 13.9; 14.9; 2.0; 2.0), в незначній мірі відрізняється від їх спостережуваного числа: (6; 36; 45; 532; 821; 926; 12; 6; 12; 11; 210; 152; 137; 6; 3; 3; 30; 29; 85; 2; 3; 9; 14; 15; 2; 2).

Якість моделювання марковського ланцюга послідовності землетрусів, тобто міри відхилення теоретичних середніх (4.29) від спостережених значень, оцінюється вибірковою статистикою тестового критерію:

$$\chi^2_q = \sum_{i=1}^{R_s} \frac{(g_N^i - N\pi_i^+)^2}{N\pi_i^+} \quad (4.30)$$

яка має χ^2 розподіл, з (R_s-1) ступенями свободи. Для сукупності землетрусів з каталогу ROMPLUS, спостережене незначне, на 5%-ному рівні значущості з 25 ступенями свободи, відхилення $\chi^2_q=0.21$ теоретичного середнього числа станів від їх спостережуваного значення.

Таким чином, одиниця виявилася максимальною власним значенням кратності один; вектор граничних розподілів станів є позитивним; всі рядки степеней матриці перехідних ймовірностей є однаковими і рівними вектору граничних розподілів станів вкладеного марковського ланцюга. Кожен з перерахованих факторів вже є достатнім для ергодичності, вкладеного марковського ланцюга умовою [Джадж и др., 1977; Карлин, 1971; Романовский, 1949; Langrock, Jahn, 1979].

У разі, коли однорідний марковський ланцюг є адекватною моделлю послідовності землетрусів, спільний розподіл параметрів n землетрусів по станам визначиться за формулою:

$$P\{\Sigma_{i_1}, \Sigma_{i_2}, \dots, \Sigma_{i_n}\} = \pi_{i_1} \pi_{i_1 i_2} \dots \pi_{i_{n-1} i_n} \quad (4.31)$$

При дослідженні марковського ланцюга виникла проблема, пов'язана з порядком матриці перехідних ймовірностей $R_s = R_\phi R_\psi R_h R_m$. З цього виразу видно, що порядок матриці різко зростає при збільшенні однієї з величин R_ϕ, R_ψ, R_h, R_m на одну одиницю. Якщо діапазони параметрів розбити на більше число інтервалів, то однорідність даних в ЕДЗ і детальність оцінок збільшується, але зменшується частота спостережених станів і матриця перехідних ймовірностей виходить сильно розрідженою, через збільшення порядку матриці. Таким чином, доводиться прийняти компромісне рішення між адекватністю ймовірнісних моделей, наглядністю вибірок і порядком матриць.

У дослідженні марковської умови потоку землетрусів застосовуються методи матричної алгебри. Наприклад, перевірка адекватності марковської моделі проводиться на основі спектрального аналізу перехідної і інфінітезимальної матриць, що пов'язано із застосуванням чисельних методів. Параметри марковської моделі визначаються за допомогою методів обчислювальної математики, які має в своєму розпорядженні процедури наближених обчислень. Як відомо, часто вихідні дані містять похибки, які є систематичними. У нашому випадку, це похибки у визначенні параметрів землетрусів, які не можуть вплинути на якість марковської моделі потоку спостережених землетрусів. Чисельні методи також є джерелом похибок. Якщо матриці погано обумовлені: число обумовленості близько нуля – внаслідок неминучих похибок округлення на обчислювальній машині, отримані наближені рішення і результати досліджень можуть виявитися нереальними. Теоретично чисельні похибки можуть бути зменшені, наприклад, відповідною організацією алгоритму проведення арифметичних операцій – зміною порядку обчислень, що важко що важко зробити практично. При комп'ютерних операціях над числами з плаваючою точкою обчислювальний процес може перерватися через переповнення і зникнення порядку. Переповнення – коли число стає настільки великим, що не поміщається в розрядну сітку, а зникнення порядку – коли результат арифметичної операції стає настільки близьким до нуля і порядок числа виходить за межі розрядної сітки.

Причиною цього є розрідженість матриць і те, що елементи матриці ймовірностей переходу менше одиниці, а елементи інфінітезимальної матриці є близькими до нуля. У деяких випадках матриці: ймовірностей переходу та інфінітезимального порядку, наприклад, більше 30, тобто коли число ЕДЗ перевищує 30, виявилися погано зумовленими. У таких випадках пропонується підібрати число ЕДЗ, при якому число обумовленості $cond(A)$: кількісна оцінка обумовленості стане близьким до одиниці.

4.3.1. Оцінка параметрів неперервного марковського ланцюга

Відносно спостережених послідовностей землетрусів вогнища Вранча висувається гіпотеза H_0 про те, що дана послідовність є реалізацією однорідного неперервного марковського ланцюга з простором станів $\Xi^s = \{\Sigma_v\}$ (табл. 4.1). Випадковий потік подій, який управляє переходами системи між станами, є пуассонівським, тобто розподіл інтервалів часу між переходами відповідає експоненціальному розподілу..

Однорідний неперервний марковський ланцюг з R_s станами повністю визначається завданням $(R_s)^2$ чисел u_{ij} , які відповідають умові:

$$\sum_{j=1}^{R_v} u_{ij} = 0; \quad u_{ij} > 0; \quad j \neq i; \quad u_{ii} = -\sum_{j \neq i} u_{ij}; \quad \lim_{\tau \rightarrow 0} \frac{p_{ij}(\tau)}{\tau} = u_{ij}; \quad \lim_{\tau \rightarrow 0} \frac{1 - p_{ii}(\tau)}{\tau} = u_{ii} \quad (4.32)$$

де u_{ij} : $i \neq j$ інтенсивність переходу ланцюга зі стану Σ_i в стан Σ_j . При малих значеннях Δt , $u_{ij}\Delta t$, при $i \neq j$ являє собою щільність ймовірності переходу з Σ_i в стан Σ_j за час Δt .

Характерна властивість неперервних марковських ланцюгів полягає в тому, що за малий проміжок часу ймовірність зміни стану набагато менше ніж навпаки.

Якщо на інтервалі часу $(0, t)$ відбулося n_{ij} переходів, і зареєстрований час перебування в стані Σ_i склав t_i , то оцінкою інтенсивності переходів служить [Карлин, 1971]:

$$u_{ij}(t) = \frac{n_{ij}(t)}{t_i} \quad (4.33)$$

Ці величини зв'язані співвідношеннями (4.32), і утворюють інфінітезимальну матрицю U неперервного марковського ланцюга:

$$U = \begin{pmatrix} u_{11} & u_{12} & \cdot & \cdot & u_{1R_v} \\ u_{21} & u_{22} & \cdot & \cdot & u_{2R_v} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ u_{R_v 1} & \cdot & \cdot & \cdot & u_{R_v R_v} \end{pmatrix} \quad (4.34)$$

Нас цікавить наступні характеристики процесу:

1. стаціонарна ймовірність того, що в момент часу t_0 марковський процес буде перебувати в стані Σ_i , тобто після землетрусу в елементарній зоні Σ_i новий землетрус ще не відбувся
2. середнє число землетрусів в ЕДЗ Σ_i
3. середній час перебування в деякому стані Σ_i , тобто середня тривалість відсутності землетрусу після події з параметрами із зони Σ_i землетрусу
4. спільне розподілу кінцевої сукупності землетрусів.

Для визначення характеристик і дослідження ланцюга використовуються спектральні властивості матриці (4.34). Серед власних значень нуль виявився максимальним власним значенням кратності один. Матриця $P(\tau) = \{p_{ij}(\tau)\}$, $i, j = 1, \dots, R_s$ – ймовірностей переходу зі стану Σ_i в стан Σ_j за час τ , виражається через інфінітезимальну матрицю за формулою [Валтер, 1971]:

$$P(\tau) = (p_{ij}(\tau)) = e^{u\tau}; i, j \in (1, R_s) \quad (4.35)$$

Зручним для практичних застосувань є уявлення:

$$P(\tau) = X\Lambda(\tau)X^{-1} \quad (4.36)$$

де X – матриця, стовпцями якої є повна система правих власних векторів матриці U , X^{-1} – матриця, обернена до X (повна система лівих власних векторів), а $\Lambda(\tau)$ – діагональна матриця виду:

$$\Lambda(\tau) = \begin{pmatrix} e^{\lambda_1\tau} & \cdot & \cdot & \cdot & \cdot \\ \cdot & e^{\lambda_2\tau} & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & e^{\lambda_{R_s}\tau} \end{pmatrix} \quad (4.37)$$

Дійсні частини відмінних від нуля власних значень виявилися негативними, отже, з виразів (4.35) і (4.36) випливає, що ймовірнісна функція $p_{ij}(\tau)$ прагне до своєї межі з показовою швидкістю:

$$\lim_{\tau \rightarrow \infty} p_{ij}(\tau) = p_j \quad (4.38)$$

З цього випливає, що неперервний марковський ланцюг є ергодичним. Отже, процес накопичення енергії пружних напруг і їх миттєвого вивільнення знаходиться в сталому режимі. Стаціонарний граничний розподіл станів системи визначається за формулою [Рытов, 1976; Langrock, Jahn, 1979]:

$$p_j = \frac{D_{jj}}{\sum_{i=1}^{R^s} D_{ii}} \quad (4.39)$$

де D_{ii} – головний мінор визначника $|-U|$, отриманий викреслюванням рядка i стовпця з номером i . Вектор граничних розподілів при 10 станах: $p=(0.2529; 0.2821; 0.0503; 0.1393; 0.0394; 0.1553; 0.0448; 0.001; 0.0025; 0.0325)$ є позитивним, отже, марковський процес, реалізацією якого є спостереження послідовності землетрусів Вранчі, є ергодичним.

Граничні ймовірності неперервного марковського ланцюга характеризують частку часу t_j , протягом якого система перебуває в деякому стані Σ_j , тобто середній період часу очікування наступного землетрусу.

Для неперервних ланцюгів Маркова існує властивість, аналогічне властивості нескорочуваності і відсутності періодичності дискретного марковського ланцюга, а саме, для будь-яких двох станів Σ_i, Σ_j існує сукупність чисел

$$u_{ik_1}, u_{k_1k_2}, \dots, u_{k_nj} \quad (4.40)$$

що задовольняють умові:

$$u_{ik_1}, u_{k_1k_2}, \dots, u_{k_nj} > 0 \quad (4.41)$$

Це означає, що марковський ланцюг за кінцевий час з позитивною ймовірністю може перейти з будь-якого стану в будь-який стан [Валтер, 1971]. Мовою сейсмології даний факт означає, що за кінцевий час, після землетрусу в Σ_i , з позитивною ймовірністю може статися землетрус в будь-якій іншій зоні.

Для ергодичного неперервного марковського ланцюга математичне очікування тривалості перебування $\eta_i(t)$ в стані Σ_i ($i=1, \dots, R_s$), за час t складе [Барлоу, Прошан, 1969; Карлин, 1971; Hilston, 2009]:

$$E(\eta_i(t)) = tp_i + O(1) \quad (4.42)$$

Нехай l_{ij} середня тривалість часу до першого переходу зі стану Σ_i в стан Σ_j , тобто середній час очікування до першої появи землетрусу в ЕДЗ Σ_j , після землетрусу в зоні Σ_i . А η_{ij} – середнє показникового розподілу часу між виникненням землетрусів з параметрами з станів Σ_i і Σ_j . Тоді має місце вираз:

$$l_{ij} = \sum_{k \neq j} \pi_{ik} (\mu_{ik} + l_{kj}) + \pi_{ij} \eta_{ij} \quad (4.43)$$

Середній час, що минув від моменту виникнення землетрусу в зоні Σ_i , до виникнення будь-якого землетрусу (середня довжина часу відсутності землетрусу) визначиться за формулою:

$$\mu_i = \sum_{k=1}^{R^s} \pi_{ik} \mu_{ik} \quad (4.44)$$

Усреднюючи по всім станам, отримаємо середню тривалість сейсмічного затишшя:

$$t_{ave} = \sum_{i=1}^{R^s} p_i \mu_i \quad (4.45)$$

При 28 ЕДЗ в сейсмічній зоні Вранча, середній час відсутності землетрусів з магнітудою більше 3, становить 70 днів. Спостережений час сейсмічного затишшя в ЕДЗ становить: (76; 1428; 252; 3276; 5054; 5499; 141; 52; 380; 85; 6008; 1824; 1207; 120; 33; 111; 1935; 624; 3859; 5; 458; 1916; 1495; 3124; 7; 0; 1314; 0), і незначно відрізняється від теоретичного часу відсутності землетрусів в ЕДЗ: (76.3; 1432.7; 253.1; 3289; 5085.1; 5525.4; 141.7; 52.3; 380.3; 85.1; 5995.5; 1831.5; 1211.7; 120.6; 33.2; 110.5; 1926.2; 623; 3803.5; 5; 456.1; 1915.3; 1490.2; 3112.8; 7; 1.0; 1320.3; 1.0).

Середній період повторення землетрусів з параметрами зі стану Σ_i (середній час повернення) визначиться з виразу [Langrock, Jahn, 1979]:

$$l_{jj} = \frac{1}{\pi_j^+} \sum_{i=1}^{R^s} \pi_i^+ \mu_i \quad (4.46)$$

Для визначення середнього числа переходів зі стану в стан застосовується формула [Gunkel, 2008]:

$$\gamma_{ij}(t) = \frac{t}{l_{ij}} - 1 + \frac{1}{l_{ij}} \sum_{k=1}^{R^s} p_k l_{kj} \quad (4.47)$$

з виразу:

$$\gamma_j(t) = \sum_{k=1}^{R_s} p_k \gamma_{kj}(t) \quad (4.48)$$

визначається безумовне середнє число виникнення землетрусів з параметрами з елементарної зони Σ_j .

Неперервні марковські ланцюги є математичними моделями фізичних систем, де переходи між станами Σ_i і Σ_j – відбуваються під дією потоку подій з інтенсивністю λ_{ij} , рівною відповідній щільності ймовірності (інтенсивності) переходу зі стану Σ_i в стан Σ_j . Переходи між станами, без урахування часу очікування переходу, утворюють вкладений марковський ланцюг. Таким чином, функціонування системи з R_s станами, складається з $(R_s)^2$ потоків подій. Якщо всі потоки подій є пуассонівськими, то процес в системі буде марковським [Winkler, 2000].

4.3.2. Статистика марковських ланцюгів

Якщо ми спостерігаємо деякі послідовності подій, щодо яких передбачається, що вони є реалізаціями процесів Маркова, то математична статистика має множина методів перевірки гіпотези за допомогою числа переходів між станами. Вони засновані на показнику якості прогнозу, виміряного за критерієм χ^2 .

Нехай гіпотеза H_1 полягає в тому, що послідовність сейсмічних подій, спостережених в моменти часу $t_0, t_1, t_2, \dots, t_n$, є реалізацією однорідного простого марковського ланцюга. Розглянемо альтернативні гіпотези:

- H_2 – послідовність подій є реалізацією послідовності випадкових подій, що утворюють марковський ланцюг другого порядку
- H_0 – послідовність подій є реалізацією послідовності незалежних випадкових подій.

Ймовірності переходу за два кроки виражаються співвідношенням [Langrock, Jahn, 1979]:

$$P_{ijk} = \frac{v_{ijk}}{\sum_l^{R_s} v_{ijl}} \quad (4.49)$$

де v_{ijk} число точок n , для яких на n -му кроці спостережену стан Σ_i , на $(n+1)$ -м кроці Σ_j , і на $(n+2)$ кроці Σ_k . Правдоподібність гіпотези H_1 визначається значенням тестового критерію:

$$\chi^2 = \sum_{i,j,k=1}^{R_s} n_{ij} \frac{(\pi_{ijk} - \pi_{jk})^2}{\pi_{jk}} \quad (4.50)$$

Застосуємо критерій χ^2 для перевірки правдоподібності гіпотези H_1 проти конкуруючої гіпотези H_2 : $p_{1jk} = p_{2jk} = \dots = p_{Rjk} = p_{jk}$. Рівень значущості α покладемо рівним $\alpha = 0.01$. Число ступенів свободи визначимо вирахуванням з $R_s(R_s-1)^2$ числа ненульових елементів в матриці (4.13).

Правдоподібність гіпотези H_1 , про те, що послідовність сейсмічних подій є реалізацією однорідного простого марковського ланцюга, підтверджується значенням тестового критерію (4.50) $\chi_q = 1237.4$, яке менше значення $\chi_{0.01,2484} = 2374.7$, відповідного 5%-не рівня значущості зі 2484-ю ступенями свободи.

У статистичних додатках часто виникає завдання перевірки гіпотези про незалежність ознак. Правдоподібність цієї гіпотези перевіряється за допомогою χ^2 тесту на незалежність. Для цього складається $(R_s \times R_s)$ таблиця спряженості ознак, де на перетині рядка і стовпця знаходиться частота спостереження, і обчислюється статистика тестового критерію χ^2 [Langrock, Jahn, 1979]:

$$\chi^2 = \sum_{i,j=1}^{R_s} \frac{\left(n_{ij} - \frac{n_i n_j}{n} \right)^2}{\frac{n_i n_j}{n}}; \quad n_i = \sum_{j=1}^{R_s} n_{ij}; \quad n_j = \sum_{i=1}^{R_s} n_{ij} \quad (4.51)$$

Вибіркова статистика (4.51) асимптотично χ^2 розподілена з $(R_s-1)^2$ степенями свободи. Якщо необхідно досягти задовільного наближення, розмір вибірки повинен бути досить великим, і очікувана частота спостережень в непересічних інтервалах не менше 5. При цьому допустиме число елементів таблиці спряженості зі значенням менше 5 не повинно перевищувати 20% [Sachs, Hedderich, 2006].

Статистика тестового критерію (4.50) $\chi_q = 984.6$ набагато перевершує граничне значення $\chi_{0.05,92} = 70.88$, що відповідає рівню значущості $\alpha = 0.05$ с 92 ступенями

свободи, що є підставою для відхилення гіпотези $H_0: p_{1j} = p_{2j} = \dots = p_{Rj} = p_j$, на користь гіпотези H_1 .

При $N \rightarrow \infty$ вибіркоий розподіл статистики:

$$\chi^2 = \sum_{i,j=1}^{R_s} \frac{(n_{ij} - N\pi_i\pi_j)^2}{N\pi_i\pi_j} \quad (4.52)$$

де π_i, π_j відносні частоти станів Σ_i і Σ_j при виникненні N землетрусів, сходиться до розподілу суми незалежних випадкових величин [Сираждинов, 1955]:

$$\chi_{(R_s-1)^2}^2 + 2\chi_{(R_s-1)}^2 \quad (4.53)$$

Вибіркове значення $\chi_q = 1120.2$ статистики (4.52) перевищує порогове значення $\chi_{0.05} = 647.44$. Ще одним критерієм незалежності є, асимптотично χ^2 розподілена з R_s ступенями свободи, вибіркова статистика [Башарин, 1957]:

$$\chi_q^2 = \frac{1}{N} \sum_{i=1}^{R_s} \frac{(n_{ii} - N\pi_i^2)^2}{\pi_i^2(1-2\pi_i)} + \frac{3}{\left(1 - 3 \sum_{i=1}^{R_s} \frac{\pi_i^2}{1+2\pi_i}\right) N} \left(\sum_{i=1}^{R_s} \frac{n_{ii} - N\pi_i^2}{1+2\pi_i} \right) \quad (4.54)$$

Таким чином, за трьома тестовим критеріям (4.51), (4.52), (4.54) гіпотеза H_0 відхиляється на користь гіпотези H_1 .

Правдоподібність припущення про те, що потік землетрусів є реалізацією неперервного марковського ланцюга, залежить від значення тестового критерію [Hilston, 2009; Langrock, Jahn, 1979]:

$$\chi_q^2 = \sum_{i=1}^{R_s} \frac{(n_i(t) - \gamma_i(t))}{\gamma_i(t)} \quad (4.55)$$

де $n_i(t)$ і $\gamma_i(t)$ спостереження і теоретична частоти виникнення землетрусів з параметрами зі стану $\Sigma_i, i=1, \dots, R_s$. Наприклад, застосовуючи формулу (4.55) для моделі вогнища Вранча отримано значення $\chi_q = 8.69$, яке менше порогового значення $\chi_{0.05;11} = 19.7$, відповідного ймовірності значущості $\alpha = 0.05$ с 11-а ступенями свободи, що підтверджує гіпотезу про те, що неперервний марковський ланцюг може розглядатися як ймовірнісна модель потоку вранчських землетрусів.

Вектор спостережуваного часу перебування в станах (середня тривалість сейсмічного затишшя): $t_0 = (10147; 11185; 2023; 5586; 1596; 6419; 1823; 39; 104; 1315)$

незначно відхиляється від теоретичних значень (4.42): $t_r = (10186.6; 1364.1; 2027.3; 5610.4; 1587.8; 6257.3; 1803.6; 39.1; 101.4; 1307.4)$.

4.3.3. Спільний розподіл кінцевої сукупності параметрів землетрусів

Введемо випадкові величини $N_{ij}(t)$, $N_i(t)$, які означають число переходів зі стану Σ_i в стан Σ_j і число землетрусів з параметрами зі стану Σ_i тому за час t . Вище було сказано, що неперервний марковський ланцюг з L станами є суперпозицією L^2 пуассонівських процесів. В такому випадку, випадкова величина $N_{ij}(t)$ є пуассонівською з параметром λ_{ij} , а величина $N_i(t)$ будучи сумою R випадкових величин $N_{ij}(t)$ розподілених за законом Пуассона також має пуассонівський розподіл з параметром λ_i .

Нехай $N(t)$ – випадкова величина, яка визначається як загальне число землетрусів за час t . тоді маємо:

$$\lambda = \sum_{i=1}^{R^s} \sum_{j=1}^{R^s} \lambda_{ij}; \quad N_i(t) = \sum_{j=1}^{R^s} N_{ij}(t); \quad N(t) = \sum_{i=1}^{R^s} N_i(t) \quad (4.56)$$

Як відомо, сума випадкових величин, розподілених за пуассонівським законом, теж має пуассонівський закон розподілу [Durner, 1999; Voit, 2003; Winkler, 2000]. Припустимо, що величина N має розподіл Пуассона з невідомим параметром λ . Достатньою статистикою для оцінки параметра λ є спостережуване число сейсмічних подій k відбулися за час t :

$$\hat{\lambda} = \frac{\sum_{k=1}^k n_k}{t} \quad (4.57)$$

Для неперервного марковського ланцюга має місце рівність: $\lambda_i = \lambda \pi_i$; $\lambda_{ij} = \lambda \pi_{ij}$.

Розподіл Пуассона з параметром λ :

$$P(N) = \frac{e^{-\lambda} \lambda^n}{n!} \quad (4.58)$$

де $n=0,1,2,\dots$ дискретним розподілом ймовірностей випадкової змінної (N), що представляє кількість подій, що відбуваються в заданому інтервалі часу або заданої області простору. Якщо λ не надто мало, то вихідний розподіл величини N

наближається до нормального розподілу, і вибіркоче значення асимптотична нормально розподілене:

$$N(\lambda, \sqrt{\lambda}) \quad (4.59)$$

З множині існуючих методів обчислення довірчих інтервалів для параметра λ – середнє число подій в даному інтервалі часу або області простору – ми зупинимося на методі Вальда [Tanusit, 2012]:

$$\hat{\lambda} - c_{\frac{\alpha}{2}} \sqrt{\frac{\hat{\lambda}}{n}} \leq \lambda \leq \hat{\lambda} + c_{\frac{\alpha}{2}} \sqrt{\frac{\hat{\lambda}}{n}} \quad (4.60)$$

де $c_{\frac{\alpha}{2}}$ – 100(1- α /2) квантиль стандартного нормального розподілу.

Нехай, припущення H_0 про те, що число землетрусів $N(t)$ має пуассонівський закон розподілу, є робочою гіпотезою. Для перевірки цієї гіпотези можна скористатися критерієм згоди для порівняння [Гайский, 1970]: розподілу частот числа подій в інтервалах часу постійної довжини h з пуассонівським розподілом і розподілу частот довжин інтервалів між наступними землетрусами з показниковим законом. Цей критерій, однак, нечутливий до змін інтенсивності виникнення сейсмічних подій. У каталогах спостерігається збільшення інтенсивності виникнення землетрусів, пов'язане з поліпшенням наглядності каталогу.

Період спостереження за землетрусами зони Вранча був розбитий на інтервали шириною в 10 діб. З числа сейсмічних подій в цих інтервалах складалася вибірка, застосовувався критерій згоди Пірсона:

$$\chi^2 = \sum_{k=1}^K \frac{(\eta_k - n_k)^2}{n_k} \quad (4.61)$$

де n_k – спостережена частота для кожного інтервалу k ; η_k – теоретична частота для кожного інтервалу. Вибіркова статистика критерію (4.61), відповідає ймовірності значущості $\alpha_q = 0.12$, яка є досить високою. Отже, спостережені частоти землетрусів відповідають пуассонівському закону розподілу.

Якщо відбувається землетрус, то його параметри з ймовірністю π_v , $v = 1, 2, \dots, R_s$ належать деякому станом Σ_v . Тоді, умовний розподіл n землетрусів по R станам елементарним джерел землетрусів) відповідає поліноміальній схемі:

$$P\{N_1(t) = n_1, N_2(t) = n_2, \dots, N_{R^s}(t) = n_{R^s} / N(t) = n\} = \frac{n!}{n_1! \dots n_{R^s}!} \pi_1^{n_1} \pi_2^{n_2} \dots \pi_{R^s}^{n_{R^s}} \quad (4.62)$$

Щоб отримати маргінальний розподіл величин $R^0 < R$:

$$\{N_1(t) = n_1, N_2(t) = n_2, \dots, N_{R^0}(t) = n_{R^0}\} \quad (4.63)$$

потрібно підсумувати ймовірність (4.63) по всіх можливих значеннях змінних:

$$n_{R^0+1}, n_{R^0+2}, \dots, n_{R^s}.$$

Множенням ймовірності (4.62) на ймовірність:

$$p_t(n) = e^{-\lambda t} \frac{(\lambda t)^n}{n!} \quad (4.64)$$

виникнення n сейсмічних подій за час t , визначається безумовна спільна ймовірність розподілу n землетрусів по R_s станам:

$$P\{N_1(t) = n_1, N_2(t) = n_2, \dots, N_{R^s}(t) = n_{R^s}\} = e^{-\lambda t} \frac{(\lambda t)^n}{n!} \frac{n!}{n_1! \dots n_{R^s}!} \pi_1^{n_1} \pi_2^{n_2} \dots \pi_{R^s}^{n_{R^s}} \quad (4.65)$$

Вище було сказано, що сейсмічна зона представляється в формі об'єднання елементарних зон, усередині яких осередки розподіляються порівняно рівномірно. Крім того, такий підхід позбавляє від проблем, пов'язаних з поняттям дискретного або неперервного розподілу глибини проміжних землетрусів в зоні Вранча, наприклад.

4.4. Ймовірнісна модель балканського сейсмічного регіону

Згідно каталогу балканських землетрусів [Shebalin et al., 1998] Балканський сейсмічний регіон складається з 18 сейсмічних зон: *ROM; CRO; BUL; BH; ALB; CZH; D; SLK; MON; HUN; SLO; SR; UKR; MAC; A; GR; TUR; POL*. Коли, в деякій зоні, наприклад, *SLK* відбувається землетрус, то це буде означати, що відбулася подія *SLK*, відомості про яких містяться в каталозі. Тобто є спостереження послідовність сейсмічних подій, ймовірнісна структура якої може характеризувати статистичний зв'язок між сейсмічними зонами. Балканський регіон можна розглянути, як фізичну систему, що в випадкові моменти часу міняє свій стан. Під зміною станів мається на увазі виникнення землетрусу в зоні X після землетрусу в зоні Y . Як стохастичною

моделлю Балканської сейсмічної зони пропонується марковська модель. Це означає, що спостереження послідовності зон, в яких відбувалися землетруси, є реалізацією дискретного марковського ланцюга з S станами.

Нехай подія $A_i^k : i=1, \dots, N; k=1, \dots, S$ означає, що i -е землетрус стався в k -й зоні з групи S сейсмічних зон, тобто фізична система знаходиться в k -м стані. Щодо даної послідовності подій висувуються три гіпотези:

H^0 : між сейсмічними зонами відсутній статистичний зв'язок

H^1 : послідовність випадкових подій є реалізацією простого дискретного марковського ланцюга

H^k : регіон, де можливо наступний землетрус визначається результатом до попередніх подій.

Гіпотеза H^0 : означає, що виникнення землетрусів в різних зонах є незалежними подіями. В цьому випадку, в матриці одно крокових ймовірностей переходу всі рядки будуть однакові:

$$P\left(\frac{A_i^k}{A_{i-1}^j}\right) = P(A_i^k) \quad (4.66)$$

Припущення H^1 означає, що можна передбачити регіон майбутнього землетрусу. Оцінками максимальної правдоподібності елементів матриці перехідних ймовірностей є:

$$(\Pi = \pi_{ik}); \quad i, k \in (1, S); \quad \pi_{ik} = \frac{n_{ik}}{n_i} \quad (4.67)$$

де S – число станів, n_{ik} – число випадків, коли після події A_i спостерігалось подія A_k , n_i число подій A_i в вибірці обсягу N . Дослідження марковського ланцюга проводилося із застосуванням матричних методів [Langrock, Jahn, 1979]. Гіпотеза про те, що послідовність випадкових подій $A_i^k : i=1, \dots, N; k=1, \dots, S$ утворює ергодичний марковський ланцюг, виявилася правдоподібною.

Отже, по матриці перехідних ймовірностей можна передбачити регіон, де може статися наступний землетрус. Гіпотеза H^k означає, що сейсмічні події мають пам'ять k -го порядку. Для перевірки правдоподібності гіпотез застосовувався критерій χ^2 . В каталозі [Shebalin et al., 1998] містяться дані про землетруси в зонах: ROM; CRO; BUL; BH; ALB; CZH; D; SLK; MON; HUN; SLO; SR; UKR; MAC; A; GR; TUR; POL.

Початковий розподіл землетрусів по регіонах: $\vec{\pi}_0 = (0.0774, 0.1380, 0.1297, 0.0687, 0.1584, 0.0247, 0.0190, 0.0190, 0.0473, 0.0407, 0.0457, 0.0977, 0.0087, 0.0827, 0.0033, 0.0123, 0.0177, 0.0090)$ обчислюється за формулою:

$$P_k = n_k / N \quad (4.68)$$

де n_k – число землетрусів в k -му регіоні, N – загальне число землетрусів.

Граничні розподілу землетрусів по регіонах $\vec{\pi}_i = (0.0770, 0.1385, 0.1297, 0.0687, 0.1584, 0.0247, 0.0190, 0.0190, 0.0474, 0.0407, 0.0457, 0.0977, 0.0087, 0.0827, 0.0033, 0.0123, 0.0177, 0.0090)$ в незначній мірі відрізняються від їх початкових значень.

Статистика критерію, використана для оцінки відхилення теоретичних частот від спостережених значень, рівна $\chi_q = 17,366$ набагато менше порогового значення $\chi_{0.05,17} = 25,587$, що відповідає рівню значущості $\alpha = 0.05$ с 17-а ступенями свободи. Таким чином відхилення вектора теоретичних частот землетрусів в регіонах $n_t = (243.3, 439.2, 412.3, 214.8, 506.1, 81.6, 64.5, 52.1, 145.2, 120.2, 138.2, 309.8, 20.1, 258.7, 6.3, 30.8, 52.1, 19.2)$ від спостережених значень $n_0 = (232, 414, 389, 206, 475, 74, 57, 57, 142, 122, 137, 293, 26, 248, 10, 37, 53, 27)$ є незначним. За критерієм χ^2 виявлено правдоподібність односторонньої гіпотези H^1 , при конкуруючих гіпотезах H^0 і H^2 .

Отже, землетрус в деякому регіоні може зумовити землетрус в інших регіонах. Поясненням цьому факту може служити те, що причиною виникнення землетрусів є глобальні процеси в надрах Землі.

4.4.1. Дослідження статистичного зв'язку між сейсмічною активністю

Балканських сейсмічних зон

Нехай y_j означає число землетрусів в j -му році на Балканах, а x_{ij} – число землетрусів в i -й сейсмічній зоні в j -му році, $j = 1900 \dots, 1998$. Використовувалися відомості про Балканські землетруси за 20-е століття з каталогу [Shebalin et al., 1998]. Досліджується багатовимірна регресія числа землетрусів за роками в Балканському сейсмічному регіоні y_j , на число землетрусів в різних Балканських сейсмічних зонах

x_{ij} : ROM; CRO; BUL; BH; ALB; CZH; D; SLK; MON; HUN; SLO; SR; UKR; MAC; A; GR; TUR; POL.

Рівняння регресії в векторній формі має вигляд:

$$y = X\beta + \varepsilon \quad (4.69)$$

Таким чином, стовпцями матриці регресорів є число землетрусів за роками в деякій сейсмічній зоні. Оцінка найменших квадратів векторних параметрів в матричній формі має вигляд:

$$\bar{\beta} = (X'X)^{-1} X'Y \quad (4.70)$$

Залишкова сума квадратів обчислюється з виразу:

$$\varepsilon'\varepsilon = Y'Y - \beta'X'X\beta \quad (4.71)$$

Для перевірки значущості регресії, тобто гіпотези $H^0: \beta_1 = \dots = \beta_{14} = 0$ застосовується F -критерій. Спостережуване значення $F_q=84.781$ перевершує табличне значення $F(14;76;0.05)=1.83$, відповідає рівню значущості $\alpha=0.05$ при (14.76) ступенях свободи. Отже, нульова гіпотеза H^0 відкидається на користь альтернативної гіпотези про значущість коефіцієнтів регресії: $\beta=(1.26;1.40;0.97; 1.47;0.94,1.01,1.83;0.66; 0.791; 0.567;1.04;0.909;0.500;0.442)$.

Високі значення коефіцієнтів детермінації і кореляції при регресійному рівнянні (4) вказують на те, що 93.98% зміни числа землетрусів на Балканах визначається множинною регресією.

Для виявлення статистичного зв'язку між стовпцями регресійної матриці обчислюється кореляційна матриця:

$$R = (X^H)^T X^H \quad (4.72)$$

а) значення критерію χ^2 , обчислене за формулою:

$$\chi^2 = -(n-1 - \frac{1}{6}(2m+5)) \ln|R| \quad (4.73)$$

$\chi_q = 133.248$, перевершує табличне значення $\chi_{\text{tab}} = 114.265$, що відповідає рівню значущості $\alpha = 0.05$ с числом ступенів свободи $df=91$. Отже, в сейсмічній активності Балканських сейсмічних зон, за даними каталогу [Shebalin et al., 1998], існує мультиколінеарність.

б) для виявлення колінеарності між будь-якими двома сейсмічними зонами визначається зворотна матриця $D=R^{-1}$, де R – кореляційна матриця. Вибіркові значення F – критерію Фішера визначаються з виразу [Sachs, Hedderich, 2006; Voit, 2003].

$$F_k = (d_{kk} - 1) \frac{n-m}{m-1} \quad (4.74)$$

де d_{kk} – діагональні елементи матриці D . Вибіркові значення F -статистики $F = (1.59; 0.985; 0.797; 1.97; 1.91; 1.61; 1.09; 1.74; 2.22; 0.784; 1.70; 0.921; 1.71; 1.25)$ порівнюються з табличним значенням $F_q = 2.28$, відповідним рівнем значущості $\alpha = 0.05$ F -критерію зі ступенями свободи: $df_1=77$, $df_2=13$.

Порівняння показує, що жоден з регресорів не колінеарний з іншими.

с) Для перевірки гіпотези про значущість кореляції між кожною парою регресорів використовується t - критерій:

$$t_{ij} = \frac{|r_{ij}| \sqrt{n-m}}{\sqrt{1-r_{ij}^2}} \quad (4.75)$$

Деякі з спостережених значень t -критерію: $t_{1,6} = 2.179$; $t_{1,13} = 3.292$; $t_{4,9} = 2.887$; $t_{5,11} = 3.691$; $t_{6,13} = 2.127$; $t_{7,9} = 2.878$; $t_{7,8} = 1.978$; $t_{8,14} = 3.153$; більше табличного значення $t_{0,05,71} = 1.96$ [Леман, 1979], що відповідає рівню значущості $\alpha = 0.05$ с числом ступенів свободи $df = 71$. Це означає, що між активністю сейсмічних зон: *ROM; CRO; BUL; BH; ALB; CZH; D; SLK; MON; HUN; SLO; SR; UKR; MAC; A; GR; TUR; POL* - існує значуща кореляція.

Таблиця 4.2

Нумерація сейсмічних зон

1	2	3	4	5	6	7	8	9	10	11	12	13	14
<i>ROM</i>	<i>CRO</i>	<i>BUL</i>	<i>BH</i>	<i>ALB</i>	<i>MON</i>	<i>HUN</i>	<i>SLO</i>	<i>SR</i>	<i>UKR</i>	<i>MAC</i>	<i>GR</i>	<i>TUR</i>	<i>POL</i>

4.4.2. Таблиці спряженості ознак

Щоб мати уявлення про спільний розподіл випадкових величин складаються таблиці спряженості ознак: кореляційний таблиця - спостережених частот з r рядками і s стовпцями [Крамер, 1975; Fisz, 1966]. Нехай гіпотеза H_0 означає, що

число землетрусів в різних Балканських сейсмічних зонах є взаємно незалежними випадковими величинами. Діапазони числа землетрусів в сейсмічних зонах розбиваються на r, s інтервалів відповідно. Тоді елементами таблиці спряженості n_{ij} є деякі значення з i -го і j -го інтервалів. Наприклад, для зон *ROM, MON*: $n_{58} = 4$ означає, що 4 рази спостерігалася спільнеаподія, коли в зоні *ROM* відбувалося 5, а в зоні *MON* 8 землетрусів відповідно. Для перевірки гіпотези H_0 обчислюється статистика критерію:

$$\chi_q^2 = N \sum_{i=1}^r \sum_{j=1}^s \frac{(n_{ij} - \frac{n_{i.} n_{.j}}{N})^2}{n_{i.} n_{.j}} \quad (4.76)$$

де χ_q^2 – χ^2 розподілена з $(r-1)(s-1)$ ступенями свободи випадкова величина. На практиці часто доводиться оцінити взаємозв'язок між окремими ознаками за допомогою непараметричних коефіцієнтів зв'язку, значення яких знаходяться в інтервалі від 0 до 1: 0 – немає зв'язку, 1 – зв'язок дуже сильний.

Для оцінки тісноти зв'язку в таблицях будь-якого розміру використовується коефіцієнт спряженості ознак:

$$C = \sqrt{\frac{\chi^2}{\chi^2 + N}} \quad (4.77)$$

Таблиця 4.3

Коефіцієнти спряженості ознак

Зона	<i>ROM</i>	<i>CRO</i>	<i>BUL</i>	<i>BH</i>	<i>ALB</i>	<i>CZH</i>	<i>D</i>	<i>SLK</i>	<i>MON</i>	<i>HUN</i>	<i>SLO</i>	<i>SR</i>	<i>UKR</i>	<i>MAC</i>	<i>A</i>	<i>GR</i>	<i>TUR</i>	<i>POL</i>
<i>ROM</i>	1.0	0.32	0.47	0.5	0.38	0.39	0.38	0.32	0.53	0.50	0.41	0.41	0.35	0.32	0.28	0.35	0.53	0.39
<i>CRO</i>	0.32	1.0	0.56	0.53	0.47	0.72	0.45	0.20	0.14	0.47	0.52	0.36	0.28	0.39	0.31	0.26	0.25	0.37
<i>BUL</i>	0.47	0.56	1.0	0.22	0.36	0.24	0.35	0.06	0.06	0.37	0.72	0.38	0.12	0.14	0.20	0.28	0.16	0.44
<i>BH</i>	0.5	0.53	0.22	1.0	0.52	0.60	0.71	0.53	0.37	0.73	0.49	0.6	0.48	0.53	0.41	0.44	0.42	0.47
<i>ALB</i>	0.38	0.47	0.36	0.52	1.0	0.35	0.1	0.6	0.52	0.51	0.43	0.51	0.52	0.66	0.39	0.33	0.13	0.24
<i>CZH</i>	0.39	0.2	0.24	0.60	0.35	1.0	0.77	0.05	0.58	0.73	0.30	0.12	0.13	0.57	0.09	0.32	0.09	0.13
<i>D</i>	0.38	0.45	0.35	0.71	0.1	0.77	1.0	0.03	0.03	0.71	0.29	0.08	0.08	0.07	0.06	0.22	0.06	0.08
<i>SLK</i>	0.32	0.20	0.06	0.53	0.6	0.05	0.03	1.0	0.03	0.22	0.34	0.5	0.59	0.44	0.21	0.11	0.06	0.08
<i>MON</i>	0.53	0.14	0.06	0.37	0.52	0.58	0.03	0.03	1.0	0.71	0.26	0.08	0.40	0.07	0.71	0.35	0.06	0.08
<i>HUN</i>	0.50	0.47	0.32	0.73	0.51	0.73	0.71	0.22	0.71	1.0	0.53	0.71	0.32	0.36	0.22	0.35	0.20	0.18

<i>SLO</i>	0.41	0.52	0.72	0.49	0.42	0.30	0.29	0.34	0.26	0.53	1.0	0.29	0.49	0.31	0.26	0.30	0.25	0.75
<i>SR</i>	0.41	0.36	0.38	0.1	0.51	0.12	0.08	0.5	0.08	0.71	0.29	1.0	0.35	0.67	0.10	0.19	0.13	0.36
<i>UKR</i>	0.35	0.28	0.12	0.48	0.52	0.13	0.08	0.59	0.40	0.32	0.49	0.35	1.0	0.30	0.49	0.15	0.42	0.41
<i>MAC</i>	0.32	0.39	0.14	0.53	0.66	0.57	0.07	0.44	0.07	0.36	0.31	0.67	0.30	1.0	0.12	0.23	0.27	0.15
<i>A</i>	0.28	0.31	0.20	0.41	0.39	0.09	0.06	0.21	0.71	0.22	0.26	0.10	0.49	0.12	1.0	0.33	0.25	0.31
<i>GR</i>	0.35	0.26	0.28	0.44	0.33	0.32	0.22	0.11	0.35	0.35	0.30	0.19	0.15	0.23	0.33	1.0	0.51	0.36
<i>TUR</i>	0.53	0.25	0.16	0.42	0.13	0.09	0.06	0.06	0.06	0.20	0.25	0.13	0.2	0.27	0.25	0.51	1.0	0.36
<i>POL</i>	0.39	0.37	0.44	0.47	0.24	0.13	0.08	0.08	0.08	0.18	0.75	0.36	0.41	0.15	0.31	0.36	0.36	1.0

Коефіцієнт спряженості Крамера пов'язаний в такий спосіб до критерію хі-квадрат [Крамер, 1975]:

$$V = \sqrt{\frac{\chi^2 / N}{\min((r-1), (s-1))}} \quad (4.78)$$

З таблиць (табл. 4.3, 4.4) видно, що між сейсмічною активністю деяких зон існує слабкий і середній статистичний зв'язок.

4.4.3. Вибіркове кореляційне відношення

Для оцінки тісноти нелінійного кореляційного зв'язку використовується кореляційне відношення η_{yx} [Маркович, 1963]. Вибірковим кореляційним відношенням називається відношення міжгрупового середнього є відхилення до загального середньоквадратичного відхилення:

$$\eta_{y/x} = \frac{\sigma(\bar{y}_x)}{\sigma_y}, \quad \sigma^2(\bar{y}_x) = \frac{1}{N} \sum n_x \bar{y}_x^2 - \bar{y}^2 \quad (4.79)$$

де σ_y – середньоквадратичне відхилення змінної y .

Кореляційне відношення є мірою тісноти будь-якого зв'язку. У цьому його перевага перед коефіцієнтом кореляції, який оцінює ступінь тісноти тільки лінійного зв'язку. Щоб мати уявлення про тісноту зв'язку можна скористатися таблицею Чеддока (табл. 4.5).

Таблиця 4.4

Коефіцієнти спряженості ознак Крамера

Зона	ROM	CRO	BUL	BH	ALB	CZH	D	SLK	MON	HUN	SLO	SR	UKR	MAC	A	GR	TUR	POL
ROM	1.0	0.14	0.22	0.23	0.17	0.17	0.17	0.14	0.25	0.24	0.18	0.18	0.15	0.14	0.12	0.155	0.25	0.17
CRO	0.14	1.0	0.28	0.26	0.22	0.43	0.21	0.09	0.06	0.22	0.25	0.16	0.12	0.17	0.13	0.11	0.11	0.16
BUL	0.22	0.28	1.0	0.09	0.15	0.10	0.15	0.03	0.03	0.14	0.42	0.17	0.05	0.06	0.09	0.12	0.06	0.20
BH	0.24	0.26	0.09	1.0	5.247	0.31	0.41	0.26	0.16	0.44	0.23	0.30	0.22	0.26	0.18	0.20	0.19	0.22
ALB	0.17	0.23	0.16	0.25	1.0	0.15	0.04	0.30	0.25	0.24	0.19	0.24	0.25	0.36	0.17	0.14	0.06	0.10
CZH	0.17	0.43	0.10	0.31	0.15	1.0	0.5	0.02	0.29	0.44	0.13	0.05	0.05	0.28	0.04	0.14	0.04	0.05
D	0.17	0.21	0.15	0.41	0.04	0.5	1.0	0.01	0.01	0.42	0.12	0.03	0.03	0.03	0.02	0.09	0.02	0.03
SLK	0.14	0.09	0.03	0.26	0.30	0.02	0.01	1.0	0.01	0.09	0.15	0.23	0.30	0.20	0.09	0.05	0.02	0.03
MON	0.25	0.06	0.03	0.16	0.25	0.29	0.01	0.01	1.0	0.41	0.11	0.03	0.18	0.03	0.41	0.15	0.02	0.03
HUN	0.24	0.22	0.14	0.44	0.24	0.44	0.42	0.09	0.41	1.0	0.26	0.41	0.14	0.16	0.09	0.15	0.08	0.07
SLO	0.18	0.25	0.42	0.23	0.19	0.13	0.12	0.15	0.11	0.26	1.0	0.12	0.23	0.13	0.11	0.13	0.10	0.47
SR	0.18	0.16	0.17	0.30	0.24	0.05	0.03	0.23	0.03	0.41	0.12	1.0	0.15	0.37	0.04	0.08	0.05	0.16
UKR	0.15	0.12	0.05	0.22	0.25	0.05	0.03	0.30	0.18	0.14	0.23	0.15	1.0	0.13	0.23	0.06	0.19	0.18
MAC	0.14	0.17	0.06	0.26	0.36	0.28	0.03	0.20	0.03	0.16	0.13	0.37	0.13	1.0	0.05	0.09	0.11	0.06
A	0.12	0.13	0.09	0.18	0.17	0.04	0.02	0.09	0.41	0.09	0.11	0.04	0.23	0.05	1.0	0.14	0.10	0.13
GR	0.15	0.11	0.12	0.20	0.14	0.14	0.09	0.05	0.15	0.15	0.13	0.08	0.06	0.09	0.14	1.0	0.24	0.16
TUR	0.25	0.11	0.06	0.19	0.06	0.04	0.02	0.02	0.02	0.08	0.10	0.05	0.19	0.11	0.10	0.24	1.0	0.16
POL	0.17	0.16	0.20	0.22	0.10	0.05	0.03	0.03	0.03	0.07	0.47	0.16	0.18	0.06	0.13	0.16	0.16	1.0

Таблиця 4.5

Співвідношення значень коефіцієнта до описової шкали

Значення коефіцієнта	Інтерпретація
$0 < V \leq 0.2$	Дуже слабкий зв'язок
$0.2 < V \leq 0.5$	слабкий зв'язок
$0.5 < V \leq 0.7$	Середній зв'язок
$0.7 < V \leq 0.9$	сильний зв'язок
$0.9 < V \leq 1$	Дуже сильний зв'язок

Таблиця 4.6

Коефіцієнти кореляційних відносин

Зона	ROM	CRO	BUL	BH	ALB	CZH	D	SLK	MON	HUN	SLO	SR	UKR	MAC	A	GR	TUR	POL
ROM	1.0	0.24	0.31	0.1	0.18	0.21	0.25	0.28	0.23	0.23	0.21	0.33	0.12	0.17	0.25	0.17	0.48	0.09
CRO	0.24	1.0	0.29	0.34	0.18	0.76	0.34	0.15	0.11	0.15	0.23	0.12	0.08	0.07	0.25	0.15	0.07	0.19
BUL	0.31	0.29	1.0	0.13	0.11	0.05	0.21	0.05	0.05	0.09	0.33	0.10	0.1	0.08	0.15	0.1	0.05	0.09
BH	0.1	0.34	0.34	1.0	0.29	0.37	0.76	0.55	0.29	0.3	0.31	0.40	0.22	0.23	0.33	0.18	0.19	0.36
ALB	0.18	0.18	0.11	0.29	1.0	0.15	0.09	0.28	0.31	0.32	0.14	0.22	0.24	0.34	0.31	0.14	0.1	0.12

<i>CZH</i>	0.21	0.76	0.05	0.37	0.15	1.0	0.1	0.04	0.1	0.6	0.15	0.1	0.1	0.40	0.09	0.13	0.07	0.1
<i>D</i>	0.25	0.34	0.21	0.76	0.09	0.1	1.0	0.02	0.03	0.26	0.1	0.06	0.07	0.06	0.06	0.1	0.04	0.06
<i>SLK</i>	0.28	0.15	0.05	0.55	0.28	0.04	0.02	1.0	0.03	0.05	0.15	0.22	0.44	0.16	0.17	0.08	0.04	0.06
<i>MON</i>	0.23	0.11	0.05	0.29	0.31	0.1	0.03	0.03	1.0	0.56	0.15	0.06	0.27	0.06	0.58	0.12	0.04	0.06
<i>HUN</i>	0.23	0.15	0.09	0.3	0.32	0.6	0.26	0.05	0.56	1.0	0.25	0.71	0.22	0.24	0.16	0.21	0.16	0.12
<i>SLO</i>	0.21	0.23	0.33	0.31	0.14	0.15	0.1	0.15	0.15	0.25	1.0	0.11	0.22	0.13	0.19	0.11	0.16	0.48
<i>SR</i>	0.33	0.11	0.10	0.40	0.22	0.1	0.06	0.22	0.06	0.71	0.11	1.0	0.23	0.37	0.08	0.14	0.10	0.11
<i>UKR</i>	0.12	0.08	0.1	0.22	0.24	0.10	0.07	0.44	0.27	0.22	0.22	0.23	1.0	0.11	0.34	0.12	0.10	0.07
<i>MAC</i>	0.17	0.07	0.08	0.23	0.34	0.40	0.06	0.16	0.06	0.24	0.13	0.37	0.11	1.0	0.1	0.09	0.07	0.11
<i>A</i>	0.25	0.25	0.15	0.33	0.31	0.09	0.06	0.17	0.58	0.16	0.19	0.08	0.34	0.1	1.0	0.13	0.22	0.27
<i>GR</i>	0.17	0.15	0.1	0.18	0.14	0.13	0.1	0.08	0.12	0.21	0.11	0.14	0.12	0.09	0.13	1.0	0.27	0.23
<i>TUR</i>	0.48	0.07	0.05	0.19	0.1	0.07	0.04	0.04	0.04	0.16	0.16	0.10	0.10	0.07	0.22	0.27	1.0	0.07
<i>POL</i>	0.09	0.19	0.09	0.36	0.12	0.1	0.06	0.06	0.06	0.12	0.48	0.11	0.07	0.11	0.27	0.23	0.07	1.0

4.5. Марковська модель сейсмічних зон

Землетруси відбуваються у випадкові моменти часу в випадкових точках геофізичної середовища. Тобто сейсмічний режим видається випадковим потоком подій $E_{j_0}, E_{j_2}, \dots, E_{j_n}$. Подія E_{j_s} означає виникнення землетрусу в s -й сейсмічній зоні в момент t_j . Випадковий потік визначається заданням спільної скінченновимірної функції розподілу. Обчислення спільного розподілу, за винятком деяких тривіальних випадків, є практично неможливим. Але в класі марковських процесів спільний розподіл виражається через одно і двовимірні функції розподілу. В теорії простих ланцюгів Маркова майбутнє залежить лише від справжнього стану, але не від минулої історії. У ланцюзі Маркова k -го порядку майбутнє залежить і від справжнього, і від k попередніх станів. У нашому випадку, коли відбувається землетрус в s -й зоні, то це буде означати, що спостережено s -ий стан ланцюга. Якщо після землетрусу в r -му стані відбувається землетрус в s -му стані, то це означає, що відбувся перехід ланцюга з r -го стану в s -е. Спільний розподіл станів такого ланцюга визначається рівністю [Langrock, Jahn, 1979]:

$$P\{E_{j_0}, E_{j_1}, \dots, E_{j_n}\} = p_{j_0} \cdot \pi_{j_0 j_1} \cdot \pi_{j_1 j_2} \dots \pi_{j_{n-1} j_n} \quad (4.80)$$

де ймовірність початкового стану, а $\pi_{j_1 j_r}$ – ймовірність переходу зі стану E_{j_1} в стан E_{j_r} . Ймовірності π_{ij} утворюють матрицю перехідних ймовірностей ланцюга Маркова $\pi=(\pi_{ij}), i, j=1, \dots, S$, S – число станів (табл. 4.1).

4.5.1. Марковська модель румунських сейсмічних зон

Згідно кластерному аналізу, на території Румунії, на основі каталогу ROMPLUS, виділено 13 сейсмічних зон. Сейсмічні події випадковим чином, і в випадкові моменти часу, відбуваються в різних сейсмічних зонах. Спостережені в сейсмічних зонах землетрусу з каталогу ROMPLUS утворюють тимчасову послідовність випадкових подій, по ймовірнісній структурі якої можна визначити наявність взаємозв'язку між сейсмічними зонами. Наведений вище алгоритм застосовується для перевірки виконання умови марковості для даної послідовності.

Нехай випадкова подія Z^s , означає виникнення землетрусу в s -ой сейсмічній зоні Z^s , ($s=1, \dots, 13$). Дослідженню підлягає послідовність $Z_{s_1}, Z_{s_2}, Z_{s_3}, \dots, Z_{s_N}$ сейсмічних подій, де N – спостережуване число землетрусів за період 1984–2016 рр., охопленого каталогом ROMPLUS. Передбачається, що дана послідовність є реалізацією ергодичного одновимірного марковського ланцюга з простором станів, що складається з сейсмічних зон Z^s , ($s=1, \dots, 13$). Щодо даної послідовності висувуються гіпотези H_0 ; H_1 і H_2 , розглянуті вище. Так, гіпотеза H_1 означає, що послідовність є реалізацією однорідного дискретного марковського ланцюга з простором станів Z^s , ($s=1, \dots, 13$). Правдоподібність гіпотези H_1 , при конкуруючій гіпотезі H_2 , підтверджується значенням тестового критерію (4.50) $\chi_q = 1682.46$, яке менше значення $\chi_{0.01, 1872} = 1973.77$, відповідним 5% рівнем значущості χ^2 розподіленої випадкової величини зі 1872 я ступенями свободи. Статистика тестового критерію (4.51) $\chi_q = 16620.0$ набагато перевершує граничне значення $\chi_{0.05, 144} = 173.0$, що відповідає рівню значущості $\alpha = 0.05$ зі 144 ступенями свободи. Отже, гіпотеза H_0 , повинна бути відхилена на користь конкуруючій гіпотези H_1 , про те, що дана послідовність є реалізацією простого марковського ланцюга. Також за критерієм χ^2 (4.52), гіпотеза H_0 відхиляється на користь конкуруючій гіпотези H_1 , так як,

вибіркове значення $\chi_q=8863.2$ перевищує порогове значення $\chi_{0.05} = 215.054$. Ще одним критерієм перевірки правдоподібності гіпотези H_0 при конкуруючій гіпотезі H_1 , є вибірка статистика $\chi_q=11143.48$ - χ^2 розподіленої випадкової величини (4.54), яка перевершує табличне значення $\chi_{0.05,13}=22.36$, що відповідає рівню значущості $\alpha = 0.05$ при 13 ступенях свободи, тобто гіпотеза H_0 відхиляється на користь конкуруючої гіпотези H_1 .

Для дослідження ланцюга Маркова застосовувалися матричні методи. Серед власних значень матриці перехідних ймовірностей (табл. 4.7), одиниця виявилася максимальною за модулем власним значенням кратності один: $\Lambda=\{(1.0,0.0); (0.398,0.0); (0.291,0.0); (0.232,0.0); (0.171,0.005); (0.171,-0.005); (0.133,0.0); (0.115,0.0); (0.093,0.0); (0.075,0.0); (0.057,0.0); (0.034,0.0); (0.011,0.0)\}$. Цей факт є достатньою для ергодичності ланцюга Маркова умовою. Вище було показано, що граничний розподіл ймовірностей виникнення землетрусів в сейсмічних зонах (спостереження станів ланцюга) можна визначити з однієї досить довгої її реалізації за формулою (4.21). Дійсно, вектор початкових розподілів землетрусів по 13 сейсмічних зонах на території Румунії, обчислений за формулою (4.18): $\gamma_0=(0.043; 0.035; 0.023; 0.12; 0.045; 0.038; 0.064; 0.092; 0.364; 0.009; 0.037; 0.11; 0.021)$ незначно відрізняється від вектора граничних розподілів розрахований методами за формулою (4.28): $\gamma=(0.043; 0.035; 0.0226; 0.12; 0.045; 0.038; 0.0644; 0.092; 0.364; 0.009; 0.037; 0.109; 0.021)$. З вектора граничних розподілів станів можна визначити, в якій зоні, і з якою ймовірністю може статися землетрус.

Матриця ймовірностей переходу вказує на статистичний зв'язок між сейсмічними зонами (табл. 4.7). На перетині i -го рядка і j -го стовпця знаходиться ймовірність переходу за один крок з i -го стану в j -е стан - тобто після виникнення землетрусу в i -й сейсмічній зоні, наступний землетрус станеться в j -й зоні. Найбільші значення ймовірностей переходу розташовані в 9-му стовпці, який відповідає зоні Вранча, отже, наступна, після землетрусу в сейсмічних зонах сейсмічне подія відбудеться в зоні Вранча, так як ймовірності $\pi_i, 9$, переходу за один крок з станів $i = 1 \dots, 13$ в 9-е стан є найбільшими, за винятком сейсмічних зон з

номерами: 1; 8; 9; 11, для яких більш ймовірно є виникнення наступного землетрусу в тих же зонах. У рядках: 1; 8, 9 і 11 найбільшими є значення: $\pi_{1,1} = 0.256$; $\pi_{8,8} = 0.376$; $\pi_{9,9} = 0.559$; $\pi_{11,11} = 0.184$, тобто, якщо в цих зонах відбувся землетрус, то наступний землетрус найімовірніше відбудеться там же. У класі ланцюгів Маркова це означає, що система має тенденцію залишатися в тому ж стані. Кластерний аналіз виділив 8-ю сейсмічну зону, яка відповідає Східній Вранча. Ймовірність того, що наступний, після землетрусу в Східній Вранча, землетрус станеться у Вранча становить $\pi_{8,9} = 0.303$, а ймовірність зворотної події дорівнює $\pi_{9,8} = 0.075$. Відомості з матриці перехідних ймовірностей можна використовувати для передбачення сейсмічної зони, де можуть виникнути землетруси.

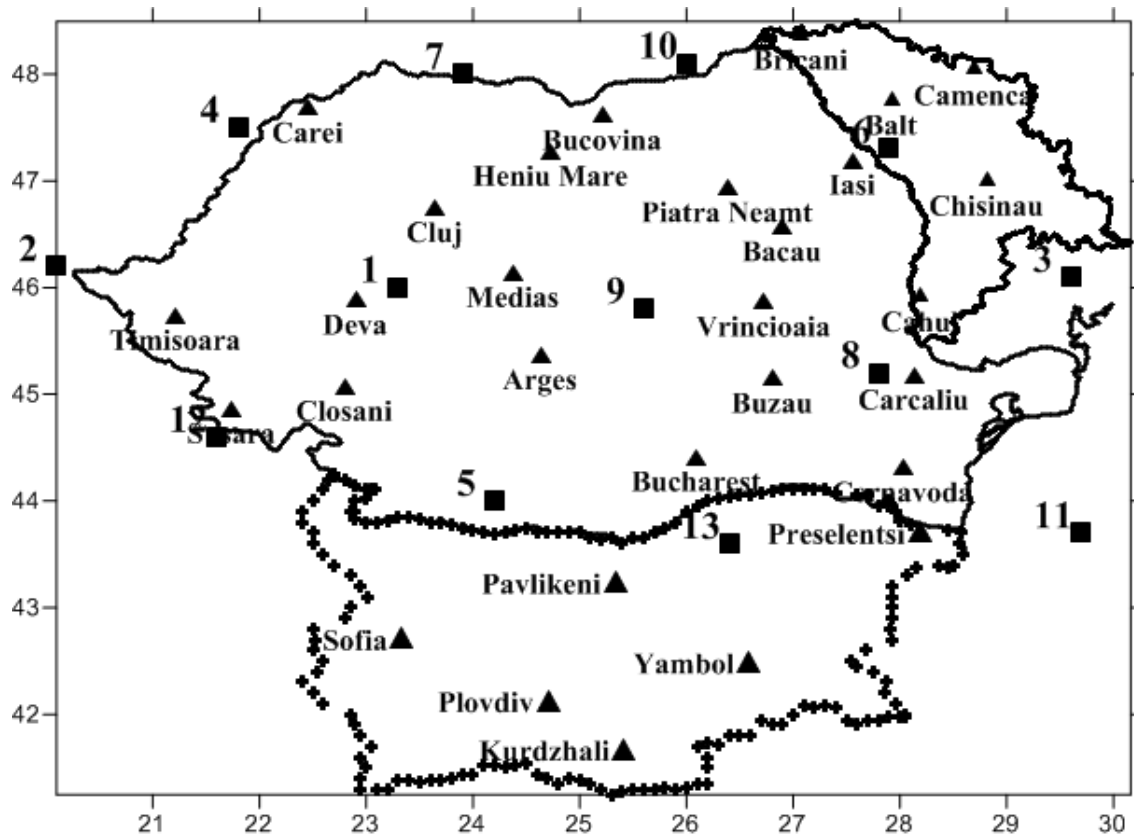


Рис. 4.2. Карта центрів сейсмічних станцій і зон на території Румунії. Символ (■) – означає центр сейсмічних зон, а (▲) – сейсмічні станції.

Таблиця 4.7

Матриця ймовірностей переходу π_{ij} між румунськими сейсмічними зонами

Зона	6	2	13	5	3	12	11	8	9	7	10	1	4
6	.272	.027	.021	.071	.023	.03	.035	.106	.315	.005	.016	.06	.02
2	.037	.114	.017	.119	.057	.056	.067	.056	.23	.012	.056	.146	.034
13	.3	.022	.159	.095	.029	.027	.05	.079	.432	.005	.013	.043	.016
5	.026	.034	.02	.227	.065	.027	.076	.053	.264	.011	.049	.128	.021
3	.031	.032	.024	.15	.086	.036	.097	.054	.284	.013	.054	.113	.025
12	.035	.041	.013	.103	.034	.206	.064	.052	.259	.005	.033	.122	.033
11	.021	.037	.008	.156	.069	.041	.184	.058	.175	.007	.05	.158	.036
8	.047	.024	.022	.059	.026	.018	.036	.376	.303	.005	.017	.057	.009
9	.037	.026	.023	.092	.034	.024	.039	.075	.559	.005	.018	.057	.012
7	.018	.046	.018	.092	.041	.032	.051	.046	.235	.18	.028	.147	.065
10	.032	.046	.024	.143	.051	.046	.074	.043	.23	.01	.118	.151	.033
1	.024	.046	.014	.126	.055	.05	.084	.042	.194	.012	.064	.256	.034
4	.019	.05	.013	.163	.044	.056	.09	.042	.195	.01	.082	.19	.046

Звівши матрицю перехідних ймовірностей в n -ю ступінь, можна передбачити зону, в якій відбудеться n -ий землетрус. Рядки всіх степеней матриці перехідних ймовірностей (табл. 4.27), починаючи з четвертого, виявилися однаковими і рівними вектору граничних розподілів станів (виникнення землетрусу в сейсмічних зонах). Для сейсмічності території Румунії, яка моделюється марковським ланцюгом, даний факт означає, що сейсмічність Румунії знаходиться в стаціонарному режимі, тобто виникнення землетрусу в деякій зоні не залежить від часу його виникнення. Отже, сейсмічний процес на території Румунії стабілізувався.

4.5.2. Марковська модель Балканських сейсмічних зон

Розглядається послідовність сейсмічних подій в 18 сейсмічних зонах Балканського регіону: *ROM; CRO; BUL; BH; ALB; CZH; D; SLK; MON; HUN; SLO; SR; UKR; MAC; A; GR; TUR; POL*. В каталозі балканських землетрусів [Shebalin et

al., 1998] наводяться дані про 3003 землетрусів, що сталися на Балканах. Щодо спостереженої послідовності подій $\{E_{j0}, E_{j1}, E_{j2}, \dots, E_{jn}\}$ висувається гіпотеза H_0 : про те, що дана послідовність є реалізацією послідовності випадкових подій, що утворюють ергодичний ланцюг Маркова. Для дослідження ланцюга Маркова застосовувалися матричні методи. Одиниця виявилася максимальною за модулем власним значенням кратності один матриці перехідних ймовірностей: $\Lambda = (1.000; 0.681; 0.404; 0.397; 0.302; 0.297; 0.297; 0.29; 0.242; 0.223; 0.184; 0.184; 0.183; 0.183; 0.144; 0.1; 0.077; 0.012)$. Цей факт є достатнім для ергодичності ланцюга Маркова умовою. Ергодичні ланцюги Маркова володіють чудовою для практичного застосування властивістю. Граничний розподіл такого ланцюга, з будь-яким ступенем точності, можна визначити з однієї досить довгої її реалізації. Дійсно, вектор початкових розподілів землетрусів по Балканським сейсмічним зонам: $P_0 = (0,0774; 0,1380; 0,1297; 0,0687; 0,1584; 0,0247; 0,0190; 0,0190; 0,0473; 0,0407; 0,0457; 0,0977; 0,0087; 0,0827; 0,0033; 0,0123; 0,0177; 0,0090)$ мало відрізняється від обчисленого матричними методами граничного розподілу: $\pi^+ = (0,077; 0,1385; 0,1297; 0,0687; 0,1584; 0,0247; 0,0190; 0,0190; 0,0474; 0,0407; 0,0457; 0,0977; 0,0087; 0,0827; 0,0033; 0,0123; 0,0177; 0,0090)$.

З вектору граничних розподілів можна визначити, в якій зоні, і з якою ймовірністю може відбутися очікуваний на Балканах землетрус. Матриця ймовірностей переходу вказує на статистичний зв'язок між сейсмічними зонами (табл. 4.7). Максимальні значення ймовірностей переходу розташовані по діагоналі матриці. Тобто, якщо в деякій зоні відбувається землетрус, то наступний землетрус з більшою ймовірністю відбудеться там же. У класі ланцюгів Маркова це означає, що система має тенденцію залишатися в тому ж стані. Звівши матрицю перехідних ймовірностей в n -ю ступінь, можна визначити ймовірність в якій зоні трапиться n -ий землетрус. При підвищенні ступеня перехідної матриці всі рядки стають однаковими. Це говорить про те, що процес став стабільним. Граничне розподіл таких процесів можна обчислити і з рішення матричного рівняння:

$$p\pi = p \quad (4.81)$$

де p – вектор граничних розподілів ймовірностей виникнення землетрусів в тій чи іншій зоні, π – матриця ймовірностей переходу.

Таблиця 4.8

Матриця ймовірностей переходу π_{ij}

Зона	ROM	CRO	BUL	BH	ALB	MON	HUN	SLO	SR	UKR	MAC	GR	TUR	POL
ROM	0.224	0.108	0.112	0.056	0.099	0.026	0.065	0.039	0.099	0.004	0.069	0.013	0.026	0.000
CRO	0.092	0.366	0.101	0.058	0.096	0.034	0.031	0.029	0.082	0.002	0.043	0.005	0.007	0.005
BUL	0.049	0.113	0.422	0.044	0.082	0.026	0.031	0.041	0.093	0.013	0.039	0.013	0.005	0.010
BH	0.063	0.141	0.107	0.248	0.136	0.049	0.044	0.049	0.083	0.005	0.029	0.010	0.015	0.010
ALB	0.074	0.080	0.040	0.063	0.438	0.032	0.013	0.042	0.082	0.011	0.067	0.008	0.019	0.008
MON	0.056	0.127	0.028	0.070	0.085	0.415	0.035	0.014	0.042	0.007	0.049	0.014	0.000	0.021
HUN	0.098	0.098	0.123	0.057	0.098	0.025	0.254	0.074	0.082	0.008	0.033	0.008	0.008	0.000
SLO	0.066	0.153	0.131	0.022	0.153	0.022	0.051	0.175	0.073	0.000	0.073	0.015	0.022	0.022
SR	0.061	0.116	0.099	0.068	0.096	0.024	0.034	0.048	0.297	0.014	0.078	0.010	0.007	0.010
UKR	0.038	0.077	0.192	0.115	0.077	0.000	0.115	0.077	0.000	0.192	0.038	0.000	0.038	0.000
MAC	0.028	0.065	0.073	0.056	0.177	0.032	0.020	0.036	0.081	0.000	0.379	0.016	0.004	0.004
GR	0.158	0.105	0.211	0.026	0.079	0.053	0.000	0.026	0.026	0.000	0.105	0.132	0.026	0.000
TUR	0.075	0.019	0.113	0.019	0.132	0.000	0.075	0.019	0.057	0.000	0.094	0.038	0.340	0.019
POL	0.000	0.111	0.037	0.074	0.259	0.074	0.000	0.111	0.037	0.000	0.074	0.037	0.037	0.111

Висновки до розділу 4

Сейсмічна зона охоплює деякий простір геофізичного середовища $[\Phi \times \Psi \times H]$: Φ ; Ψ ; H – діапазони можливих значень широти, довготи і глибини землетрусів відповідно. Декартове множення даного простору на діапазон M – можливих значень магнітуди утворює $\Xi^s = (\Phi \times \Psi \times H \times M)$ – чотиривимірний простір, точки якого можуть виявитися можливими значеннями параметрів землетрусів.

Розбиттям даного простору на 4-х мірні непересічні підпростори, утворюються елементарні джерела землетрусів (ЕДЗ). Виникнення землетрусу з параметрами з деякого ЕДЗ Σ_j в момент часу t_n означає спостереження сейсмічного події $\Sigma_j(t_n)$ в момент часу t_n . На сучасному рівні розвитку геофізики, моменти часу виникнення землетрусів, і значення їх параметрів розглядаються як випадкові величини. Це означає, що спостереження послідовності сейсмічних подій $\Sigma_j(t_n)$, $(j=1, \dots, R;$

$n=1, \dots, N$), є реалізацією деякого випадкового процесу. Зона вогнищ землетрусів розглядається як фізична система, яка в випадкові моменти часу змінює свій стан, реалізацією, якої є дана послідовність. Вичерпним способом визначення випадкового процесу є завдання його спільної скінченновимірної функції розподілу, визначення якої, за винятком деяких тривіальних випадків, є завданням практично нездійсненним. Однак існує клас марковських процесів, де спільні скінченновимірні функції розподілу виражаються через одно і двовимірні розподілу. У цьому випадку, стає можливим визначення спільного розподілу параметрів кінцевої сукупності землетрусів. Вогнища землетрусів розподілені випадково в деякому обсязі геофізичного середовища, де становище кожного вогнища може залежати від положення попереднього і від положення деякого числа попередніх землетрусів. Останнє твердження означає, що ряд сейсмічних подій може мати довготривалу пам'ять. Гіпотеза про те, що порядок опцій землетрусів є реалізацією ергодичного простого марковського ланцюга, виявилася правдоподібною. Ергодичні марковські ланцюги володіють чудовою для практичних застосувань властивістю: стаціонарні граничні розподіли станів такого ланцюга з будь-яким ступенем точності можна визначити з однієї досить довгої її реалізації. Це означає, що ми зможемо передбачити розподіл координат епіцентру, глибини і магнітуди майбутніх землетрусів. Марковська модель сейсмічних зон може допомогти виявити міграцію осередків землетрусів уздовж сейсмічних поясів і наявності зв'язку між сейсмічними регіонами. Виявлено статистично значущий зв'язок між активністю деяких сейсмічних регіонів Балкан. Динаміка сейсмічної активності балканського регіону – чергування вогнищ землетрусів в регіонах – відстежується елементами матриці ймовірностей переходу. Виявлено, що землетруси мають тенденцію відбуватися там, де воно вже відбулися. Наприклад, по матриці ймовірностей переходу видно, що після землетрусу в Румунії, наступний землетрус, швидше за все, трапиться теж в Румунії, тому що ймовірність переходу $\pi_{11}=0.224$ зі стану *ROM* в *ROM* найбільше. Другим за величиною є ймовірність $\pi_{13}=0.112$ переходу зі стану *ROM* в *BUL*, тобто, наступний після Румунії землетрус може статися, але з меншою ймовірністю в

Болгарії. Ймовірність переходу $\pi_{114}=0.0$ зі стану *ROM* в *POL* дорівнює нулю. Це означає, що сейсмічні зони Румунія і Польща не пов'язані. З вектора граничних розподілів сейсмічних зон можна визначити зону, де і з якою ймовірністю може відбутися очікуваний землетрус. Критерії таблиці спряженості ознак також вказують на наявність середньої тісноти взаємозв'язку сейсмічної активності деяких зон. Марковська модель сейсмічних зон може допомогти виявити міграцію осередків землетрусів уздовж сейсмічних поясів.

Список використаних джерел

- Барлоу Р., Прошан Ф. Математическая теория надежности. Москва, Советское радио, 1969, 321.
- Башарин Г.П. Об использовании критерия согласия Хи-квадрат в качестве критерия независимости испытаний. Теория вероятностей и ее применение, Вып. I.1957,89–90.
- Богданов В.В., Павлов А.В., Полюхова А.Л. Вероятностная модель сейсмичности на примере камчатских землетрясений. Вулканология и сейсмология, 2010, 64–74.
- Буртиев Р.З., Исследование стационарности потока Карпатских землетрясений. Известия АН МССР. Сер. физ.–тех. и мат. наук. Кишинев, 1983, №2, 32–36.
- Буртиев Р.З., Онофреш Н.И. О статистическом подходе к определению параметров сейсмичности. Вопросы инженерной сейсмологии. Задание сейсмических воздействий, Москва. 1993. 4.
- Валтер Я. Стохастические модели в экономике. Москва, Мир, 1971, 231.
- Вентцель Е.С., Овчаров Л.А., Теория случайных процессов и ее инженерные приложения, Высшая Школа, Москва, 2000, 384.
- Естева Л. О моделях сейсмичности. Сейсмический риск и инженерные решения, Москва, Наука, 1968, 375.

- Гайский В.Н., Статистические исследования сейсмического режима Наука, Москва, 1970, 121.
- Гантмахер Ф. Р. Теория матриц. Наука, Москва, 1967. 467.
- Гнеденко Б.В., Курс теории вероятностей. Изд. Физмат литературы. Москва, 1961, 406.
- Джадж Д., Ли И., Зельнер М. Оценивание параметров марковских моделей по агрегированным временным рядам. Статистика, Москва, 1977, 221.
- Карлин С., Основы теории случайных процессов. Москва. Мир. 1971. 327.
- Кнопов Л., Маркушевич В. М. Стационарность и устойчивость марковской последовательности землетрясений. Математическое моделирование и интерпретация геофизических данных. Наука, Москва, 1984. 21–31.
- Крамер Гаральд. Математические методы статистики. Москва. Мир. 1975. 648.
- Курош А. Г., Курс высшей алгебры. Наука, Москва, 1965, 431.
- Ланкастер П. Теория матриц. Наука, Москва, 1978. 269.
- Маркович Э. С., Курс высшей математики. Росвузиздат, 1963, 407.
- Романовский В.И. Дискретные цепи Маркова. Гостехиздат, 1948.
- Рытов С. М. Случайные процессы. Наука, Москва, 1976, 424.
- Сираждинов С. Х. Предельные теоремы для однородных цепей Маркова. Ташкент, Изд-во АН УзССР, 1955.
- Смирнов, И.В. Дунин-Барковский Н.В., Курс теории вероятностей и математической статистики. Наука, Москва, 1965, 511.
- Соколов Н.П. Введение в теорию многомерных матриц. Наукова Думка, Киев, 1972, 175.
- Эрисоглу М., Калис Н., Сэрви Т., Эрисоглу У., Топаксу М., Смешанные модели распределения повторяемости сейсмических событий. Геология и геофизика, т. 52, № 7, 2011, 940–949.
- Шумила В.И. Статистические модели интерпретации поля сейсмической интенсивности. Кишинев. Штиинца. 1985. 128.

- Bottari A., Caccamo D. and Neri G., An example of seismic occurrence modelling by a several–state random process. *Geophys. Journal, Int.* 108, 1992, 267–272.
- Burtiev R.Z. Seismic Hazard Mapping Based on Multi–dimensional Markov Model of a Seismic Regime // *Gerlands Beitrage zur Geophysik, Leipzig*, 95, 1986, 499– 508.
- Burtiev Rashid. Evaluation of seismic hazards from several seismic zones. *Environmental Engineering and Management Journal*. 12, 2012. 32–42.
- Delic Milica, Radojicic Zoran. Seismicity as a multidimensional stochastic process. The 7th Balkan conference on operational research “Bacor05” in Constanta, 2005, Romania. 11.
- Durner Wolfgang. Einführung in die Statistik. Skript zur Vorlesung an der Universität Bayreuth SS, 1999, 73.
- Groß Jürgen. Markov–Prozesse und Markov–Netzwerke. Carl von Ossietzky Universität, Oldenburg, 2009, 32.
- Gunkel Christina. Die Markov Ketten. Monte Carlo Methode zum Testen stochastischer Unabhängigkeit. Kassel, 2008, 106.
- Hillston Jane. <http://www.inf.ed.ac.uk/teaching/courses/pm/PM–lecture1.pdf>, 2009.
- Huisinga Wilhelm, Meerbach Eike. Markov Processes for Everybody. Introduction to the Theory of continuous time Markov processes. Fachbereich Mathematik und Informatik. Freien Universität Berlin & DFG Research Center Matheon, Berlin, 2005, 34.
- Kagan, Y. and Knopoff, L., Earthquake risk prediction as a stochastic process. *Phys. Earth Planet. Inter.*, 14, 1977, 97–108.
- Langrock P., Jahn W., Einführung in die Theorie der markoffschen Ketten und ihre Anwendungen. BSB B. G. Teubner Verlagsgesellschaft. Leipzig, 1979, 192.
- Lombardi Anna Maria. Development and Application of Stochastic Models of Earthquake Occurrence. Universita degli Studi di Bologna, Facolta di Scienze Matematiche, Fisiche e Naturali. Dottorato di Ricerca in Geofisica – XIX Ciclo, 2007, 138.

- Ashok S. Patwardhan, Ram B. Kulkarni and Don Tocher. A semi Markov model for characterising Recurrence of Great Earthquakes. Bull. Seism, Soc. Am. 1980, 323–347.
- Sachs L.und Hedderich J., Angewandte Statistik. Springer Verlag. 2006, 702.
- Shebalin N.V., Leydecker G., Mokrushina N.G., Tatevossian R. E. and Vassiliev V. YU.: Earthquake Catalogue for Central and Southeastern Europe 342 BC–1990 AD. European Commission, Report No. ETNU CT 93–0087, Brussels. 1998.
- Vere–Jones David, How to educate yourself as a statistical seismologist, Community Online Resource for Statistical Seismicity Analysis, doi: 10.5078/corssa–17728079, available at <http://www.corssa.org>. 2010, 14.
- Voit M. Wahrscheinlichkeitsrechnung und mathematische Statistik für Informatiker. 2003. 54.
- Winkler Gerhard. Stochastische Prozesse in der statistischen Modellierung. Institut für Biomathematik und Biometrie. Oberschleißheim. 2000, 161.

РОЗДІЛ 5

СТАТИСТИЧНИЙ АНАЛІЗ МАКРОСЕЙСМІЧНОГО ПОЛЯ

Відомо, що загасання інтенсивності сейсмічних струшувань з відстанню визначається геометричним розсіюванням фронту хвилі, розсіюванням енергії на неоднорідностях і поглинаннях [Корольов, Лазаренко, 2003; Alvarez-Rubio et al., 2010 року; Ambraseys, 1985; Bakun, Scotti, 2006; Bune, 1979; Cecic, Musson, 2004; Constantinos, Papazachos, 1992; Cua et al., 2010; Gomez, Augusto, 2006; Gupta, Nuttli, 1976; Hosh, Mahajan, 2011; Howell, Schultz, 1975; Papoulia, Stavrakakis, 1990; Peruzza, 1996; Sørensen et al., 2009; Sponheuer, 1960; Tselentis, Danciu, 2010]. Особливості сейсмічного впливу кожного землетрусу визначаються такими його характеристиками, як тектоніка, глибина вогнища, механізм, напрямок і хід процесу розриву гірських порід. Картина макросейсмічного поля є відображенням впливу всіх цих факторів і локальних геологічних особливостей на розподіл сейсмічного ефекту на даній поверхні. Загасання інтенсивності сейсмічних впливів є одним з факторів, що визначає якість аналізу сейсмічної небезпеки. Якщо знати де, якої магнітуди, і як часто відбувались землетруси, то за допомогою функції загасання можна відобразити картину сейсмічного впливу в точках земної поверхні. Вище було показано, що пророкування розташування і магнітуди очікуваних сейсмічних подій можна здійснити на основі ймовірнісної моделі сейсмічності. Наступним кроком в ЙАСН, після визначення моделі сейсмічного режиму, стане визначення закону загасання інтенсивності сейсмічних впливів з відстанню.

5.1. Рівняння загасання

Рівняння загасання, яке використовується для аналізу сейсмічної небезпеки, передбачає ізотропність геофізичного середовища і рівномірність поширення сейсмічної енергії. В такому випадку, форма теоретичних ізосейст визначалася б геометрією вогнища в ближній зоні і формою близькою до кола в далекій. Однак, в реальності, від епіцентру землетрусів ізосейсти розходяться у вигляді овалів або

вигнутих ліній довільної форми. На форму цих ізоліній впливають фактори: як загальні: геометрія вогнища, швидкості хвиль поштовхів, так і регіональні – склад матеріалу, товщина ґрунтового шару, рівень ґрунтових вод, відмінність швидкості в корінних породах і в верхніх шарах, фізичні властивості середовища, тип ґрунтів, особливості геологічної будови середовища і т.д. Зазвичай ізосейсти спостережених землетрусів апроксимуються еліпсом [Раутиан, 1982; Різниченко, 1974; Шебалин, 1961, 1997].

Аналіз макросейсмічних і інструментальних даних землетрусів проміжної глибини вогнища Вранча виявив особливості впливу землетрусів (Ismail-Zadeh et al., 2007): вплив землетрусів на великі площі з переважною орієнтацією північно схід-південно захід; амплітуди сейсмічних зсувів ґрунту в більшій мірі залежать від місцевих і регіональних геологічних умов, ніж від магнітуди і відстані до вогнища; параметри сильних рухів ґрунту виявляють велику мінливість. Ізолінії відображають топографію поверхні.

У ближній зоні: $r \sim h$, (r -епіцентрально відстань, h -глибина вогнища землетрусу) геометрія вогнища землетрусів має визначальний вплив на конфігурацію макросейсмічного поля [Шебалин, 1961; Сейсмічне районування території СРСР, 1980]. Наприклад, при дослідженні сейсмічності Киргизстану виявлено, що макросейсмічний ефект протяжності вогнища зникає для землетрусів з $M=6$ на відстанях 25–35 км, для землетрусів з $M=7$ на відстанях 80–110 км. [Джанузаков, 2013].

Розрахункова міра сейсмічної небезпеки в значній мірі залежить від функції загасання інтенсивності сейсмічного ефекту. При аналізі сейсмічної небезпеки і при побудові карт сейсмічного районування території Республіки Молдова використовувалися функції загасання, запропоновані різними сейсмологами. Однак, деякі закони загасання дають завищені значення інтенсивності в балах, тому що, сума квадратів відхилень теоретичних значень інтенсивності струшувань від їх спостережених значень є статистично значущою. Існуюча відмінність в картах сейсмічної небезпеки,

обумовлена відмінностями використаних їх розробниками функцій загасання, спонукало автора на проведення власного аналізу загасання інтенсивності сейсмічних впливів. Існує безліч варіантів функції загасання, проте всі вони засновані на законі загасання Коваслігеті: [Гусев, Шумилина, 1999; Шебалин, 1968; Alvarez-Rubio et al., 2010; Ambraseys, 1985; Bakun, Scotti, 2006; Constantinos, Papazachos, 1992; Cua et al., 2010; Chyi-Tyi Lee et al., 2012; Gomez, Augusto, 2006; Gupta, Nuttli, 1976; Hosh, Mahajan, 2011; Howell, Schultz, 1975; Lapajne, 1988; Papoulia и Stavrakakis, 1990; Rodler, 2012; Sponheuer, 1960; Tselentis, Danciu, 2010].

Таблиця 5.1.

Рівняння загасання

Коваслігетігеті (1906)	$I = I_0 - a \log\left(\frac{R}{h}\right) - b(D - h)$
Блейк (1941)	$I = I_0 - a \log\left(\frac{R}{h}\right) + b$
Sponheuer(1962)	$I_0 - I = kq \lg\left(\frac{D}{h}\right)^a + k\alpha(\lg e)(D - h)$
Shebalin (1968)	$I = am - b \lg R + c$
Howel, Schutz (1975)	$\ln\left(\frac{I}{I_0}\right) = 0.364 - 0.130 \ln R - 0.0019R$
Gupta, Nutti (1976)	$I = I_0 + 3.7 - 2.7 \lg D - 0.0011D, D \geq 20km.$
Chandra (1979)	$I = I_0 + 6.453 - 4.96 \lg(D + 20) - 0.00121D$
Sbar, Du Bois (1984)	$I = I_0 + 3.2 - 1.5 \ln(D) - 0.0015D$
Ambraseys (1985)	$I_0 - I = -0.22 + 0.0024(R - h) + 2.85 \log(R/h)$
Dugue (1989)	$I_0 - I = 0.2 \ln(D - d) + 0.04(D - d), d > 10$
Dowrick (1992)	$I = 2.18 + 1.41m - 1.18 \ln R - 0.0044R$
Berardi et al. (1993)	$I_0 - I = -0.729 + 1.122\sqrt[3]{D}$

Gasperini (2001)	$I_0 - I = \begin{cases} 0.445 + 0.59R, R \leq 45km \\ 0.445 + 0.59 \cdot 45 + 0.0207(R - 45), \\ R > 45km \end{cases}$
Albarelo, D'Amico(2004)	$I = a + bR + c \ln R + dI_0$
Гомец Капера (2006)	$I - I_0 = a\sqrt[3]{R} + b$
Bakun, Wentworth (2006)	$I = am + b \lg R + cR + d$

де a, b, c шукані коефіцієнти, I – інтенсивність в балах макросейсмічної шкали, m – магнітуда землетрусів, I_0 – інтенсивність в епіцентрі, D – епіцентральна відстань, R – гіпоцентральна відстань, h – глибина вогнища землетрусів.

5.2. Математичний опис еліпса

Для дослідження використовуються макросейсмічні дані землетрусів, що сталися 10.11.1940; 7.04.1977; 31.08.1986; 30.05.1990 і 31.05.1990 рр. [Kronrod, 2012]. Досить великий обсяг макросейсмічних спостережень за всіма азимутними напрямками землетрусів 1977 і 1986 рр., 4088 і 2119 *IDP* точок, і відповідно дозволяє провести детальний аналіз макросейсмічного поля. Макросейсмічне поле характеризується зонами однакової інтенсивності та лініями, що розділяють зони зміни бальності, які бувають досить хвилясті. Ширина зон однакової інтенсивності збільшується в міру віддалення від епіцентру землетрусів (рис. 5.2–5.6). Наприклад, ширина 3-х і 4-х бальних зон доходять до декількох сотень кілометрів, тоді як зони високої бальності мають ширину в кілька десятків км. Аналіз макросейсмічного поля проводиться для обґрунтування апроксимації еліпсом зон однакової бальності.

Апроксимація еліпсом, хоча і є спрощеною моделлю реальної картини, проте, виявляється прийнятним для ймовірнісного аналізу сейсмічної небезпеки підходом. Крім того, передбачається, що лінії, що розділяють області зміни бальності є кривими другого порядку. У ближній зоні вони мають форму близьку до еліптичної, що визначає, в основному, геометрією вогнища, наприклад, горизонтальну

протяжність вогнища. По мірі збільшення епіцентральної відстані форма ізоліній змінюється відповідно до особливостей геологічної будови середовища. Аналітично криві другого порядку описуються рівнянням другого степеня:

$$ax^2 + bxy + cy^2 + dx + ey + f = 0 \quad (5.1)$$

Залежно від співвідношень між коефіцієнтами це рівняння описує коло, параболу, гіперболу або еліпс [Глаголев, Солнцева, 1971; <http://mathworld.wolfram.com/Ellipse.html>].

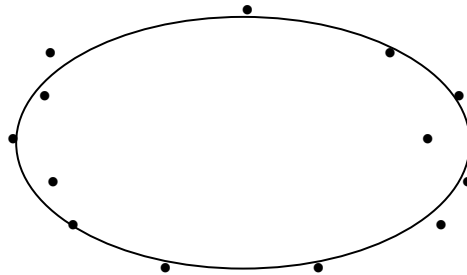


Рис. 5.1. Схема апроксимації точок еліпсом.

Канонічне рівняння еліпса [Смирнов, 1974]:

$$\frac{x_n^2}{a^2} + \frac{y_n^2}{b^2} = 1 \quad (5.2)$$

Якщо точки знаходяться на еліпсі, то їх координати задовольняють канонічне рівняння (рис. 5.1), і значення осей еліпса a , b визначаються з мінімуму суми квадратів відхилень [Вылегжанин, Рыбалка, 2006; LI et al., 2005; Harkera et al., 2008; Dörfler, Wieners, 2009; Niehoff, 2006]:

$$\sum_{n=1}^N \left(\frac{x_n^2}{a^2} + \frac{y_n^2}{b^2} - 1 \right)^2 = \min \quad (5.3)$$

Таким чином, база даних являє собою множину точок макросейсмічних спостережень (x_n, y_n, I^k) , $n=1, \dots, N$; $k=1, \dots, 12$, где x_n и y_n – географічні координати *IDP* точок (рис. 5.2–5.6). Потрібно апроксимувати ці точки еліпсом, тобто знайти координати його фокусів і осі еліпса a , b . Ясно, що апроксимацію слід проводити окремо для кожної зони бальності $k=3, 4, 5, \dots, I_{\max}$ спостереженого макросейсмічного поля. У деяких випадках, центр еліпса поєднується зі середньоарифметичною

широтою і довготою *IDP* точок. У нашому випадку, центр еліпса і початок координат сполучений з епіцентром вранчських землетрусів. Введемо позначення:

$$\alpha = \frac{1}{a^2}, \beta = \frac{1}{b^2} \quad a_{11} = \sum_{n=1}^N x_n^2 x_n^2, \quad a_{12} = a_{21} = \sum_{n=1}^N x_n^2 y_n^2, \quad (5.4)$$

$$a_{22} = \sum_{n=1}^N y_n^2 y_n^2, \quad c_1 = \sum_{n=1}^N x_n^2, \quad c_2 = \sum_{n=1}^N y_n^2$$

далі отримаємо:

$$\sum_{n=1}^N \left(\frac{x_n^2}{a^2} + \frac{y_n^2}{b^2} - 1 \right)^2 = a_{11}\alpha^2 + 2a_{12}\alpha\beta + a_{22}\beta^2 - 2c_1\alpha - 2c_2\beta - N \quad (5.5)$$

З рішення системи лінійних рівнянь:

$$\begin{cases} a_{11}\alpha + a_{12}\beta = c_1 \\ a_{21}\alpha + a_{22}\beta = c_2 \end{cases} \quad (5.6)$$

і використовуючи вирази (5.4), визначимо значення довжин піввісей еліпса, апроксимуючого лінію, що з'єднує точки (x_n, y_n) рівної інтенсивності.

Таблиця 5.2

Довжини піввісей еліпсів, апроксимуючих зони бальності

	10.11.1940		4.03.1977		31.08.1986		30.05.1990		31.05.1990	
Бал	Вісь <i>a</i>	Вісь <i>b</i>	Вісь <i>a</i>	Вісь <i>b</i>	Вісь <i>a</i>	Вісь <i>b</i>	Вісь <i>a</i>	Вісь <i>b</i>	Вісь <i>a</i>	Вісь <i>b</i>
3	1477.2	837.06	1684.4	816.55	984.0	643.03			637.0	532.0
4	2147.3	533.81	806.09	633.58	641.1	622.79			504.5	478.2
5	439.7	789.84	466.99	483.71	281.9	379.81	409.9	417.86	210.2	279.50
6	378.3	457.07	328.24	420.14	225.0	297.49	245.8	300.53	126.1	143.34
7	230.4	378.72	251.78	313.17	157.7	178.63	170.4	191.90	32.12	29.91
8	214.9	173.85	253.20	382.52	225.92	129.37	88.7	25.9		
9	144.7	85.15								

Канонічне рівняння еліпса передбачає, що центром еліпса є початок координат, і вісі еліпса розташовуються на осях координат.

$$\sqrt{(x+c)^2 + y^2 + (x-c)^2 + y^2} = 2a \quad (5.7)$$

де c – половина фокусної відстані. Для досліджуваних макросейсмічних полів умова (5.7) виконується. Фокальна піввісь орієнтована в залежності від інтенсивності (табл. 5.2). Спостерігається цікавий факт, в канонічному рівнянні еліпса для 5, 6 і 7 бальних зон, всіх досліджуваних землетрусів, фокальна піввісь направлена по осі ординат (табл. 5.2). Це означає, що 5, 6 і 7 бальні зони орієнтовані в напрямку північ-південь, тобто витягнуті уздовж меридіанів, і в цьому напрямку енергія сейсмічних коливань затухає повільніше. Фокальна вісь еліпсів, апроксимуючих інші зони рівної інтенсивності, знаходиться на вісі абсцис і спрямована по паралелі. Цей факт є ще одним підтвердженням анізотропного середовища.

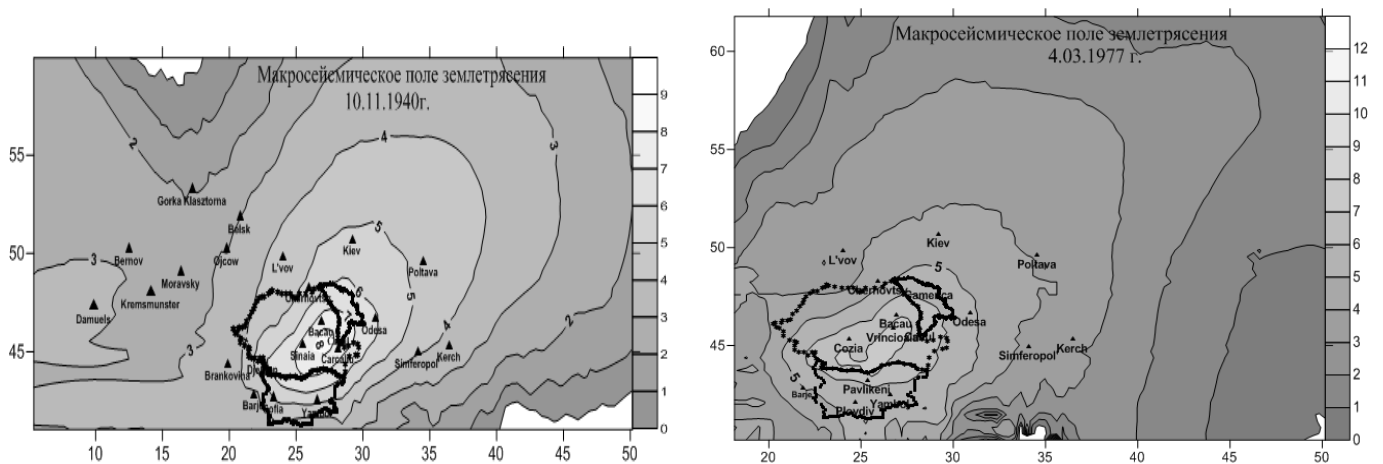


Рис. 5. 2. Макросейсмічне поле землетрусу: а) 10.11.1940; б) 4.03.1977

Легенда карт, представлених на малюнках 5.2–5.6:

- зліва і внизу вказані географічні широти і довготи з кроком в 5 градусів, а праворуч показана чисельна і колірна шкала відмінностей.
- на ізолініях (лініях рівного значення відмінностей) вказані значення макросейсмічних балів.
- сейсмічні станції, із зазначенням їх назв, позначені символом (▲).
- державні кордони позначені символами (+).

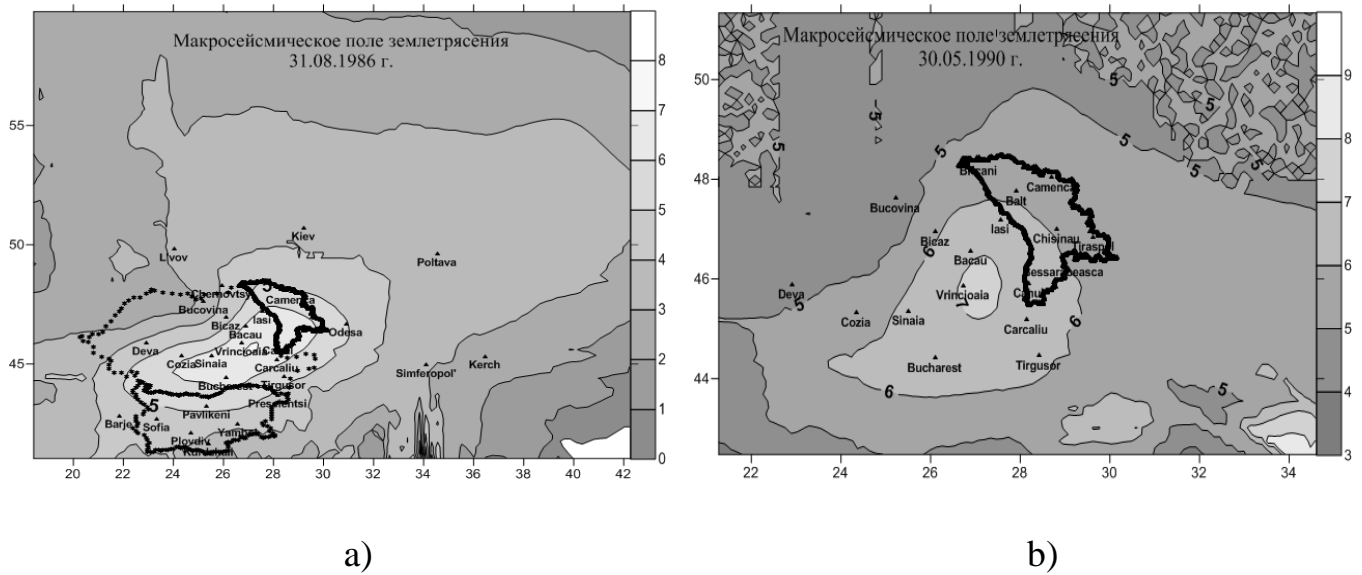


Рис. 5.3. Макросейсмічне поле землетрусів: а) 30.08.1986; б) 30.05.1990.



Рис. 5.4. Макросейсмічне поле землетрусу 31.05.1990

Для визначення параметрів еліпса (центру еліпса, довжин піввісей і їх орієнтації щодо обраної системи координат) використовується загальне рівняння кривих другого порядку:

$$F(x, y) = ax^2 + bxy + cy^2 + dx + ey + f = 0 \quad (5.8)$$

У рівнянні (5.8) коефіцієнти a , b , c , d , e входять лінійно, отже, їх можна обчислити методом найменших квадратів [Harkera et al., 2008; Dörfler, Wieners, 2009]. Це завдання можна також розглядати як процедуру наближення багатьох точок еліпсом

[Agin, 1981; Burtiev, 2014; Eberly, 2004; Fitzgibbon, 1995; Fitzgibbon et al., 1999; Van Loan, 2007]. Для вирішення цього завдання застосовується метод апроксимації, запропонований Фітцгіббон [Fitzgibbon, 1995], який представляє стійкий не ітеративний алгоритм апроксимації еліпсом багатьох точок. Підхід заснований на методі найменших квадратів, і оптимальне рішення обчислюється безпосередньо, без ітерації. Метод дозволяє підібрати еліпс для розсіяних, що містять перешкоди даних. Рівняння другого ступеня (5.8) визначає еліпс, якщо виконується умова:

$$4ac - b^2 > 0 \quad (5.9)$$

Підстановкою в поліном (5.8) координат точок (x_n, y_n) , $n=1, \dots, N$, визначається алгебраїчна відстань від точок до еліпса. Причому відстань береться не за віссю Y (чи X), – як в звичайних задачах на *МНК*, а в сенсі найкоротшого шляху від точки спостереження до найближчої до неї точки на еліпсі. Ввівши позначення:

$$\vec{a} = [a, b, c, d, e, f], \quad \vec{x} = [x^2, xy, y^2, x, y, 1]^T \quad (5.10)$$

можна отримати векторну форму виразу (5.10):

$$F_a(x) = \vec{a} \cdot \vec{x} = 0 \quad (5.11)$$

Нехай N_k – число точок в зоні інтенсивності макросейсмічного поля бальності k . Апроксимацією еліпсом проводиться для кожної зони точок $(x_n, y_n)_k$ для $n=1, \dots, N_k$, $k=1, \dots, K$, $N_1 + \dots + N_K = N$, рівній бальності, шляхом обчислення мінімуму суми квадратів алгебраїчних відстаней [Eberly, 2004]:

$$\min_a \sum_{n=1}^{N_k} F(x_n, y_n)^2 = \min_a \sum_{n=1}^{N_k} (F_a(\vec{x}^i))^2 = \min_a \sum_{n=1}^{N_k} (\vec{x}^i \vec{a})^2 \quad (5.12)$$

При обчисленнях, умова (5.9) замінюється рівністю [Fitzgibbon, 1995; Agin, 1981]:

$$4ac - b^2 = 1 \quad (5.13)$$

Рішення завдання апроксимації точок еліпсом, за методом найменших квадратів, у векторній формі, з урахуванням обмежень (5.13), знаходиться з рівняння:

$$\min_a |Da|^2 = 1 \quad (5.14)$$

$$4ac - b^2 = \vec{a} \cdot C \cdot \vec{a}^T = 1 \quad (5.15)$$

Матриця C - матриця обмежень, а матриця D , розмірності $(N_k \times 6)$, визначається координатами точок (x_n, y_n) :

$$C = \begin{pmatrix} 0 & 0 & 2 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \quad (5.16)$$

$$D = \begin{pmatrix} x_1^2 & x_1 y_1 & y_1^2 & x_1 & y_1 & 1 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ x_j^2 & x_j y_j & y_j^2 & x_j & y_j & 1 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ x_{N_i}^2 & x_{N_i} y_{N_i} & y_{N_i}^2 & x_{N_i} & y_{N_i} & 1 \end{pmatrix} \quad (5.17)$$

Найменше значення функції (5.14), за умови, що змінні зв'язані співвідношеннями (5.15), визначається за допомогою способу множників Лагранжа [Dörfler, Wiener, 2009]:

$$S \cdot \vec{a} = \lambda C \cdot \vec{a}, \quad \vec{a} \cdot C \cdot \vec{a}^T = 1 \quad (5.18)$$

де матриця S дорівнює:

$$S = D^T D = \begin{pmatrix} S_{x^4} & S_{x^3 y} & S_{x^2 y^2} & S_{x^3} & S_{x^2 y} & S_{x^2} \\ S_{x^3 y} & S_{x^2 y^2} & S_{xy^3} & S_{x^2 y} & S_{xy^2} & S_{xy} \\ S_{x^2 y^2} & S_{xy^3} & S_{y^4} & S_{xy^2} & S_{y^3} & S_{y^2} \\ S_{x^3} & S_{x^2 y} & S_{xy^2} & S_{x^2} & S_{xy} & S_x \\ S_{x^2 y} & S_{xy^2} & S_{y^3} & S_{xy} & S_{y^2} & S_y \\ S_{x^2} & S_{xy} & S_{y^2} & S_x & S_y & S_1 \end{pmatrix} \quad (5.19)$$

Елементи матриці S визначаються формулою [Eberly, 2004]:

$$S_{x^l y^m} = \sum_{n=1}^{N_i} x_n^l y_n^m \quad (5.20)$$

У системи (5.18) існують до шести рішень (λ_j, a_j) . З умови:

$$\|Da\|^2 = a_T D_T D a = a_T S a = \lambda a_T C a = \lambda \quad (5.21)$$

визначається власний вектор a_k , відповідний позитивному найменшому своїм значенням λ_k . Далі визначається оптимальне рішення рівняння (5.18), що задовольняє умові:

$$\vec{a}_k^T \cdot C \cdot \vec{a}_k = 1 \quad (5.22)$$

яке є найкращою апроксимацією точок еліпсом [Halir, Flusser, 2008]. Завдання вирішується за алгоритмом Фітцгиббон:

- Матриця D (5.17) розкладається на дві матриці D_1, D_2 , що містять його квадратні і лінійні частини:

$$D = (D_1 | D_2) \quad (5.23)$$

$$D_1 = \begin{pmatrix} x_j^2 & x_j y_j & y_j^2 \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ x_{N_i}^2 & x_{N_i} y_{N_i} & y_{N_i}^2 \end{pmatrix} \quad D_2 = \begin{pmatrix} x_j & y_j & 1 \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ x_{N_i} & y_{N_i} & 1 \end{pmatrix} \quad (5.24)$$

- Далі, матриця S (5.19) записується в блоковій формі:

$$S = \begin{pmatrix} S_1 & S_2 \\ S_2^T & S_3 \end{pmatrix}; \quad S_1 = D_1^T D_1; \quad S_2 = D_1^T D_2; \quad S_3 = D_2^T D_2 \quad (5.25)$$

Блокова форма матриці обмежень C (5.16) має вигляд:

$$C = \begin{pmatrix} C_1 & 0 \\ 0 & 0 \end{pmatrix}; \quad C_1 = \begin{pmatrix} 0 & 0 & 2 \\ 0 & -1 & 0 \\ 2 & 0 & 0 \end{pmatrix} \quad (5.26)$$

- Вектор коефіцієнтів рівняння еліпса (5.10) також розкладається на два вектора:

$$a = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}; \quad a_1 = \begin{pmatrix} a \\ b \\ c \end{pmatrix}; \quad a_2 = \begin{pmatrix} d \\ e \\ f \end{pmatrix}; \quad (5.27)$$

З виразу (5.18) випливає матричне рівняння:

$$\begin{pmatrix} S_1 & S_2 \\ S_2^T & S_3 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = \lambda \begin{pmatrix} S_1 & S_2 \\ S_2^T & S_3 \end{pmatrix} \quad (5.28)$$

еквівалентне системі:

$$\begin{cases} S_1 a_1 + S_2 a_2 = \lambda C_1 a_1 \\ S_2^T a_1 + S_3 a_2 = 0 \end{cases} \quad (5.29)$$

Матриця S_3 :

$$S_3 = D_2^T \cdot D_2 = \begin{pmatrix} S_{x^2} & S_{xy} & S_x \\ S_{xy} & S_{y^2} & S_y \\ S_x & S_y & S_1 \end{pmatrix} \quad (5.30)$$

є регулярною, так як точки макросейсмічного поля не лежать на одній прямій. З цього випливає [Eberly, 2004]:

$$a_2 = -S_3^{-1} S_2^T a_1 \quad (5.31)$$

Підставляючи знайдене вираз (5.31) в (5.29) отримаємо:

$$(S_1 - S_2 S_3^{-1} S_2^T) a_1 = \lambda C_1 a_1 \quad (5.32)$$

Визначник матриці C_1 відмінний від нуля (матриця регулярна). Отже, (5.32) набуде вигляду:

$$C^{-1} (S_1 - S_2 S_3^{-1} S_2^T) a_1 = \lambda a_1 \quad (5.33)$$

З огляду на форму визначення матриці C (5.26) отримаємо [Halir, Flusser, 2008]:

$$a_1^T C_1 a_1 = 1 \quad (5.34)$$

Остаточне рішення, що задовольняє умовам (5.18), виражається набором формул:

$$M a_1 = \lambda a_1; a_1^T C_1 a_1 = 1; \bar{a}_2 = -S_3^{-1} S_2^T \bar{a}_1; \bar{a} = \begin{pmatrix} \bar{a}_1 \\ \bar{a}_2 \end{pmatrix} \quad (5.35)$$

де M наведена і розріджена матриця третього порядку [Halir, Flusser, 2008]:

$$M = C_1^{-1} S_1 - S_2 S_3^{-1} S_2^T \quad (5.36)$$

Наведеною називається матриця, якщо в кожному її ненульовому рядку є ненульовий елемент такий, що всі інші елементи стовпця, що містять цей елемент, дорівнюють нулю [Ланкастер, 1978]. Матриця, що складається з відносно великого числа нульових елементів називається розрідженою матрицею. У таблиці показані результати роботи обчислювальної програми реалізації описаного алгоритму на мові програмування ФОРТРАН (табл. 5.3). За винятком деяких зон рівних бальностей макросейсмічних полів землетрусів 30.05. 1990 і 31.05.1990 рр. коефіцієнти загального рівняння (5.14) задовольняють умові еліптичності кривої (5.15). У

рівнянні еліпса коефіцієнт $b \neq 0$, тобто в рівнянні міститься член з змінними x і y . Поворотом осей координат на деякий кут α можна досягти того, що член при одній з них зникне [Глаголев, Солнцева, 1971]. Для вирішення складаються матриці, де елементи матриці G_0 пов'язані з коефіцієнтами рівняння еліпса [Niehoff, 2006]:

$$\gamma_{11} = a(1); \quad \gamma_{12} = a(2)/2; \quad \gamma_{22} = a(3); \quad \gamma_{13} = a(4)/4 \quad (5.37)$$

$$\gamma_{23} = a(5)/2; \quad \gamma_{33} = a(6); \quad \gamma_{31} = \gamma_{13}; \quad \gamma_{32} = \gamma_{23}; \quad \gamma_{21} = \gamma_{12}$$

$$G_0 = \begin{pmatrix} -0.8862 & 0.2251 \\ 0.2251 & -0.1097 \end{pmatrix}; \quad G = \begin{pmatrix} -0.8862 & 0.2251 & -0.6285 \\ 0.2251 & -0.1097 & 0.1312 \\ -0.6285 & 0.1312 & 0.4585 \end{pmatrix} \quad (5.38)$$

Знаходимо центр еліпса [<http://mathworld.wolfram.com/Ellipse.html>]:

$$F_0 = (\gamma_{12}\gamma_{23} - \gamma_{13}\gamma_{22})/\det(G_0); \quad \Lambda_0 = (\gamma_{13}\gamma_{12} - \gamma_{11}\gamma_{23})/\det(G_0) \quad (5.39)$$

де $\det(G_0)$ визначник матриці G_0 . Для обчислення довжин вісей еліпса і фокусної відстані застосовуються формули:

$$\eta = \sqrt{\frac{-\det(G)}{\det(G_0)\lambda_1}}; \quad \mu = \sqrt{\frac{-\det(G)}{\det(G_0)\lambda_2}}; \quad \nu = \sqrt{\eta^2 - \mu^2} \quad (5.40)$$

Власні значення матриці ранжовані відповідно до нерівності [Dörfler, Wieners, 2009]:

$$|\lambda_1 - a| < |\lambda_1 - c|; \quad |\lambda_2 - c| < |\lambda_2 - a| \quad (5.41)$$

Кут α між великою віссю еліпса і віссю X , орієнтований проти напрямку руху годинникової стрілки, визначиться формулою [Глаголев, Солнцева, 1971]:

$$\tan 2\alpha = \frac{2\gamma_{12}}{\gamma_{11} - \gamma_{22}} \quad (5.42)$$

і тангенс кута визначиться з виразу:

$$k = \sqrt{\frac{(\tan^2 2\alpha + 1) - 1}{\tan 2\alpha}} \quad (5.43)$$

Координати фокусів залежать від орієнтації осей еліпса. Якщо велика вісь спрямована вздовж осі Y , то тангенс кута нахилу приймає значення [Eberly, 2004]:

$$k = -1/k, \quad p = \nu\sqrt{1+k^2} \quad (5.44)$$

Центри апроксимуючих макросейсмічне поле еліпсів, обчислені за формулою (5.39), і інструментальні епіцентри для деяких землетрусів віддалені на значну відстань (Таб. 5.3). Причиною даного феномена може бути якість макросейсмічних даних, механізм вогнища, а також геометричні особливості осередків землетрусів.

Таблиця 5.3

Координати центрів еліпсів, апроксимуючих макросейсмічне поле

Дата	10.11.1940		4.03.1977		31.08.1986		30.05.1990		31.05.1990	
координати	широта	довгота	широта	довгота	широта	довгота	широта	довгота	широта	довгота
макросейсмічні	44.63	24.41	45.46	26.44	46.38	26.55	45.68	26.65	47.00	28.51
інструментальні	45.75	26.93	45.85	26.97	45.55	26.59	45.89	26.99	45.88	26.98

Координати фокусів знаходяться з виразів:

$$\begin{aligned} Fx_1 &= F_0 - p, & Fy_1 &= L_0 - kp \\ Fx_2 &= F_0 + p, & Fy_2 &= L_0 + kp \end{aligned} \quad (5.45)$$

Сам кут дорівнює:

$$\psi = \begin{cases} \frac{1}{2} \arctan k, & b \neq 0, b < c \\ \frac{\pi}{2} + \frac{1}{2} \arctan(k), & b \neq 0, b > c \end{cases} \quad (5.46)$$

Як відомо ексцентриситет еліпса:

$$e = \sqrt{1 - \left(\frac{b}{a}\right)^2} \quad (5.47)$$

завжди менше одиниці [Маркович, 1963] і служить для визначення величини стиснення еліпса, тобто характеризує еліптичність кривої другого порядку. Величина (5.47) для кривих апроксимуючих макросейсмічне поле різної бальности вранчських землетрусів виявилася менше одиниці, тобто криві, якими апроксимували спостережені точки макросейсмічною поля, є еліпсом.

Таблиця 5.4

Коефіцієнти загального рівняння еліпса

	4.03.1977							10.11.1940	31.08.1986	30.05.1990	31.05.1990
Бал	<i>a</i>	<i>b</i>	<i>c</i>	<i>d</i>	<i>e</i>	<i>f</i>	$4ac-b^2$	$4ac-b^2$			
3	-0.74	0.26	-0.4	-0.5	0.01	0.43	1.14	0.78	1.32		0.24
4	-0.87	0.22	-0.11	-0.63	0.13	0.45	0.34	0.97	1.86		0.29
5	0.67	-0.35	0.25	0.37	-0.16	-0.23	0.54	1.71	0.1	0.91	0.49
6	-0.46	0.04	-0.88	-0.35	-0.36	0.44	1.63	0.19	0.24	0.48	-0.66
7	-0.18	0.20	-0.89	-0.05	0.24	0.15	0.61	0.51	0.32	-0.35	-0.25
8	0.34	-0.42	0.43	0.07	-0.13	0.002	0.41	0.33	0.16	0.39	
9								0.38			

5.3. Визначення епіцентральної відстані

Відстань є однією з основних характеристик сейсмічної небезпеки. Зазвичай при розрахунку СН, використовується епіцентральної відстань для площинних джерел і Джойнер-Бура відстань для розломних джерел, яке є найкоротшою відстанню від пункту $Q(\varphi, \psi)$ до проекції площини розриву на дану поверхню [Левін, 2012; Chui-Tui Lee et al., 2006]. Землетруси в зоні Вранча відбуваються в межах земної кори і в проміжному шарі на глибинах 80–200 км., і неясно, що визначає епіцентр майбутнього землетрусу, і де він буде знаходитися, адже епіцентральної площина зони Вранча займає площу $60 \times 80 \text{ км}^2$ [Onicescu, 1997]. Площа епіцентральної площини деякого землетрусу залежить від його магнітуди і глибини. Залишається неясним, що визначає відстань від точки $Q(\varphi, \psi)$ до епіцентру конкретного вранчського землетрусу і як воно визначається.

Існують два підходи до моделювання глибинної сейсмічності зони Вранча – рівномірна і дискретна. Рівномірна сейсмічність передбачає можливість виникнення великої сейсмічної події на будь-якій глибині інтервалу 80–200 км, а дискретна модель допускає появи землетрусу певної магнітуди в певних інтервалах глибин [Алказ, 2007]. Осередок землетрусів займає певний обсяг геофізичної середовища, проте в сейсмічних відомостях представляється точкою в геофізичному просторі, що

є спрощенням реальної ситуації. Якщо при проведенні ЙАСН в далекій зоні геометричними розмірами вогнища землетрусів можна знехтувати, вважаючи його матеріальною точкою, то в ближній зоні подібне спрощення може відбитися на якості аналізу СН.

Щоб передбачити інтенсивність майбутніх сейсмічних впливів у пункті $Q(\varphi, \psi)$ досить передбачити 4-вимірної інтервал, що містить параметри майбутнього землетрусу. Тобто, при аналізі СН немає необхідності в точковій оцінці розподілу параметрів землетрусу, досить визначити ймовірнісний розподіл сейсмічності по елементарним зонам. Розбиття зони вогнищ на ЕДЗ дозволить підвищити точність визначення відстаней, і відпадуть неясності, пов'язані з визначенням глибини та епіцентральної відстані. При визначенні обсягу ЕДЗ слід вибирати між детальністю визначення параметрів і достатньою наглядністю сейсмічності елементарних зон.

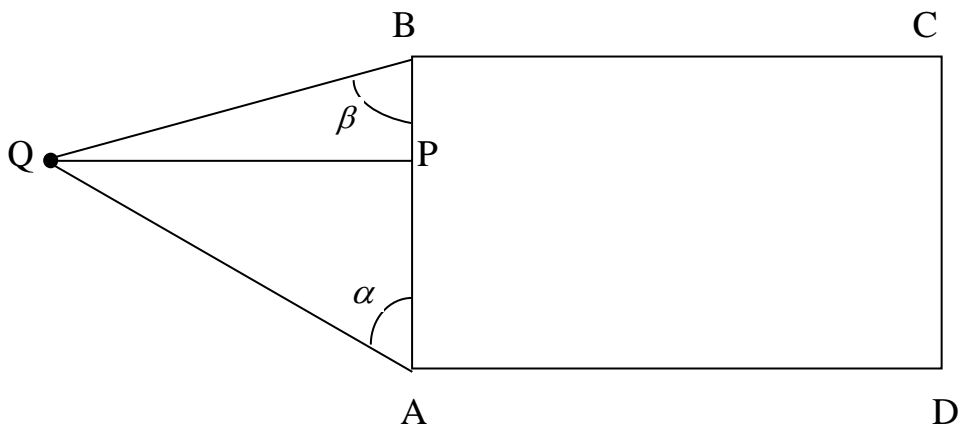


Рис. 5.7. Схема визначення мінімальної відстані від точки до проекції ЕДЗ.

Проекцією ЕДЗ на поверхню Землі є прямокутник. Для досліджуваних землетрусів використовується епіцентральної відстань – відстань від $Q(\varphi, \psi)$ до середини прямокутників, і відстань Джойнер-Бура – найкоротша відстань від точки $Q(\varphi, \psi)$ до епіцентрального прямокутника $ABCD$ (рис. 5.7) – проекції ЕДЗ на дану поверхню. Це відстань дорівнюватиме довжині відрізка прямої, що з'єднує задану точку з найближчою точкою, розташованої на прямокутнику. Нехай прямокутник заданий координатами вершин $A(x_A, y_A)$; $B(x_B, y_B)$; $C(x_C, y_C)$; $D(x_D, y_D)$. Мінімальною відстанню від точки до прямокутника є найменша відстань від цієї точки до сторін

прямокутника, а найкоротша відстань $Q(\varphi, \psi)$, від точки до відрізка прямої AB визначається за схемою:

1. з чотирьох вершин визначаються дві найближчі. Вони зададуть сторону, до якої буде обчислюватися мінімальна відстань
2. довжина перпендикуляра, опущеного з точки на пряму лінію, є найкоротшою відстанню між ними. Якщо основа перпендикуляра P буде перебувати на найближчій до точки Q стороні AB , то довжина перпендикуляра і є найменшою до прямокутника відстанню
3. якщо ж основа перпендикуляра знаходиться лівіше точки A , або правіше точки B , то мінімальним є менше з двох відстаней QA , QB . Для цього обчислюється скалярний добуток векторів QA , AB по їх декартових координатах:

$$P = (\overrightarrow{Q - A})(\overrightarrow{B - A}) \cos \alpha = (\varphi - x_A)(x_B - x_A) + (\psi - y_A)(y_B - y_A) \quad (5.48)$$

Косинус кута α визначиться з рівності:

$$\cos \alpha = \frac{(\varphi - x_A)(x_B - x_A) + (\psi - y_A)(y_B - y_A)}{\sqrt{(\varphi - x_A)^2 + (\psi - y_A)^2} \sqrt{(x_B - x_A)^2 + (y_B - y_A)^2}} \quad (5.49)$$

аналогічно визначається $\cos \beta$, косинус кута при вершині B .

4. якщо $\cos \alpha$ менше нуля, то основа перпендикуляра QP розташоване лівіше вершини A , і найменшою є відстань AQ . Якщо ж $\cos \beta$ негативний, основа знаходиться правіше точки B і мінімальною відстанню від точки Q до відрізка прямої AB є BQ . При позитивних значеннях $\cos \alpha$ і $\cos \beta$ основа перпендикуляра QP знаходиться в межах інтервалу $[AB]$, і найменшою відстанню є довжина перпендикуляра: $|PQ| = |AQ| \cdot \sin \alpha$.

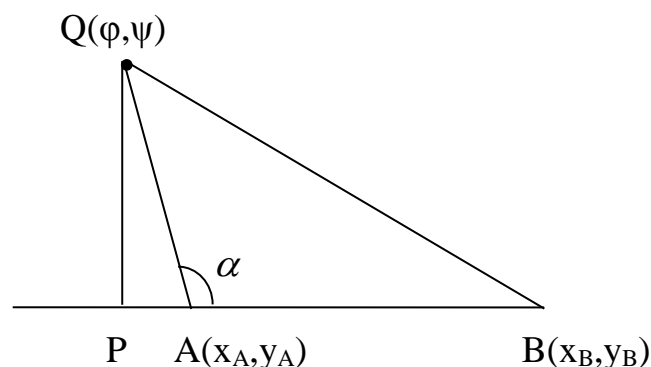


Рис. 5.8. Схема обчислення найкоротшої відстані від точки до відрізка прямої лінії.

5.4. Обчислення значень коефіцієнтів рівняння макросейсмічного поля

В каталозі румунських землетрусів для кількох сотень подій наводиться і епіцентральна інтенсивність землетрусів I_0 в балах шкали *MSK-64* [Oncescu, 1997]. Залежність епіцентральної інтенсивності I_0 від магнітуди MW і глибини h цих землетрусів виражається формулою лінійної регресії:

$$I_0 = 1.49M_w - 0.01h - 0.93 \quad (5.50)$$

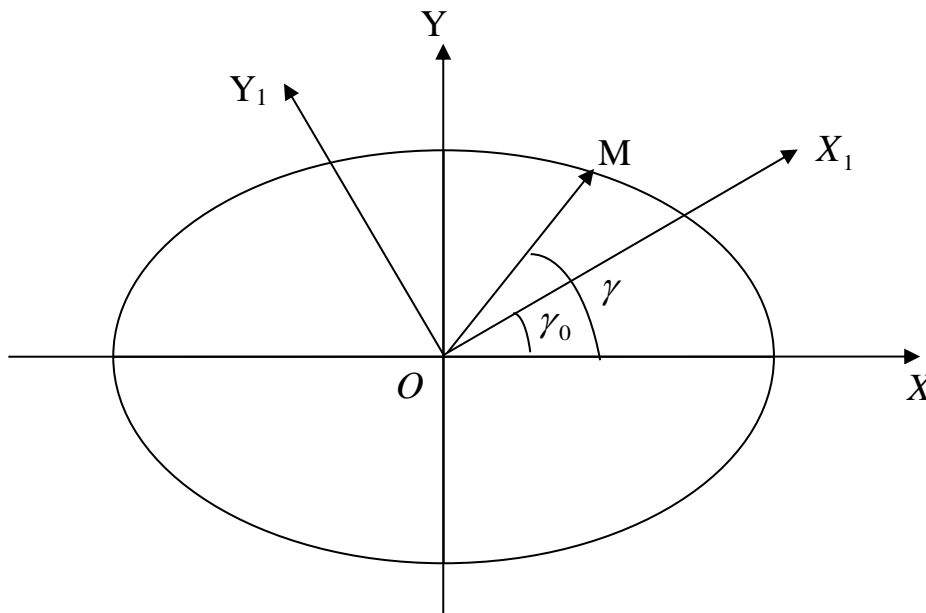


Рис. 5.9. Схема визначення еліптичної залежності загасання інтенсивності сейсмічних впливів

Макросейсмічне поле досліджуваних землетрусів мають витягнуту з південного міру на північний схід конфігурацію, що нагадує форму еліпса, велика вісь якого повернена на деякий кут γ_0 щодо позитивного напрямку осі абсцис (рис. 5.2–5.6). Для обліку азимутальної залежності загасання інтенсивності струшувань в рівняння макросейсмічного поля, в формі множника або доданка, додається параметр, що залежить від нахилу великої вісі еліпса загасання щодо позитивного напрямку вісі абсцис. Загальне рівняння макросейсмічного поля, що зв'язує

інтенсивність в шкалі балів в певній точці (φ_i, ψ_i) з магнітудою m , епіцентральною відстанню Δ і глибиною вогнища h , має вигляд [Шебалин, 1961]:

$$I = am - b \log \sqrt{\Delta^2 + h^2} + c \quad (5.51)$$

де a, b, c – константи.

Для обліку азимутальної залежності рівняння макросейсмічного поля від параметрів землетрусу значення коефіцієнта b визначається формулою:

$$b = \frac{b_{\max} b_{\min}}{\sqrt{b_{\min}^2 \cos^2(\gamma - \gamma_0) + b_{\max}^2 \sin^2(\gamma - \gamma_0)}} \quad (5.52)$$

де b_{\max}, b_{\min} велика і мала осі еліпса, γ кут між напрямом на точку $M(\varphi_i, \psi_i)$ і позитивним напрямом осі абсцис, γ_0 кут повороту системи координат. Поворотом системи координат на кут γ_0 можна досягти поєднання великої осі еліпса, апроксимуючого макросейсмічне поле точок, з орієнтацією цього поля. Кут між напрямом OM від епіцентру на точку M і віссю OX в старій системі координат XOY становить γ , а в новій системі X_1OY_1 дорівнює $\gamma - \gamma_0$. Нашим завданням є оптимальна апроксимація макросейсмічного поля землетрусів 10.11.1940, 7.04.1977, 31.08.1986, 30.05.1990 та 31.05.1990 рр., функцією загасання (5.51).

Оптимізація тісно пов'язана з моделюванням, тобто методи оптимізації застосовуються до математичних моделей, для визначення значень параметрів моделі або ж функцій, при яких цільова функція, при заданих обмеженнях, була б зведена до мінімуму.

Невідомі коефіцієнти в рівнянні загасання визначаються методом найменших квадратів (в дальнішому – МНК), який призводить до системи лінійних, щодо невідомих коефіцієнтів, рівнянь. Цим методом обчислюються такі значення параметрів, при яких сума квадратів відхилень спостережених значень інтенсивностей від їх теоретичних значень мінімальна:

$$S = \sum_{n=1}^N \sum_{k=1}^{K_n} (I_{nk} - am_n + b \log \sqrt{\Delta_{nk}^2 + h_n^2} - c)^2 \rightarrow \min \quad (5.53)$$

де I_{nk} – інтенсивність в k -й точці (φ_k, ψ_k) , $k=1, \dots, K_n$; при n -му землетрусі $n=1, \dots, N$;

Δ_{nk} – епіцентральна відстань, h_n – глибина n -го землетрусу, K_n – число точок макросейсмічного поля n -го землетрусу. Нерегулярний характер картини макросейсмічної поля (рис. 5.2-5.6) пояснюється неоднорідністю геологічного середовища і похибками оцінки інтенсивності. Ширина зон однакової бальності, для низьких значень інтенсивності, іноді становить кілька сотень кілометрів, і містити незначну кількість *IDP* точок. Тому, для підбору найкращої моделі поля інтенсивності сейсмічних впливів, необхідно спеціальним чином аналізувати макросейсмічне поле спостережених землетрусів: визначити мінімальну і максимальну епіцентральні відстані; визначити нижній поріг наглядності макросейсмічного поля інтенсивності. Макросейсмічне спостереження необхідно обробити так, щоб по можливості точно відобразити загальну тенденцію загасання в залежності від параметрів землетрусів, і згладити випадкові відхилення, пов'язані з неминучими похибками спостережень і визначення макросейсмічної інтенсивності. Стандартних критеріїв визначення мінімальних і максимальних епіцентральних відстаней для обробки макросейсмічних даних не існує. У нашому випадку найбільша епіцентральна відстань визначається з міркувань наглядності даних, тобто це найбільша довжина радіуса, поза яким знаходиться незначна кількість точок макросейсмічного поля. Для досліджуваних землетрусів значення R_{\max} рівні: 845; 805; 685; 385; 445 км, відповідно.

МНК є найкращим методом згладжування макросейсмічної поля, так як дозволяє порівняно простим математичним способом знаходити шукані значення коефіцієнтів рівняння (5.54) і відповідає ймовірнісній структурі відхилень теоретичних значень від спостережених балів. Якщо відхилення розподілені нормально, то оцінки, отримані, за методом максимальної правдоподібності та МНК збігаються. Принцип найменших квадратів може бути застосований і в разі, коли похибки не є випадковими величинами, а можуть мати різну природу.

Мінімум диференціала функції багатьох змінних $S=f(a, b_{\max}, b_{\min}, \gamma_0, c)$ можна визначити, досліджуючи її значення в критичних точках, які визначаються з рішення системи диференціальних рівнянь [Турчак, 1987; Бахвалов, 1973]:

$$\frac{dS}{da} = 0; \frac{dS}{db_{\max}} = 0; \frac{dS}{db_{\min}} = 0; \frac{dS}{dc} = 0; \frac{dS}{d\gamma_0} = 0 \quad (5.54)$$

Цільова функція (5.53) є диференційованою, але в виразі функції (5.53) числові параметри b_{\max} , b_{\min} , γ_0 входять нелінійно. У такому випадку рішення системи нелінійних рівнянь (5.54) може виявитися складним і трудомістким [Sauer, 2003]. Крім того, може визначитися точка локального екстремуму цільової функції, а не глобального. Майже всі методи вирішення оптимізаційних задач мають цей недолік. Його можна усунути, якщо вдається обґрунтувати вибір початкового наближення, що знаходиться поблизу глобального екстремуму. Оптимальні значення для коефіцієнтів пропонується знаходити за допомогою методу сіток, перебираючи всі комбінації значень коефіцієнтів на сітці значень [Бирюков и др., 2010; Gerdt, 2013; Gnüchtel, 2012; Pöschko, 2009], тому що інтервали значень коефіцієнтів рівняння (5.51) обмежені. Критерієм порівняння при цьому виступає величина сума квадратів відхилень (5.53). Розбиттям діапазону можливих значень коефіцієнтів на рівні інтервали отримуємо сітку з певним кроком. Тобто, інтервали $[a_i; b_i]$, $i=1, \dots, 5$; що визначають область пошуку екстремуму функції I (5.53), діляться на рівні частини довжиною $h_i = (b_i - a_i) / n_i$, $i=1, 2, \dots, 5$. У всіх точках $(a_1 + j_1 \cdot h_1, a_2 + j_2 \cdot h_2, \dots, a_5 + j_5 \cdot h_5)$, $j_1=0, 1, \dots, n_1; \dots; j_5=0, 1, \dots, n_5$, обчислюються значення апроксимуючої функції, і вибирається точка $(a^*, b_{\max}^*, b_{\min}^*, \gamma_0^*, c^*)$, якій відповідає мінімальне значення суми квадратів відхилень (5.53). Точність оцінки точки мінімуму цільової функції (5.53) залежить від вибору значень n_i . При малих значеннях n_i можна втратити точку мінімуму між вузлами великої сітки, а при великих значеннях зростає час обчислень.

В даному випадку точки макросейсмічного поля відомі, і в цих точках за формулою (5.53), методом перебору всіх значень коефіцієнтів рівняння, обчислюються різниці модальних і реальних значень інтенсивності в балах. Застосування цього методу пов'язане з великим обсягом обчислень, однак в залежності від кроку дискретності, при заданні значень коефіцієнтів, визначається глобальний екстремум функції загасання в заданій області. Оптимальними, по мінімуму цільової функції (5.53), є коефіцієнти: $a=1.6$; $c=7.2$; найбільшого $b_{\max}=5.6$ і

найменшого $b_{\min}=4.9$ загасання інтенсивностей струшувань; кут між великою віссю еліпса загасання і позитивним напрямом осі абсцис $\gamma_0=51^\circ$. Отже, загасання інтенсивності сейсмічних впливів підкорових землетрусів зони Вранча визначиться рівнянням:

$$\bar{I} = 1.6M_w - \frac{b_{\max} b_{\min}}{\sqrt{b_{\min}^2 \cos^2(\gamma - 51^\circ) + b_{\max}^2 \sin^2(\gamma - 51^\circ)}} \lg \sqrt{h^2 + r^2} + 7.2 \quad (5.55)$$

Коефіцієнт детермінації $r^2=0.567$, і значення кореляції Спірмена $R_s=0.86$, характеризують високий ступінь відповідності спостережених макросейсмічних даних апроксимуючої функції загасання інтенсивності сейсмічних впливів. Теоретичні значення інтенсивності в балах пояснює 56.7% макросейсмічних спостережень. Якість лінійної регресії (5.51) теоретичних оцінок інтенсивності на спостережених значеннях пунктів-балів становить 12.47%, що також вказує на значущість регресії.

Оцінка параметрів затухання для корових землетрусів не проводилася, так як відсутні відомості про макросейсмічне поле румунських корових землетрусів. При аналізі сейсмічної небезпеки загасання інтенсивності корових землетрусів оцінювалося по середньосвітових значеннях коефіцієнтів рівняння макросейсмічного поля [Шебалин, 1961]:

$$\hat{I} = 1.5m - 3.5 \log \sqrt{\Delta^2 + h^2} + 3.0 \quad (5.56)$$

Як було сказано вище, апроксимація еліпсом макросейсмічного поля є спрощеною моделлю реальної картини, яка є мозаїчною, а лінії, що розділяють зони рівної бальності – хвилястими. Ці факти і труднощі, які виникли при вирішенні системи нелінійних рівнянь (5.54), спонукало на оцінку коефіцієнтів загасання за наступною схемою (таб. 5.4):

- початок координат поєднується з епіцентром досліджуваних землетрусів;
- у кожному дотичному азимутному створі ширини θ з кроком дискретності ω ($\Omega < \theta$), визначаються коефіцієнти рівняння загасання (5.51) за допомогою регресійного аналізу.

Таблиця 5.4

Значення коефіцієнтів функції загасання

	Ліва межа азимутального інтервалу	Права межа азимутального інтервалу	Коефіцієнти загасання		
			<i>a</i>	<i>b</i>	<i>c</i>
1	0	90	1.52	-4.74	6.79
2	90	180	1.63	-5.80	8.24
3	180	270	2.10	-6.94	8.07
4	270	360	1.41	-5.40	8.11

5.5. Дослідження неоднорідності геофізичного середовища за макросейсмічними даними

Інтенсивність сейсмічних впливів на деякій території визначається параметрами вогнища землетрусів, особливостями геологічної будови середовища на шляху поширення сейсмічних хвиль і ґрунтовими умовами під досліджуваною територією. Як відомо, загасання сейсмічних хвиль пов'язане з явищами поглинання, геометричної розбіжності і розсіювання енергії хвиль при поширенні їх в геологічному середовищі. У розглянутому нами випадку осередки вранчських землетрусів з глибиною 80–140 км розташовані неглибоко, і геологічні умови поширення сейсмічних хвиль розрізняються головним чином у верхніх шарах. Азимутна анізотропія, яка характерна для форми макросейсмічного поля землетрусів даного вогнища, обумовлена, в числі інших причин, наявністю численних геологічних зрушень, які формують тріщини в земній корі, а також механізмом прояву активності даного вогнища. Азимутна неоднорідність загасання інтенсивності сейсмічних коливань від землетрусів вогнища Вранча вже привертало увагу фахівців [Пустовитенко и др., 2002]. У даній роботі проведений дисперсійний аналіз макросейсмічного поля інтенсивностей останніх сильних землетрусів з метою виявлення характеру неоднорідності середовища.

Дослідження азимутальної анізотропії на основі макросейсмічних даних вже застосовувалося для території Італії [Pasolini, 2008]. Опубліковані також роботи по

дослідженню неоднорідності верхньої мантії Румунії та суміжних територій [Буртиєв, 1995; 2003]. У даній роботі дослідження неоднорідності зроблено за наступною схемою:

- проводяться концентричні кільця з центром в епіцентрі землетрусу (рис.10а). Обчислюється вибіркова середня арифметична величина η_i і середнє квадратичне відхилення σ_i спостереженої сейсмічної інтенсивності в точках, що знаходяться в межах кільця ρ_l , $l=1, \dots, L$, де L – число концентричних кілець координатної сітки.
- далі проводяться з центру промені, з азимутом β_k , $k=1, \dots, K$. геометричні осередки Δ_{lk} , утворені перетином кільця ρ_l і сектора β_k , матимуть форму криволінійних трикутників і трапецій (рис. 5.10b). Обчислюються числові характеристики η_{lk} і σ_{lk} інтенсивності в точках, що належать осередку, що є перетином l -го кільця і k -го азимутального променя.

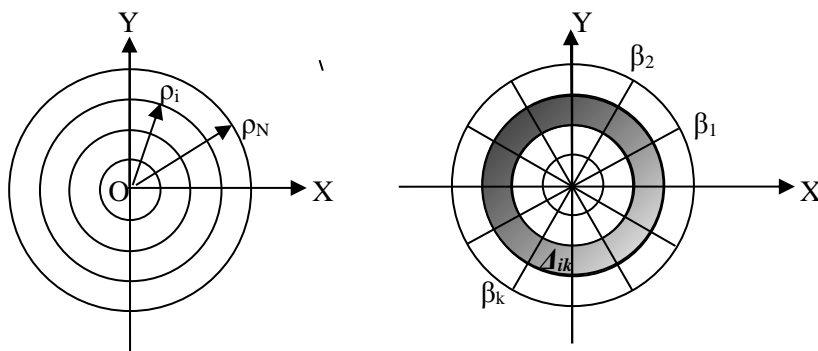


Рис. 5.10. а) Система концентричних кілець навколо епіцентру,
б) концентричні кільця і азимутальні сектори.

За макросейсмічними картками, складеним на основі даних реєстрацій землетрусів, видно, що ширина зон зміни інтенсивності в балах збільшується в міру віддалення від епіцентру землетрусу. Це факт повинен впливати на вибірку дисперсію точок IDP в осередках, розташованих в межах деякого азимутального нахилу на різній відстані від епіцентру. Тобто, розсіювання значень інтенсивності навколо середньоарифметичного значення в осередках повинно зменшуватися в міру віддалення від епіцентру. Передбачається, що значення інтенсивності в

осередках Δ_{lk} мають нормальний розподіл з параметрами (η_{lk}, σ_{lk}) , з постійною вибірковою дисперсією в межах кільця: $\sigma_{11} = \sigma_{12} = \dots = \sigma_{1K}$.

Опираючись на макросейсмічні дані, перевіряється нульова гіпотеза H_0 про рівність середніх величин інтенсивностей у всіх осередках кільця $H_0: \eta_{11} = \eta_{12} = \dots = \eta_{1K}$, при конкуруючій гіпотезі $H_1: \eta_{1i} \neq \eta_{1j}$ – про розбіжності хоча б при одному $i \neq j$, $i, j = 1, \dots, K$. Правдоподібність нульової гіпотези означає, що немає значимої розбіжності між вибірковими середніми значеннями інтенсивності в осередках, однаково віддалених від центру і розташованих від епіцентру під різними азимутними кутами. Якщо гіпотеза вірна для всіх епіцентральної відстаней (для всіх концентричних кілець), то це буде свідчити про однорідність геологічного середовища. Дисперсійний аналіз дозволяє оцінити ступінь і достовірність відмінності декількох вибірових середніх одночасно, тобто вивчити вплив одного контрольованого фактора на результат шляхом оцінки його відносної ролі в загальній мінливості цієї ознаки, викликаній впливом всіх факторів. Нашим завданням є аналіз залежності статистичних характеристик інтенсивності струшувань від азимута *IDP* точок.

Під фактором розуміється азимут β_k , а під рівнем ознаки значення інтенсивності I_{lkj} точок в осередках Δ_{lk} . У методі дисперсійного аналізу загальна сума квадратів відхилень від загальної середньої розкладається на суму квадратів відхилень групових середніх від загальної середньої і на суму квадратів відхилень спостережень від групових середніх [Феллер, 1967; Новицкий, Зограф, 1991].

У разі справедливості припущення про рівність дисперсій спостережень в осередках, перевірка правдоподібності гіпотези H_0 , проводиться за критерієм Фішера [Шеффе, 1980]:

$$F_s = \frac{S_1 / (K - 1)}{S_2 / (N_l - K)}; \quad S_1 = \sum_{l=1}^L n_l (\bar{I}_l - \bar{I})^2; \quad S_2 = \sum_{l=1}^L \sum_{r=1}^{n_l} (I_{rl} - \bar{I}_l)^2 \quad (5.57)$$

Середнє значення інтенсивності точок в деякому кільці l , і в деякій комірці k кільця l , визначається за формулами:

$$\bar{I}_l = \frac{1}{N_l} \sum_{j=1}^{N_l} I_{lj}; \quad \bar{I}_{lk} = \frac{1}{N_{lk}} \sum_{j=1}^{N_{lk}} I_{lkj} \quad (5.58)$$

K – число осередків в концентричних кільцях, L – число кіл, N_l – число точок в кільці l , N_{lk} – число точок в осередку Δ_{lk} , а $N_{1l} + N_{2l} + \dots + N_{Kl} = N_l$ визначає число точок в кільці з номером l , $N_1 + N_2 + \dots + N_L = N$ – загальне число спостережень.

Для практичних застосувань тестового критерію χ^2 рекомендується визначити розмір осередків так, щоб число спостережень точок IDP , що потрапили в осередок, було не менше 10 [Афифи, Ейзен, 1982; Hillston, 2009].

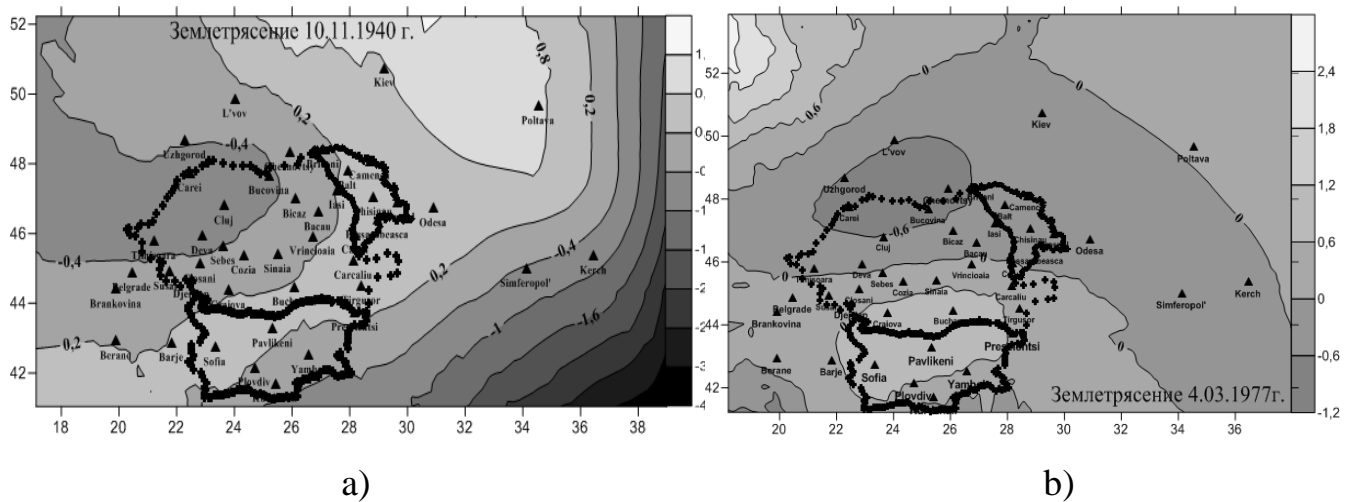


Рис. 5.11. Карта азимутальних неоднорідностей по макросейсмічними даним землетрусів: а) 10.11.1940; б) 4.03.1977

Нульова гіпотеза про рівність середніх арифметичних значень точок IDP у всіх осередках концентричних кілець для макросейсмічних даних землетрусу, що стався в 1940 р, не підтвердилася. Дана гіпо $\Delta I_{rl} = I_{rl} - I_l$ илася лише в деяких випадках, для окремих кілець. Однак, для більшої частини досліджуваної території припущення про рівність середніх значень, відповідно припущення про однорідність середовища, не підтвердилося. Як правило, вибіркова статистика (5.60) перевищує табличне значення F -розподілу, що відповідає рівню значущості α 0.05. Для дослідження однорідності середовища будувалися концентричні кільця шириною в 20, 40 і 60 км, а промені, що йдуть від епіцентру, проводилися через кожні 20 градусів (рис. 5.10.б). Далі будувалися карти відмінностей бальності в осередках – різниці значень балів в осередках з їх арифметичним середнім (рис. 5.11–5.15):

(5.59)

Легенда карт, представлених на малюнках 5.11–5.13, наступна:

- зліва і внизу вказані географічні широти і довготи з кроком в 2 градуси, а праворуч показана чисельна і колірна шкала відмінностей
- на изолініях (лініях рівного значення відмінностей) вказані усереднені значення відмінностей в розрахункових осередках
- сейсмічні станції, із зазначенням їх назв, позначені символом (▲).
- державні кордони позначені символом (+).

Передбачається, що відмінності мають нормальний розподіл з невідомим математичним очікуванням μ і дисперсією σ^2 . Довірчий інтервал із заданою надійністю α для параметра μ , при невідомій дисперсії σ^2 визначається за формулою [Худсон, 1970]:

$$\gamma = P\left(\bar{x} - t_{\alpha} \frac{S}{\sqrt{n-1}} < \mu < \bar{x} + t_{\alpha} \frac{S}{\sqrt{n-1}}\right) \quad (5.60)$$

де \bar{x} - вибіркове середнє арифметичне балів і S – вибіркове середнє квадратичне відхилення. Середньоарифметичне значення точок *IDP* в деяких осередках виявилось в межах 95% -них довірчих інтервалів.

У тих випадках, коли величина відмінності знаходиться в межах відповідного довірчого інтервалу, слід вважати правдоподібним припущення про рівність нулю відмінностей i , відповідно, про однорідність геологічного середовища. Внаслідок цього, як видно з таблиці (таб. 5.5) і як показано на рисунках (рис. 5.11–5.13), частина досліджуваної території слід вважати однорідною. У таблиці (таб. 5.5) наведені довірчі інтервали для всіх проаналізованих землетрусів.

Стандартні розміри шаблонів (однакова ширина кілець і азимутальних секторів) дозволяє порівнювати вплив локальних геологічних умов на затухання інтенсивності струшувань при різних землетрусах.

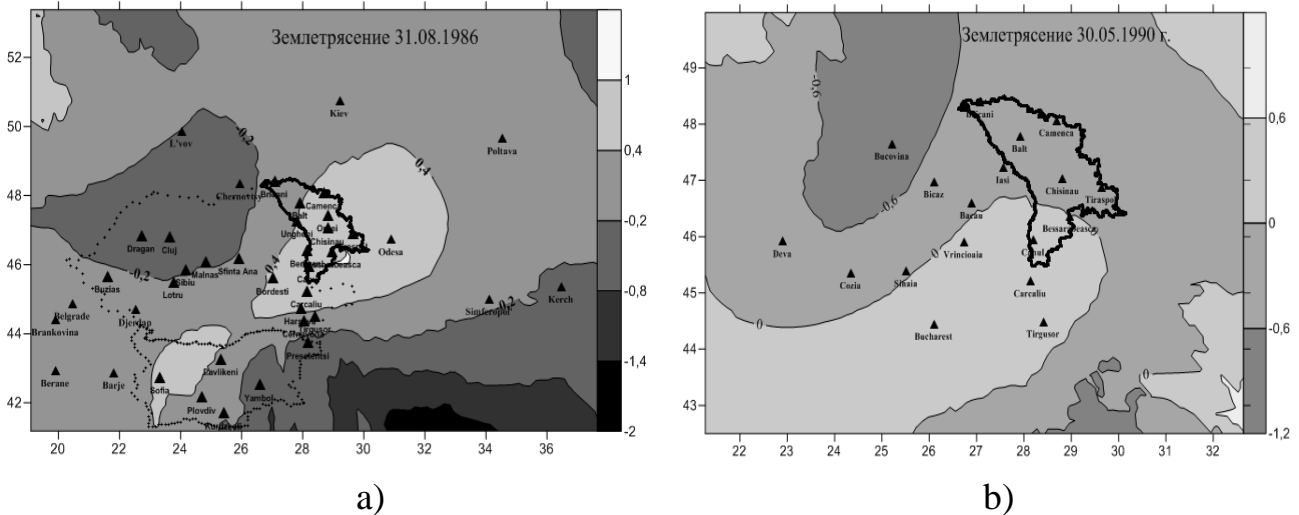


Рис. 5.12. Карта азимутальних неоднорідностей по макросейсмічними даним землетрусів: а) 31.08.1986; б) 30.05.1990.

Якби причиною загасання сейсмічних хвиль було поглинання, геометрична розбіжність і розсіювання енергії хвиль при поширенні їх в горизонтально шаруватому однорідному геологічній середовищі, то дисперсії в осередках відрізнялися б слабо. Тому в якості робочої (нульової) гіпотези взято припущення про рівність дисперсій в осередках, тобто H_0 :

$$\sigma_{11}^2 = \sigma_{12}^2 = \sigma_{21}^2 \dots = \sigma_{LK}^2 \quad (5.61)$$

Нехай n_{kl} , $k=1,2,\dots,K$; $l=1,2,\dots,L$ число точок в осередках Δ_{lk} . Незміщена вибіркова дисперсія спостережень осередку Δ_{lk} обчислюється за формулою:

$$\widehat{S}_{lk}^2 = \frac{1}{n_{lk} - 1} \sum_{j=1}^{n_{lk}} (I_{lkj} - \bar{I}_{lk})^2 \quad (3.62)$$

Для перевірки гіпотези H_0 про однорідність (рівність) декількох дисперсій, отриманих за вибірками різного об'єму, використовується критерій Бартлетта, для цього розраховується середня арифметична незміщена оцінка математичного очікування:

$$\bar{I}_{lk} = \frac{1}{n_{lk}} \sum_{j=1}^{n_{lk}} I_{lkj} \quad (5.63)$$

Таблиця 5.5

Верхні (В) і нижні (Н) границі 95%-них довірчих інтервалів для відмінностей інтенсивності

Відст. (км)	10.11.1940.		4.03.1977.		31.08.1986.		30.05.1990.		31.05.1990.	
	Н	В	Н	В	Н	В	Н	В	Н	В
60			-1.229	1.229	-0.041	0.041				
120			-0.029	0.029	-0.217	0.217			-0.056	0.056
180	-0.604	0.604	-0.037	0.037	-0.229	0.229			-0.045	0.045
240	-0.242	0.242	-0.048	0.048	-0.183	0.183			-0.194	0.194
300	-0.252	0.252	-0.079	0.079	-0.171	0.171	-0.036	0.036	-0.151	0.151
360	-0.351	0.351	-0.098	0.098	-0.207	0.207	-0.035	0.035	-0.347	0.347
420	-0.214	0.214	-0.068	0.068	-0.140	0.140	-0.058	0.058	-0.343	0.343
480	-0.088	0.088	-0.057	0.057	-0.129	0.129	-0.053	0.053	-0.506	0.506
540	-0.142	0.142	-0.043	0.043	-0.030	0.030	-0.054	0.054	-0.337	0.337
600	-0.116	0.116	-0.037	0.037	-0.038	0.038	-0.053	0.053	-0.179	0.179
660	-0.080	0.080	-0.039	0.039	-0.046	0.046	-0.051	0.051	-0.154	0.154
720	-0.103	0.103	-0.020	0.020	-0.089	0.089			-0.171	0.171
780	-0.620	0.620	-0.022	0.022	-0.125	0.125			-0.146	0.146
840	-1.028	1.028	-0.021	0.021	-0.136	0.136				
900			-0.081	0.081	-0.110	0.110				

і середня незміщена оцінка дисперсій, зважена по числу ступенів свободи:

$$\hat{S}_{ave}^2 = \frac{\sum_{k=1}^K n_{lk} \hat{S}_{lk}^2}{\sum_{k=1}^K n_{lk}} \quad (5.64)$$

Критерій Бартлетта для перевірки нашої гіпотези має вигляд [Мхитарян и др., 2008]:

$$\chi_q^2 = \frac{n_l \ln \hat{S}_{ave}^2 - \sum_{k=1}^K n_{lk} \ln \hat{S}_{lk}^2}{1 + \frac{1}{3(l-1)} \left[\sum_{k=1}^K \frac{1}{n_{lk}} - \frac{1}{n_l} \right]} \quad (5.65)$$

При підтвердженні нульової гіпотези H_0 і при $n_{lk} > 3$ вибіркова статистика (5.65) має розподіл χ^2 з $n_l - 1$ ступенями свободи [Худсон, 1970]. Критерій Бартлетта застосовувався для перевірки нульової гіпотези про рівність дисперсій в концентричних кільцях, проведених з радіусами кратними 20, 40 і 60 км, тобто H_0 :

$$\sigma_1^2 = \sigma_2^2 = \dots = \sigma_L^2 \quad (5.66)$$

Робоча гіпотеза H_0 повинна бути відкинута для макросейсмічних даних всіх землетрусів. вибіркові статистики χ_q критерію (5.65) 99.7; 65.7; 53.1; 46.4; 72.6 для точок макросейсмічного поля від всіх землетрусів виявилися більше порогового значення χ^2 – розподіленої випадкової величини, яка відповідає рівню значущості $\alpha=0.05$.

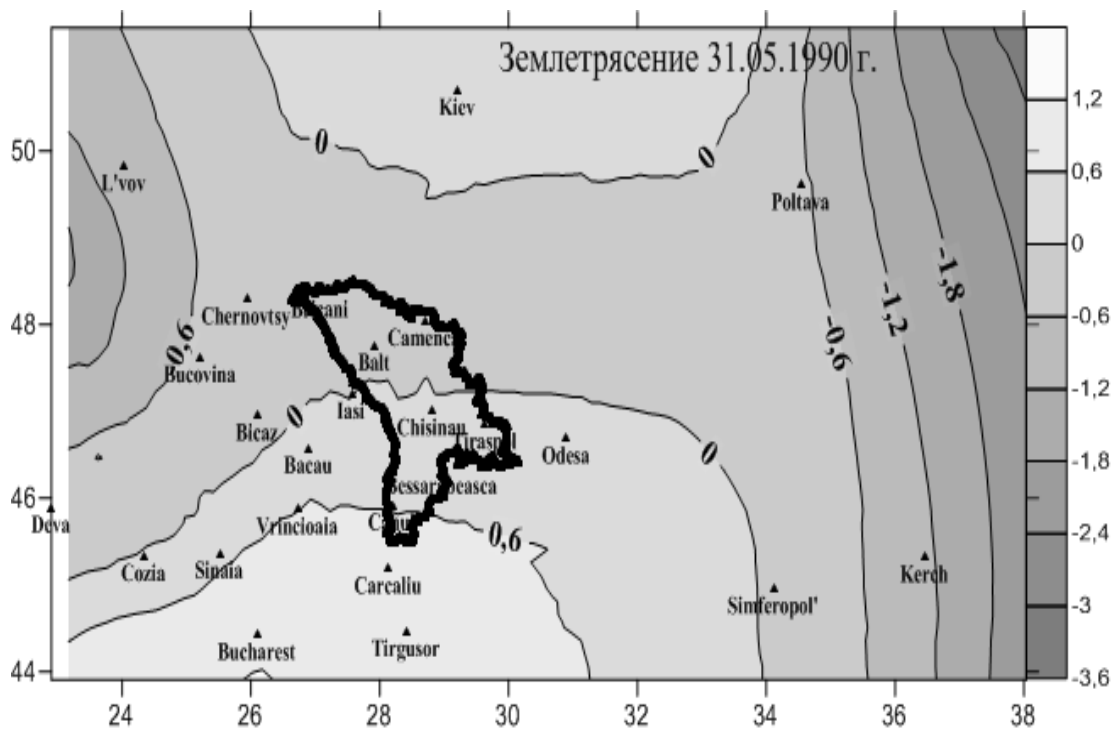


Рис. 5.13. Карта азимутальних неоднорідностей за макросейсмічними даними землетрусу 31.05.1990 р.

Таким чином, гіпотеза про рівність середніх арифметичних і дисперсій спостережених значень інтенсивності в балах за всіма шаблонами відкидається, тобто геофізичне середовище, в якому поширюються сейсмічні коливання, є азимутально неоднорідне.

Висновки до розділу 5

Проведено статистичний аналіз макросейсмічних даних землетрусів, що сталися 10.11.1940, 7.04.1977, 31.08.1986, 30.05.1990 та 31.05.1990 рр. Дослідження показують, що прийнятною згладжуючою лінією макросейсмічного поля є еліпс. З канонічного рівняння еліпса слідує, що фокальна вісь 5, 6 і 7 бальних зон, всіх

досліджуваних землетрусів, спрямована по меридіану. Це означає, що 5, 6 і 7 бальні зони витягнуті в напрямку північ-південь, тоді як інші зони витягнуті уздовж паралелі, що свідчить про неоднорідність геологічного середовища, де поширювалися сейсмічні хвилі від досліджуваних землетрусів. Апроксимація еліпсом багатьох точок рівної інтенсивності проведена за методом Фітцгиббон. Визначено параметри еліпсів, що апроксимують різні зони рівної бальності макросейсмічного поля. Коефіцієнт лінійної регресії епіцентральної інтенсивності I_0 на магнітуду, для Румунських землетрусів, вийшов рівним $b = 1.5$. Оптимальними значеннями коефіцієнтів рівняння макросейсмічного поля виявилися: $a = 1.6$; $c = 7.2$; найбільшого $b_{\max} = 5.6$ і найменшого $b_{\min} = 4.9$ загасання інтенсивності струшувань; а кут між великою віссю еліпса загасання і позитивним напрямом вісі абсцис рівним $\gamma_0 = 51^\circ$. Загасання сейсмічних впливів оцінюється на основі вранчських землетрусів.

На основі макросейсмічних даних виявлена азимутальна неоднорідність геологічних структур Молдови і суміжних територій. Застосований в даній роботі алгоритм, при використанні рівновеликих шаблонів, дозволяє порівняти карти неоднорідності, побудовані за макросейсмічними даними різних землетрусів. Гіпотеза про рівність середніх арифметичних і дисперсій спостережених значень інтенсивності в балах в межах кожного окремого концентричного кільця відкидається, що свідчить про азимутальні неоднорідності геофізичного середовища, в якому поширюються сейсмічні коливання.

Список використаних джерел

- Алказ В.Г. Основы прогноза сейсмической опасности и сейсмического риска территории Республики Молдова. Кишинев, 2007, 229.
- Афифи А., Ейзен С., Статистический анализ. Мир, Москва, 1982, 487.
- Бахвалов Н.С., Численные методы. Москва, «Наука», 1973, 631.
- Бирюков Р.С., Городецкий С.Ю., Григорьева С.А., Павлючонок З.Г., Савельев В.П. Методы оптимизации в примерах и задачах. Учебно-методическое пособие. Нижний Новгород, 2010, 101.

- Буртиев Р.З. Исследование горизонтальных неоднородностей верхней мантии Румынии и Молдовы *Academiei de Stiinte a Republicii Moldova. Buletinul Fizica și Technica*, 1995, №1, 76–82.
- Буртиев Р.З. Скоростное строение верхней мантии Румынии и сопредельной территории. *Analele ATIC. Vol. I(II)*. Chisinau, 2003, 102–109.
- Джанузаков К.Д. Региональные особенности затухания интенсивности сотрясений сильных землетрясений Кыргызстана и прилегающих к нему районов. *Вестник Института сейсмологии НАН КР* No. 2. Бишкек. 2013, 11–16.
- Вылегжанин О.Н., Рыбалка С. А., Приближение множества точечных наблюдений эллипсом. *Известия Томского политехнического Университета*. Т. 309. № 7, 2006, 14.
- Глаголев А.А., Солнцева Т.В., Курс высшей математики. Москва, «Высшая школа», 1971, 653.
- Гусев А. А., Шумилина Л. С., Моделирование связи балл–магнитуда– расстояние на основе представления о некогерентном протяженном очаге, *Вулканология и Сейсмология* No. 4–5, 1999, 29–40.
- Королев В.А, Лазаренко М.А. Параметры уравнения поля сейсмической интенсивности сильных подкоровых землетрясений зоны Вранча для территории Украины. *Геофизический журнал*. №1 Киев, 2003, 27–33.
- Ланкастер П. Теория матриц. Наука, Москва, 1978. 269.
- Левин Б.В., Ким Ч.У., Соловьев В.Н., Оценка сейсмической опасности и результаты детального сейсмического районирования для городов о. Сахалин. *Тихоокеанская Геология*, том. 31, №5, 2012, 93–103.
- Маркович Э. С., Курс высшей математики. Росвузиздат, 1963, 407.
- Новицкий П.В., Зограф И.А. Оценка погрешностей результатов измерений. Л.: Энергоатомиздат, 1991, 303.
- Пустовитенко Б.Г., Кульчицкий В.Е., Капитанова С.А. О затухании сейсмической энергии землетрясений зоны Вранча // *Сейсмологический бюллетень Украины* за 2001 год. Севастополь: Изд. НПЦ «ЭКОСИ– Гидрофизика», 2002, с. 80–89.

- Раутиан Т.Г. и др. Макросейсмическое проявление землетрясений и очаговые спектры. Вопросы инженерной сейсмологии; вып. 23. Москва. Наука. 1982. 74–82.
- Ризниченко Ю.В. Энергетика макросеймики. Изв. АН СССР, Физика Земли, Москва, No. 8, 1974. 3–10.
- Смирнов В.И., Курс высшей математики. Т. 1. Наука, Москва, 1974, 479.
- Турчак Л. И., Основы численных методов. Москва, «Наука», 1987, 318.
- Худсон Д., Статистика для физиков, Мир, Москва, 1970, 295
- Шебалин Н.В. Бальность, магнитуда и глубина очага землетрясений. Землетрясения в СССР. Москва. Наука. 1961, 126–138.
- Шебалин Н.В. Методы использования инженерно–сейсмологических данных в сейсмическом районировании. В кн.: Сейсмическое районирование СССР. Ч.1, гл. 6. Москва. Наука. 1968, 54–69.
- Шебалин Н.В. Сильные землетрясения. Избранные труды. Москва, 1997, 542.
- Шеффе Г. Дисперсионный анализ. Наука, Москва, 1980, 511.
- Феллер В., Введению в теорию вероятностей и ее приложения. Мир, Москва, 1967, 498.
- Agin, Gerald Jacob. Fiting ellipses and general second–order curves. Robotics Institute, 1981. <http://repository.cmu.edu/robotics>, 487.
- Alvarez-Rubio Sonia, Kästli Philipp, Donat Fäh, Sellami Souad. Calibration of historical earthquakes for the earthquake catalogue of Switzerland. ECOS–Earthquake Catalogue of Switzerland (ECOS-09), March 2010, 31.
- Ambraseys N., Intensity attenuation and magnitude–intensity relationships for northwest European earthquakes, Earthquake Eng. Struct. Dyn., 13, 1985, 733–778.
- Bakun H. William, Scotti Oona. Regional intensity attenuation models for France and the estimation of magnitude and location of historical earthquakes. Geophys. J. Int. 2006, 596–610.

- Bune V.I., (1979), The Methodology for Mapping Regional Seismic Zones, Meeting of the Working Group 4.3 DAPG, Bratislava, 1979, 95–103.
- Burtiev Rashid. Fitting an Ellipse to the set of intensity data points of Vrancea earthquakes. *Journal of Engineering Studies and Research – Volume 20 (2014) No. 4*, Bacau, Romania, 2014, 15– 24.
- Cecic Ina, Musson Roger. *Macroseismic Surveys in Theory and Practice*. *Natural Hazards* 31: 2004, 39–61.
- Constantinos B. Papazachos. Anisotropic Radiation Modelling of Macroseismic Intensities for Estimation of the Attenuation Structure of the Upper Crust in Greece. *PAGEOPH*, Vol. 138, No. 3, Birkhäuser Verlag, Basel, 1992, 1–25.
- Cua G., Wald D. J., Allen T. I., Garcia D., Worden C. B., Gerstenberger M., Lin K., Marano K., “Best Practices” for Using Macroseismic Intensity and Ground Motion Intensity Conversion Equations for Hazard and Loss Models in GEM1. *GEM Technical Report 2010*, 67.
- Chui–Tui Lee, Bow–Shan Hsieh, Chih–Hsuan Sung and Po–Shen Lin. Regional Arias Intensity Attenuation Relationship for Taiwan Considering V_{S30} . *Bulletin of the Seismological Society of America*, Vol. 102, No .1, 2006, 129–142.
- Dörfler Willy und Wieners Christian, *Mathematik beschreibt die Welt*. Universität Karlsruhe, Institut für Angewandte und Numerische Mathematik, 2009, 25.
- Eberly David. *Distance from a Point to an Ellipse, an Ellipsoid, or a Hyperellipsoid*. Geometric Tools, LLC. Academic Press, Inc., New York NY, 2004. 28.
- Fitzgibbon, A. W.: Set of MATLAB files for ellipse fitting. Dept. of Artificial Intelligence, The University of Edingburgh, <ftp://ftp.dai.ed.ac.uk/pub/vision/src/demofit.tar.gz>, September 1995. 17.
- Fitzgibbon W. Andrew, Pilu Maurizio, and Fisher B. Robert, Direct Least Squares Fitting of Ellipses. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 21(5), 1999, 476–480.

- Gerdts Matthias. Einführung in die lineare und nichtlineare Optimierung. Institut für Mathematik und Rechneranwendung. Universität der Bundeswehr München, München, 2013, 107.
- Gnüchtel Stefan. Methoden der Optimierung von Energiesystemen. Institut für Energietechnik, Dresden, 2012, 79.
- Gomez C, Augusto A. Seismic Hazard Map for the Italian Territory Using Macroseismic Data. *Earth Sci. Res. J.* [online]. Vol.10, No.2, 2006, 67–90.
- Gupta, I. N., and Nuttli O. W., Spatial attenuation of intensities for Central U.S. earthquakes, *Bull. Seism. Soc. Am.* 66, No. 3, 1976, 743–751.
- Halır Radim, Flusser Jan, Numerically stable direct least squares fitting of ellipses. Prague, Czech Republic, 2008, 17.
- Harkera Matthew, O’Leary Paul, Zsombor–Murray Paul. Direct type–specific conic fitting and eigenvalue bias correction. *Image and Vision Computing* 26, 2008, 372– 381.
- Hillston Jane. <http://www.inf.ed.ac.uk/teaching/courses/pm/PM–lecture1.pdf>, 2009.
- Hosh G. K., Mahajan A. K., Interpretation of Intensity Attenuation Relation of 1905 Kangra Earthquake with Epicentral Distance and Magnitude in the Northwest Himalayan Region. *Journal Geological Society of India*, Vol.77, 2011, 511–520.
- Howell B.F. Jr., Schultz T.R., Attenuation of Modified Mercalli Intensity with Distance from the Epicenter, *Bull. Seism. Soc. Am.*, Vol. 65, No. 3, 1975, 651–665.
- <http://mathworld.wolfram.com/Ellipse.html>
- Ismail-Zadeh, A., Sokolov, V. and Bonjer, K., Tectonic stress, seismicity, and seismic hazard in the southeastern Carpathians, *Nat. Hazards*, Springer, 2007, 1–22.
- Kronrod Tatiana, Panza Giuliano, Radulian Mirchea: Integrated transnational macroseismic data set for the strongest earthquakes of Vrancea (Romania). <http://issues.pangaea.de/browse/PDI–3917>, *Tectonophysics*, 2012, 271–282
- Lapajne Janez. Generalization of the Kövesligethy equation for non–circular macroseismic fields. *Geologija* 31, 32, Ljubljana, 1988/89, 577–580.
- Niehoff Christian. Verfolgung von Partikelbewegungen und Robuste Schätzung von Mittelwerten in Konfokal–Mikroskopischen Messdaten. Julich, 2006, 83.

- Oncescu C. Mihnea, Marza I. Vasile, Rizescu Mihaela. The Romanian earthquake catalogue between 984–1997. Vrancea earthquakes: Tectonics, Hazard and Risk Mitigation. Kluwer Academic Publishers, 1997, 43–47.
- Papoulia J.E., Stavrakakis G.N., Attenuation Laws and Seismic Hazard Assessment, Kluwer Academic Publishers. Natural Hazards. Netherlands. 1990, 49–58.
- Pasolini Chiara. The attenuation of seismic intensity. Settore scientifico disciplinare: Geo/10 Bologna, Marzo 2008, 96.
- Peruzza Laura. Attenuating Intensities. Annali di Geofisica. Vol.XXXIX, No. 5, 1996, 1–15.
- Pöschko Jan (in German). Mathematische Optimierung. TU Graz, 2009, 108.
- Rodler Fee–Alexandra. Bestimmung von Erdbebenparametern aus makroseismischen Daten von Österreich. Universität Wien. Wien, 2012, 87.
- Sauer Tomas (in German). Einführung in die Optimierung für Hörer aller Fachbereiche. Lehrstuhl für Numerische Mathematik. Justus–Liebig–Universität Gießen, 2003, 97.
- Sørensen M., Stromeyer D., Grünthal G., Attenuation of Macroseismic Intensity: A New Relation for the Marmara Sea Region, Northwest Turkey. Bulletin of the Seismological Society of America, Vol. 99, No. 2A, 2009, 538 – 553.
- Sponheuer W., Methoden zur Herdtiefenbestimmung in der Makroseismik Freiburger Forschungshefte C88. Berlin: Akademie Verlag. 1960. 117.
- Tselentis G.A., Danciu L., Probabilistic seismic hazard assessment in Greece – Part 1: Engineering ground motion parameters. Nat. Hazards Earth Syst. Sci., 10, 2010, 25–39.
- Van Loan F. Charles, Using the Ellipse to Fit and Enclose Data Points. Department of Computer Science Cornell University. 54.

РОЗДІЛ 6

ЧАСОВІ РЯДИ У ДОСЛІДЖЕННІ СЕЙСМІЧНОГО РЕЖИМУ

Спостереження за активністю сейсмічних зон показують, що тимчасові характеристики сейсмічності не є стаціонарними. Періоди варіацій кількості землетрусів (місяці, роки) розкидані хаотично на вісі часу. За графіком щомісячного числа землетрусів вогнища Вранча, не можна впевнено говорити про закономірності в тривалості періодів варіацій числа землетрусів і в чергуванні періодів сейсмічного затишшя з періодом високої сейсмічної активності. Стимулом для проведення даного дослідження було бажання аналізувати структуру ряду формальними методами для пошуку статистичної закономірності в варіаціях параметрів сейсмічності в часі.

Сейсмічність характеризується:

- частотою виникнення землетрусів
- статистичним розподілом сили поштовхів (магнітуди).
- просторовим розподілом осередків.
- макросейсмічними спостереженнями сильних сейсмічних подій

Для дослідження сейсмічності можна застосувати просторові моделі та моделі часових рядів. Просторова модель описує сукупність параметрів сейсмічності в даний момент часу.

Часовий ряд – це ряд регулярних спостережень над деякими параметром сейсмічності в послідовні моменти часу або на проміжках часу. У даній роботі, модель часового ряду застосовується для дослідження ймовірнісної структури потоку сейсмічних подій. Загалом, метою дослідження часового ряду є виявлення закономірностей в зміні рівнів ряду і побудова його моделі з метою прогнозування та дослідження взаємозв'язків між явищами.

Теорія часових рядів використовується для вирішення наступних основних завдань: визначення природи ряду; визначення основних параметрів ряду; передбачення майбутніх значень (щомісячного числа землетрусів) часового ряду

за наявними даними. Часові ряди можуть містити такі компоненти: тренд; циклічну компоненту; сезонну компоненту; випадкову складову. Часовий ряд складається з детермінованої і випадкової частин. Детермінована частина: тренд; циклічна і сезонна компоненти – використовуються для передбачення майбутніх значень ряду.

Для правильного відображення часовим рядом реального процесу вибірка даних повинна задовольняти умовам: порівнянності; однорідності; стійкості і достатності обсягу даних.

6.1. Часовий ряд числа місячних землетрусів

Теорія часових рядів може бути корисною в дослідженні тимчасового режиму сейсмічності, вказати на періоди сейсмічного затишшя, також на періоди помірної і підвищеної активності землетрусів. Трендом називають не випадкову функцію, що формується під дією загальних або довгострокових тенденцій. Циклічною компонентою також є не випадкова функція, обумовлена дією довготривалих циклів різної природи. Сезонна компонента використовується для позначення не випадковою функції. Ця функція формується на основі періодично повторюваних в певну пору року коливань досліджуваного ряду. У нашому випадку, передбачається, що форма взаємодії перерахованих компонент може бути адитивною (6.1), мультиплікативною (6.2) і змішаною (6.3, 6.4) [Андерсон, 1971; Бриллинджер, 1981; Кендал, 1981; Хеннан, 1974]:

$$x_t = m_t + c_t + s_t + \varepsilon_t \quad (6.1)$$

$$x_t = m_t c_t s_t \varepsilon_t \quad (6.2)$$

$$x_t = m_t c_t s_t + \varepsilon_t \quad (6.3)$$

$$x_t = m_t c_t + s_t + \varepsilon_t \quad (6.4)$$

m_t – тренд, стійка довгострокова тенденція

c_t – циклічна складова, досить довгі нерегулярні коливання, з періодом більше одного року

s_t – сезонна складова, досить регулярні періодичні коливання, що відбуваються в інтервалі часу, що не перевищує одного року

ε_t – випадкова складова (похибка).

Ряд $g_t = m_t + c_t$, що складається з суми тренда і циклічної компоненти називається тренд – циклічної компоненти (кон'юнктурним циклом). Наявність тренда означає присутність довготривалої систематичної складової в тимчасовому ряді, яка описує основну тенденцію динаміки часового ряду. Тренд як гладка крива виражає вплив довготривалих неперіодичних чинників. Формально під трендом розуміється послідовність умовних середніх значень, обумовлених тимчасовою змінною. Періодичні коливання можна поділяти на сезонні коливання, у яких період не перевищує одного року, викликані кліматичними умовами і циклічні з періодом коливань кілька років. Періодичність означає $x_{t+kp} = x_t$; $k = 1, 2, \dots$, і є сезонною при $p = 1$ рік, циклічною при $p = 2, 3, \dots$ років.

Випадковий шум ускладнює виявлення регулярних компонент. Зазвичай методи дослідження часових рядів включають різні алгоритми фільтрації шуму, що дозволяють більш чітко визначити регулярну складову. Головною метою дослідження часових рядів є прогноз майбутніх значень за допомогою детермінованої компоненти часового ряду. Аналіз випадкової компоненти проводиться для визначення прогнозованого потенціалу моделі, який ґрунтується на статистичному аналізі різниці між модельними і спостереженими значеннями.

6.2. Аналіз даних

Аналізу підлягає ряд щомісячного числа землетрусів у вогнищі Вранча за час з 1978 по 2011 рр. Рівнями ряду є щомісячне число землетрусів. Для аналізу динаміки процесу важливо, щоб часовий ряд мав достатню довжину (з урахуванням конкретних цілей дослідження). Наприклад, при вивченні періодичних коливань бажано мати інформацію не менше ніж за три повних періоди коливань. На практиці 50 значень за період спостережень вважається достатнім [Бокс, Дженкінс, 1974].

При аналізі сезонних коливань на базі рядів місячної або квартальної динаміки бажано мати інформацію, не менше ніж за 3 роки. Використання ж більш тонкого статистичного апарату для дослідження періодичності вимагає більшого об'єму інформації – не менше п'яти повних періодів коливань. Наприклад, для використання регресійного аналізу необхідно мати тимчасові ряди, довжина яких в кілька разів перевершує кількість незалежних змінних. У тимчасових рядах не повинні міститися пропущені рівні. В цьому випадку, для використання, даного часового ряду в подальшому аналізі, необхідно відновити пропущені рівні одним з відомих способів відновлення пропусків. Рівні рядів динаміки можуть містити аномальні значення або "викиди".

На етапі первинної обробки даних проводиться виявлення, вилучення "викидів" і заміна їх істинними або розрахунковими значеннями. Відповідність вибірки зазначеним вимогам проводиться статистичними процедурами пакету SPSS-20 на етапі попереднього аналізу часового ряду.

Статистичний пакет SPSS має окремий модуль, в якому представлено повне математичне забезпечення і набір статистичних процедур по дослідженню часових рядів: описові методи; методи згладжування (просте дотичне середнє, Брауна, Холта і ін.).

Мультиплікативними і адитивними методами розкладання часових рядів на умовні складові є: спектральний аналіз; регресійний аналіз і методи прогнозування. Визначення ймовірнісних характеристик випадкового процесу є складною процедурою і проводиться поетапно: досліджується стаціонарність процесу; обчислюються статистичні характеристики; усереднюються отримані оцінки; досліджується ергодичність процесу. Властивість ергодичності має важливе для практичного застосування значення, тому що статистичні характеристики ергодичних процесів можуть бути отримані з однієї досить довгої його реалізації.

Попереднім етапом статистичної обробки повинен бути етап перевірки однорідності вибірки в сенсі незмінності ймовірнісних характеристик. Мірою однорідності статистичної сукупності є коефіцієнт варіації [Крамер, 1975]:

$$V = \frac{\sigma}{\bar{a}} 100\%, \quad (6.5)$$

де σ – середньоквадратичне відхилення, \bar{a} – середнє арифметичне вибірки. Сукупність вважається однорідною, якщо коефіцієнт варіації не перевищує 33%. Вибіркове значення $V_q=77\%$ не узгоджується з припущенням про однорідність вибірки. За критерієм Колмогорова-Смирнова гіпотеза про відповідність пуассонівського закону розподілу місячного числа землетрусів відхилена.

6.3. Дослідження структури ряду

Перше уявлення про природу часового ряду можна скласти за його графіком. Наприклад, іноді можна визначити характер сезонних коливань: адитивний або мультиплікативний. Аддитивна модель відрізняється тим, що амплітуда сезонних коливань, що відображає відхилення від тренда або середнього, залишається приблизно постійною в часі. Графічним аналізом можна визначити компоненти часових рядів, а також з'ясувати структуру і характер сезонних коливань ряду. Якщо присутність тренда в часі ряду не простежується, то спочатку потрібно з'ясувати існування тенденції в досліджуваному процесі.

На підставі графіка часового ряду (рис. 6.1) можна стверджувати, що спостерігається повільне стійке зростання рівнів ряду. Отже, складові тимчасового ряду компоненти входять в нього мультиплікативно (6.2). Мультиплікативні моделі відрізняються від адитивних моделей тим, що в адитивній моделі сезонні коливання не залежать від рівня ряду, а в мультиплікативній моделі амплітуда сезонних флуктуацій змінюється в залежності від значень ряду. Часто число підземних поштовхів просторової і тимчасової околиць сильних землетрусів збільшується, утворюючи рій землетрусів, що включають попередні (форшоки) і наступні (афтершоки) землетрусу. Мультиплікативну модель можна пояснити наявністю рою землетрусів, а недостатність сейсмічних даних наявністю плавного зростання числа щомісячних землетрусів (тренд). На графіку видно, що ряд містить циклічну компоненту, і характер щомісячних землетрусів повторюється. Різко виділені

значення означають періоди підвищення сейсмічної активності. Однак для повного аналізу структури ряду слід побудувати корелограму.



Рис. 6.1. Графік часового ряду

Таблица 6.1

Статистика груп

		N	середнє	стандартне відхилення	стандартна похибка середнього
Групи	1	204	15,32	12,925	0,905
	2	204	24,39	16,437	1,151

На стаціонарність в широкому сенсі може вказати незалежність параметрів ряду від початку відліку часу. Для цього ряд розбивається на дві групи - складаються ряди даних за періоди 1978–1994 і 1995–2011 рр., і перевіряється гіпотеза про рівність середніх значень і дисперсії груп. Одне з припущень класичного дисперсійного аналізу полягає в тому, що дані в кожній групі підкоряються розподілам з однаковими дисперсіями. Часто для перевірки цього припущення використовується критерій Лівіньо, оскільки він досить стійкий до відхилень від нормальності (ϵ робастним). У цьому критерії використовуються відхилення (без знака) кожного спостереження від групового середнього, як в дисперсійному аналізі.

Критерій рівності дисперсії Лівінью. Він застосовується для того, щоб визначити, чи відрізняється розсіювання в різних групах. Нульова гіпотеза означає, що дисперсії двох груп рівні. F -статистику для критерію Лівіні отримують, проводячи одно факторний дисперсійний аналіз для модулів відхилення кожного спостереження від групового середнього. Якщо спостережувана ймовірність значущості незначна (наприклад, менше 0.05), то для середніх значень слід використовувати t -критерій з роздільними дисперсіями. У нашому випадку F -статистика дорівнює 15.085 із значущістю (значення p) 0.0, тобто гіпотезу про рівність дисперсій можна відкинути. Тому t -критерій з об'єднаними дисперсіями (Equal variances assumed) для порівняння середніх значень тут не підходить. У процедурі порівняння середніх обчислюються два критерії, критерій, названий «Передбачається рівність дисперсій», t -критерієм з об'єднаними дисперсіями. Він передбачає, що дисперсії сукупностей, з яких набираються члени груп, рівні, тобто їх розподіли мають однакову форму. Другий критерій, названий «Рівність дисперсій не передбачається», не вимагає рівності дисперсій і називається t -критерієм з роздільними дисперсіями. В t -критерії парних вибірок гіпотеза $H_0: \bar{u}=0$. означає рівність нулю середнього різниць спостережень в двох групах $u_i = x_i - y_i$, $i=1, \dots, n$.

Застосовуючи t - критерій:

$$t = \frac{\bar{u}}{\frac{\sigma}{\sqrt{n}}}; \quad \bar{u} = \frac{1}{n} \sum_{i=1}^n (x_i - y_i); \quad \sigma^2 = \frac{1}{n-1} \sum_{i=1}^n u_i^2 \quad (6.6)$$

з роздільними дисперсіями (рівність дисперсій не передбачається) отримана статистика $t = -9.074$. Це значення перевищує табличне значення, що відповідає рівню значущості $\alpha=0.01$, с $(n-1)$ ступенями свободи, тобто гіпотеза про рівність середніх значень двох груп з різними дисперсіями неправдоподібна. Таким чином, припущення про стаціонарність ряду за середнім значенням і по дисперсії відкидається. Щоб отримати можливість одностороннього критерію потрібно розділити ймовірність значення t на 2.

Таблиця 6.2

Критерії для незалежних вибірок

Критерії рівності дисперсії Лівінью		<i>t</i> -критерій порівняння групових середніх							
Дисперсії	<i>F</i>	Ймовірність значущості	<i>t</i>	Ст.св.	Ймовірність значущості (2-стороння)	Різниця середніх	Стандартна похибка різниці	95% -ній довірчий інтервал різниці	
								Нижній	Верхній
рівні	15.08	0.00	-6.918	406	0.0	-9.074	1.464	-6.196	-11..951
рівні			-6.918	384, 606	0.0	-9.074	1.464	-6.196	-11..951

Критерій Лівінью застосовується, коли потрібно порівняти умови, в яких проводяться спостереження і числові характеристики часових рядів на різних відрізках часу. Статистичні моменти, обчислені за різними інтервалами ряду, є незалежними випадковими величинами. Параметричні методи перевірки стаціонарності ряду застосовуються, коли відома ймовірна структура ряду. Коли відсутня інформація про частотну природу ряду, застосовуються непараметричні методи. Ці методи, наприклад, метод інверсій, виявляють тенденції у тимчасових рядах і виявляють статистичну незалежність в вибірках [Бендат, Пирсол, 1989; Дуброва, Архипова, 2004; Марно, 2008; Кисляк, 2007; Spezialvorlesung, 2006]. Багато статистичних тестів припускають, що порядок, в якому були зібрані дані, не має значення. Якщо це не так, то вибірка не є випадковою, і дані спостережень не є незалежними. Для перевірки гіпотези про випадковість вибірки застосовується також критерій серій.

Випадковість, в разі тимчасових рядів, означає відсутність тренда. Для застосування критерію серій кожен рівень порівнюється з величиною M – медіана, мода або ж середня тимчасового ряду [Карманов, 2007; Никитин, Сосунова, 2003; Шанченко, 2008]. Якщо рівень ряду x_i більше M , то цьому значенню присвоюється знак (+), якщо менше, то (-), якщо ж однакові, то символ не присвоюється. Серією називається послідовність наступних один за одним знаків: (+) і (-). Вибірka вважається випадковою, якщо виконуються обидві нерівності [Бендат, Пирсол, 1989; Дуброва, Архипова, 2004; Карманов, 2007]:

$$L_{\max} < 3,31 \lg(n+1); v_t < 0,5(n+1 - u_{1-\frac{\alpha}{2}} \sqrt{n-1}) \quad (6.7)$$

Гіпотеза про випадковий характер ряду відкидається, тому що відповідна ймовірність значущості менше ймовірності рівня значущості $\alpha=0.05$.

U-тест за методом Манна-Уїтні – це тестовий критерій непараметричного порівняння двох незалежних вибірок [Смирнов, Дунин-Барковский, 1965]. Він заснований на використанні однієї загальної послідовності значень обох вибірок. З двох рядів складається один варіаційний ряд. Нульова гіпотеза означає, що значення однієї вибірки будуть рівномірно розподілені серед значень іншої вибірки. Альтернативною гіпотезою буде відповідати ситуація, коли значення однієї з вибірок будуть переважати на одному з кінців об'єднаного ряду [Агафонова, 2010; Володин, 2006; Горбацевич, 2000; Сеньо, 2007]. В цьому випадку перетин двох рядів буде мінімальним. Для великих вибірок m, n (m -довжина першого ряду, а n – довжина другого ряду), $m > 20, n > 20$, статистика критерію приблизно нормально розподілена з математичним очікуванням $a=mn/2$, $\sigma^2=mn(m+n+1)/12$. Інтервал $(a-1.96\sigma; a+1.96\sigma)$ є областю прийняття гіпотези про випадковість ряду на 5% -ном рівні значущості. За методом Манна-Уїтні гіпотеза про випадковості ряду відкидається.

При проведенні тесту Мозеса перший з двох рядів розглядається як контрольний. Значення обох рядів розташовуються в порядку статистику, і їм присвоюються відповідні рангові місця. У контрольній групі підраховується розмах між цими ранговими місцями. За тестом Мозеса гіпотеза про те, що обидва ряди є складовими одного ряду, подвоєної довжини, також відкидається.

Основою тесту Колмогорова-Смирнова є розрахунок максимальної різниці між кумулятивними частотами обох рядів. Ця різниця позначається величиною z , на підставі якої, обчислюється ймовірність значущості. За критерієм Колмогорова-Смирнова нульова гіпотеза також відкидається [Сеньо, 2007; Bernhardt, 2007; SPSS Trends 14.0, 2006].

Якщо вірно припущення про випадковий характер коливань спостережень в досліджуваному ряду, то не повинно бути зв'язку між рівнями ряду. Альтернативне припущення допускає наявність залежності між послідовними рівнями, тобто, в

цьому випадку, часовий ряд не випадковий. Для оцінки зв'язку між послідовними рівнями ряду може бути використана функція автокореляцій рядів. Однак, в цьому випадку, коефіцієнт автокореляції може бути оцінений і для нестационарного ряду, але в цьому випадку його ймовірнісна інтерпретація втрачається [Андерсон, 1971; Володин, 2006; Teusch, 2006].

6.4. Автокореляційні функції рядів

Метою аналізу часових рядів є визначення моделі реалізації ряду. Кореляційний аналіз дозволяє виявити структуру ряду, тобто визначити присутність в ряді тієї чи іншої періодичної компоненти заздалегідь не відомої частоти. Автокореляційна функція визначається за формулою [Бриллинджер, 1980]:

$$ACF(\tau) = \frac{\sum_{k=\tau+1}^n (y_k - \bar{y})(y_{k-\tau} - \bar{y})}{\sum_{k=1}^n (y_k - \bar{y})^2} \quad (6.8)$$

де $\tau=1,2,\dots,n-1$ величина зсуву, звана лагом, визначає порядок коефіцієнта кореляції. Хоча автокореляційна функція визначена тільки для стаціонарних процесів, але обчислити її можна для будь-яких рядів і аналізувати природу ряду. Так, якщо найбільш високим виявився коефіцієнт автокореляції першого порядку, то досліджуваний ряд містить тільки тенденцію. Якщо найбільш високим виявився коефіцієнт автокореляції порядку $\tau > 1$, то ряд містить циклічні коливання з періодичністю, в τ моментів часу [Андерсон, 1971; Бриллинджер, 1980]. Якщо жоден з коефіцієнтів автокореляції не є значущим, то це може означати:

- ряд не містить тенденції і циклічних коливань і має випадкову структуру – "білий шум"
- ряд містить сильну нейтральну тенденцію, для виявлення якої потрібно провести додатковий аналіз.

Для визначення виду і порядку процесів, що породжують стаціонарний часовий ряд, використовується апарат автокореляційних функцій: звичайної (АКФ) та приватної (ЧАКФ).

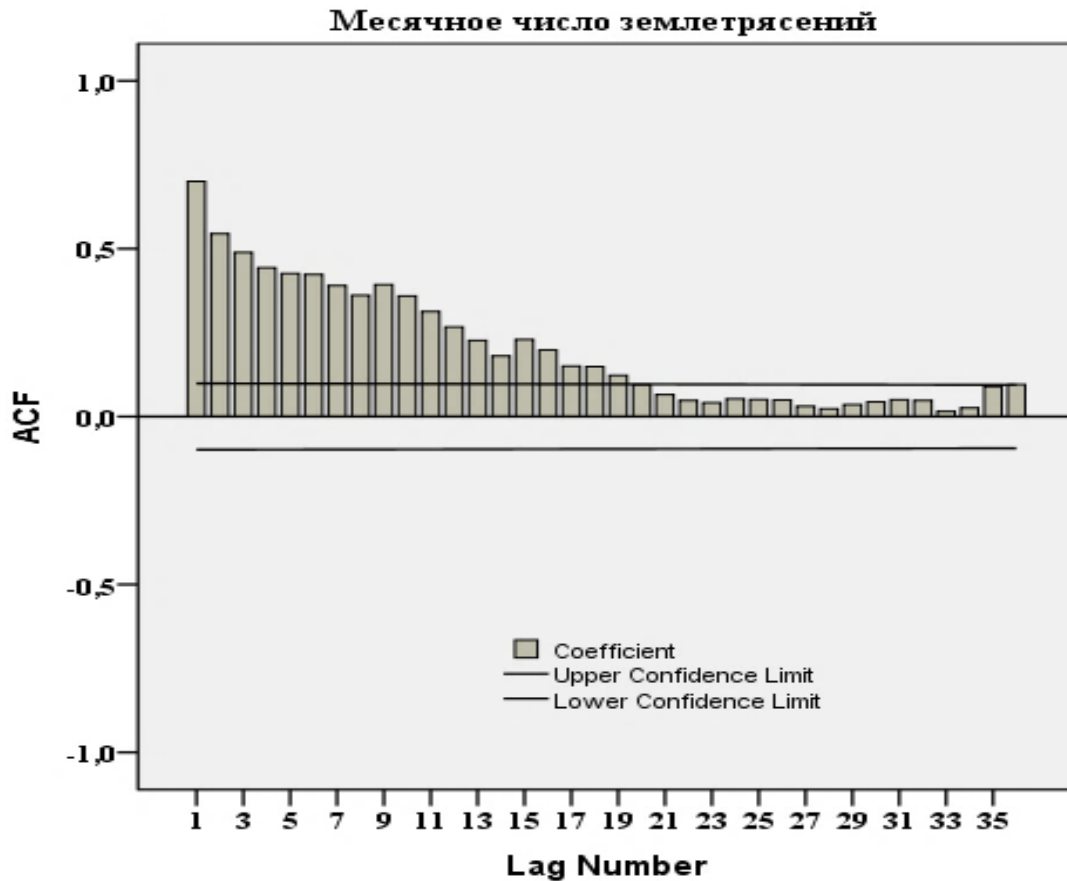
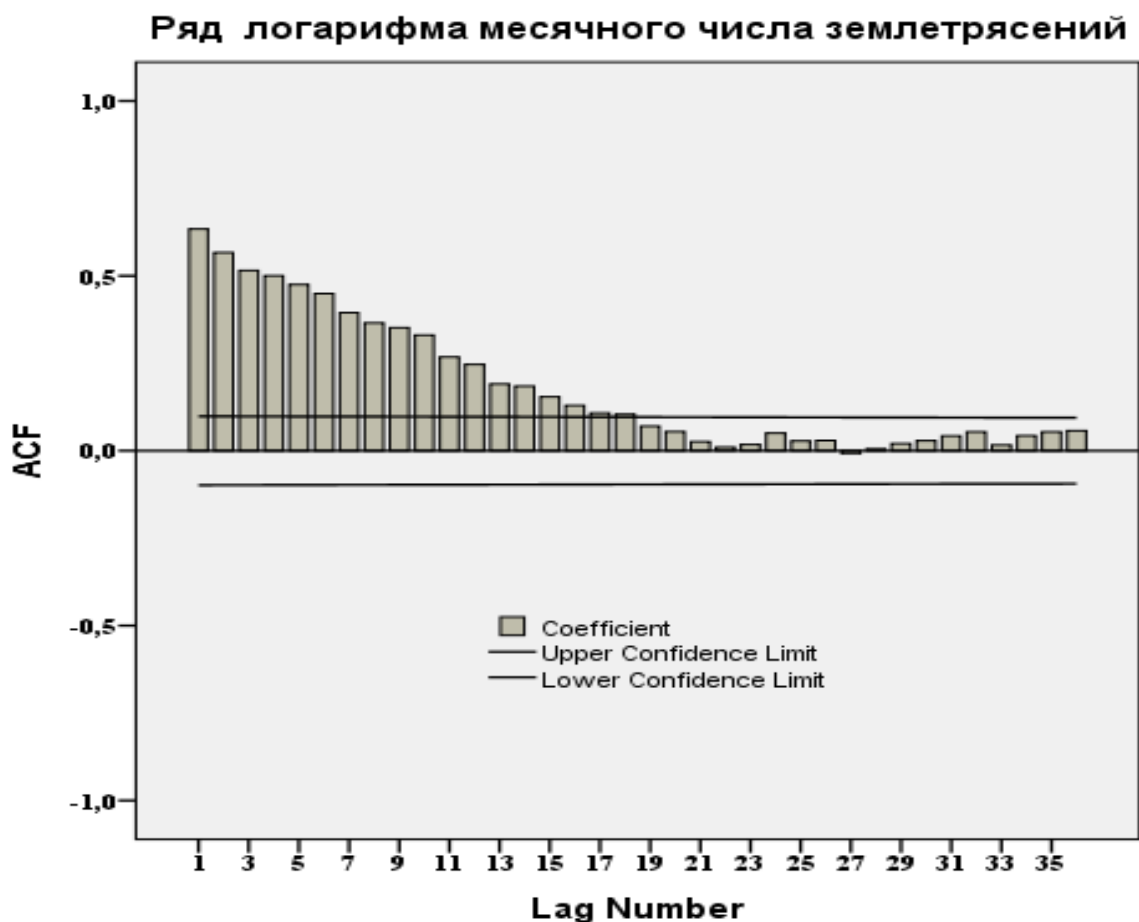


Рис. 6.2. Автокореляційна функція ряду
(Місячне число землетрусів)

Якщо часовий ряд близький до білого шуму, то корелограма осилує близько до горизонтальної вісі, а її значення близькі до 0. При великих значеннях τ оцінка $ACF(\tau)$ для коефіцієнта автокореляції містить похибки. Це пов'язано з частковим підсумовуванням, тому що, з підсумовування відкидаються спостереження. Тому корелограма при великих значеннях τ не відображає справжню структуру ряду. Для стаціонарного ряду $ACF(\tau)$ швидко зменшується з ростом τ . При наявності тренда автокореляційна функція набуває вигляду повільно спадаючої кривої. У разі сезонної періодичності в графіку ACF присутні піки для лагів, кратних періоду сезонності. Але ці піки можуть бути приховані присутністю тренда або великою дисперсією випадкової компоненти. Для заданого рівня ймовірності α (за замовчуванням береться зазвичай $\alpha=5\%$) можна обчислити межі

довірчого інтервалу, в якому знаходження значення функції автокореляції при заданому лагу τ з ймовірністю $1-\alpha$ який суперечить припущенню про відсутність кореляції перетинів з цим лагом. При графічному зображенні функції автокореляції або її оцінки (корелограма) ці інтервали дають дві граничні криві (вище і нижче основного графіка). Вихід за ці граничні криві розглядається як вказівка на значимість кореляції з відповідним лагом.



Автокореляційна функція ряду (рис. 6.2) показує значимий пік на лагу 1, який є і максимальним $\rho(1) = 0.7$. Тобто ряд містить тільки тренд. Крім того, наявність статистично значущого коефіцієнта автокореляції вказує, що ряд не випадковий, і між послідовними спостереженнями існує певний зв'язок. Значний негативний коефіцієнт свідчить про високочастотну осциляцію рівнів.

Лінійні коефіцієнти автокореляції характеризують тісноту тільки лінійного зв'язку поточного і попередніх рівнів ряду. Тому за коефіцієнтами автокореляції

можна судити тільки про наявність чи відсутність лінійної залежності. Для перевірки ряду на наявність нелінійної тенденції обчислювалися лінійні коефіцієнти автокореляції для тимчасового ряду, що складається з логарифмів вихідних рівнів. Відмінні від нуля значення коефіцієнтів автокореляції свідчать про наявність нелінійної тенденції. Корелограма (рис. 6.3) показує, що рівні ряду пов'язані також нелінійною залежністю.

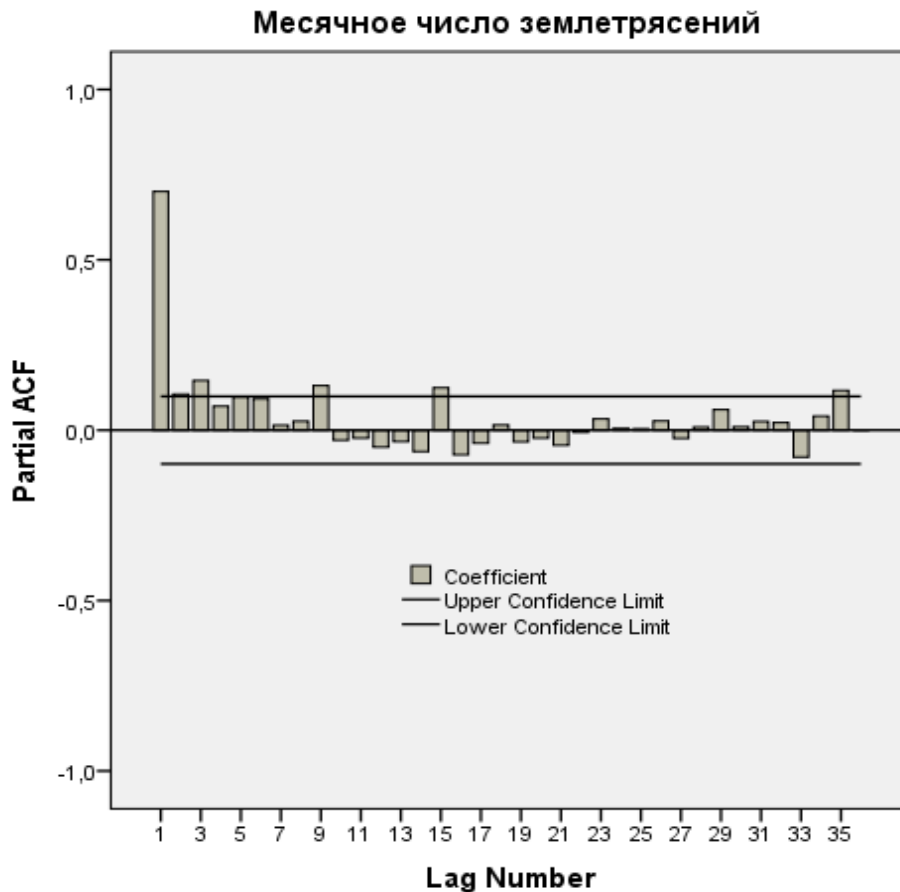


Рис. 6.4. Приватна автокореляційна функція ряду місячного числа землетрусів

Інший корисний метод дослідження періодичності полягає в дослідженні ЧАКФ (рис. 6.4), що представляє собою поглиблене поняття звичайної автокореляційної функції. У ЧАКФ усувається залежність між проміжними спостереженнями. Іншими словами, приватна автокореляція на даному лагу аналогічна звичайній автокореляції, за винятком того, що при обчисленні з неї видаляється вплив автокореляцій з меншими лагами [Андерсон, 1971; Бриллинджер,

1980]. На лагу 1 (коли немає проміжних елементів всередині лага), приватна автокореляція дорівнює, очевидно, звичайній автокореляції. Як зазначалося вище, періодична складова для даного лага k може бути видалена вибором різниці відповідного порядку. Це означає, що з кожного i -го елемента ряду віднімається $(i-k)$ -й елемент.

Є два аргументи на користь розглянутих перетворень:

- 1) по-перше, таким чином можна визначити приховані періодичні складові ряду. Нагадаємо, що автокореляції на послідовних лагах залежні. Тому видалення деяких автокореляцій змінить інші автокореляції, які, можливо, придушували їх, і зробить деякі інші сезонні складові більш помітними
- 2) по-друге, видалення сезонних складових перетворює тимчасовий ряд в стаціонарний ряд, що є необхідним для застосування *ARIMA* і інших методів, наприклад, спектрального аналізу.

Для перевірки значущості коефіцієнтів автокореляції застосовується критерій Боксу-Пірса [Шанченко, 2008]:

$$Q = n \sum_{k=1}^m r_k \quad (6.9)$$

де r_k – коефіцієнт автокореляції з лагом k , m – найбільший лаг. Вибіркова статистика критерію (6.9) є χ^2 розподіленої випадкової завбільшки з m ступенями свободи.

6.5. Аналіз тренда в часових рядах

Формальними методами визначення тренда є: метод найменших квадратів; змінного середнього. Для моделювання тенденції часового ряду будується аналітична функція, що характеризує залежність рівнів ряду від часу. Цей метод називається аналітичним вирівнюванням тимчасового ряду. Для опису тренда використовуються наступні функції: лінійна; гіперболічна; експоненціальна; степенева; поліноміальна. Наприклад, для виявлення тренда в формі степеневі функції можна використовувати регресію рівнів часового ряду на час:

$$x_t = a + \sum_{j=1}^p \beta_j t^j + \varepsilon_t, \quad (6.10)$$

$\varepsilon_t \sim iid(0, \sigma^2)$. Якщо регресія незначуща, то це означає, що тенденція в тимчасовому ряді відсутня. Розглянемо модель лінійної регресії рівнів ряду на час $x_t = \alpha + \beta t + \varepsilon_t$. Таким чином, гіпотеза про відсутність тренду означає правдоподібність гіпотези $H_0: \alpha=0, \beta_j=0, j=1, \dots, p$ —про рівність нулю коефіцієнтів регресії. Для перевірки гіпотези H_0 обчислюється статистика F -критерію [Сеньо, 2007]:

$$F = \frac{(RSS_H - RSS)/(p-1)}{RSS/(n-p)} \quad (6.11)$$

де n — число спостережень, p — число параметрів. якщо $F > F_{p-1, n-p}^\alpha$, то гіпотеза H_0 відкидається, і це означає, що регресія значуща, і всі регресори слід врахувати.

В даному випадку, гіпотеза H_0 F -критерію вказує на статистично значущий регресивний зв'язок рівнів ряду з часом по всіх моделях тренда. Тобто в ряді місячного числа землетрусів існує тенденція зростання рівнів ряду з часом. Наявність тренда в ряді місячного числа землетрусів можна пояснити тим, що:

- поліпшується робота системи сейсмічних спостережень. Адже каталоги землетрусів не відображають природу сейсмічності і не містять відомості про всі сейсмічні події, що трапилися в осередку землетрусів. Як відомо поріг мінімальної магнітуди комплектності каталогів землетрусів знижується з часом. Тобто спостережена тенденція не пов'язана з природою сейсмічності, а викликана відсутністю представницького обсягу відомостей про землетруси
- період повторення великих сейсмічних подій становить десятки, іноді сотні років, тоді як ряди спостережень охоплюють більш короткі періоди часу. Наприклад, в даному випадку досліджується ряд місячного числа землетрусів, що сталися з 1978 по 2011 рр.

З огляду на досить поважний вік вогнищ землетрусів, який обчислюється мільйонами років, довжина ряду повинна перевищувати період повторення найбільшого для даного вогнища землетрусу, і крок квантування за часом повинен становити не тільки місяць, але і більш довгі інтервали часу. Під довжиною ряду мається на увазі час, що минув від першого до останнього спостереження в ряді.

Одним з непараметричних методів виявлення тренда M_t є метод змінного середнього, при якому випадкові відхилення погашаються. При згладжуванні цим методом значення рівнів ряду замінюються середніми значеннями, які характеризують серединну точку періоду дотику [Бриллинджер, 1980; Кисляк, 2007; Христиановский, 2011]. При відніманні від значень ряду значень, змінного середнього $\varepsilon_t = x_t - m_t$ можна виділити випадкову компоненту ряду, яка використовується для побудови моделі авторегресії з метою прогнозу.

Існує безліч критеріїв для вибору типу оптимальних моделей часових рядів. Якість прогнозу має вирішальне значення при виборі моделі. Критерії реалізовані в процедурі аналізу часових рядів пакета SPSS [SPSS Trends 14.0., 2006].

6.5.1. Вибір оптимальних моделей часових рядів

R-квадрат для стаціонарної частини [Spezialvorlesung Zeitreihenanalyse, 2006; SPSS Trends 14.0, 2006]. Показник порівняння стаціонарної частини моделі (без тренда) з простою моделлю середнього значення. Цей захід краще звичайного R -квадрата, при наявності тренда або сезонної компоненти. R -квадрат для стаціонарної частини може приймати і негативні значення, $(-\infty, 1]$. Негативні значення означають, що дана модель гірше, ніж проста модель. Позитивні значення означають, що дана модель краще, ніж проста модель. Можна отримати негативне значення R -квадрат для рівнянь, які не містять вільний член. У цьому випадку R -квадрат не може бути розтлумачено як квадрат кореляції. Такі ситуації показують, що постійний член повинен бути в моделі [Володин, 2006].

R-квадрат, ця оцінка всієї варіації $ЧР$, яка пояснюється моделлю. Цей захід важливий, якщо $ЧР$ є стаціонарною. R -квадрат може бути і негативною, приймати значення з інтервалу $(-\infty, 1]$. Негативні значення означають, що дана модель гірше, ніж проста модель, і що модель містить змінні, які не беруть участі в прогнозі. Позитивні значення означають, що розглянута модель краще, ніж проста модель.

RMSE – Root Mean Square Error, корінь з середньоквадратичної похибки. Цей захід показує, як сильно відхиляється залежний *ЧР* від його модельних значень, хоча виражається в тих же одиницях виміру, що і залежний ряд.

MAPE – Mean Absolute Percentage Error, середня складова відносної похибки. Цей захід величини відхилення, залежного *ЧР* від його модельних значень. Це значення не залежить від застосовуваних одиниць вимірювання і може застосовуватися для порівняння *ЧР* виміряних в різних одиницях. Для оцінки якості підгонки для всього ряду краще використовувати середню абсолютну відносну похибку. Часто цей захід більш виразніший, ніж середньоквадратическая похибка. Наприклад, знання того, що точність прогнозу $\pm 5\%$, є корисною.

MAE – Mean Absolute Error, Середня абсолютна похибка. Обчислюється як середнє абсолютних похибок. Вона вимірює, як сильно *ЧР* відхиляється від передбачених значень рівня ряду. **MAE** задається в тих же одиницях, що вихідний ряд. Якщо вона дорівнює нулю, то маємо досконалу адекватну модель. У порівнянні із середньою квадратичною похибкою, цей захід "не надає надто великого значення" розсіюванням.

MaxAPE Maximum Absolute Percentage Error, максимальна складова відносної похибки. Це найбільша передбачена похибка, виражена у відсотках. Цей захід допомагає уявити сценарій прогнозу Worst-Case (найбільш несприятливий варіант).

MaxAE – Maximum Absolute Error, максимальна складова відносної похибки. Це найбільша передбачена похибка, виражена в тих же одиницях виміру, як і залежний *ЧР*. Як і **MaxAPE** цей захід допомагає уявити сценарій прогнозу Worst-Case. Максимальна складова похибки і максимальна складова відносної похибки можуть з'явитися в різних місцях *ЧР*, наприклад, якщо вклад похибки для деякого великого значення *ЧР* в незначній мірі перевищує, ніж вклад похибки для деякого мінімального значення *ЧР*. В цьому випадку, найбільший внесок похибки спостерігається для більших значень *ЧР*, і відносний найбільший внесок похибки для найменших значень *ЧР*.

Normalized BIC – Normalized Bayesian information criterion, нормалізований інформаційний критерій Байеса (*BIC*). Загальна міра досягнутої в цілому величини відповідності, яка повинна, так само враховувати складність моделі. Вона є величиною, яка визначається середньо-квадратичної похибкою і штрафом за параметри моделі і довжину ряду. Штраф компенсує перевага моделі з великим числом параметрів і перетворює статистику в гарну можливість порівняння різних моделей ЧР.

Percentile показує поріг, нижче якого знаходяться значення, відповідні відсотку передбачених моделлю величин.

Для дослідження структури часових рядів застосовуються різні перетворення вихідних рядів. Одним з поширених методів аналізу рядів є просте експоненціальне згладжування, яке застосовується для згладжування процесів з постійним трендом і для рядів з сезонною складовою. Метод був запропонований Броуном і Холтом, і залишається найпоширенішим методом прогнозування часових рядів.

6.5.2. Метод експоненціального згладжування і прогнозування часових рядів

Теорія часових рядів має велику кількість методів передбачення рівня рядів, які реалізують схему екстраполяції. Тобто досліджується ряд, і передбачається, що властивості його не змінюватися в майбутньому. Одним з поширених методів є метод простого експоненціального згладжування. Незважаючи на простоту використаного математичного апарату, прогнозний потенціал методу не поступається методам, де застосовуються більш глибокі математичні методи екстраполяції. Метод експоненціального згладжування відноситься до непараметричних методи аналізу часових рядів, тому що його застосування не залежить від виду розподілу випадкової компоненти. Метод експоненціального згладжування дає можливість отримати оцінку параметрів тренда, що характеризують не середній рівень ряду, а тенденцію, що склалася до моменту останнього спостереження.

Часто, для прогнозування нестационарних часових рядів застосовується найпоширеніша модель простого експоненціального згладжування:

$$S_t = \alpha x_t + \beta S_{t-1} \quad (6.12)$$

де S_t – значення згладженого і x_t вихідного ряду в момент t ; α – параметр згладжування, $0 < \alpha < 1$; $\beta = 1 - \alpha$. Це формула застосовується рекурсивно – кожне нове значення обчислюється як зважене середнє значення даного спостереження (яке є також прогнозом) і згладженого ряду. Результат згладжування залежить від параметра α . Коли $\alpha = 1$ попередні спостереження повністю ігноруються. При $\alpha = 0$ ігноруються поточні спостереження. Коли α приймає значення з інтервалу $(0,1)$, то отримуємо проміжні результати. У практичних застосуваннях рекомендується вибирати значення параметра згладжування від 0.1 до 0.3 [Кисляк, 2007; Шанченко, 2008]. Для ряду місячного числа землетрусів були випробувані моделі несезонні та сезонні [Backhaus, 2011; Nuxol, 2007; Marques de Sá, 2007; Teusch, 2006]:

- несезонні моделі – проста, Холта і Брауна (лінійний тренд).
- сезонні моделі – проста сезонна, адитивна Уінтерса, мультиплікативна Уінтерса.

Коли в ряді міститься нелінійний тренд, то є сенс перетворити дані. У пакеті SPSS для перетворення даних застосовуються дві функції: натуральне логарифмування і функція квадратного кореня. Для прогнозу прологарифмованого тимчасового ряду щомісячних землетрусів оптимальними, за величиною вибіркового статистичного критерію перевірки якості моделі, виявилися проста сезонна для прологарифмованого ряду, і мультиплікативна сезонна модель Уінтерса для ряду, що складається з значень квадратних коренів рівнів вихідного ряду (табл. 6.3).

Якщо ряд містить сезонну складову, то має сенс експоненціально згладити цю складову, з деяким додатковим параметром δ . Проста сезонна модель експоненціального згладжування відрізняється від моделі простого експоненціального згладжування, тим, що в адитивній моделі в прогноз додається сезонна компонента, а в мультиплікативній моделі множиться. У адитивній моделі прогноз будується за формулою:

$$P_t = S_t + I_t \quad (6.13)$$

де I_t – згладжений сезонний індекс в момент $t-p$, який задається виразом:

$$I_t = I_{t-p} + \delta(1-\alpha)e_t \quad (6.14)$$

У мультиплікативної моделі додавання заміниться множенням:

$$P_t = S_t I_t \quad (6.15)$$

$$I_t = I_{t-s} + \delta(1-\alpha)\frac{e_t}{S_t} \quad (6.16)$$

де δ параметр сезонного згладжування, і приймає значення між 0 і 1. S_t – значення експоненціально згладженого рівня ряду в момент t , і I_{t-p} позначає згладжений сезонний фактор в момент t мінус s (s – кількість періодів в сезонному циклі), а e_t спостережене мінус передбачене значення ряду в момент t .

Мультиплікативна модель Уінтерса застосовується для ряду з лінійним трендом і сезонною варіацією, що змінюється з величиною ряду. У мультиплікативної моделі Уінтерса прогноз на k кроків вперед визначається формулою:

$$Y'_{t+p} = (L_t + pT_t)S_{t-s+p} \quad (6.17)$$

де $Y'_{(t+p)}$ передбачений на p періодів вперед рівень ряду. Складова L_t описує згладжений ряд, T_t є значенням тренда в момент t , а S_t служить для оцінки сезонності і обчислюється за формулами:

- згладжування вихідного ряду:

$$L_t = \frac{\alpha y_t}{S_{t-s}} + (1-\alpha)(L_{t-1} + T_{t-1}) \quad (6.18)$$

- згладжування тренда:

$$T_t = \beta(L_t - L_{t-1}) + (1-\beta)T_{t-1} \quad (6.19)$$

- Оцінка сезонної складової:

$$S_t = \gamma \left(\frac{y_t}{L_t} + (1-\gamma)S_{t-s} \right) \quad (6.20)$$

де α , β , γ - параметри згладжування, які приймають значення з інтервалу $[0,1]$.

Метод заснований на застосуванні експоненціального зваженого середнього для сезонних рядів. Передбачені за обраними моделями значення місячного числа землетрусів за 2012 р. в незначній відрізняються (другий і третій рядок таблиці. 6.3). У першому рядку (табл. 6.3) наводяться відомості про спостереження місячного

числа землетрусів в 2011р. Числа 1, 2, ..., 12 відповідають назві місяців: 1-й січень; 2-й лютий і так далі.

Якість прогнозу (табл. 6.4) визначається по максимуму коефіцієнта детермінації R^2 , і за оптимальними значеннями числових характеристик ряду залишків: $RMSE$; $MAPE$; MAE ; $MaxAPE$; $MaxAE$; BIC (Табл. 6.4). Якщо модель адекватна, то ряд залишків є "білим шумом". Для перетвореного ряду були побудовані моделі: проста; проста сезонна; Холта; Брауна; Уінтерса [Андерсон, 1971; Дуброва, Архипова, 2004; Карманов, 2007; Шанченко, 2008]. Оптимальними виявилися проста сезонна і мультиплікативна Уінтерса моделі. Вибіркові значення ймовірності значущості 0.194 і 0.15 (табл. 6.4) статистики Боксу-Льюнг, менше рівня значущості $\alpha=0.05$. Графіки автокореляційної і приватної автокореляційної функцій ряду залишків не мають значимих піків, і знаходяться в межах 95% -ного довірчого інтервалу на всіх лагах.

Таблиця 6.3

Передбачені значення ряду за моделями сезонного експоненціального згладжування і мультиплікативної моделі Уінтерса.

Місяць	1	2	3	4	5	6	7	8	9	10	11	12
2011	24	21	24	20	16	27	21	27	24	29	16	21
Проста сезонна. 2012	20	18	23	27	29	26	26	34	31	33	26	31
Уінтерса. 2012	21	19	24	28	30	27	27	34	32	34	27	32

Динаміка часових рядів може мати різні форми, і існує велика кількість моделей, які можуть більш-менш достовірно описати часовий ряд. Вибір оптимальної моделі, з метою прогнозу є складним і основним завданням теорії часових рядів. У формуванні ряду беруть участь закономірні і випадкові чинники, і дослідження міри зв'язку між ними може виявити структуру ряду і визначити його компоненти.

6.6. Розкладання рядів на компоненти

Очищений від сезонної складової тимчасової ряд краще придатний для аналізу тренда. Перше уявлення про наявність річних періодичних складових дають графіки рядів (рис. 6.1.). Візуально, за корелограмою ряду (рис. 6.2) неможливо однозначно підтвердити наявність річних сезонних коливань в ряді щомісячних землетрусів. Застосуємо статистичний аналіз для дослідження компонентного складу ряду.

Позначимо через $X_{j,k}$ значення метеорологічного показника в k -й місяць j -го року. Припускаючи постійність тренда протягом року, обчислимо його значення за формулою [Никитин, Сосунова, 2003; Сеньо, 2007; Spezialvorlesung Zeitreihenanalyse, 2007]:

$$\hat{t}_j = \frac{1}{12} \sum_{k=1}^{12} x_{j,k} \quad (6.21)$$

Сезонні складові для кожного місяця $k=1, \dots, 12$ обчислюються за формулою:

$$\hat{s}_k = \frac{1}{m} \sum_{j=1}^m (x_{j,k} - \hat{t}_j) \quad (6.22)$$

Залишки виражаються через співвідношення:

$$\hat{\varepsilon}_{j,k} = x_{j,k} - \hat{t}_j - \hat{s}_k \quad (6.23)$$

$j = 1, \dots, m$ і $k = 1, \dots, 12$. Таким чином, отримаємо ряд залишків:

$$\varepsilon_t = x_t - m_t - s_t \quad (6.24)$$

Для виявлення і визначення сезонної складової, що присутня в тимчасовому ряді застосовувалася процедура "Сезонне розкладання" статистичного пакета SPSS. Дана процедура повторюється для ряду залишків, поки він буде відрізнятися від "білого шуму". Дана процедура створює такі нові ряди [Bernhardt, 2007; SPSS Trends 14.0, 2006]:

SAF – Рівнями ряду є значення щомісячної сезонної складової. Ці значення показують вплив окремих періодів на рівні ряду

SAS – Очищений від сезонної компоненти ряд. Тобто ряд рівнів, отриманих після, видалення відповідних значень місячної сезонної складової

STC – Згладжені значення циклічних і трендових складових ряду. Ці значення показують тренд і циклічну поведінку в ряді $g_t = m_t + c_t$

ERR – Залишки. Значення рівнів ряду, що залишилися після видалення тренда, сезонної і циклічної складових.

Автокореляційна функція ряду залишків показує, що в ньому ще існують тренд і флуктуації. Повторне застосування процедури не змінює природу ряду залишків. У таких випадках рекомендується побудувати модель змінного середнього часового ряду і проводити більш глибокий аналіз періодичних коливань [Spezialvorlesung Zeitreihenanalyse, 2007]. Критерієм адекватності моделі тимчасового ряду є непомітність ряду залишків від процесу "білого шуму" [Андерсон, 1971; Брилінджер, 1980; Бокс, Дженкінс, 1974; Марно, 2008]. Статистика Боксу-Льюнга, обчислена для перевірки значущості автокореляцій менше критичного значення відповідної ймовірності значущості $\alpha=0.05$.

Таблиця 6.4

Значення місячної сезонної складової

Місяць	1	2	3	4	5	6	7	8	9	10	11	12
<i>SAF</i>	0.91	0.8	0.98	1.06	1.1	0.98	0.96	1.17	1.08	1.11	0.88	0.99
<i>SAS</i>	26.5	26.35	24.5	18.85	14.58	27.65	22.05	23.18	22.15	26.13	18.27	21.14
<i>STC</i>	26.73	25.2	22.75	20.96	20.36	22.03	22.72	23.52	22.82	22.62	21.85	21.46
<i>ERR</i>	0.99	1.04	1.08	0.91	0.72	1.26	0.97	0.99	0.97	1.16	0.84	0.99

Більш глибокий аналіз періодичних коливань, з метою виявлення прихованих періодичностей, проводиться за допомогою спектрального аналізу часового ряду.

6.7. Спектральний аналіз часового ряду

Спектральний аналіз застосовується для визначення періодичної компоненти при відомій довжині періоду. По суті це лінійна регресія, де залежною змінною є

рівні ряду, а регресорами служать функції синусів і косинусів. Спектральний аналіз визначає кореляцію регресанта різної частоти за спостереженими даними. Якщо виявлена значуща кореляція, то це означає, що існує стійка періодичність на відповідній частоті в ряді щомісячних землетрусів. Відома теорема [Бокс, Дженкінс, 1974; Колмогоров, Фомін, 1976; Піскунов, 1964], згідно з якою серед усіх тригонометричних многочленів порядку n найменше середнє квадратичне відхилення має многочлен, шуканими коефіцієнтами якого є коефіцієнти Фур'є. На застосуванні одновимірних рядів Фур'є заснований один з методів моделювання сезонних і циклічних коливань. Ряди Фур'є – евристичний алгоритм, і є одним різновидів спектрального аналізу. За допомогою спектрального аналізу в структурі часового ряду визначається пік відхилень від тренд, що дозволяє розрахувати тривалість періодичної компоненти ряду. При застосуванні спектрального аналізу випадковий стаціонарний процес представляється у вигляді суми гармонійних коливань різних частот, які називаються гармоніками. Спектр описує розподіл амплітуд випадкового стаціонарного процесу по різних частотах. Дослідження частотної структури ряду виконується процедурою "Спектральний аналіз" пакета SPSS. Як відомо, майже будь-яку періодичну функцію можна апроксимувати рядом Фур'є, сумою синусів і косинусів [Бокс, Дженкінс, 1974; Піскунов, 1964; Смирнов, Дунин-Барковский, 1965; Backhaus, 2011; Spezialvorlesung Zeitreihenanalyse, 2006]:

$$x_t = a \cos(\omega t) + b \sin(\omega t); \quad 0 \leq \omega \leq \pi \quad (6.25)$$

Спектр Фур'є містить частоти, присутні у вихідному ряді. Отже, сезонна компонента тимчасового ряду щомісячних землетрусів може бути розкладена в ряд Фур'є. Для виявлення періодичних коливань на тлі білого шуму розроблені періодограми [Bernhardt, 2007]. Відхилення залишків від білого шуму можуть свідчити про не враховані сезонні коливання.

Розглянутий ряд охоплює період $T=408$ місяців. Синусоїдальні коливання поширюються на весь період спостережень, коли $\omega=2\pi/408$. При частоті 2ω отримаємо два коливання на всьому інтервалі спостережень, тому що, в цьому випадку довжина періоду становить $2\pi/2\omega = (2\pi)408/4\pi=204$ місяці. Частота $k\omega$,

$k=1, \dots, 204$ означає існування k коливань, кожен з періодом $408/k$ місяців відповідно. Таким чином, часовий ряд можна розкласти на суму синусоїдальних коливань різної частоти і амплітуди:

$$x_t = \sum a_k \cos(k\omega t) + b_k \sin(k\omega t), \quad k=1, 2, 3, \dots \quad (6.26)$$

де a_k і b_k коефіцієнти Фур'є – некорельовані випадкові величини з математичним очікуванням рівним нулю і однаковими дисперсіями. Сума кінцевого числа членів ряду є наближеною величиною, яка, зі збільшенням K , з будь-яким ступенем точності наближається до значень ряду.

Графічне представлення відносних значень коливань різних частот називається періодограма:

$$I(\omega_i) = \frac{2}{K} \left[\left(\sum_{k=1}^K a_k \cos(k\omega_i) \right)^2 + \left(\sum_{k=1}^K b_k \sin(k\omega_i) \right)^2 \right] \quad (6.27)$$

де $\omega_i = 2\pi i / 408$ – частота. Значення періодограми оцінюються на основі вибірки та пов'язані з випадковими коливаннями. Для видалення випадкових коливань періодограма згладжується методом зваженого дотичного середнього. Ширина вікна, змінного середнього береться рівною непарному числу $m=5$. Значення спектральної щільності обчислюється як середнє $m/2$ попередніх і наступних значень періодограми.

Періодограма – це графік розподілу дисперсії ряду Фур'є по частотах. Якщо на графіку тимчасового ряду видно нерегулярні піки, то це означає можливість прихованих періодичностей в тимчасовому ряді. Пік, що з'явився на тлі слабкої інтенсивності спектральної щільності в деякому діапазоні частот, показує, що досягнута частота періодичної компоненти ряду. Сукупність піків визначає набір гармонійних компонент ряду [Андерсон, 1971; Spezialvorlesung Zeitreihenanalyse, 2007]. Періодограма зображає послідовність піків, якщо найбільший з них припадає на частоту $0.083 = 1/12$, то це означало б присутність в ряді щомісячних землетрусів сезонної складової з періодом 12 місяців.

Передбачені за обраними моделями значення місячного числа землетрусів на 2012р. дещо відрізняються (другий і третій рядок таблиця. 6.4). У першому рядку

(табл. 6.4) наводяться відомості про спостереження у місячному числі землетрусів в 2011р. Числа 1, 2, ..., 12 відповідають назві місяців: 1-й січень; 2-й лютий і так далі.

При моделюванні тимчасового ряду сумою синусів і косинусів синусоїдальні періодичні компоненти проявляються на періодограмі в формі окремих вершин. Несинусоїдальні періодичні компоненти з'являються на періодограмі в формі ряду рівновіддалених вершин різної висоти. Вершина, відповідна найменшій частоті, вказує на частоту періодичної компоненти в тимчасовому ряді [Spezialvorlesung Zeitreihenanalyse, 2007]. Характер інших вершин каже, про те, що форма річної періодичної компоненти не є синусоїдальною. Для нестационарних рядів з гладким трендом періодограма містить різкий підйом в області низьких частот. При моделюванні тимчасового ряду сумою синусів і косинусів синусоїдальні періодичні компоненти проявляються на періодограмі в формі окремих вершин. Несинусоїдальні періодичні компоненти з'являються на періодограмі в формі ряду рівновіддалених вершин різної висоти. Вершина, що відповідає найменшій частоті, вказує на частоту періодичної компоненти в тимчасовому ряді [Spezialvorlesung Zeitreihenanalyse, 2007]. Характер інших вершин говорить, про те, що форма річної періодичної компоненти не є синусоїдальною. Для нестационарних рядів з гладким трендом періодограма містить різкий підйом в області низьких частот.

За графіком тимчасового ряду щомісячних землетрусів (рис. 6.1.) видно, що ряд може містити приховані періодичні коливання. Періодограма (рис. 6.5.) зображає послідовність хаотичних піків, і неможливо говорити про суттєві періодичні цикли різних частот. Форма періодограми вказує на тренд [Burtiev, 2014]. Взагалі, аналіз можливих періодичностей краще проводити по згладженій періодограмі: функції спектральної щільності (рис 6.6). У частотній області тренд можна розглянути, як коливання з нескінченно великим періодом і відповідно з дуже малою частотою. Даний тренд може вплинути на значення функції спектральної щільності на лівому кінці діапазону частот. Отже, для коректного застосування спектрального аналізу ряд слід розкласти на компоненти і видалити тренд. Різні підходи до виявлення

тренда дають різні його варіанти, відповідно різні і значення рівнів очищеного від тренду ряду.

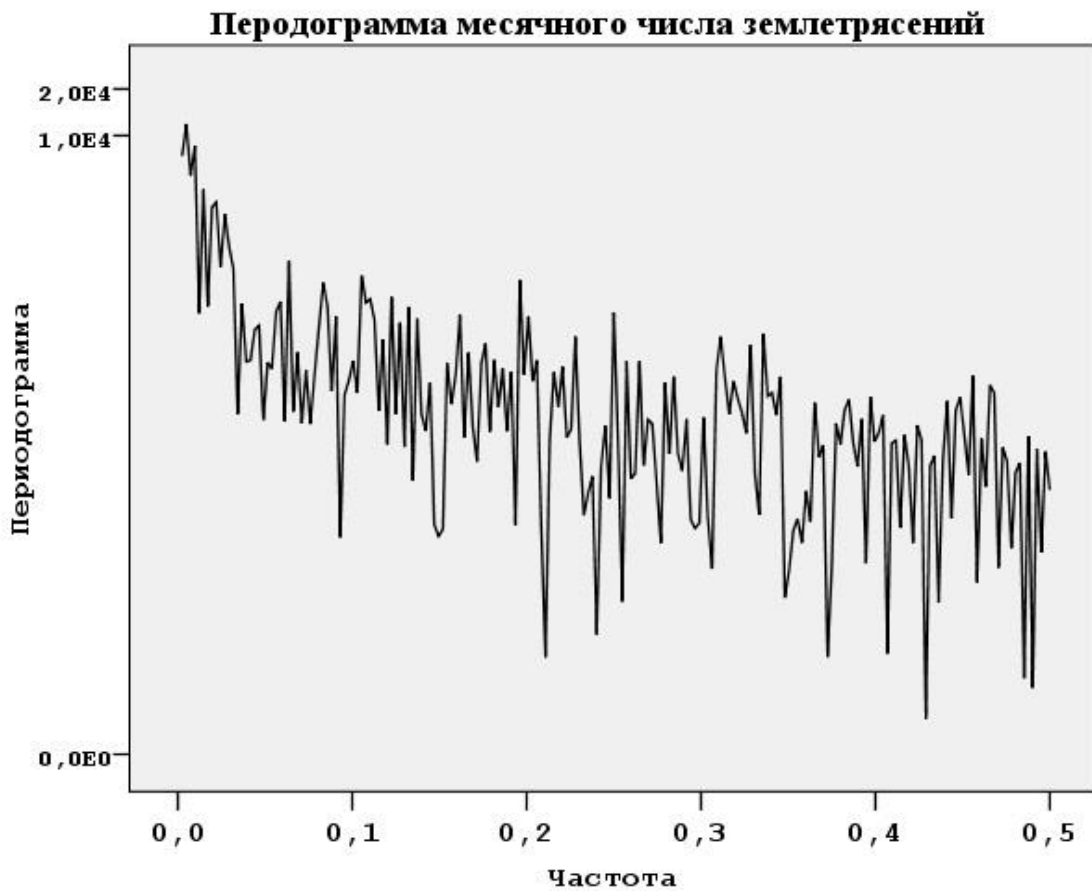


Рис. 6.5. Періодограма тимчасового ряду місячного числа землетрусів

Цей факт використовувався для отримання ряду, відповідного гармонійним коливанням з деякого діапазону частот. Для цього виключалися коливання з частотами поза заданого діапазону, і для суми залишилися гармоніки методом зворотного перетворення Фур'є і обчислювалися рівні відповідного ряду. Застосування даного способу вимагає знання частотної структури ряду. Такий метод очищення ряду від складових, з не бажаними частотами, що не поліпшили якість прогнозу.

Для моделювання нестационарних часових рядів застосовується модель авторегресії про інтегрованої змінної середнього, яка називається *ARIMA* в

англомовній літературі. Дана модель це один з методів оцінки невідомих параметрів і прогнозування часових рядів.

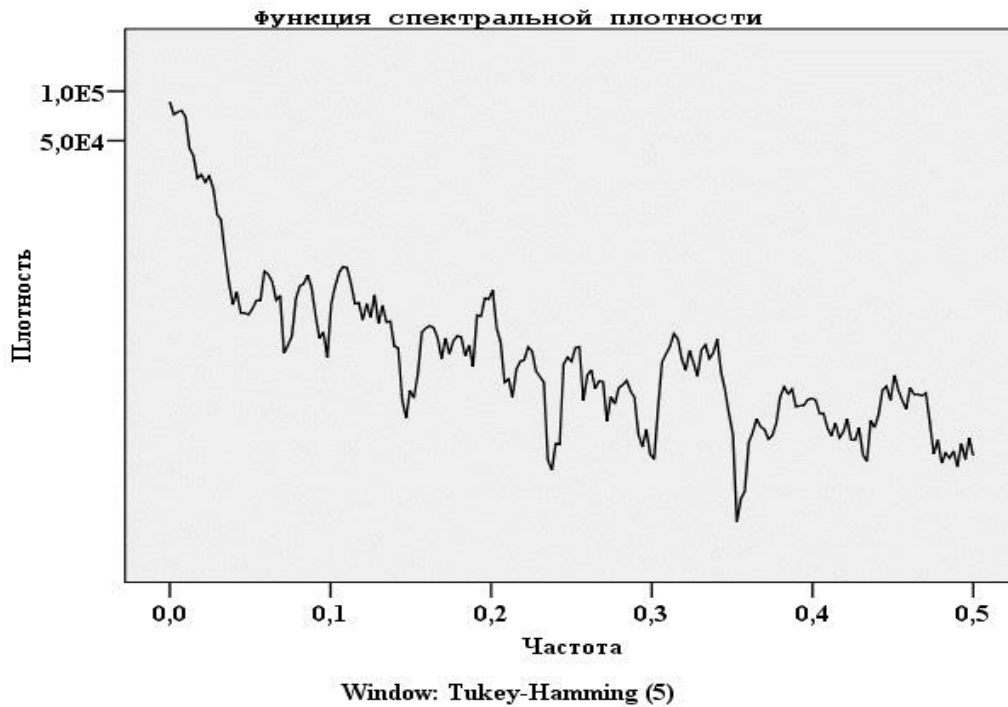


Рис. 6.6. Функція спектральної густини ряду

6.8. Лінійні моделі часових рядів

Лінійні моделі стаціонарних часових рядів: моделі авторегресії; моделі змінного середнього; моделі авторегресії і дотичного середнього. Будь-який слабо стаціонарний випадковий процес може бути як завгодно точно представлений процесом авторегресії (AR) – змінного середнього (MA) $ARMA(p,q)$ кінцевого порядку (p,q) [Бокс, Дженкінс, 1974; Дуброва, Архипова, 2004; Spezialvorlesung Zeitreihenanalyse, 2007]. Слабка стаціонарність: стаціонарність в широкому сенсі, в даному випадку означає, що часовий ряд характеризується постійністю математичного очікування, дисперсії і автокореляції.

Рівень тимчасового ряду x_t представленою моделлю авторегресії порядку p , представляється в такий спосіб [Nuxol, 2007; Spezialvorlesung Zeitreihenanalyse, 2007]:

$$x_t = \alpha_0 + \sum_{j=1}^p \alpha_j x_{t-j} + \varepsilon_t \quad (6.29)$$

де p – порядок моделі авторегресії, α_j – коефіцієнти моделі авторегресії, що підлягають оцінюванню, ε_t – білий шум.

Процеси $AR(p)$ і $ARMA(p,q)$ слабо стаціонарні тоді і тільки тоді, коли корені характеристичного рівняння менше одиниці [Карманов, 2007], $MA(q)$ процес завжди слабо стаціонарний (математичне очікування, дисперсія і автокореляція не залежать від зсуву за часом, тобто постійні). У моделі змінного середнього рівень часового ряду виражається алгебраїчною сумою членів ряду білого шуму з числом доданків q (порядок змінного середнього).

Модель авторегресії порядку p позначається як $AR(p)$. Значення ряду в даний момент часу є лінійною функцією його минулих значень. Модель авторегресії $AR(p)$ пов'язує змінну з власним минулим і "випадковим блуканням". кореляція r_k між рівнями x_t і x_{t-k} , де k – лаг, відповідний коефіцієнту авторегресії $\alpha_k = r_k/r_0$. Перевірка стаціонарності ряду здійснюється на основі критерію Дікі-Фуллера на одиничні корені [Кисляк, 2007; Toutenburg, Neumann, 2008]:

$$t_\alpha = \frac{\alpha - 1}{s_\alpha} \quad (6.30)$$

де α – коефіцієнт авторегресії, s_α – стандартна похибка оцінки. За цим критерієм гіпотеза про існування одиничного кореня відхиляється. З цього випливає, що часовий ряд стаціонарний. Адекватність $AR(p)$ і $ARMA(p,q)$ моделей часового ряду за спостереженими даними підтверджується за критерієм χ^2 .

Вид ACF і $PACF$ залежить від довжини тимчасового ряду. Якщо ряд містить кілька сотень спостережень, то форма автокореляційних функцій виявляється чітко вираженою, модель ряду і параметри моделі визначаються досить точно [Spezialvorlesung Zeitreihenanalyse, 2007]. Порядок AR визначається по поведінці $PACF$, а по поведінці ACF – порядок MA (Рис. 6.3–6.4). При наявності закономірності в ACF в модель включається тільки AR , а при наявності закономірності в $PACF$ в модель включається тільки MA . Що стосується порядків, то вони визначаються номером лага, на якому відбувається вихід за межу довірчого інтервалу.

Вище було показано, що часовий ряд є нестационарним. У деяких випадках нестационарний ряд неможливо описати ні трендом чи, ніякою конкретною формою сезонності. У цих випадках для перетворення ряду до стаціонарного ряду застосовується так зване зрушення ряду з фіксованим лагом. Найчастіше застосовується зрушення з лагом рівним 1. В цьому випадку ряд перетворюється в такий спосіб: $Z_t = Y_t - Y_{t-1}$ для всіх значень t . У літературі, присвяченій дослідженню часових рядів це перетворення позначається так: $Z_t = D(-1)Y_t$. Зрушення ряду з лагом 1 можна використовувати і для того, щоб виключити лінійний тренд, а кілька таких послідовних зрушень допоможуть виключити і статичні нелінійні тренди. Зрушення ряду з великим лагом можна використовувати для виключення сезонного фактору. Наприклад, якщо сезонні коливання мають річний період, а дані представлені щомісяця, то для виключення сезонності потрібно використовувати зрушення з лагом 12 - $D(-12)$.

Таким чином, виявляється, що будь-який часовий ряд можна перетворити в стаціонарний за допомогою операцій зсуву, а будь-який стаціонарний ряд описується за допомогою процесів авторегресії і дотичного середнього. Опис тимчасового ряду в термінах зсуву, авторегресії і дотичного середнього часто іменується процедурою Боксу-Дженкінса.

Нестационарні часові ряди можуть моделюватися як *ARIMA* моделі, засновані на 3-х складових процесу: авторегресії порядку p ; змінного середнього порядку q ; "Випадкового блукання" $\xi_t \sim N(0, \sigma^2)$. Моделі часових рядів будувалися за допомогою процедури "Конструктор моделей" пакета SPSS. Математична форма моделі має вигляд [Spezialvorlesung Zeitreihenanalyse, 2007]:

$$\Delta^d x_t = \alpha_0 + \sum_{j=1}^p \alpha_j \Delta^d x_{t-j} - \sum_{n=1}^q \beta_n \varepsilon_{t-n} + \xi_t \quad (6.31)$$

ARIMA модель застосовується для стаціонарних рядів, тому на першому етапі, за допомогою автокореляційної і приватної автокореляційної функцій ряду залишків, перевіряється гіпотеза про стаціонарність ряду в широкому сенсі. Залишки представляють собою різниці вихідного часового ряду і модальних

значень. Якщо ряд нестационарний, то він зводиться до слабо стаціонарному ряду, взяттям послідовних різниць деякого порядку:

$$\Delta x_j = x_j - x_{j-1}; \quad \Delta^2 x_k = \Delta x_k - \Delta x_{k-1}; \quad \Delta^d x_l = \Delta^{d-1} x_l - \Delta^{d-1} x_{l-1} \quad (6.32)$$

Зворотна дія – відновлення вихідного ряду – здійснюється d -кратним підсумовуванням проінтегрованого ряду. Пакет SPSS містить процедуру інтегрування тимчасового ряду і вибору відповідних порядків *ARIMA* – процесу (p , d , q) [Marques de Sá SPSS, 2007; Trends 14.0., 2006].

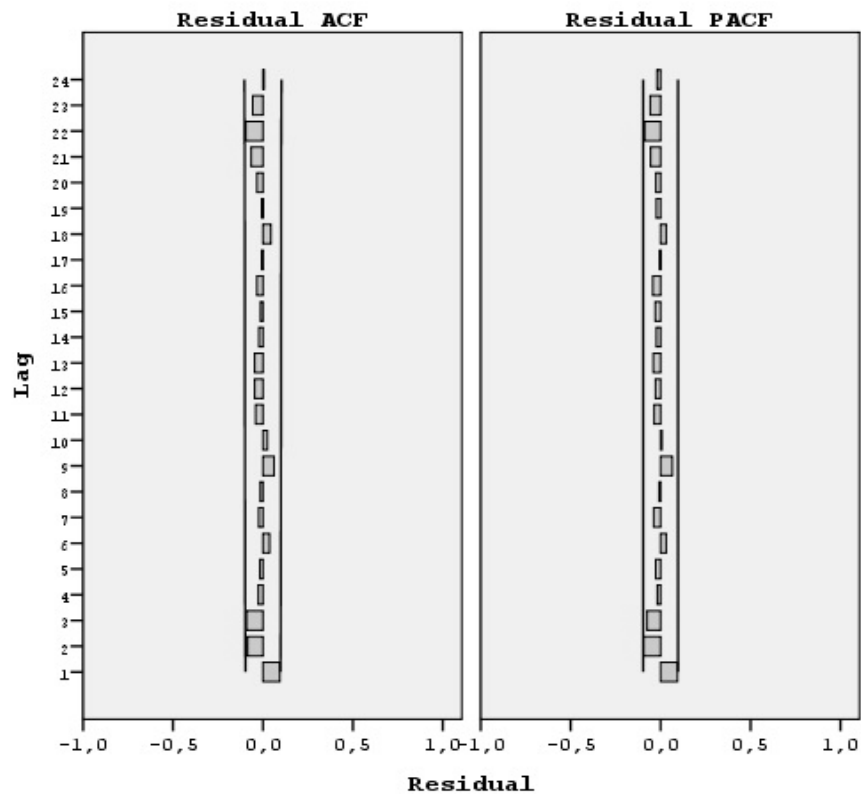


Рис. 6.7. Автокореляційна і приватна автокореляційна функції залишків.

Діагностична перевірка адекватності моделей заснована на аналізі ряду залишків. Якщо модель є адекватною, то залишки є "білим шумом". Це означає, що коефіцієнти кореляції між рівнями ряду залишків дорівнюють нулю, і нормована кумулятивна періодограма мало відрізняється від прямої лінії, що з'єднує точки (0;

0) і (0.5; 1) [Шанченко,2008]. У конструкторі моделей часових рядів розглядалися $(p,d,q) \times (P,D,Q)_s$ моделі при різних значеннях звичайних і сезонних параметрів. Для тимчасового ряду щомісячного числа землетрусів в конструкторі моделей пакета SPSS визначилася $ARIMA(0,1,2) \times (1,0,1)_s$ модель.

Найкращі моделі рядів визначаються по максимуму коефіцієнта детермінації, по оптимальних значеннях числових характеристик ряду залишків: $RMSE$; $MAPE$; MAE ; $MaxAPE$; $MaxAE$; BIC ; за значенням критерію Льюнг-Боксу; по автокореляційній і приватній автокореляційній функціях залишків (табл. 6.5). Правдоподібність нульової гіпотези H_0 про те, що коефіцієнти автокореляції між рівнями ряду залишків дорівнюють нулю, перевіряється за критерієм Льюнг-Боксу [Spezialvorlesung Zeitreihenanalyse, 2007; Teusch, 2006]:

$$Q_{LB} = T(T+2) \sum_{\tau=1}^m \frac{r_{\tau}^2}{T-\tau} \sim \chi^2(m) \quad (6.33)$$

Вибіркові статистики (6.33) менше критичного значення χ^2 розподілу з $(k-p-q)$ ступенями свободи, що відповідає рівню значущості $\alpha=0.05$. Величина спостережуваної ймовірності значущості 0.302, перевищує ймовірність похибки першого роду $\alpha=0.05$.

Графік автокореляційної і приватної автокореляційної функцій залишків знаходяться в межах 95%-ного довірчого інтервалу (рис. 6.7). Отже, модель $ARIMA(0,1,2) \times (1,0,1)_s$ є адекватною ряду даних щомісячного числа землетрусів. Графіки АКФ і ЧАКФ ряди залишків, отримані за допомогою простої сезонної і мультиплікативної моделі Уінтерса, виходять за межі 95%-ного довірчого інтервалу.

Таким чином, оптимальними по всьому набору тестових критеріїв перевірки якості моделі, оптимальними виявилися моделі: проста сезонна; мультиплікативна Уінтерса і $ARIMA(0,1,2) \times (1,0,1)_s$. З порівняння передбачених за трьома моделями значень можна стверджувати, що прогноз по моделі $ARIMA$ ближче до спостережуваних значень (табл. 6.6).

Таблиця 6.5

Вибіркові статистики критеріїв перевірки якості моделі

Model	Model fit statistics								Ljung-Box Q(18)		
	Stationary R-squared	R-squared	RMSE	MAPE	MAE	MaxAPE	MaxAE	Normalized BIC	Statistics	df	Sig.
Simple seasonal	0.624	0.490	11.037	61.05	7.028	5848.347	68.358	4.832	20.610	16	0.194
Winters' Additive	0.623	0,487	11,086	61,762	7,051	6004,718	68,258	4,856	20,598	15	0.150
$ARIMA(0,1,2) \times (1,0,1)_s$	0.216	0.468	11.309	66.082	7.121	6878.756	71.526	4,896	17,296	15	0.302

Таблиця 6.6

Передбачені значення ряду по моделі $ARIMA(3,1,5) \times (0,1,1)_s$

Місяць	1	2	3	4	5	6	7	8	9	10	11	12
2011	24	21	24	20	16	27	21	27	24	29	16	21
Простая сезонна. 2012	20	18	23	27	29	26	26	34	31	33	26	31
Мульти. Уінтерса. 2012	21	19	24	28	30	27	27	34	32	34	27	32
$ARIMA(0,1,2) \times (1,0,1)_s$	22	21	24	27	29	27	27	32	31	32	28	31

Зазвичай прогноз не повинен виходити за межі встановленого часу одного циклу [Кендал, 1981; Шанченко 142, 2008; Burtiev, 2014; Spezialvorlesung Zeitreihenanalyse, 2006]. Проведений аналіз тимчасового ряду і отриманий результат свідчать про доцільність застосування лінійних моделей на практиці і високому прогнозному потенціал $ARIMA$ моделі.

Висновки до розділу 6.

У ряді щомісячного числа землетрусів присутня довгострокова систематична складова, у автокореляційної функції найбільшим виявився пік на лагу 1. Припущення про стаціонарності за середнім значенням і по дисперсії відкидається. Наявність статистично значущих коефіцієнтів автокореляції означає, що часові ряди

не випадкові, і між послідовними рівнями існує певний зв'язок. Для прогнозування ряду застосовувався метод експоненціального згладжування. Найкращими, за вибірковими значеннями статистики критеріїв перевірки якості моделі, виявилися проста сезонна, для прологарифмованого ряду, і адитивна модель Уінтерса для ряду з квадратного кореня рівнів вихідного ряду. Для прогнозу застосовувалися також лінійні моделі авторегресії. Параметри моделей рядів визначалися з поведінки автокореляційної і приватної автокореляційної функцій. Авторегресійні методи (модель *ARIMA*) можна застосувати для аналізу і прогнозу однорідних нестационарних часових рядів з нелінійним трендом, типу полінома невисокого степеня. Найкращою моделлю для прогнозу часових рядів виявилася $ARIMA(0,1,2) \times (1,0,1)_s$ модель, підібрана конструктором моделей пакета SPSS. Час прогнозу не повинен перевищувати декількох періодів циклічних коливань. Сейсмічний процес – це циклічний процес. Результати досліджень були б більш достовірними, якби довжина ряду перевищувала, хоча б період одного сейсмічного циклу.

З огляду на поважний вік вогнищ землетрусів, який обчислюється мільйонами років, довжина ряду повинна перевищувати період повторення найбільшого, для даного вогнища землетрусу і крок квантування за часом повинен становити не тільки місяць, але і більш довгі інтервали часу. Під довжиною ряду мається на увазі час, що минув від першого до останнього спостереження в ряді. Однак, період інструментальних сейсмічних спостережень почався в кінці XIX століття, який тільки, в деяких випадках, охоплює період повторення найбільших землетрусів. Труднощі, пов'язані з нестачею відомостей про сейсмічні спостереження, можна вирішити створенням штучних каталогів землетрусів, за допомогою Монте-Карло моделювання.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

Агафонова Н. Ю. Анализ временных рядов. Саратов, 2010, 27.

Андерсон Т., Статистический анализ временных рядов. Мир, Москва, 1971, 746.

- Бендат Дж., А. Пирсол. Прикладной анализ случайных данных, Москва, Мир, 1989, 540.
- Бокс Дж., Дженкинс Г., Анализ временных рядов. Прогноз и управление. Вып. I. Москва. Мир. 1974. 406.
- Бриллинджер Д., Временные ряды. Обработка данных и теория. Мир, Москва, 1980, 532.
- Володин И. Н., Лекции по теории вероятностей и математической статистике, Казань, 2006, 270.
- Горбацевич В.В. Анализ и прогнозирование временных рядов МАТИ им. К.Э.Циолковского, МОСКВА, 2000, 25.
- Дуброва Т. А., Архипова М.Ю. Статистические методы прогнозирования в экономике: Москва. 2004. 136.
- Карманов В.С. Анализ временных рядов. Новосибирск, 2007, 28.
- Кендал М., Временные ряды. Москва. 1981. 198.
- Кисляк Н. В. Эконометрика. Екатеринбург. 2007, 157.
- Колмогоров А.Н., Фомин С. В., Элементы теории функций и функционального анализа. Москва, Наука, 1976, 542.
- Крамер Гаральд. Математические методы статистики. Москва. Мир. 1975. 648.
- Марно Вербик. Путеводитель по современной эконометрике. Москва 2008, 616.
- Никитин А.Я., Сосунова И. А., Анализ и прогноз временных рядов в экологических наблюдениях и экспериментах. Иркутск. 2003. 88.
- Сеньо П. С., Теорія ймовірностей та математична статистика. Київ, Знання, 2007, 558.
- Смирнов, И.В. Дунин-Барковский Н.В., Курс теории вероятностей и математической статистики. Наука, Москва, 1965, 511
- Пискунов Н.С., Дифференциальное и интегральное исчисление. Москва, Наука, 1964, 312.
- Хеннан Э., Многомерные временные ряды. Мир, Москва, 1974, 575.

- Христиановский В.В. Анализ временных рядов в экономике: практика применения: 2011, 127.
- Цапаева С. А. Ряды Фурье, Великий Новгород, 2011.
- Шанченко Н. И., Лекции по эконометрике. Ульяновск. 2008, 139.
- Ющенко Д.П., О. В. Якубович, Математический Анализ. Ряды Фурье, Гомель, 2008, 148.
- Backhaus K. et al., Multivariate Analysemethoden, Springer–Verlag Berlin Heidelberg, 2011, 120–154.
- Bernhardt Christine. Modellierung von Elektrizitätspreisen durch lineare Zeitreihenmodelle und Value–at–Risk–Schätzung mittels Methoden aus der Extremwerttheorie Technische Universität München. Zentrum Mathematik. München. 2007, 97.
- Burtiev R. Z., Time Series in the Study of Seismic Regime of Vrancea (Romania) Seismic Zone. The Global Environmental Engineers, 2014, Volume. 1. N2, Karachi, Pakistan, 54–63.
- Huxol Stephan, Trendanalyse von Zeitreihen der Komponenten des Wasserkreislaufes im Einzugsgebiet der Dreisam zur prozessorientierten Beurteilung hydrologischer Klimafolgen, Institut für Hydrologie der Albert-Ludwigs-Universität Freiburg i. Br., 2007, 108.
- Marques de Sá P., Joaquim, Applied Statistics Using SPSS, STATISTICA, MATLAB and R, Springer, 2007, 520.
- Spezialvorlesung Zeitreihenanalyse–Mit Beispielen in Mathematica Institut für Stochastik, Johannes Kepler Universität Linz. Linz, 2006, 277.
- SPSS Trends 14.0. Copyright © 2005 SPSS Inc. Chicago. 2006, 165.
- Teusch Annette. Einführung in die Spektral– und Zeitreihenanalyse mit Beispielen aus der Geodäsie. Verlag der Bayerischen Akademie der Wissenschaften München, 2006, 167.

Toutenburg Helge, Heumann Christian, Deskriptive Statistik. Eine Einführung in Methoden und Anwendungen mit R und SPSS. Springer-Verlag Berlin-Heidelberg. 2008. 127.

РОЗДІЛ 7

ЙМОВІРНІСНИЙ АНАЛІЗ СЕЙСМІЧНОЇ НЕБЕЗПЕКИ

ЙАСН включає в себе чотири основні етапи. На першому етапі, як і в ДАСН, визначаються сейсмічні зони, і передбачається, що вогнища землетрусів розподілені рівномірно в геофізичному просторі, яке займає зона вогнищ землетрусів. У цьому полягає основна відмінність від ДАСН, де фіксується найближчий до об'єкту активний розлом, тобто розлом, в якому з ймовірністю, що дорівнює одиниці може виникнути осередок землетрусу.

На другому етапі будується марковська модель сейсмічного режиму. На основі ергодичного неперервного марковського ланцюга можна передбачити магнітуду, місце виникнення землетрусів і їх середнє число в фіксований період часу. У ДАСН цьому відповідає закон повторюваності Гуттенберга-Ріхтера.

Третім кроком є визначення параметрів загасання сейсмічного ефекту. В даному випадку аналіз сейсмічної небезпеки проводиться на основі бала макросейсмічною шкали і по будівельним нормам *EC8*. Передбачається, що відхилення (помилка вимірювань) спостережених макросейсмічних значень інтенсивності сейсмічних впливів від їх теоретичних значень є випадковою величиною. На величину помилок вимірювань впливає безліч неврахованих, незалежних факторів, і жоден з них не домінує – кожен вносить незначний по величині внесок в загальну помилку. Відповідно до теорії похибок, вони носять випадковий характер, розподілені за нормальним законом і піддаються обробці за допомогою апарату математичної статистики [Тейлор, 1985].

На четвертому етапі для конкретного району, який знаходиться в радіусі дії групи сейсмічних зон, обчислюється сумарна міра сейсмічної небезпеки. Для сейсмічного районування територія, що знаходиться в зоні впливу сейсмічних зон, покривається географічної мережею з деяким кроком дискретності. У вузлах мережі

визначається міра сейсмічної небезпеки по прийнятим нормам. Фактор випадковості присутній на всіх етапах проведення ЙАСН.

Запропонований метод ймовірнісного аналізу сейсмічної небезпеки був випробуваний для територій Молдови й інших країн, а також Туреччини. Сейсмічна небезпека території Молдови, Румунії, і здебільшого сусідніх країн визначається сейсмічним впливом вогнища землетрусів Вранча.

Перші карти сейсмічної небезпеки Молдови, засновані на сейсмологічних і геологічних даних, було складено Інститутом Фізики Землі Академії наук СРСР в 1957. Південно-західна частина Молдови, біля кордону з Румунією, була віднесена в VIII бальну зону, інша частина території знаходиться в зоні впливів з інтенсивністю до 7 балів. Згідно з картою *CP-69* (Сейсмічне районування території СРСР, 1969), були здійснені деякі виправлення в попередніх картах сейсмічного районування – для північної і північно-східної частин території Молдови, де рівень сейсмічної небезпеки був знижений до 6 балів. Карта *CP-78* (Сейсмічне районування території СРСР, 1978) сейсмічного районування території республіки Молдови все ще є нормативним документом. Карта *CP-78* включає ізолінії сейсмічної інтенсивності 6, 7 і 8 балів, відповідно за сейсмічною шкалою *MSK-64* [Алказ, 2007].

Карта ймовірнісного аналізу сейсмічної небезпеки (*ЙАСН*) корисна для розробки стратегії сейсмостійкого будівництва в зоні, що знаходиться в радіусі руйнівного впливу землетрусів. Багатьма дослідниками було запропоновано різні методи розрахунку сейсмічної небезпеки [Уломов, 1993; Уломов, Богданов, 2013; Golbs, 2008; Baker, 2008]. Однак, існуючі міри сейсмічної небезпеки недостатньо інформативні для оптимального планування і будівництва в сейсмонебезпечних областях. Дослідження з аналізу сейсмічної небезпеки і за розрахунком сейсмічного ризику проводяться для того, щоб за їх результатами можна було прийняти адекватне рішення.

7.1. Оцінка сейсмічної небезпеки

Припустимо, що пункт спостереження знаходиться в радіусі впливу S сейсмічних регіонів $Z^s, s=1, \dots, S$. Число станів в регіоні з номером s позначається через Z^s . Подія, що складається в тому, що відбувається землетрус з параметрами зі стану Σ_j , позначається тим же ідентифікатором. Вище було сказано, що параметри землетрусів можуть належати одному і тільки одному *ЕДЗ* (станом), тобто, маємо повну групу несумісних подій, які в об'єднанні дають весь простір можливих розподілів параметрів землетрусів:

$$\{\Sigma_{r_1}, \Sigma_{r_2}, \dots, \Sigma_{r_N}\}, \quad (\Sigma_i \cap \Sigma_j = 0, i \neq j), \quad \Xi_s = \sum_{i=1}^{R_s} \Sigma_i \quad (7.1)$$

Крім того, припускаємо, що подія $B^k = \{I^k - 0.5 \leq I < I^k + 0.5\}$, що складається у виникненні сейсмічних поштовхів з інтенсивністю I^k в деякій точці земної поверхні, може статися спільно з кожним з R^s взаємно виключаючих подій $\Sigma_r, r=1, \dots, R^s$. У цьому випадку має місце вираз:

$$B^k = \sum_{r=1}^{R_s} B^k \Sigma_r \quad (7.2)$$

де події $B_k E_i$ і $B_k E_j$ з різними індексами i та j є несумісними. Подія B_k і повна група подій (7.1) пов'язані формулою повної ймовірності:

$$P(B^k) = \sum_{j=1}^{R_s} P\left(\frac{B^k}{\Sigma_j}\right) P(\Sigma_j) \quad (7.3)$$

де умовна ймовірність:

$$p_r^k = P(B^k / \Sigma_r) = P\{(I^k - 0.5 \leq I < I^k + 0.5) / \Sigma_r\} \quad (7.4)$$

інтенсивності сейсмічного впливу I , викликаного землетрусом з *ЕДЗ* Σ_r відповідає нормальному закону. Введемо вірогідну матрицю (7.5), елементами якої є умовні ймовірності (7.3). В цьому випадку множенням зліва вектор розподілів станів вкладеного марковського ланцюга на матрицю (7.4) визначиться вектор розподілів інтенсивності сейсмічних впливів:

$$P_s = \begin{pmatrix} P_1^1 & P_1^2 & \cdot & \cdot & P_1^{12} \\ P_2^1 & P_2^2 & \cdot & \cdot & P_2^{12} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ P_{R_s}^1 & P_{R_s}^2 & \cdot & \cdot & P_{R_s}^{12} \end{pmatrix} \quad (7.5)$$

$$\vec{p}_s = \vec{\pi}_s P_s = \left(\sum_{r=1}^{R_s} \pi_r P_r^1, \sum_{r=1}^{R_s} \pi_r P_r^2, \dots, \sum_{r=1}^{R_s} \pi_r P_r^{12} \right) = (p_1^s, p_2^s, \dots, p_{12}^s) \quad (7.6)$$

Землетрус, з ймовірністю γ_s , може статися в одному з S сейсмічних регіонів Z^s . Випадкові події, що складаються у виникненні землетрусу в сейсмічних зонах, утворюють повну групу несумісних подій. Якщо точка (φ, ψ) знаходиться в межах сейсмічного впливу групи регіонів Z^s , $s=1, \dots, S$, то сумарна ймовірність інтенсивності струшувань може бути оцінена за формулою повної ймовірності:

$$\beta_k = \sum_{s=1}^S p_k^s \gamma_s; \quad \vec{\beta} = (\beta_1, \dots, \beta_{12}) \quad (7.7)$$

Нехай в пункті спостереження (φ, ψ) за час t з ймовірністю $p(t, n)$ відбуваються n сейсмічних поштовхів. Тоді умовна ймовірність появи m_1 поштовхів з інтенсивністю I^1 , m_2 раз з інтенсивністю I^2, \dots , і m_{12} випадків з інтенсивністю I^{12} , $(m_1 + m_2 + \dots + m_{12} = n)$ відповідає поліноміальній схемі [Гнеденко, 1961]:

$$p_n(m_1, m_2, \dots, m_{12} / n) = \frac{n!}{m_1! m_2! \dots m_{12}!} \beta_1^{m_1} \beta_2^{m_2} \dots \beta_{12}^{m_{12}} \quad (7.8)$$

А виникнення m_k сейсмічних поштовхів з інтенсивністю I^k , при спостереженні n сейсмічних подій, буде відповідати біноміальній схемі [Birkner, 2009; Luchsinger, 1997; Peichl, 1999]:

$$p_n(m_k) = \frac{n!}{m_k! (n - m_k)!} \beta_k^{m_k} (1 - \beta_k)^{n - m_k} \quad (7.9)$$

Множенням ймовірності (7.9) на $p(t, n)$ прийдемо до виразу для безумовної ймовірності:

$$P_{\varphi,\psi}(t, n, m, I^k) = p(t, n) \frac{n!}{m!(n-m)!} \beta_k^m (1-\beta_k)^{n-m} \quad (7.10)$$

того, що в географічній точці $Q(\varphi, \psi)$ за час t трапиться n сейсмічних поштовхів, m з яких з інтенсивністю I^k . Якщо відбувається n землетрусів, то середнє число сейсмічних поштовхів m_k інтенсивності I^k , $k=0,1,\dots,12$ задається формулою [Гнеденко, 1961]:

$$E(m_k) = n\beta_k \quad (7.11)$$

При великих значеннях n и m_k , згідно локальній граничній теоремі Муавра-Лапласа, для ймовірності (7.9) має місце асимптотичний вираз:

$$P_n(m) = \frac{n!}{m!(n-m)!} \beta_k^m (1-\beta_k)^{n-m} = \frac{1}{\sqrt{2\pi n\beta_k(1-\beta_k)}} e^{-\frac{x^2}{2}} \quad (7.12)$$

де:

$$x = \frac{m - n\beta_k}{\sqrt{n\beta_k(1-n\beta_k)}} \quad (7.13)$$

Інтегральна гранична теорема Муавра-Лапласа застосовується для обчислення ймовірності виникнення в пункті спостереження не менше m_1 і не більше m_2 сейсмічних струшувань з інтенсивністю I^k [Гнеденко, 1961; Колемаев, Калинина, 1997]:

$$P\{m_1 \leq m \leq m_2\} = P\{x_1 \leq x \leq x_2\} = \frac{1}{\sqrt{2\pi}} \int_{x_1}^{x_2} e^{-\frac{x^2}{2}} dx = \Phi(x_2) - \Phi(x_1) \quad (7.14)$$

де:

$$x_1 = \frac{m_1 - n\beta_k}{\sqrt{n\beta_k(1-n\beta_k)}}; \quad x_2 = \frac{m_2 - n\beta_k}{\sqrt{n\beta_k(1-n\beta_k)}} \quad (7.15)$$

Наприклад, ймовірність того, що в Кишиневі за 475 років трапиться $n=43477$ сейсмічних поштовхів, і 350 з них з інтенсивністю 5 балів становить: $P(475,43477,350,5)=0.00041$. Якість аналізу сейсмічної небезпеки в значній мірі залежить від функції загасання сейсмічного ефекту. Характер загасання інтенсивності сейсмічних впливів визначається за результатами макросейсмічних досліджень наслідків землетрусів, які містять похибки - випадкові і систематичні.

Для аналізу сейсмічної небезпеки за нормами ЕС8, визначається річна

інтенсивність перевищень $\lambda_k = \lambda(I^k)$ порога інтенсивності I^k і деякого рівня параметрів зсувів ґрунту a_k . Для обчислення ймовірності перевищення порога I^k , інтенсивності сейсмічних впливів у балах шкали *MSK-64*, використовується вираз:

$$p_r^k = P(I \geq I^k / \Sigma_r) = \frac{1}{\sigma \sqrt{2\pi}} \int_{I_k}^{12} e^{-\frac{(I-I)^2}{2\sigma^2}} dI \quad (7.16)$$

а в параметрах пікових прискорень:

$$p_r^k = P\left(\ln pga_h \geq \ln a_k / \Sigma_r\right) = 1 - \frac{1}{\sigma \sqrt{2\pi}} \int_{\ln a_k}^{\ln a_{\max}} e^{-\frac{(\ln a - \ln pga_h)^2}{2\sigma^2}} dpga \quad (7.17)$$

Застосовуючи вищенаведений алгоритм (7.2–7.8) отримаємо ймовірність (7.18), того що в точці (ϕ, ψ) ефект сейсмічного впливу, викликаного землетрусами з групи *S* сейсмічних зон, перевищить поріг інтенсивності I^k або деякий рівень параметра зсувів ґрунту:

$$\beta_k = \sum_{s=1}^S p_k^s \gamma_s, \quad (7.18)$$

Інтенсивність перевищень рівня інтенсивності I^k і параметра зсуву a_k за час t визначається за формулою:

$$\lambda_k = \lambda t \beta_k \quad (7.19)$$

де β_k задано формулою (7.18). Передбачається, що виникнення сейсмічних впливів в деякій точці з деякою інтенсивністю I^k або ж із значенням параметра зсувів ґрунту a , що перевищує заданий рівень a_k , являє собою пуассонівський процес:

$$p(t, n) = e^{-\lambda_k} \frac{(\lambda_k)^n}{n!} \quad (7.20)$$

Ймовірність того, що за 50 років в деякій сейсмічній зоні хоча б один раз трапиться землетрус, який в деякій точці (ϕ, ψ) викличе вплив, що перевищує поріг інтенсивності I^k балів, або ж рівень параметра зсувів ґрунту a_k , виражається формулою:

$$1 - P(50, 0) = 1 - e^{-\lambda_k 50} \quad (7.21)$$

7.2. Розподіл інтенсивності сейсмічних впливів

Якщо до теоретичного значення інтенсивності (\bar{I}) додати похибку вимірювань ε , ми отримаємо реальне значення інтенсивності як суму двох незалежних змінних:

$$I = \bar{I} + \varepsilon \quad (7.22)$$

де ε – похибка вимірювань. Складовими похибки є багато факторів: похибка експерта при огляді пошкоджень; похибка при класифікації типу будівель; похибки при визначенні геологічних умов середовища, внесок неврахованих факторів і т.д. Функція загасання інтенсивності є апроксимацією реального значення сили поштовхів. Передбачається, що сумарна похибка, отже, і оцінка I , є нормально розподіленою випадковою величиною $N(a, \sigma^2)$.

Рівняння макросейсмічної поля:

$$\bar{I} = am - b \lg \sqrt{h^2 + r^2} + c \quad (7.23)$$

визначає криву, яка апроксимує спостережені точки макросейсмічної поля [Шебалин, 1961]. Статистичний аналіз був проведений автором на спостережених даних наслідків вранчських землетрусів на території республіки Молдова (табл. 7.1). Ми припускаємо, що дані похибок вимірювання ($\varepsilon_1; \varepsilon_2; \dots; \varepsilon_n$) є незалежними і однаково (*iid*, $N(\mu; \sigma)$) розподіленими випадковими величинами.

Завдання розрізнення розподілів є однією з основних задач математичної статистики. Одним з ключових інструментів при вирішенні даного завдання є лема Неймана-Пірсона, яка дозволяє знайти рівномірно найбільш потужний критерій для розрізнення двох довільних розподілів при використанні відношення правдоподібності. Розглянемо нульову гіпотезу $H_0: \mu = \mu_0, \sigma = \sigma_0$, при простій альтернативній гіпотезі $H_1: \mu = \mu_1, \sigma = \sigma_1$. Перевірка правдоподібності робочої гіпотези H_0 проводиться на основі леми Неймана-Пірсона [Кендалл, Стьюарт, 1973; Уилкс, 1967; Soong, 2004]. Відповідно до цієї леми, найкраща критична область задається множиною ($\varepsilon_1; \varepsilon_2; \dots; \varepsilon_n$) [Birkner, 2009; Luchsinger, 1997]:

$$T = \frac{L(\varepsilon / H_0)}{L(\varepsilon / H_1)} \leq k_\alpha \quad (7.24)$$

для всіх:

$$(\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n \in C) \quad (7.25)$$

Найкраща критична область містить всі вибірки, для яких значення відношення правдоподібності, відповідно до кожної гіпотези, вище певного порогу k_α , який залежить тільки від рівня значущості.

Ми скористаємося методом відношення правдоподібності для перевірки гіпотез [Кендалл, Стьюарт, 1973; Luchsinger, 1997]:

t -критерій, коли дисперсія невідома

F - критерій, коли середнє невідомо.

Область відхилення гіпотези $H_0: \mu=0$ при альтернативній гіпотезі $H_1: \mu \neq 0$ визначається виразом [Jackman, 2006; Soong, 298]:

$$A = \{|T|\} > t_{\frac{\alpha}{2}} \quad (7.26)$$

Формула (7.26) містить вибіркоче значення t -статистики (7.24):

$$T = \frac{\bar{X} - \mu_0}{S / \sqrt{n}} \quad (7.27)$$

Якщо вибіркоче значення t -статистики $abs(T_q)=0.12$ менше критичного значення $t=1.96$, двостороннього t -критерію з $df = 572$ ступенями свободи, що відповідає рівню значущості $q=0.025$, то гіпотеза $H_0: E(x)=0$ є правдоподібною.

$H_0: \sigma=\sigma_0, H_1: \sigma \neq \sigma_0, \mu=\bar{u}=-0.016, s=\sigma=0.67.$

$$T = \frac{(n-1)S^2}{\sigma_0} \quad (7.28)$$

Критична область – область відхилення гіпотези H_0 задається формулою:

$$A = \left\{ T > \chi_{\frac{\alpha}{2}}^2 \right\} \cup \left\{ T < \chi_{1-\frac{\alpha}{2}}^2 \right\} \quad (7.29)$$

Якщо вибіркова статистика $\chi_q=623.5$ менше критичного значення $\chi_{1,0.025}=507.6$, χ^2 розподілу відповідає рівень значущості $q=0.025$, с $df=573$ ступенями свободи, або більше величини 640.17, χ^2 критерію, то гіпотеза H_0 : про те, що дисперсія дорівнює одиниці, повинні бути відхилені.

Таблиця 7.1

Параметри землетрусів

Дата	широта	довгота	глибина	магнітуда (mb)	обсяг вибірки
4.03.1977	45.34	26.30	109	7.2	205
30.08.1986	45.53	26.47	132	7.0	238
30.05.1990	45.82	26.90	91	6.7	130
31.05.1990	45.83	26.89	79	6.1	49

Правдоподібність нульової гіпотези $H_0: \mu=0, \sigma=1$, засвідчується за допомогою леми Неймана-Пірсона при простій альтернативній гіпотезі $H_1: \mu_0 \neq 0, \sigma_0 \neq 1$. Критична область C має потужність α , якщо похибка I роду становить α . Для великих значень величина $-2\ln T$ приблизно χ^2 розподілена з числом ступенів свободи $r_0 - r$ [Luchsinger, 1997]. Число ступенів свободи граничного χ^2 розподілу дорівнює різниці [Ferber, 2010; Grachyova, Müller, 2006; Peichl, 1999]:

- числа вільних параметрів чисельника r_0 , t - статистики T
- числа вільних параметрів r знаменника статистики T .

З відношення ймовірностей правдоподібності отримаємо нерівність для визначення критичної області прийняття нульової гіпотези [Ferber, 2010; Grachyova, Müller, 2006; Jackman, 2006; Luchsinger, 1997; Peichl, 1999]:

$$T = \frac{L(u/H_0)}{L(u/H_1)} = \left(\frac{\sigma_1}{\sigma_0}\right)^n \exp \left\{ - \left[\sum_{i=1}^n \frac{(u_i - \mu_0)^2}{\sigma_0^2} - \sum_{i=1}^n \frac{(u_i - \mu_1)^2}{\sigma_1^2} \right] \right\} \leq k_\alpha \quad (7.30)$$

Підставами вибірових значень $\mu_1 = \bar{u}$; $\sigma_1 = s$ в формулу для відношення правдоподібності (7.24), яке після перетворення приймає форму [Burtiev, 2012]:

$$\bar{u}^2 - \log s^2 + s^2 - 1 \geq \frac{-2 \log k_\alpha}{n} \quad (7.31)$$

Число ступенів свободи величини $-2\ln T$, при альтернативній гіпотезі $H_1: \mu=0.016$; $\sigma=0.67$ становить $r_0 - r = 2$, так як, параметри нормального розподілу $\mu=0.016$ і $\sigma=0.67$ є вибіровими значеннями $r_0=2$ мірної випадкової величини (\bar{u}, s) з простору параметрів θ , і $r=0$.

Вибіркове значення критерію відносно максимальної правдоподібності $T=9.4$ перевищує порогове значення $\chi_0=0.01$, χ^2 розподілу. Отже, гіпотеза H_0 : про те, що $E(\varepsilon)=0$, и $D(\varepsilon)=1$ повинна бути відкинута на користь гіпотези H_1 : $\mu=0.016$; $\sigma=0.67$.

Густина спільного розподілу (\bar{u}, s) має вигляд [Кендалл, Стьюарт, 1973; Уилкс, 1967; Fisz, 1966]:

$$h(\bar{u}, s) = \frac{\sqrt{n}}{\sigma\sqrt{2\pi}} e^{-\frac{n\bar{u}^2}{2\sigma^2}} \frac{n^{\frac{n-1}{2}} s^{n-2} e^{-\frac{ns^2}{2\sigma^2}}}{2^{\frac{n-3}{2}} \Gamma\left(\frac{n-1}{2}\right) \sigma^{n-1}} \quad (7.32)$$

7.3. Метод побудови карти сейсмічної небезпеки

Осередок землетрусів охоплює деякий простір геофізичної середовища. Якщо цей простір і діапазон магнітуди землетрусів розбити на непересічні 4-мірні осередки, то параметри кожного землетрусу будуть належати одній, і тільки одній клітинці. Назвемо ці осередки елементарними джерелами землетрусів (ЕДЗ).

Для сейсмічного районування територія, що знаходиться в зоні впливу сейсмічних зон, покривається географічної мережею з деяким кроком дискретності. У вузлах мережі визначається міра сейсмічної небезпеки по прийнятим нормам. Наприклад, в якості запобіжного сейсмічної небезпеки пропонується ймовірність:

$$P_{(\varphi, \psi)}(t, n, m, I^k) \quad (7.33)$$

того, що в точці земної поверхні $Q(\varphi, \psi)$, з географічними координатами (φ, ψ) , за час t трапиться n сейсмічних струсів, m з яких з інтенсивністю не менше I^k балів [Burtiev, 1986, 2012, 2016, 2017]. Складно будувати карти сейсмічного районування ґрунтуючись на міру (7.32), тому що вона є неперервною величиною. Для спрощення цього завдання вогнищева зона розглядається як об'єднання ЕДЗ.

Далі будується марковська модель потоку сейсмічних подій кожного джерела землетрусів. В цьому випадку ймовірнісна модель сейсмічного режиму осередкової зони складається з суперпозицій ймовірнісних моделей ЕДЗ. Проекція цих зон на земну поверхню матиме форму прямокутника. Закон загасання інтенсивності пов'язує з сейсмічністю деякої зони Z^s інтенсивність поштовхів в точках земної

поверхні. Об'єднання всіх ЕДЗ представляє зону вогнищ землетрусів $\Xi^s = \{\Sigma_v\}$, $v=1, \dots, R^s$, де R^s – число ЕДЗ. Далі, простір Ξ^s можна розбити на два підпростори $\Xi^s = \Xi^1 \cup \Xi^2$, де Ξ^2 є об'єднанням джерел землетрусів, здатних викликати в точці $Q(\varphi, \psi)$ сейсмічний поштовх з інтенсивністю не менше балів I^k шкали інтенсивності.

$$I_{\Sigma_v}(I \geq I^k) = \begin{cases} 1, I \geq I^k \\ 0, I < I^k \end{cases} \quad (7.34)$$

$$(7.35)$$

де (7.35) – функція Хевісайда; R^s – число ЕДЗ; $k=0, 1, \dots, 12$. Далі, виникнення в точці $Q(\varphi, \psi)$ m сейсмічних поштовхів з інтенсивністю не менше I^k , за їх загальної кількості n буде відповідати біноміальній схемі. Умовна ймовірність даної події дорівнює:

$$P(n, m, I^k) = C_n^m p_k^m (1 - p_k)^{n-m} \quad (7.36)$$

Помноживши на ймовірність $P(t, n)$ отримаємо вираз для обчислення безумовної ймовірності:

$$P_{(\varphi, \psi)}(t, n, m, I^k) = P(t, n) P(n, m, I^k) \quad (7.37)$$

$$P_{(\varphi, \psi)}(t, n, m, I^k) = P(t, n) \frac{n!}{m!(n-m)!} p_k^m (1 - p_k)^{n-m} \quad (7.38)$$

Карти сейсмічної небезпеки (рис. 7.1) побудовані на основі мір (7.34), де ізолініями розділені зони зміни інтенсивності в балах. Карти СН, побудовані на основі мір (7.34), підходять для простого ранжування території за ступенем інтенсивності сейсмічних впливів майбутніх землетрусів. Додатково для точок досліджуваної території розрахована і більш інформативна оцінка СН (7.1).

Запропонований метод аналізу сейсмічної небезпеки дозволяє враховувати власний закон загасання інтенсивності струсів для кожної сейсмічної зони і територій.

Сейсмічне зонування території Румунії здійснювалося за допомогою кластерного аналізу, який має безлич методів оптимізації кластерного рішення. Рішення в 13 кластерів (сейсмічних зон) виявилось оптимальним домінуючим рішенням для території Румунії на основі каталогу ROMPLUS. Тобто проводиться

сейсмічне районування території, що знаходиться в радіусі дії групи з 13 сейсмічних зон.

Загасання інтенсивності сейсмічних поштовхів в балах шкали *МСК-64*, викликаних всіма підкоровими землетрусами на досліджуваній території, оцінюється по функції загасання вранчських землетрусів (5.58):

$$\bar{I} = 1.6M_w - \frac{b_{\max} b_{\min}}{\sqrt{b_{\min}^2 \cos^2(\gamma - 51^\circ) + b_{\max}^2 \sin^2(\gamma - 51^\circ)}} \lg \sqrt{h^2 + r^2} + 7.2 \quad (7.39)$$

де $b_{\max}=5.6$, $b_{\min}=4.9$, γ – поточне значення кута, який утворює радіус-вектор точки (φ , ψ), з позитивною піввіссю абсцис .

Розрахунок СН за нормами *ЕС8* здійснюється на основі функції загасання пікових прискорень [Lungu et al., 1997, 2002; Sokolov et al., 2007]:

$$\hat{a}_h = \ln pga_h = c_1 + c_2 M_w + c_3 \ln \sqrt{r^2 + h^2} + c_4 h + \varepsilon \quad (7.40)$$

де: pga_h – горизонтальна складова максимального прискорення ґрунту; M_w – моментна магнітуда; r – епіцентрально відстань; h – глибина вогнища; c_1, c_2, c_3, c_4 – коефіцієнти; ε – випадкова величина з нульовим середнім і стандартним відхиленням $\sigma_\varepsilon = \sigma_{\ln PGA}$ (стандартне відхилення змінної $\sigma_{\ln PGA}$). Величина ε дорівнює нулю при оцінці середніх значень, і стандартного відхилення $\sigma_{\ln PGA}$ при оцінці суми середніх значень і стандартного відхилення. Параметри моделі загасання були оцінені з використанням різних наборів даних, які представляють сектори, розташованих в трьох напрямках із зони Вранча: Молдова (азимут $<90^\circ$, північний сектор); Черновода (азимут $90^\circ-180^\circ$, південно-східний сектор); Бухарест (азимут $180^\circ-270^\circ$, південно-західний сектор). Загасання пікових прискорень для землетрусів у межах земної кори виражається функцією [Sokolov et al., 2007]:

$$\hat{a}_h = \ln pga_h = -1.39 + 0.266M_w - 0.922 \ln \sqrt{r^2 + 3.5^2} + \varepsilon \quad (7.41)$$

Таблиця 7.2

Значення коефіцієнтів функції загасання пікового прискорення

Коефіцієнти	Всі дані	0^0-90^0	90^0-180^0	180^0-270^0
c_1	5.571	4.150	8.136	6.470
c_2	0.937	0.913	0.876	0.923
c_3	-1.256	-0.962	-1.657	-1.403
c_4	-0.0069	-0.006	-0.0076	-0.007
$\sigma_{\ln PGA}$	0.398	0.415	0.348	0.366

Початок системи координат поєднується з середньою точкою ЕДЗ - середня точка тривимірних інтервалів $\{F_i x \Psi_j x H_k\}$, $i=1, \dots, R_\varphi$; $j=1, \dots, R_\psi$; $k=1, \dots, R_h$. На основі отриманих значень побудована карта сейсмічної небезпеки території Молдови, Румунії та Болгарії (рис. 7.1). Карти сейсмічного районування (рис. 7.1, 7.2(а), 7.3, 7.4, 7.5 (а), 7.6) складаються з точок, в яких ймовірність (7.38) позитивна.

Легенда карт сейсмічної небезпеки, представлених на малюнках (7.1–7.6):

- зліва і внизу вказані географічні широти і довготи з кроком в 1 градус для Молдови, Румунії, Болгарії та 5 градусів для Туреччини. Справа показана числова шкала відмінностей
- на ізолініях вказані значення макросейсмічних балів або значення пікових прискорень в см/сек^2
- сейсмічні станції позначені символом (▲)
- державні кордони позначені символами (+)
- центри кластерів (сейсмічних зон), з номером внизу, позначені знаком (■)
- k -а зона інтенсивності, в балах шкали *МСК-64*, складається з точок, в яких інтенсивність сейсмічних впливів I приймає значення з інтервалу $[I^k-0.5, I^k+0.5]$. Наприклад, 7 бальна зона складається з точок, в яких розрахункова інтенсивність задовольняє умові: $6.5 \leq I \leq 7.5$.

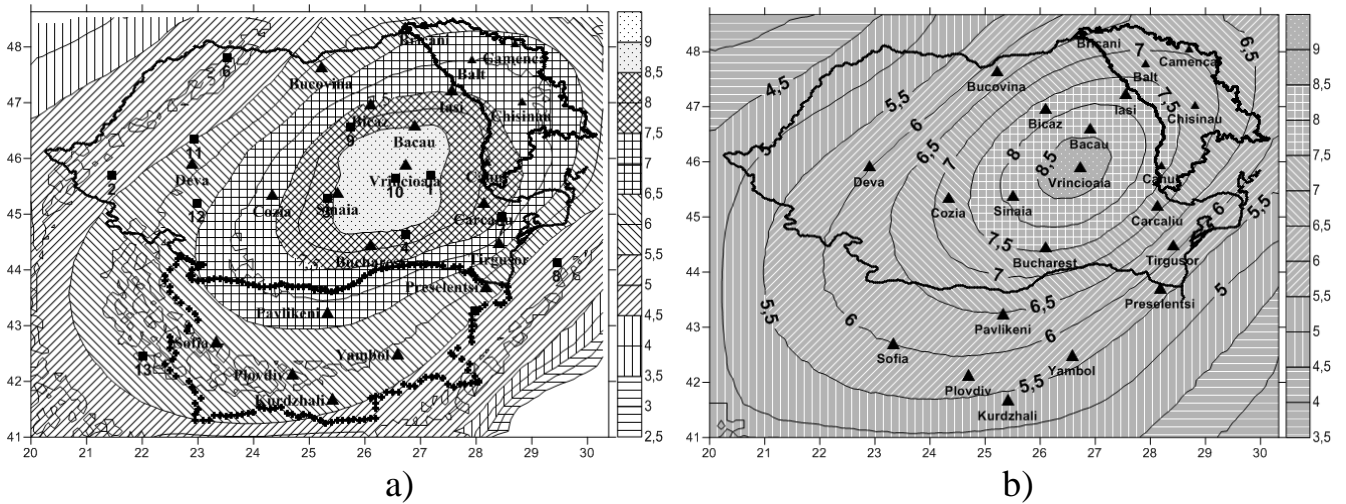


Рис. 7.1. Карта сейсмічної небезпеки території, яка підлягає впливу румунських землетрусів: а) в балах шкали *МСК-64*; б) вранчських землетрусів

При розрахунку СН, для точок в ближній зоні обчислюється відстань Джойнера-Бура, і гіпоцентральною відстань для точок в далекій зоні.

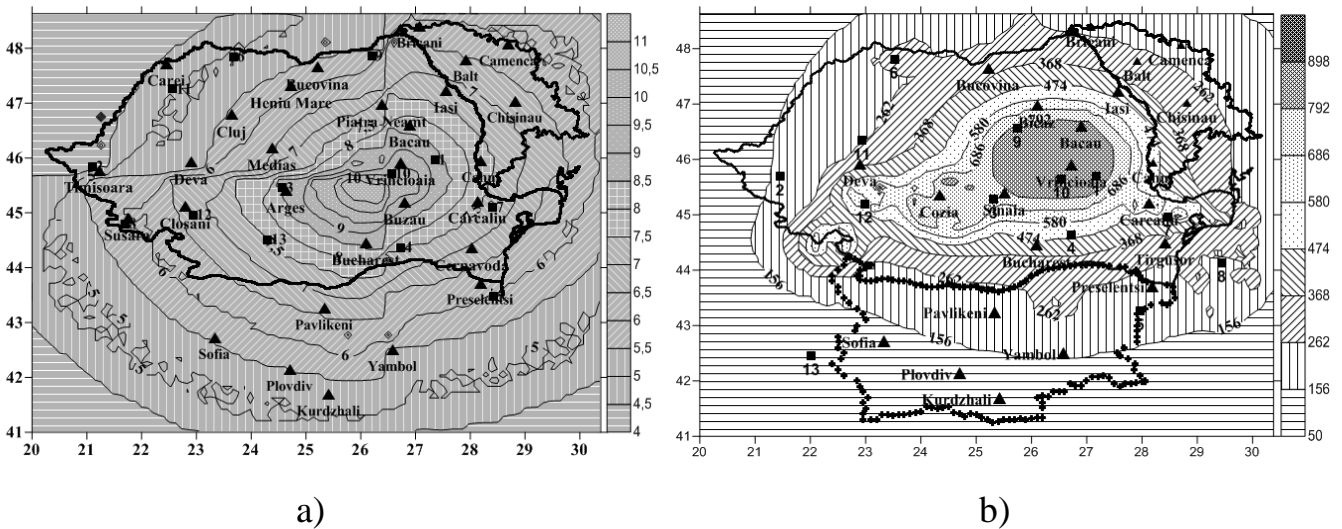
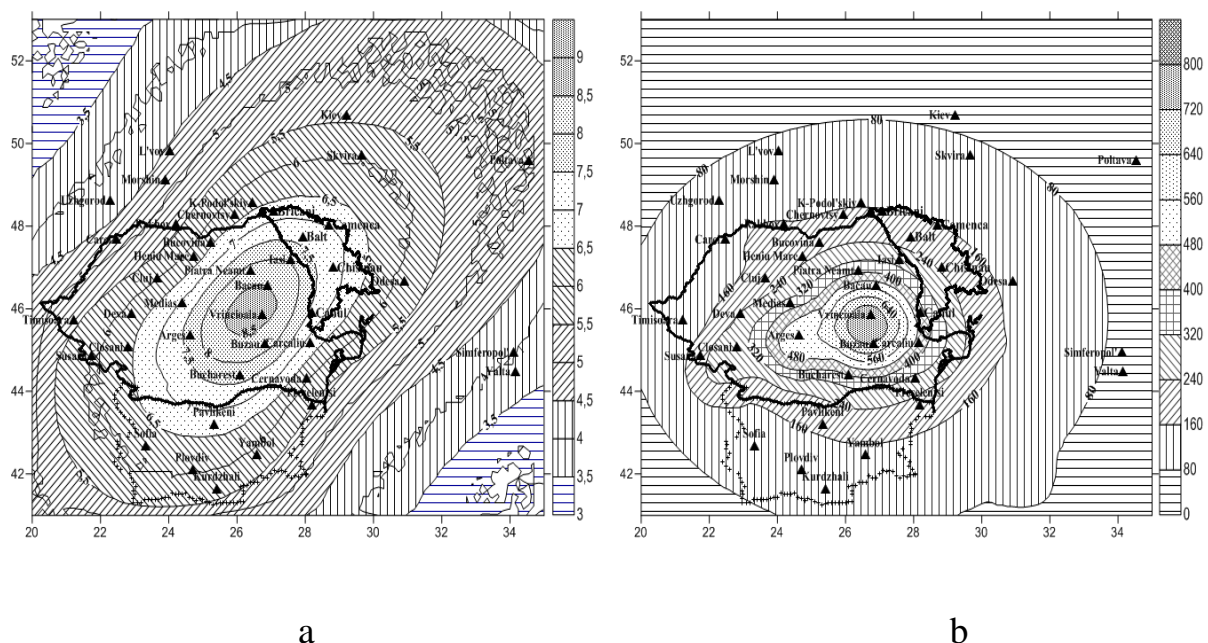


Рис. 7.2. а) Карта сейсмічної небезпеки території Молдови, Румунії та Болгарії в балах шкали *МСК-64*, за значеннями коефіцієнтів функції загасання (таб. 5.4), обчислені за методом змінного середнього. б) карта сейсмічної небезпеки, побудована за нормами *ЕС8*.

При значеннях $\theta=180^\circ$ були оцінені значення коефіцієнтів (табл. 5.4), і побудована карта сейсмічної небезпеки (рис. 7.2а). Геометрична форма малюнка

(рис. 7.2а), побудована за коефіцієнтами загасання, визначені за допомогою регресійного аналізу в дотичному азимутному створі, також демонструє еліптичності теоретичної карти сейсмічної небезпеки і північно-східну спрямованість великої осі еліпса загасання.

При розрахунках СН від підкорових землетрусів в балах шкали *MSK-64* застосовані значення коефіцієнтів загасання: $a=1.6$; $c=7.2$; $b_{\max}=5.6$; $b_{\min}=4.9$ і при значенні $\gamma_0=51^\circ$, кута між великою віссю еліпса загасання і позитивним напрямом осі абсцис. Сейсмічність Румунії детально досліджена, і існує ряд робіт, присвячених аналізу сейсмічної небезпеки території, яка підлягає впливу румунських землетрусів [Ardeleanu et al., 2005, 2007; Enescu et al., 2004; Ismail-Zadeh, 2007; Marmureanu et al., 2011; Moldovan et al., 2008; Musson, 2000; Sokolov et al. 2015, 2016, Leydecker et al. 2008]. При порівнянні авторських карт сейсмічної небезпеки (7.1, 7.2, 7.3 а, б) з картами: (рис. 7.3с) [Leydecker, et al. 2008] і (рис. 7.3d) [Sokolov, et al. 2007], виявлено незначну розбіжність значень сейсмічної небезпеки в точках, з однаковими координатами, досліджуваної території. Отримані з авторських карт СН значення інтенсивності сейсмічного ефекту, з періодом повторюваності 475 років, в точках досліджуваної території, схожі з результатами інших авторів [Ardeleanu et al., 2007; Ismail-Zadeh et al. 2007; Sokolov et al., 2007].



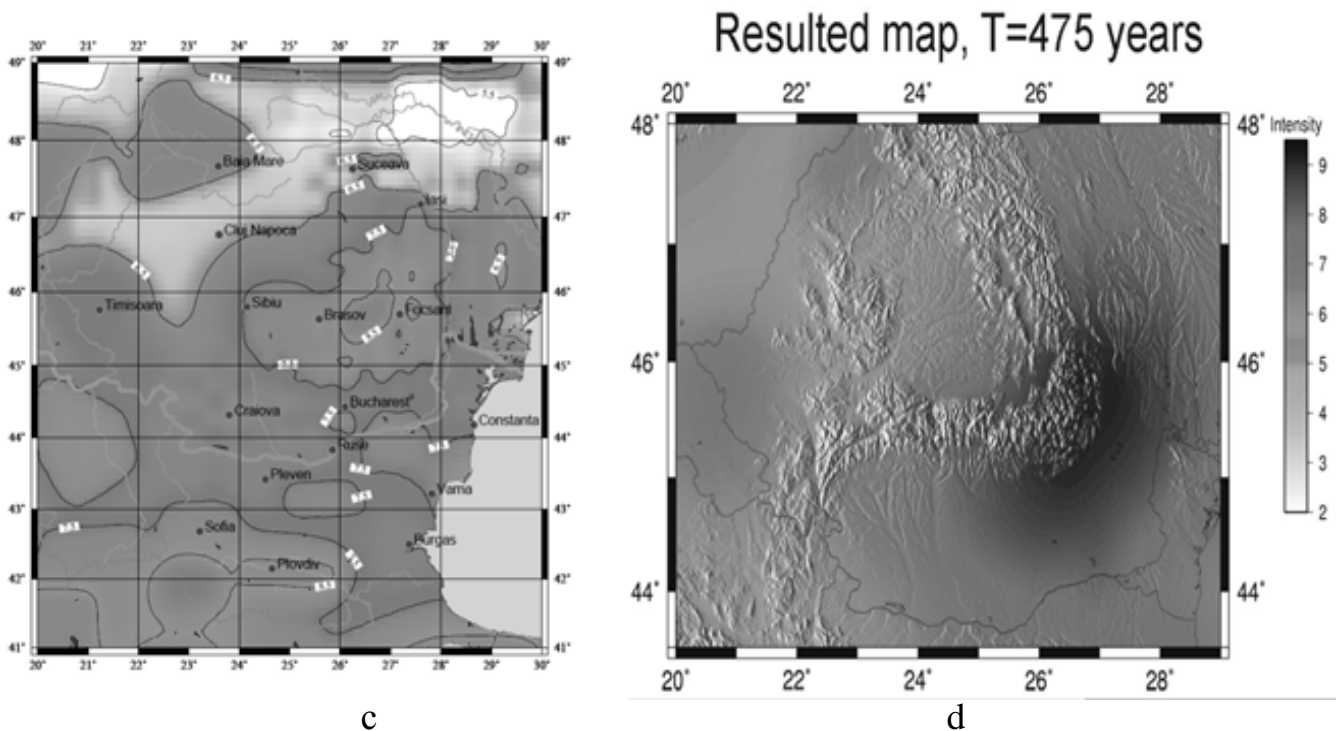


Рис. 7.3. Карта сейсмічної небезпеки території Болгарії, Молдови, Румунії та України для періоду повторюваності в 475 років: а) в балах шкали *MSK-64*; б) за нормами *EC8*; с) карта ЙАСН Румунії [Leydecker et al. 2008]; д) [Sokolov et al. 2007]

Аналіз сейсмічної небезпеки території Туреччини виконувалися на базі відомостей з каталогу турецьких землетрусів, що охоплює період 1900–2010рр. Методом кластерного аналізу на території Туреччини виділені 50 сейсмічних зон. Розрахунок мір сейсмічної небезпеки проводився в балах шкали інтенсивності *MSK-64* і по нормам *EC8* на основі ймовірнісної міри (7.37). Використовувалися закони загасання інтенсивності сейсмічних впливів у балах (7.39), і прискорень зміщення ґрунту для румунських землетрусів (7.40), а також закони загасання, встановлені за макросейсмічними даними турецьких землетрусів [Erdik et al., 2004; Kalafat, 2011; Musson, 2000; Yüce men et al., 2006; Sorensen et al., 2008]. Відповідно до будівельних норм *EC8*, аналіз сейсмічної небезпеки проводиться на основі розрахунку міри (7.21). Умовна ймовірність перевищення рівня I^k інтенсивності сейсмічних впливів визначиться виразом [Burtiev et al., 2016, 2017]:

$$p_r^k = P(I \geq I^k / E_r) = \frac{1}{\sigma\sqrt{2\pi}} \int_{I_k}^{12} e^{-\frac{(I-\bar{I})^2}{2\sigma^2}} dI \quad (7.42)$$

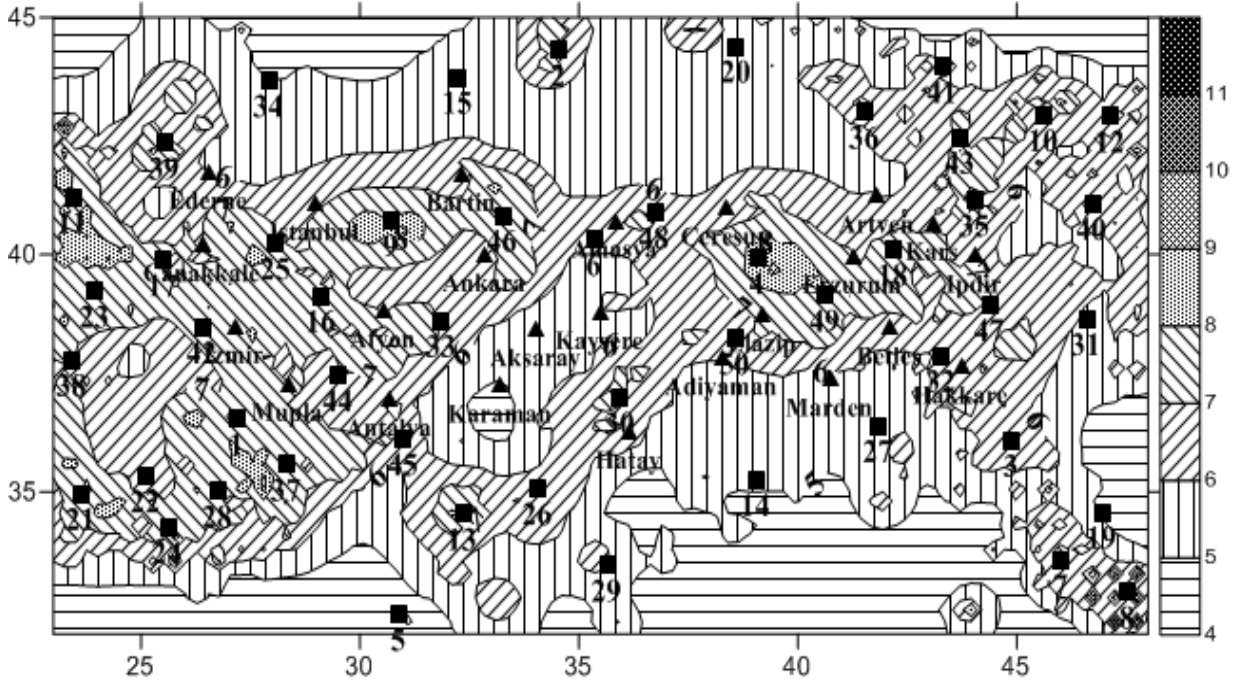


Рис. 7.4. Карта сейсмічної небезпеки Туреччини в балах шкали МСК-64

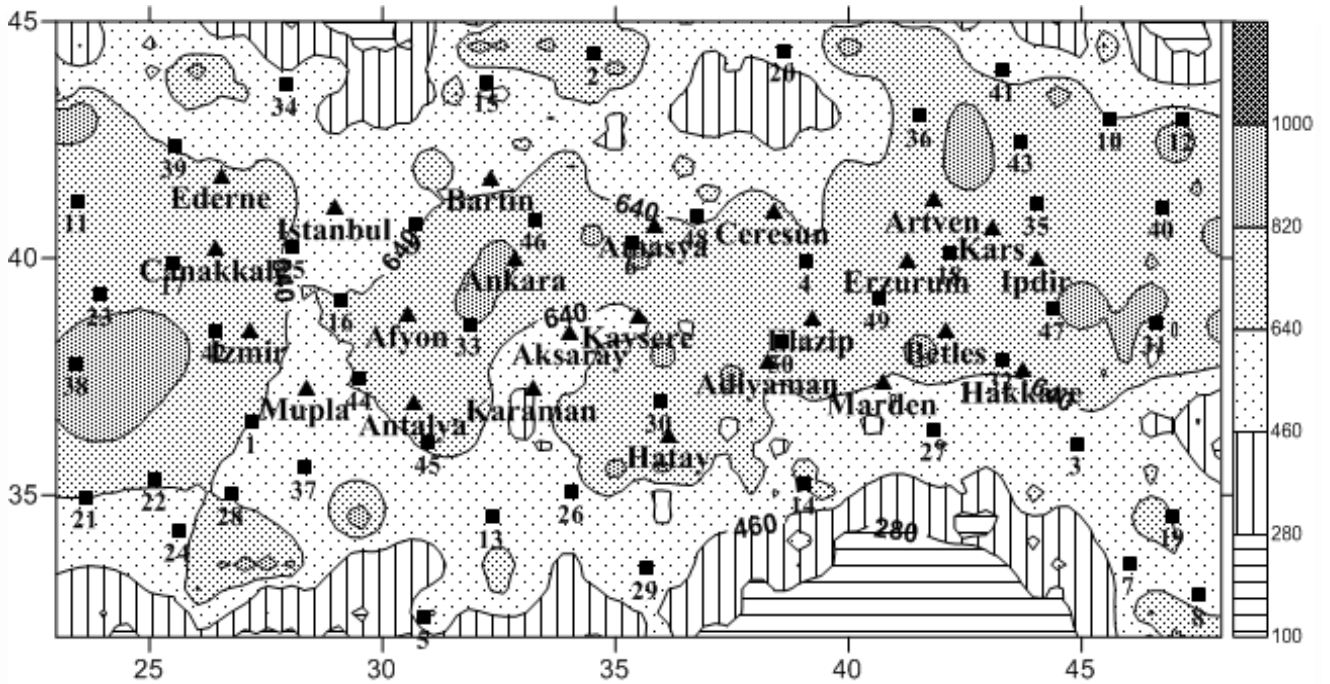


Рис. 7.5. а) Карта сейсмічної небезпеки Туреччини побудована за стандартами EC8 з урахуванням 50 сейсмічних зон. На ізолініях вказані значення пікових

прискорень cm/sec^2 .

Далі, застосовуючи формули 7.1–7.7 обчислюється α_k (7.48), річна інтенсивність перевищень рівня I^k інтенсивності в балах шкали *MSK-64* і характеристик зсувів ґрунту a_k . Карти сейсмічної небезпеки (рис.7.1–7.4) побудовані на основі мір (7.21).



Рис. 7.5. b) Нормативна карта сейсмічної небезпеки Туреччини, побудована за стандартами *EC8* [Solomos et al., 2008].

Картина ранжирування території Туреччини за ступенем сейсмічного ефекту на обох картах має схожу тенденцію, однак, на картах *CP* (рис. 7.4, 7.5.a, b), площі полів рівної бальності відрізняються. Це пояснюється тим, що при побудові авторських карт *CP* (рис. 7.4, 7.5.a), використовувалися коефіцієнти загасання інтенсивності сейсмічних впливів, визначених для території Румунії.

Найбільш вірогідні значення біноміального розподілу (7.36), при цілому значенні величини $np-(1-p)$, є:

$$m = np - (1 - p); \quad m = np + p \quad (7.43)$$

де n – число землетрусів, m – число сейсмічних поштовхів інтенсивності I^k , які можуть статися з ймовірністю p . Якщо ж $m_0=np-(1-p)$ не є цілим числом, то найбільшого значення біноміальна ймовірність (7.36) досягає при значенні m , що дорівнює найменшому цілому числу, більшому m_0 [Гнеденко, 1961]. Нехай випадкова величина N , розподілена по пуассонівському закону і означає можливе число землетрусів в Кишиніві за 50 років. Теоретично, множина можливих значень числа сейсмічних подій в точках земної поверхні є розрахунковою, а практично діапазон можливих значень визначається інтервалом $[N_{\min}, N_{\max}]$, значення з якого мають позитивну ймовірність. Тобто число сейсмічних подій може приймати значення з інтервалу $[N_{\min}, N_{\max}]$.

$$P(t, m \geq 0, I^k) = \sum_{n=N_{\min}}^{N_{\max}} \left(\sum_{m=1}^n C_n^m \beta_k^m (1-\beta_k)^{n-m} \right) \frac{e^{-\lambda t} (\lambda t)^n}{n!} \quad (7.44)$$

Таблиця 7.3

Характеристики сейсмічної небезпеки

Кишинів						
I^k	4	5	6	7	8	9
β_k	0.0128	0.00517	0.00178	0.000841	0.000201	0
$n\beta_k$ (1000 років)	352	157	43	29	9	0.5
період повторюваності	2.8	6.4	23	34	114	1867
n	2312	2312	2312	2312	2312	0
Ймовірне значення m	31	13	6	3	2	0
$P(50, n, m_k, I^k)$	0.000591	0.000916	0.00106	0.00178	0.0004	0
$P(n, m_1 \leq m \leq m_2, I^k)$	0.004	0.0041	0.0027	0.004	0.0018	0
m_1, m_2	$30 \leq m \leq 229$	$12 \leq m \leq 211$	$5 \leq m \leq 204$	$2 \leq m \leq 201$	$1 \leq m \leq 200$	0
$I-P(50, 0, \lambda)$	1.0	1.0	0.984	0.857	0.372	0
$P(50, m \geq 0, \lambda)$	1.0	0.999	0.964	0.855	0.377	0
Бухарест						
β_k	0.024	0.009	0.0036	0.001	0.00032	0.0
$n\beta_k$ (1000 лет)	840	302	146	38	14	2

У таблиці (таб. 7.3) наведені значення характеристик сейсмічної небезпеки в Кишиневі: перший рядок містить розподіл ймовірності інтенсивності сейсмічного ефекту; другий рядок середнє число сейсмічних поштовхів з інтенсивністю I^k (сейсмічна струшуваність); в третьому рядку наведені середні періоди їх

повторюваності; четвертий рядок – середнє число землетрусів за 50 років; п'ятий рядок – ймовірне значення біноміального розподілу; шостий рядок – ймовірність виникнення за 50 років в Кишиневі n сейсмічних поштовхів, рівно m з яких матимуть інтенсивність I^k в балах шкали *MSK-64*; сьомий рядок містить значення ймовірності того, що в Кишиневі за 50 років відбудеться не менш m_1 і не більше m_2 сейсмічних подій з інтенсивністю I^k ; дев'ята і десята рядки - ймовірність виникнення в Кишиневі за 50 років хоча б однієї сейсмічної події з інтенсивністю I^k , обчислені за формулами (7.21), (7.44).

7.4. Обчислення сейсмічного ризику

На основі розробленого методу аналізу сейсмічної небезпеки проводиться оцінка сейсмічного ризику [Burtiev et al., 2016]. Нехай випадкова величина D означає величину матеріальних збитків, в деякій умовній грошовій одиниці, які може заподіяти сейсмічна подія деякому населеному пункту, наприклад, місту «Сейсмоград», з координатами (45,28). Якщо $F(d/I)$ умовна функція розподілу випадкової величини D , то інтеграл:

$$M(D/I) = \int d dF(d/I) \quad (7.45)$$

визначає умовне математичне сподівання розміру збитку від сейсмічного впливу з інтенсивністю I . Припустимо, що величина D з умовною ймовірністю $p(d/I^k)$ приймає одне з L можливих значень d_l , ($l=1, \dots, L$), при виникненні сейсмічного поштовху інтенсивності I^k . Ймовірності $p(d_l/I^k)$ утворюють (K, L) матрицю сейсмічного ризику D_r :

$$D_r = \begin{pmatrix} p\left(\frac{d_1}{I^1}\right) \dots & p\left(\frac{d_2}{I^1}\right) & \cdot & \cdot & p\left(\frac{d_L}{I^1}\right) \\ p\left(\frac{d_1}{I^2}\right) \dots & p\left(\frac{d_2}{I^2}\right) & \cdot & \cdot & p\left(\frac{d_L}{I^2}\right) \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ p\left(\frac{d_1}{I^{12}}\right) \dots & p\left(\frac{d_2}{I^{12}}\right) & \cdot & \cdot & p\left(\frac{d_L}{I^{12}}\right) \end{pmatrix} \quad (7.46)$$

Формула обчислення умовного математичного очікування обсягу шкоди D , щодо інтенсивності сейсмічних впливів, в матричній формі має вигляд:

$$M\left(\frac{D}{I^k}\right) = \sum_{l=1}^L d_l p\left(\frac{d_l}{I^k}\right) = D, \vec{d} \quad (7.47)$$

$$\vec{d} = (d_1, d_2, \dots, d_L) \quad (7.48)$$

Так як величина I приймає одне із значень I^k , ($k=1, \dots, 12$), то події $(I=I^1)$, $(I=I^2), \dots, (I=I^{12})$ утворюють повну групу несумісних подій і $F(d/(I=I^1))$, $F(d/(I=I^2)), \dots, F(d/(I=I^{12}))$ – відповідні цим подіям умовні функції розподілу випадкової величини D . З формули повної ймовірності знаходимо безумовну функцію розподілу обсягу збитку при землетрусі з інтенсивністю I^k :

$$F(d) = \sum_{k=1}^{12} P(I=I^k) F\left(\frac{d}{(I=I^k)}\right) \quad (7.49)$$

Для оцінки сумарного сейсмічного ризику служить формула, яка обчислює математичне очікування умовного математичного очікування величини збитку D , при фіксованих значеннях інтенсивності струсів I^k :

$$MD = M\{M(D/I)\} = \sum_{k=1}^{12} P(I=I^k) M\left(\frac{d}{I^k}\right) \quad (7.50)$$

Вираз (7.50) в матричній формі має вигляд:

$$MD = \sum_{k=1}^{12} P(I=I^k) M\left(\frac{d}{I^k}\right) = \sum_{k=1}^{12} \beta_k M\left(\frac{d}{I^k}\right) = \vec{\beta} D, \vec{d} \quad (7.51)$$

Припустимо, що при кожному I^k величина D приймає тільки одне значення D_k , тобто $P(D=d_k)=1$, для кожного $k=1, \dots, 12$. Отже, при виникненні події I^k може статися тільки одна подія $\{D=d_k\}$, тобто подія $\{I=I^k\}$ веде до обов'язкового здійснення події $\{D=d_k\}$, тобто $P(d_k/I^k)=1$. З рівності одиниці умовної ймовірності випливає, що спільна ймовірність того, що в результаті виникнення землетрусу інтенсивність струсів прийме значення I^k і обсяг збитків прийме значення d_k , визначається ймовірністю виникнення сейсмічного струсу з інтенсивністю I^k :

$$p(d_k, I^k) = P(D=d_k, I=I^k) = P\left(\frac{D=d_k}{I=I^k}\right) P(I=I^k) = p(I^k) \quad (7.52)$$

Таким чином, сейсмічний ризик при загрозі виникнення одного землетрусу визначиться за формулою:

$$M(d) = \bar{D} = \sum_{k=1}^{12} \beta_k d_k \quad (7.53)$$

Множенням середнього числа сейсмічних подій за час t (7.19), на величину $M(d)$ (7.53) визначається сейсмічний ризик за період часу t років:

$$\bar{D}_t = \sum_{k=1}^{12} \lambda t \beta_k d_k \quad (7.54)$$

Таблиця 7.4

Оцінки сейсмічного ризику для міста «Сейсмоград»

№	I^k	4	5	6	7	8	9
1	β_k	0.0128	0.00517	0.00178	0.000841	0.000201	0
2	d_k	25600	62500	129600	240100	409600	656100
3	$d_k \beta_k$	327	323	231	202	82	7.7
4	$\lambda \beta_k 475 D_k$	$718 \cdot 10^4$	$729 \cdot 10^4$	$506 \cdot 10^4$	$443 \cdot 10^4$	$181 \cdot 10^4$	$169 \cdot 10^3$
5	$\bar{D} = \sum_{k=1}^{12} \beta_k D_k = 1462$		$\bar{D}_{475} = \sum_{k=1}^{12} \lambda t \beta_k D_k = 321 \cdot 10^5$				

Результати реалізації алгоритму на прикладі міста «Сейсмоград» містяться в таблиці (таб. 7.4), де в першому рядку наведено розподіл інтенсивності сейсмічних поштовхів; другий рядок містить розмір збитку d_k ; третій рядок можливий розмір збитку при виникненні одного землетрусу. Обсяги середнього збитку за період часу 475 років наведені в четвертому рядку. Можливий сумарний збиток при виникненні одного землетрусу і за період часу 475 років показаний в п'ятому рядку.

7.5. Оцінка максимальної магнітуди землетрусів за статистикою Гумбеля

Магнітуда є одним з головних параметрів землетрусів, що визначає ступінь сейсмічних впливів. Період виникнення максимально можливих для конкретної сейсмічної зони землетрусів часом набагато перевершує тривалість часу щодо достовірних інструментальних сейсмічних спостережень, історія яких почалася в

кінці XIX століття. Тому для отримання оцінок максимальної магнітуди землетрусів M_{max} застосовуються всі доступні методи: геофізичні та математичні.

Часто при статистичних дослідженнях ігнорують можливість появи великих, але рідкісних подій, що лежать на хвості розподілу. Розподілу з важкими хвостами: хвіст яких не можна відрізати, тобто не можна нехтувати великими, але рідкісними сейсмічними подіями. Нехай $F(x)$ – функція розподілу випадкової величини X , тобто $F(x)=P(X<x)$, тоді функція $1-F(x)=P(X\geq x)$ є розподілом хвоста. Функція $F(x)$ називається функцією розподілу з важким хвостом, якщо математичне сподівання випадкової величини дорівнює нескінченності [Родкин, 2005; Hölzlwimmer, 2010; Trömel, 2004]:

$$E(X) = \int xf(x)dx = \infty \quad (7.55)$$

Під великими сейсмічними подіями розуміються землетруси, збиток від яких може перевершити сумарний збиток, від всіх інших землетрусів, разом узятих. У сейсмології, як і в деяких областях, екстремальні значення становлять велику важливість. Однак, в деяких випадках землетруси починаючи тільки з деякого порога магнітуди, становлять небезпеку. При проектуванні об'єктів стратегічної важливості необхідно знати можливий найбільший ступінь сейсмічних впливів на будівельному площі.

Більшість статистичних методів підбору теоретичних розподілів для вибірових даних обмежуються перевіркою адекватності моделей в центральній частині вибірки, де зосереджена основна частина спостережень. Тоді, як ймовірнісна структура відносно невеликого числа крайніх членів вибірки залишається недослідженою. Крайні і невелике число екстремальних членів ряду можуть бути спостереженнями з різних по ймовірнісній структурі генеральних сукупностей [Grieser et al., 2009; Hölzlwimmer, 2010; Raschke, 2012; Trömel, 2004].

Недоліком застосування розподілу екстремальних значень є відкидання великого обсягу даних, тому що використовуються тільки найбільші на деяких інтервалах часу спостереження. Цей факт є основним аргументом противників застосування статистики екстремальних значень в аналізі сейсмічної небезпеки,

який усувається складанням допоміжної вибірки зі спостережень, що лежать вище деякого порога.

Метою нашої роботи є розподіл усіх визначених найбільшої магнітуди землетрусів вогнища Вранча. Величина максимальної можливої магнітуди (M_{max}) оцінюється для передбачення найгіршого сценарію розвитку сейсмічної небезпеки. Для цього складається вибірка з максимальних спостережених за рік землетрусів за даними каталогу [Shebalin et al., 1998].

Вибірка повинна відповідати таким основним допущенням: достатність обсягу вибірових даних (не менше 30 значень); відсутність похибок; незалежність; наочність; відсутність викидів.

Нехай $F(m)=P(M<m)$ закон розподілу магнітуди землетрусів. Припустимо, що послідовність магнітуд M_1, M_2, \dots, M_n є послідовністю n незалежних однаково розподілених випадкових величин з функцією розподілу $F(m)$.

Тоді ймовірність того, що магнітуда всіх n землетрусів перевищить величину m , дорівнює:

$$P(M_n > m) = 1 - F^n(m) \quad (7.56)$$

При класичному підході, для підтвердження вірогідності того, що з довірчою ймовірністю f не відбудеться землетрус з магнітудою менше m необхідно, щоб магнітуда:

$$n = \frac{\ln(1-f)}{\ln F(m)} \quad (7.57)$$

землетрусів перевищувала величину m . Максимально можливі землетруси відбуваються рідко. Їх середній період повторюваності становить зазвичай сотні і тисячі років. Тоді як, каталоги землетрусів охоплюють більш короткий період часу. Тому визначення найбільшої магнітуди, на основі обмеженої сейсмостатистиці, здійснюється за допомогою статистики екстремальних значень (розподіл Гумбеля). Дана статистика дозволяє оцінити максимальну магнітуду за порівняно невеликою за обсягом вибіркою. Найбільш прийнятною моделлю для опису ймовірнісної структури максимальних магнітуд M_{max} є один з трьох типів розподілу Гумбеля

[Гумбель, 1965]:

$$\bullet F_1(x) = \exp(-\exp(-q(x-u))) \quad (7.58)$$

де u – характеристичний екстремум $F(u)=1/e$, q –параметр розподілу, застосовується для апроксимації асимптотичних розподілів максимальних значень необмежених випадкових величин.

$$\bullet F_2(x) = \exp\left(-\left(\frac{\omega-x}{x-u}\right)^k\right) \quad (7.59)$$

де u – характеристичний екстремум $F(\omega)=1/e$, ω – ліва межа випадкових величин $\omega < x$, k – параметр розподілу.

$$\bullet F_3(x) = \exp\left(-\left(\frac{\omega-x}{w-u}\right)^k\right) \quad (7.60)$$

де u – характеристичний екстремум $F(\omega)=1/e$, ω – права межа діапазону магнітуди землетрусів $x < \omega$, параметр, k – параметр розподілу.

Застосуємо статистику Гумбеля для оцінки найбільшої магнітуди землетрусів вогнища Вранча. Для цього використовуються відомості про сейсмічні події з каталогу балканських землетрусів. Для апроксимації найбільшої магнітуди, при відмінності вибіркового розподілу магнітуди землетрусів від нормального закону, використовується розподіл Гумбеля [Гумбель, 1965].

7.5.1 Перший граничний розподіл

Випадкова величина може приймати значення з інтервалу $(-\infty, \infty)$. Із сукупності значень магнітуди складається N вибірок, по n спостережень в кожній, витягнутих з однієї і тієї ж генеральної сукупності. З найбільших значень в кожній вибірці складається вибірка обсягу N . Найбільша величина серед Nn спостережень є і найбільшим і серед N максимальних спостережень. Для стійкості розподілу, найбільші значення в Nn спостереженнях і в n спостереженнях повинні прагнути до одного граничного розподілу. Математичний вираз постулату стійкості має вигляд:

$$F^n(x) = F(a_n x + b_n) \quad (7.61)$$

де параметри a_n і b_n є функціями n [Гумбель, 1965].

З основної вибірки значень моментної магнітуди MW складалися N_1 вибірка обсягу по n_1 , і N_2 вибірки з числом елементів n_2 . Стійкість означає, що вибірках обсягу і N_2 для ймовірнісного розподілу Гумбеля (y, x) , $y=q(x-u)$ відповідатимуть паралельні прямі, зміщені одна відносно другої на величину приблизно $\ln N_1 / \ln N_2$.

Оцінкою параметрів служить:

$$\frac{1}{q} = \frac{s}{\sigma_n}; \quad u = \bar{x} - \frac{y_n}{q} \quad (7.62)$$

Величини y_n , σ_n є середньоарифметичне і середнє квадратичне відхиленнями наведених випадкових величин $y=q(x-u)$, відповідно, \bar{x} і s – вибіркові середнє арифметичне і стандартне відхилення. При різних значеннях N виразом [Розин, Харитонов, 1972]:

$$y_i = -\ln\left(-\ln\left(\frac{i}{N+1}\right)\right) \quad (7.63)$$

де $i=1, \dots, N$ визначаються числові характеристики вибірки. Потім визначаються вибіркові моменти $\bar{y}_N, \bar{\sigma}_N^2$. Величини y_n, σ_n табульовані, і їх значення зі збільшенням N сходяться до математичного сподівання $a=0.57722$ і стандартного відхилення генеральної сукупності, відповідно. Значення параметрів для вибірок обсягу $N_1=25$ і $N_2=30$ вийшли рівними $q_1=2.19, u_1=5.67; q_2=2.20, u_2=5.57$. Близько значення q_1 і q_2 вказують на стійкість розподілу Гумбеля.

Для дуже малих за обсягом вибірок Ліблейн запропонував метод, в якому замість окремої оцінки параметрів u і $1/q$ розглядаються ймовірнісні точки [Гумбель, 1965]:

$$\xi_\Phi = u + \frac{y_\Phi}{q} \quad (7.64)$$

Тобто оцінка параметрів розподілу Гумбеля і прогноз значень включені в одну формулу. Метод призводить до оцінок:

$$u = \sum_{i=1}^N a_i x_i; \quad \frac{1}{q} = \sum_{i=1}^N b_i x_i \quad (7.65)$$

де $x_{(1)}, \dots, x_{(n)}$ порядкова статистика, a_i, b_i – ваги в оцінці Ліблейна. Якщо спостереження отримані в упорядкованому вигляді, то їх необхідно рангомінізувати. Значення ваги табульовані тільки для значень $N=4, 5, 6$. Для великих вибірок обсягу $N=kn$, який кратний $n=5$ або 6 , Ліблейн запропонував формувати підвибірки обсягів

m і будувати під оцінці [Гумбель, 1965]:

$$L_i = \frac{1}{k} \sum_{j=1}^n w_j x_j \quad (7.66)$$

для вибіркової повної оцінки отримаємо:

$$\tilde{L} = \frac{1}{k} \sum_i^k L_i \quad (7.67)$$

Якщо обсяг вибірки не кратний 5 або 6, то оцінки параметрів для решти 1, 2, 3, 4 спостережень використовується в повній оцінці. Застосувавши метод Ліблейна до вибірки магнітуд обсягу N , отримані значення $q=2.2$, $u=5.67$. 99.9%-ні квантілі, які оцінені за обома методами, співпадають: $M_{0.999}=8.8$.

7.5.2 Третій граничний розподіл

Цей тип розподілу справедливий для обмежених справа вихідних розподілів, якими є магнітуда:

$$F_3(x) = \exp\left(-\left(\frac{\omega - x}{w - u}\right)^k\right) \quad (7.68)$$

Оцінкою максимальної правдоподібності параметрів w , u , k є рішення векторного рівняння:

$$\left(\frac{\partial L}{\partial w}, \frac{\partial L}{\partial u}, \frac{\partial L}{\partial k}\right) = (0, 0, 0) \quad (7.69)$$

Тут L – логарифм функції максимальної правдоподібності [Розин, Харитонов, 1972; Al-Abbasi, Fahni, 1985; Bülow, Jacob, 2007; Dessekey, 1985; Hamann, 2009; Jonas, 2003; Kabluchko, 2011]:

$$L(\bar{x}, w, u, k) = \ln \left[\left(\frac{k}{w - u}\right)^N \prod_{i=1}^N \left(\frac{w - x_i}{w - u}\right)^{k-1} F_3(x_i) \right] \quad (7.70)$$

Координати точки максимуму цієї функції є оцінками максимальної правдоподібності параметрів розподілу Гумбеля. Максимум гладкої функції 3-х змінних (7.70) визначається дослідженням значень функції в критичних точках, які знаходяться з рішення системи рівнянь:

$$\begin{cases} \frac{\partial L}{\partial w} = -\frac{nk}{\omega - u} + (k-1) \sum_{i=1}^n \frac{1}{\omega - x_i} - k \sum_{i=1}^n \frac{(\omega - x_i)^{k-1} (x_i - u)}{(\omega - u)^{k+1}} = 0 \\ \frac{\partial L}{\partial u} = \frac{nk}{\omega - u} - k \sum_{i=1}^n \frac{(\omega - x_i)^k}{(\omega - u)^{k+1}} = 0 \\ \frac{\partial L}{\partial k} = \frac{n}{k} - n \ln(\omega - u) + \sum_{i=1}^n \ln(\omega - x_i) - \sum_{i=1}^n \left(\frac{\omega - x_i}{\omega - u}\right)^k \ln\left(\frac{\omega - x_i}{\omega - u}\right) = 0 \end{cases} \quad (7.71)$$

Дану систему 3-х нелінійних рівнянь не вдається рішити прямими методами, виразивши корені через деякі формули. Були застосовані чисельні методи умовної оптимізації, зокрема метод умовного градієнта. Цей метод можна застосовувати, тому що область, в якій знаходиться максимальне значення функції (7.70) має досить просту структуру:

$$Y : \begin{cases} a_1 \leq \omega \leq b_1 \\ a_2 \leq u \leq b_2 \\ a_3 \leq k \leq b_3 \end{cases} \quad (7.72)$$

інакше кажучи, множина Y є паралелепіпедом. У цьому випадку схема методу умовного градієнта:

$$y^{p+1} = y^p + \alpha_p h^p; \quad y^p = (\omega_p, u_p, k_p) \quad (7.73)$$

де $p=0,1,\dots,m$, α_p – крок на p -ої ітерації, а h^p – напрямок руху з точки y^p на кожному кроці, для визначення напрямку руху h^p потрібно вирішити задачу лінійного програмування:

$$(\text{grad}L(\omega_p, u_p, k_p), y - y^p) \rightarrow \max \quad (7.74)$$

Вибір початкового наближення впливає на роботу алгоритму. В принципі метод умовного градієнта дозволяє знайти стаціонарні точки, які вимагають подальшого дослідження.

Рівняння (7.70) вирішується ітераційним методом, збіжність якого залежить від початкового розподілу [Al-Abbasi, Fahni, 1985]:

$$\begin{pmatrix} w_n \\ u_n \\ k \end{pmatrix} = \begin{pmatrix} w_{n-1} \\ u_{n-1} \\ k_{n-1} \end{pmatrix} - \begin{pmatrix} \frac{\partial L}{\partial w} \\ \frac{\partial L}{\partial u} \\ \frac{\partial L}{\partial k} \end{pmatrix} \begin{pmatrix} \frac{\partial^2 L}{\partial w^2} & \frac{\partial^2 L}{\partial w \partial u} & \frac{\partial^2 L}{\partial w \partial k} \\ \frac{\partial^2 L}{\partial u \partial w} & \frac{\partial^2 L}{\partial u^2} & \frac{\partial^2 L}{\partial u \partial k} \\ \frac{\partial^2 L}{\partial k \partial w} & \frac{\partial^2 L}{\partial k \partial u} & \frac{\partial^2 L}{\partial k^2} \end{pmatrix} \quad (7.75)$$

З огляду на обмеженість інтервалів можливих значень параметрів розподілу (7.58) для вирішення завдання застосовувався метод сіток, який не поступається по точності рішення методу умовного градієнта, але значно простіше в реалізації. Критерієм порівняння при цьому виступає величина середньоквадратичної похибки.

Оптимальні значення коефіцієнтів знаходяться експериментальним шляхом, перебираючи всілякі комбінації значень цих коефіцієнтів у вузлах сітки. Для цього параметрам присвоюються всілякі значення з деяким кроком дискретності. Цей спосіб застосовується для пошуку початкового рішення. Дійсно, процес швидко сходиться, якщо в якості початкового наближення використовуються значення $w=8.0$, $u=5.8$, $k=4.3$, які максимізують функцію правдоподібності (7.70). Ітераційний процес припиняється при виконанні умови [Буртиев,2003;]:

$$|w_n - w_{n-1}| + |u_n - u_{n-1}| + |k_n - k_{n-1}| < h \quad (7.76)$$

де h – деяка наперед задана мала величина. Якщо x_1, \dots, x_n порядкова статистика, тобто $x_i < x_{i+1}$, то число спостережень, розташованих лівіше точки x_{i+1} одно i . Теоретичне ж число цих точок становить $NF_3(x_{i+1})$. Оцінка вектора параметрів (w, u, k) , що мінімізує функціонал:

$$\sum (i - NF_3(x_{i+1}))^2 \quad (7.77)$$

знаходиться методом перебору. Шуканими значеннями виявилися: $w=8.0$, $u=5.8$, $k=3.4$. 99.9%-ній квантиль дорівнює 7.87 для $N_1=25$ і 7.87 для $N_2=30$ вибірок відповідно, тобто, квантилі є більш стійкими оцінками, ніж абсолютні значення максимальних магнітуд. Далі складалася вибірка з щорічних максимальних магнітуд землетрусів. Рішенням ітераційного процесу є значення: $w=8.0$, $u=5.6$, $k=6.3$. Таким чином, максимальне значення магнітуди з періодом повторення 1000 років, становить $MW_{\max}=7.87$.

Поле з ізосейстами екстремального сейсмічного впливу, побудована автором (рис. 7.6а), відрізняється по геометричній формі і площі полів рівної інтенсивності від поля з ізосейстами побудованої румунськими сейсмологами (рис.7.7) [Marmureanu et al. 2011]. Наприклад, на карті (рис. 7.6а) Бухарест знаходиться в 8-ми бальній зоні, а відповідно до карти (рис.7.7) в 9-ти бальній. Існуюча відмінність пояснюється: значеннями використаних коефіцієнтів загасання, розміром осередків географічної мережі, якою покривається досліджувана територія, і в вузлах якої розраховується сейсмічна небезпека.

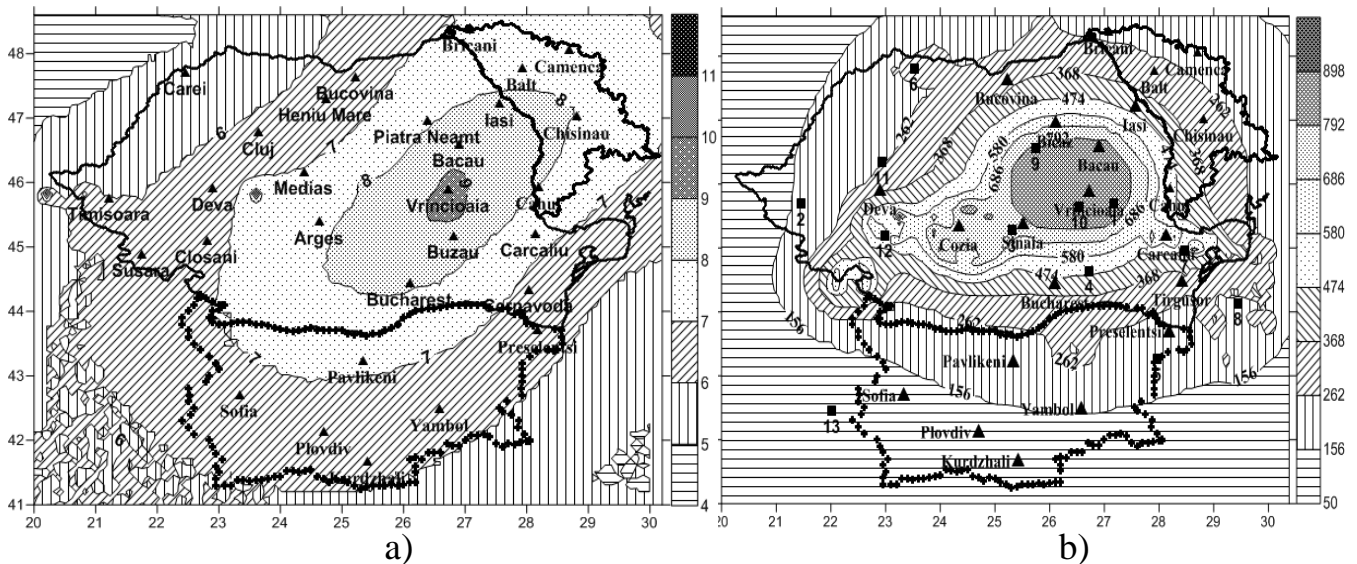


Рис. 7.6. Карта екстремальної сейсмічної небезпеки загальної території Молдови, Румунії та Болгарії: а) в балах шкали *MSK-64*; б) за нормами *EC8*. Для точок в ближній зоні використовується відстань Джойнер-Бура, і гіпоцентральна відстань для точок в дальній зоні ($MW = 7.87$)..

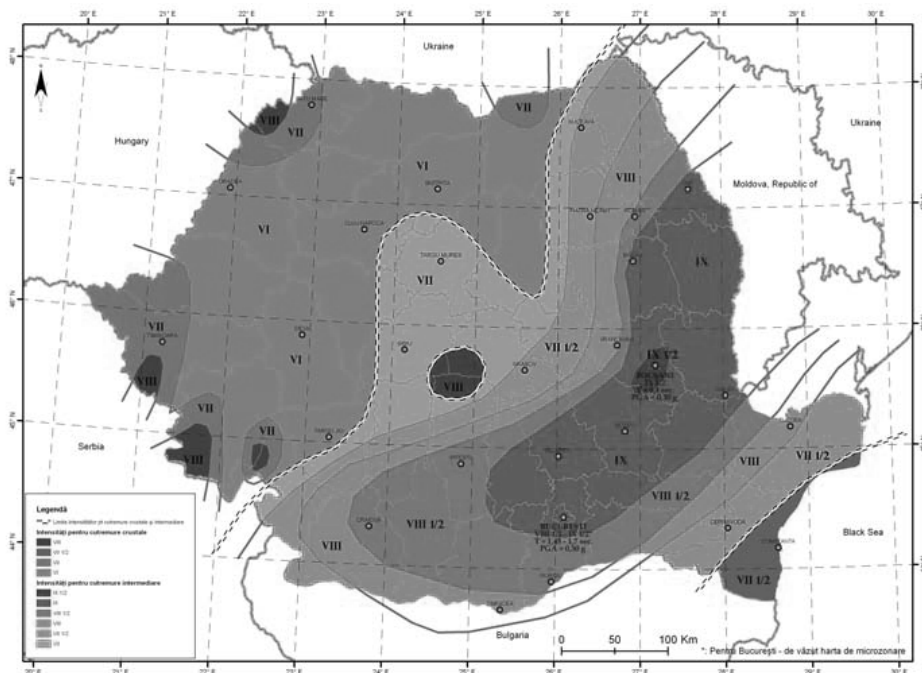


Рис. 7.7. Поле ізосейст можливого найбільшого землетрусу у вогнищі Вранча [Marmureanu et al. 2011].

7.6. Складання стохастичного каталогу землетрусів

Сучасні моделі будови Землі і теорії, що пояснюють виникнення землетрусів, базуються на непрямих даних, в основному на сейсмічних спостереженнях. Основною метою геофізичних досліджень є рішення оберненої задачі, тобто визначення будови середовища за спостереженнями над характеристиками фізичних полів. По суті, завдання сейсмології – це теж зворотні завдання. Сейсмологія займається вивченням землетрусів і пов'язаних з ними явищ. Основне завдання сейсмології – навчитися передбачити силу, час і місце виникнення землетрусів. Це завдання вирішується на базі відомостей з каталогів землетрусів: важливою продукцією сейсмології, і їх складання пов'язано з обробкою великого обсягу геолого-геофізичних і сейсмічних даних. Однак, часто, каталог не відображає всю сейсмічну обстановку в сейсмічних зонах, і не відповідає реальній структурі потоку сейсмічних подій, тому що не включає в себе значну частину землетрусів з магнітудою нижче мінімальної магнітуди комплектності, яка залежить від періоду часу спостереження. Представницькі каталоги охоплюють відносно короткий інтервал часу – близько декількох десятиліть. З огляду на досить поважний вік вогнищ землетрусів, який обчислюється мільйонами років, сейсмічний процес на відрізок часу, який представляє практичний інтерес, можна вважати стаціонарним. Дефіцит відомостей про сейсмічні події можна компенсувати складанням стохастичних каталогів за допомогою Монте-Карло моделювання (стохастичне моделювання) каталогів: генеруванням випадкових послідовностей сейсмічних подій.

Для ймовірнісного аналізу сейсмічної небезпеки складається модель сейсмічного режиму зони вогнищ землетрусів, яка з певним ступенем повноти описує просторово-часовий розподіл землетрусів в цій зоні. Для реалізації Монте-Карло моделювання для деякого регіону використовується існуюча модель сейсмічності регіону.

У широкому сенсі моделювання синтетичного каталогу (стохастичний каталог) ґрунтується на існуючій моделі сейсмічності, на сейсмічних записах, на відомостях

про історичні сейсмічності, на даних про регіональну тектоніку і про геолого-геофізичні особливості регіону.

Часовий режим сейсмічності моделюється законом повторюваності; закон Гутенберга-Ріхтера відображає співвідношення між магнітудою землетрусів і частотою їх виникнення в заданому регіоні і періоді часу, а в активній разломній зоні використовується характеристичне магнітудно-частотне співвідношення [Hardebeck et al., 2008; Ismail-Zadeh et al., 1999; Kagan, 1991; Milutinovic, 2013; Robinson et al., 2009; Sokolov, Ismail-Zadeh, 2015]. Закон повторюваності характеризує ймовірність виникнення землетрусів певної сили в деякій області за одиницю часу. Передбачається, що землетруси розподілені рівномірно в межах даної просторово – тимчасової області. Це означає, що інтенсивність виникнення сейсмічних подій не залежить від обсягу і від геометрії осередкової зони. Однак дослідження показують, що в широкому діапазоні магнітуд графік повторюваності не є лінійним, спостережені оцінки повторюваності землетрусів середньої магнітуди за різними періодами часу можуть значно відрізнятися [Гусєв, Шумилина, 2004]. Для великих сейсмічних подій оцінки за графіком повторюваності виявляються значно нижче спостережених.

Згенерувавши різні варіанти стохастичних каталогів землетрусів вдається змодельовати різні сценарії розвитку сейсмічної обстановки в зонах вогнищ землетрусів, і вибрати параметри сейсмічного режиму, при яких стохастичний каталог землетрусів буде максимально відповідати історичній сейсмічності. Стохастичний каталог може бути корисним також при перевірці якості методів прогнозу землетрусів. Основним припущенням при моделюванні каталогів землетрусів є гіпотеза про те, що на виникнення сейсмічних подій не впливають ніякі сили і відбуваються випадково, тобто потік сейсмічних подій відповідає пуассонівській моделі.

Існує безліч алгоритмів створення стохастичних каталогів [Crowley, 2005; Hardebeck et al., 2008; Ismail-Zadeh et al., 1999; Kagan, 1991; Milutinovic, 2013; Robinson et al., 2009; Zolfaghari, 2000], які, в принципі, мало чим відрізняються. У

всіх випадках каталог повинен містити відомості про параметри землетрусів, достатніх для аналізу сейсмічної небезпеки: час виникнення; розташування вогнища; магнітуда і глибина землетрусів. Значення параметрів сейсмічності, як значення випадкових величин задається конкретною функцією розподілу. Щоб згенерувати випадкові числа, відповідні необхідній функції розподілу параметрів сейсмічності, виходять від рівномірно розподілених на інтервалі (0,1) випадкових величин, які за допомогою відповідних перетворень приводяться до значень модельованих параметрів землетрусів. Таким чином, алгоритм складається з наступних кроків [Ikeda, 2010; Musson, 1999; Naylor et al., 2009; Sigbjornsson et al., 1996; Weatherill и Burton, 2006]:

- створення випадкових чисел
- перетворення випадкових чисел для приведення до виду, відповідного розподілу конкретного параметра землетрусів.

Час виникнення:

Землетруси відбуваються у випадкові моменти часу, а афтершоки виникають після великих сейсмічних подій, і їх активність зменшується з часом:

$$\lambda = \mu + \sum_{t_i < t} \frac{K \cdot 10^{\alpha(m_i - m_0)}}{(t - t_i + c)^p} \quad (7.78)$$

λ – денна інтенсивність землетрусів за час $t-t_i$, після i -го землетрусу, з магнітудою m_i .

η – денна інтенсивність фонові сейсмічності

m_0 – мінімальне значення, землетрус з магнітудою нижче якого не супроводжується афтершоками

α – постійна "продуктивності" характеризує сейсмічну активність

c – постійна величина, висловлює зміщення в часі

p – ступінь загасання інтенсивності афтершоків

K – постійна Гутенберга-Ріхтера.

Щоб використовувати формулу (7.78), на основі відомостей про час виникнення і магнітуди землетрусів в регіоні обчислюються параметри: μ , m_0 , α , c , p , K .

Розташування вогнищ:

Вогнища землетрусів повинні розташовуватися близько до тектонічних розломах, а афтершоки, в просторі, повинні прив'язуватися до головного поштовху, на деякій дистанції і напрямку від нього, які є випадковими величинами. Віддаль залежить від магнітуди головного поштовху. Далі територія розбивається на географічні осередки розміру $L \times L \text{ км}^2$, і визначається число землетрусів в осередках після видалення афтершоків. Вибір осередку здійснюється випадково, для чого генеруються випадкові числа. Епіцентри землетрусів, випадково розподілені навколо центра осередку, моделюються генератором випадкових чисел.

Магнітуда:

- розподіл сили землетрусу, як правило, описується магнітудно - частотним співвідношенням Гутенберга-Ріхтера:

$$\lg N(m) = a_t - b(m - m_t); m \geq m_t \quad (7.79)$$

де $N(m)$ число землетрусів з магнітудою не менше m , a_t і b параметри

- розподіл Гутенберга-Ріхтера (7.79) перетворено в усічений розподіл Парето з щільністю [Kagan, 2001]:

$$f(M) = \beta M_t^\beta M^{-1-\beta}; M \geq M_t \quad (7.80)$$

і з функцією розподілу ймовірності події, що складається в тому, що сейсмічний момент більше або дорівнює M [Eugster, 1999]:

$$\Phi(M) = \left(\frac{M_t}{M} \right)^\beta e^{-\frac{M_t - M}{M_c}} \quad (7.81)$$

де M_t – нижній поріг значень сейсмічного моменту; M_c – кутовий момент; $\beta = (2/3)b$.
Рішенням рівняння (7.81) щодо M , після деяких перетворень, отримана формула для генерування магнітуди [Ikeda, 2010]:

$$m = \frac{2}{3} \log \beta + m_c + \frac{2}{3} \log W \left(\frac{10^{1.5(m_t - m_c)} e^{\frac{10^{1.5(m_t - m_c)}}{\beta}}}{\beta r \frac{1}{\beta}} \right) \quad (7.82)$$

через r позначена випадкова величина (7.81), W - функція Ламберта. Методом перетворень отримано ітеративний алгоритм генерування магнітуд [Ikeda, 2010]:

$$w_{j+1} = w_j - \frac{w_j e^{w_j} - x}{e^{w_j} (w_j + 1) - \frac{(w_j + 2)(w_j e^{w_j} - x)}{2w_j + 2}} \quad (7.83)$$

Для значень $x \geq e$, функція $w(x)$ може бути апроксимована виразом $\ln x - \ln(\ln x)$. При $x < e$, для значень аргументу в межах нуля, функція Ламберта наближається поліномом четвертого степеня [Ikeda, 2010; Naylor et al., 2009]:

$$y = -0.0285x^4 + 0.1892x^3 - 0.508x^2 + 0.9138x; \quad M \geq M_1, \quad (7.84)$$

Глибина: Обчислюється середня глибина землетрусів в регіоні, і середня глибина в осередках $LxLкм^2$. Генеруються нормально розподілені випадкові величини з параметрами розподілу рівними вибіркового середнє арифметичному і стандартному відхиленню значень глибини спостережених в межах осередків землетрусів. Якщо ж в осередку не спостерігалися землетруси, то параметрами розподілу присвоювалися вибіркові регіональні значення.

У даній роботі пропонується більш простий алгоритм створення стохастичних каталогів, заснований на поданні зони вогнищ у вигляді об'єднання ЕДЗ: $\Xi^s = \{\Sigma_r\}$, $r=1, \dots, R^s$, де R^s – число ЕДЗ. Сейсмічний режим ЕДЗ стає більш однорідним в просторі, в діапазоні магнітуд і в часі. Генеруються рівномірно розподілені на інтервалі $(0,1)$ випадкові числа, які за допомогою відповідних перетворень приводяться до розподілу конкретного параметра землетрусів: магнітуда; час виникнення; географічні широта та довгота епіцентру, і глибина вогнища землетрусів. Методами математичної статистики визначається теоретична функція розподілу параметрів землетрусів, яка служить наближенням вибіркової функції розподілу відповідних параметрів.

Далі генерується синтетичний каталог для ЕДЗ. Підсумковий синтетичний каталог зони складається з об'єднання синтетичних каталогів усіх ЕДЗ. Для цього досить нанести на вісь часу моменти виникнення землетрусів в штучних каталогах і

зберегти, з урахуванням хронології, відповідні відомості. Генератор випадкових чисел інтегрований в пакет прикладних програм мови програмування ФОРТРАН.

Висновки до розділу 7

Здійснено ймовірнісний аналіз сейсмічної небезпеки територій Болгарії, Румунії та Молдови на основі розробленої автором методології. Проведено статистичний аналіз макросейсмічного поля значень інтенсивності сейсмічних впливів вранчських землетрусів. Досить великий обсяг макросейсмічних даних землетрусів 10.11.1940, 7.04.1977, 31.08.1986, 30.05.1990 та 31.05.1990 рр. вогнища Вранча, дозволяє проводити детальний статистичний аналіз. Гіпотеза про відповідність похибок – різниця між спостереженими значеннями інтенсивності і теоретичними значеннями, обчислені за рівнянням загасання, нормальним законом розподілу підтвердилася. На основі леми Неймана-Пірсона, за критерієм розрізнення двох довільних розподілів, гіпотеза H_0 : про те, що різниці між інтенсивностями, що спостерігаються і теоретичними значеннями інтенсивності, обчисленими за рівнянням загасання, відповідають стандартному нормальному розподілу $N(0,1)$ відхилення, на користь альтернативної гіпотези H_1 : яка у тому, що похибки становлять вибірку з нормально розподіленої генеральної сукупності з нульовим математичним очікуванням і відмінним від одиниці середнім квадратичним відхиленням.

Складено алгоритм оцінки сейсмічної небезпеки як ймовірності того, що в точці земної поверхні $Q(\varphi, \psi)$, з географічними координатами (φ, ψ) , за час t трапиться n сейсмічних струсів, m з яких з інтенсивністю не менше I^k балів. Побудована карта сейсмічної небезпеки на основі відомостей про румунські землетруси з каталогу ROMPLUS, території Молдови і загальної території Молдови, Румунії, Болгарії. Метод був застосований також і для аналізу сейсмічної небезпеки на території Туреччини. Обчислені характеристики сейсмічної небезпеки в Кишиневі. Запропонований метод аналізу сейсмічної небезпеки дозволяє враховувати власний

закон загасання інтенсивності струшувань для кожної сейсмічної зони і територій. Також розроблений алгоритм обчислення сейсмічного ризику.

Побудована карта найгіршого сценарію розвитку сейсмічної обстановки на даних територіях. Для цього, за допомогою статистики Гумбеля, оцінена максимально можлива магнітуда землетрусів у вогнищі Вранча, і побудована карта екстремальної сейсмічної небезпеки.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

- Алказ В.Г. Основы прогноза сейсмической опасности и сейсмического риска территории Республики Молдова. Кишинев, 2007, 229.
- Буртиев Р.З. Оценка максимальной магнитуды землетрясений очага Вранча по статистике экстремальных значений. *Analele ATIC*. Vol. I(II). Chisinau. 2003, 110–117.
- Буртиев Р. Методология оценки сейсмической опасности. LAMBERT Academic Publishing, 2017, 281.
- Гнеденко Б.В., Курс теории вероятностей. Изд. Физмат литературы. Москва, 1961, 406.
- Гумбель Э. Статистика экстремальных значений. М.: Мир, 1965, 449.
- Гусев А.А., Шумилина Л.С. Повторяемость сильных землетрясений Камчатки в шкале моментных магнитуд. *Физика Земли* №3, 2004, 34–42.
- Кендалл М. Дж., Стьюарт А., Статистические выводы и связи. Наука, Москва, 1973, 898.
- Колемаев В. А., Калинина В. Н., Теория вероятностей и математическая статистика, Москва, ИНФРА – М, 1997, 302.
- Родкин М. В., Теория экстремальных значений, распределения вероятностей для хвоста распределения для основных типов распределений, вопросы применения. Международный институт теории прогноза землетрясений и математической геофизики РАН, 2005, 27.

- Розин И. Т., Харитонов Е. В. О статистическом определении электрической прочности диэлектриков. Заводская лаборатория. Москва, Металлургия, 1972, 47–53.
- Тейлор Дж., Введение в теорию ошибок. Москва, Мир, 1985, 272.
- Уилкс С., Математическая статистика. Наука, Москва, 1967, 631.
- Уломов В.И., Сейсмичность и сейсмическое районирование Северной Евразии. Отв. Ред. Вып. 1, Москва, ИФЗ РАН, 1993, 303.
- Уломов В.И., Богданов М.И. Новый комплект карт общего сейсмического районирования территории Российской Федерации (ОСР–2012). Инженерные изыскания, 8, 2013. 8–17.
- Шебалин Н.В. Бальность, магнитуда и глубина очага землетрясений. Землетрясения в СССР. Москва. Наука. 1961, 126–138.
- Al–Abbasi J., Fahni K. Estimating maximum magnitude earthquakes in Iraq using extreme value statistics. Geophys. J. R. Ast. Soc. 82, 1985, 535–548.
- L. Ardeleanu, G. Leydecker, K.-P. Bonjer, H. Busche, D. Kaiser and T. Schmitt, Probabilistic seismic hazard map for Romania as a basis for a new building code, Natural Hazards and Earth System Sciences, 5, 2005, 679–684.
- L. Ardeleanu, G. Leydecker, T. Schmitt, K.-P. Bonjer, H. Busche, D. Kaiser, S.D. Simeonova, D.E. Solakov, PROBABILISTIC SEISMIC HAZARD MAPS IN TERMS OF INTENSITIES FOR ROMANIA AND BULGARIA, International Symposium on Strong Vrancea Earthquakes and Risk Mitigation Oct. 4–6, 2007, Bucharest, Romania.
- Baker W. Jack. An Introduction to Probabilistic Seismic Hazard Analysis (PSHA), 2008, 1–72.
- Birkner Matthias. Angewandte Stochastik. Felix Hofmann, Universität Bonn, 2009, 72.
- Bülow Katharina, Jacob Daniela. Änderung der Auftretswahrscheinlichkeiten von Extremereignissen in modellierten Temperatur– und Niederschlagszeitreihen des 21. Jahrhunderts. Max–Planck–Institut für Meteorologie, Hamburg, 2007, 7.

- Burtiev R.Z. Seismic Hazard Mapping Based on Multi-dimensional Markov Model of a Seismic Regime // *Gerlands Beitrage zur Geophysik, Leipzig*, 95, 1986, 499– 508.
- Burtiev Rashid. Evaluation of seismic hazards from several seismic zones. *Environmental Engineering and Management Journal*. 12, 2012. 32–42.
- Burtiev Rashid, Alcaz Vasile, Cardanets Vladlen, Probabilistic Seismic Hazard Analysis on the Base of the Stochastic Models of Seismicity, *Indian Journal of Applied Research*, V.6, Issue 8, 2016, India. 454 – 466.
- Burtiev Rashid, Seismic Hazard Assessment Method Based on the Stochastic Models of Seismicity, *Bulletin of the International Institute of Seismology and Earthquake Engineering*, Vol.51, 2017, Japan, 22–31.
- Crowley Helen. An investigative study on the modelling of earthquake hazard for loss assessment. An Individual Study Submitted in Partial Fulfilment of the Requirements for the Doctor of Philosophy Degree in earthquake engineering. *Universita degli Studi di Pavia*, 2005, 85.
- Dessekey M. Extreme value models for seismic hazard analysis. *Bull. of the International Institute of Seismology and Earthquake Engineering*. 1985, 23–41.
- Enescu D., Mărmureanu A., Enescu B.D., A procedure for assessing seismic hazard generated by Vrancea earthquakes and its application. II. Attenuation curves, *Romanian Reports in Physics*, 56 (1), 2004, 147–159.
- Erdik M., Demircioglu M., Sesetyan K., Durukal E., Siyahi B., Earthquake hazard in Marmara Region, Turkey. *Soil Dynamics and Earthquake Engineering*. Elsevier Ltd. 24, 2004, 605–631.
- Eugster, S., Ruttener, E. and Liechti, D. The risk premium distribution (annual average loss) with respect to earthquake magnitude. XXIV European Geophysical Society General Assembly, The Hague, The Netherlands, 1999.
- Ferger D., *Mathematical Statistics (in German)*, Technische Universität Dresden, Germany, 2010, 147.
- Fisz M., *Wahrscheinlichkeitsrechnung und Mathematische Statistik*. VEB Deutscher Verlag der Wissenschaften, Berlin, 1966, 551.

- Golbs Christian, Probabilistische seismische Gefährdungsanalysen auf der Grundlage von Epizentrendichten und ihre ingenieurpraktischen Anwendungsgebiete. Dissertation zur Erlangung des akademischen Grades, Weimar, Deutschland, 2008, 193.
- Grachyova Marina, Müller Alexander. Maximum–Likelihood Schätzung, Statistische Methoden und ihre Anwendungen, Universität Kockelkorn, Einführung in die Statistik. 2006, 17–20.
- Grieser Jürgen, Staeger Tim, Schönwiese Christian, Zur Schätzung der Wiederkehrzeiten von Starkniederschlägen. Institut für Atmosphäre und Umwelt der J. W. Goethe–Universität Frankfurt/M. 2009, 35.
- Hamann Anna. Analyse von Extremwerten, Institut für Statistik Ludwig–Maximilians–Universität München. München, 2009, 18.
- Hardebeck L. Jeanne, Felzer R. Karen and Michael J. Andrew, Improved tests reveal that the accelerating moment release hypothesis is statistically insignificant. JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 113, 2008, 19.
- Hözlwimmer Manula, Extremwerte II. Peak over Threshold. Statistische Verfahren in den Geowissenschaften, WS, 2010, 17.
- Ikeda Nancy. staffwww.fullcoll.edu/.../ikedaabstract.htm, 2010.
- Ismail–Zadeh A.T., Keilis–Borok V.I., Soloviev A.A., Numerical modelling of earthquake flow in the southeastern Carpathians (Vrancea): effect of a sinking slab, <http://www.sciencedirect.com/science/article/pii/S0031920198001666> – CORR1
- Ismail-Zadeh, A., Sokolov, V. and Bonjer, K., Tectonic stress, seismicity, and seismic hazard in the southeastern Carpathians, Nat. Hazards, Springer, 2007, 1–22.
- Jackman Simon. Maximum Likelihood. Stanford University, 2006, 24.
- Jonas Martin. Zeitlich gleitende extremwertstatistische Untersuchung von monatlichen Temperatur– und Niederschlagsreihen in Deutschland. Johann Wolfgang Goethe Universität Frankfurt/Main, 2003, 103.
- Kabluchko Zakhar Kabluchko. Extremwerttheorie. Universität Ulm, ULM, 2011, 77.

- Kagan Y., Likelihood analysis of earthquake catalogues. *Geophys. J. Int.* 1991, 106, 135–148.
- Kagan Y., EARTHQUAKE SIZE DISTRIBUTION: POWER–LAW WITH EXPONENT. Department of Earth and Space Sciences, University of California, Los Angeles, California, USA, 2001, 53.
- Kalafat D, Güneş Y, Kekovalı K, Kara M, Deniz P, Yılmaz M. (2011). Bütünleştirilmiş Homojen Türkiye Deprem Kataloğu (1900-2010; $M \geq 4.0$). İstanbul, Turkey: Boğaziçi Üniversitesi, Kandilli Rasathanesi ve Deprem Araştırma Enstitüsü, İstanbul, 2011.
- G. Leydecker, H. Busche, K.-P. Bonjer, T. Schmitt, D. Kaiser, S. Simeonova, D. Solakov and L. Ardeleanu, Probabilistic seismic hazard in terms of intensities for Bulgaria and Romania – updated hazard maps, *Nat. Hazards Earth Syst. Sci.*, 8, 1431–1439, 2008.
- Luchsinger C.J. *Statistical Methods*, Departement Mathematik, Universität Basel, 1997, 103.
- Lungu D., Cornea T., Nedelcu C., Hazard assessment and site–dependent response for Vrancea earthquakes. *Vrancea earthquakes: Tectonics, Hazard and Risk mitigation*. Edited by: F. Wenzel and D. Lungu, Co–Editor O. Novak. Kluwer Academic publishers, 1997, 251–267.
- Lungu D., Aldea A., Demetriu S. and Arion C., Probabilistic zonation of Vrancea seismic hazard. EUROCODE 8 representation of design action. Earthquake loss estimation and risk reduction. *Proceedings of the International Conference*. Bucarest, 2002, 137–150.
- Marmureanu G., Cioflan C.O., Marmureanu A., Intensity seismic hazard map of Romania by probabilistic and (neo) deterministic approaches, linear and nonlinear analyses, *Romanian Reports in Physics*, 63 (1), 2011, 226–239.
- Milutinovic Zoran, Improvements in the Harmonized Seismic Hazard Maps for the Western Balkan Countries. SFP 984374, Improvements and Harmonization of Macedonian Earthquake Catalogue. 2nd Project Workshop, Belgrade, 2013, 15.

- Moldovan I.A., Popescu E., Constantin A., Probabilistic seismic hazard assessment in Romania: application for crustal seismic active zones, *Rom. Journ. Phys.*, 53 (3–4), 2008, 575–591.
- Musson, R.M.W. Determination of design earthquakes in seismic hazard analysis through Monte Carlo simulation. *Journal of Earthquake Engineering*, 463–474, 1999.
- Musson R.M.W., Intensity-based seismic risk assesment. *Soil Dynamics and Earthquake Engineering*, ELSEVIER, 2000, 353–360.
- Naylor M., Greenhough J., McCloskey J., Bell A.F., and Main I. G., Statistical evaluation of characteristic earthquakes in the frequency magnitude distributions of Sumatra and other subduction zone regions. *GEOPHYSICAL RESEARCH LETTERS*, VOL. 36, 2009.
- Peichl H. Gunther, (1999), *Introduction to Probability and Statistics (in German)*, Institut für Mathematik Karl–Franzens–Universität Graz, 1999, 194.
- Raschke, M., Kröger, W., Bilis, V. (2011): Erdbeben und Wintersturmrisko des Schweizer Stromübertragungsnetzes. *SEV/VSE-Bulletin*, 2011, 5.
- Robinson Russell, Van Dissen Russell and Litchfield Nicola, *Using Synthetic Seismicity to Evaluate Seismic Hazard in the Wellington Region, New Zealand GNS Science*, New Zealand, 2009, 67.
- Shebalin N.V., Leydecker G., Mokrushina N.G., Tatevossian R. E. and Vassiliev V. YU.: *Earthquake Catalogue for Central and Southeastern Europe 342 BC – 1990 AD*. European Commission, Report No. ETNU CT 93 – 0087, Brussels. 1998.
- Sigbjornsson Ragnar, Baldvinsson Gunnar, Thrainsson Hjörtur. *The mapping of seismic hazard using stochastic simulation and geographic information systems*. Elsevier Science LTD paper No. 710, 1996, 1–8.
- Sokolov Vladimir, Wenzel Friedemann, Mohindra Rakesh, Grecu Bogdan, Radulian Mircea, 2007, Probabilistic seismic hazard assessment for Romania considering intermediate –depth (vrancea) and shallow (crustal) seismicity: *International Symposium on Strong Vrancea Earthquakes and Risk Mitigation*, Oct. 4–6, 2007, Bucharest, 1–16.

- Sokolov, V., and Ismail-Zadeh, A., Seismic hazard from instrumentally recorded, historical and simulated earthquakes: Application to the Tibet-Himalayan region, *Tectonophysics*, 657, 187–204, 2015 (http://www.mitp.ru/~aismail/papers/VS_AIZ_Tecto_2015.pdf).
- Sokolov, V., and Ismail-Zadeh, A. On the Use of Multiple-Site Estimations in Probabilistic Seismic Hazard Assessment, *Bull. Seismol. Soc. Am.*, 106(5), 2233–2243, 2016 (http://www.mitp.ru/~aismail/papers/2016_BSSA_Multiple-Site.pdf).
- Solomos G., Pinto A., Dimova S., A Review of The Seismic Hazard Zonation in National Building Codes in the context of Eurocode 8, European Commission Joint Research Centre, Luxembourg, 2008, 72.
- Soong T.T., *Fundamentals of probability and statistics for engineers*. State University of New York at Buffalo. New York, USA, 2004, 298.
- Sorensen M.B., Stromeyer D. and Grünthal G., Estimation of macroseismic intensity – new attenuation and intensity vs. Ground motion relations for different Parts of Europe. The 14th World Conference on Earthquake Engineering, Beijing, China, 2008, 8.
- Trömel Silke. *Statistische Modellierung von Klimazeitreihen*, Johann Wolfgang Goethe – Universität in Frankfurt am Main, 2004, 242.
- Graeme A. Weatherill and Paul W. Burton, The Application of Multiple Random Earthquake Simulations to Probabilistic Seismic Hazard Assessment in the Aegean Region, First European Conference on Earthquake Engineering and Seismology Geneva, Switzerland, 2006, 1–10.
- Yüçemen M. S, Öztürk N. Y. and Deniz A., Probabilistic Seismic Loss Estimation for Eskisehir, Turkey Probabilistic Seismic Loss Estimation for Eskisehir, Turkey. Department of Civil Engineering and Earthquake Engineering Research Center Middle East Technical University, 2006, 18.
- Zolfaghari M.R. Earthquake loss estimation model for southern Europe. Proceedings of 6th International Conference on Seismic Zonation, 2000.

ЗАГАЛЬНІ ВИСНОВКИ

Фундаментальні дослідження сейсмічності і прогнозування сейсмічної небезпеки носять неперервний характер, оскільки удосконалюються системи спостережень, надходить якісно нова інформація, розробляються нові математичні та обчислювальні моделі сейсмічності і проектування сейсмостійких споруд. У даній роботі розроблена принципово нова методологія оцінки сейсмічної небезпеки на основі ймовірнісних моделей сейсмічності, яка дозволяє розрахувати сейсмічну небезпеку в населених пунктах, розташованих в радіусі впливу групи сейсмічних зон і побудувати карту сейсмічного районування великих територій. В якості запобіжної сейсмічної небезпеки розглядається ймовірність того, що в деякій точці земної поверхні, за деякий час трапиться певна кількість сейсмічних поштовхів, відповідне число яких матиме певну інтенсивність.

Методологія включає в себе чотири основні етапи:

- 1) на першому етапі визначаються сейсмічні зони і передбачається, що вогнища землетрусів розподілені з певною ймовірністю у геофізичній просторі, який займає зона вогнищ землетрусів
- 2) на другому етапі будується марковська модель сейсмічного режиму, яка дозволяє визначити розподіл параметрів землетрусів, необхідних для аналізу сейсмічної небезпеки
- 3) третім кроком є визначення параметрів рівняння макросейсмічного поля
- 4) на четвертому етапі розробляється алгоритм оцінки сейсмічної небезпеки в населеному пункті, який знаходиться в радіусі дії групи сейсмічних зон.

Сейсмічні зони виявлялися за допомогою кластерного аналізу. В якості пробної вибірки використовувався каталог землетрусів південного сходу Європи, що охоплює період 1900–1990 рр. У цьому каталозі містяться відомості про землетруси в 18 сейсмічних зонах, виявлених експертами. При порівнянні координат епіцентрів в сейсмічних зонах, виділених алгоритмом k -середніх, з координатами епіцентрів 18 сейсмічних зон, наведених в каталозі, спостерігається 97%-ний збіг. Висока частка

збігів є показником ефективності методу. Хоча кластерний аналіз, будучи формальним методом, не може враховувати геолого-геофізичні характеристики зони вогнищ землетрусів, він цілком придатний для оперативного сейсмічного зонування (групування вогнищ землетрусів). Методом кластерного аналізу на території Румунії виділено 13 сейсмічних зон, і 50 зон на території Туреччини.

Сейсмічна зона охоплює деякий простір геофізичного середовища, положення якого характеризується декартовими координатами і діапазонами можливих значень параметрів землетрусів. Розбиттям даного простору на 4-вимірні непересічні підпростори, отримуються елементарні джерела землетрусів (ЕДЗ). Виникнення землетрусу з параметрами з деякого ЕДЗ Σ_j в момент часу t_n означає спостереження сейсмічної події $\Sigma_j(t_n)$ в момент часу t_n . На сучасному рівні розвитку геофізики, моменти часу виникнення землетрусів, і значення їх параметрів розглядаються як випадкові величини. Це означає, що спостереження послідовності сейсмічних подій $\Sigma_j(t_k)$ ($j=1, \dots, R; n=1, \dots, N$), є реалізацією деякого випадкового процесу. Зона вогнищ землетрусів розглядається як фізична система, яка в випадкові моменти часу змінює свій стан і реалізацією якої є дана послідовність. Вичерпним способом визначення випадкового процесу є задання його спільної скінченновимірної функції розподілу, визначення якої, за винятком деяких тривіальних випадків, є завданням практично нездійсненним. Однак, існує клас марковських процесів, де спільні скінченномірні функції розподілу виражаються через одно і двовимірні розподіли. В цьому випадку, стає можливим з усіх розподілів визначення спільного розподілу параметрів кінцевої сукупності землетрусів.

Вогнища землетрусів розподілені випадково в деякому обсязі геофізичної середовища, де становище кожного вогнища може залежати від стану попереднього і від положення деякого числа попередніх землетрусів. Останнє твердження означає, що ряд сейсмічних подій може володіти довготривалою пам'яттю. Гіпотеза про те, що порядок опцій землетрусів є реалізацією ергодичного простого марковського ланцюга, виявилася правдоподібною. Ергодичні ланцюги Маркова володіють

чудовою для практичних застосувань властивістю: стаціонарні граничні розподіли станів такого ланцюга з будь-яким ступенем точності можна визначити з однієї досить довгої її реалізації. Це означає, що ми зможемо передбачити розподіл координат епіцентру, глибини і магнітуди майбутніх землетрусів. Марковська модель сейсмічних зон може допомогти виявити міграцію осередків землетрусів уздовж сейсмічних поясів і наявності зв'язку між сейсмічними регіонами. Виявлено статистично значущий зв'язок між активністю деяких сейсмічних регіонів Балкан.

Динаміка сейсмічної активності балканського регіону: чергування вогнищ землетрусів в регіонах – зображається елементами матриці ймовірностей переходу. Виявлено, що землетруси мають тенденцію відбутися там, де вони вже відбулися. Наприклад, по матриці ймовірностей переходу, видно, що після землетрусу в Румунії, наступний землетрус, швидше за все, трапиться теж в Румунії, тому що ймовірність переходу $\pi_{11}=0.224$ зі стану *ROM* в *ROM* найбільша. Другою за величиною є ймовірність $\pi_{13}=0.112$ переходу зі стану *ROM* в *BUL*, тобто, наступний після Румунії землетрус може статися, але з меншою ймовірністю в Болгарії. Ймовірність переходу $\pi_{14}=0.0$ зі стану *ROM* в *POL* дорівнює нулю. Це означає, що сейсмічні зони Румунії і Польщі не пов'язані. З вектора граничних розподілів сейсмічних зон можна визначити зону, де і з якою ймовірністю може відбутися очікуваний землетрус.

Критерії таблиці спряженості ознак також вказують на наявність середньої тісноти взаємозв'язку сейсмічної активності деяких зон.

Для застосування рівняння макросейсмічною поля, при оцінці сейсмічної небезпеки, необхідно привести значення різних магнітуд до єдиного типу. Перетворення магнітуд здійснюється за допомогою лінійної багатовимірної регресії, і виконуються умови теореми Гаусса-Маркова.

Дослідження гомоскедастичності – явища гомоскедастичності ґрунтуються на значеннях: тесту рангової кореляції Спірмена, тесту Парку, методу Глейзера. У рівняння залежності магнітудних шкал включалися додаткові регресори, проте це не привело до істотного поліпшення якості регресії.

Мінімальна магнітуда комплектності $M_c=2.3$, для землетрусів, що сталися після 1978 року, оцінена по методах *МАХС* і *МВS* збігається.

Між магнітудою і глибиною землетрусів виявлений статистично слабкий зв'язок. Для землетрусів в межах земної кори зв'язок значно слабший, ніж для проміжних землетрусів. Поясненням цього факту є те, що в межах земної кори відбуваються слабкі землетруси з магнітудою до 3.8, а в проміжному шарі з магнітудою до $mb=7.2$, і область підготовки більш сильних землетрусів охоплює більший обсяг геофізичного середовища.

Ймовірніший зв'язок між магнітудою і інтервалами часу між землетрусами Вранча виявився статистично незначним, а магнітуда і інтервали часу між землетрусами, в межах земної кори, пов'язані слабким лінійним зв'язком. Для землетрусів нижче межі M , статистичні характеристики взаємозв'язку свідчать про незалежність магнітуди від інтервалів часу. Виявлено слабкий зв'язок між річним числом землетрусів в межах земної кори і нижче границі Мохо. Статистична залежність між короною і підкороною сейсмічною активністю для землетрусів з магнітудою $M \geq 2.6$ практично відсутня.

Між щомісячним числом корових і проміжних землетрусів також існує слабкий статистичний зв'язок. В рядах річного та щомісячного числа землетрусів в межах кори існує тренд, а в проміжній сейсмічності припущення про відсутність тренда підтвердилося.

Проведено статистичний аналіз макросейсмічних даних вранчських землетрусів, що сталися 10.11.1940, 7.04.1977, 31.08.1986, 30.05.1990 та 31.05.1990 рр. Дослідження показують, що прийнятною згладжуючою лінією макросейсмічного поля є еліпс. З канонічного рівняння еліпса слідує, що фокальна вісь 5, 6 і 7 бальних зон, всіх досліджуваних землетрусів, спрямована по меридіану. Це означає, що 5, 6 і 7 бальні зони витягнуті в напрямку північ-південь, тоді як, інші зони витягнуті уздовж паралелі. Проведена апроксимація еліпсом багатьох точок рівної інтенсивності. Визначено параметри еліпсів, що апроксимують зони рівної бальності макросейсмічного поля. Коефіцієнт лінійної регресії епіцентральної

інтенсивності I^0 на магнітуду, для румунських землетрусів, вийшов рівним $b=1.5$. Оптимальними значеннями коефіцієнтів рівняння макросейсмічного поля виявилися: $a=1.6$; $c=7.2$; найбільшого $b_{\max}=5.6$ і найменшого $b_{\min}=4.9$ загасання інтенсивності струшувань; а кут між великою віссю еліпса загасання і позитивним напрямом осі абсцис рівним $\gamma_0=51^0$.

На основі макросейсмічних даних виявлена азимутальна неоднорідність геологічних структур Молдови і суміжних територій.

Теорія часових рядів застосовувалась для дослідження послідовності місячного числа румунських землетрусів. Як і очікувалося, в ряді щомісячного числа землетрусів присутня довгострокова систематична складова. Припущення про стаціонарності за середнім значенням і по дисперсії відкидається. Для прогнозування ряду застосовувався метод експоненціального згладжування. Найкращими, за вибірковими значеннями статистики критеріїв перевірки якості моделі, виявилися проста сезонна, для прологарифмованого ряду і адитивна модель Уінтерса. Для прогнозу застосовувалися також лінійні моделі авторегресії. Параметри моделей рядів визначалися з поведінки автокореляційної і приватної автокореляційної функцій. Авторегресійні методи (модель *ARIMA*) застосовуються для аналізу і прогнозу однорідних нестационарних часових рядів з нелінійним трендом, типу полінома невисокого степеня, яким є досліджуваний ряд. Найкращою моделлю для прогнозу часових рядів виявилася $ARIMA(0,1,2) \times (1,0,1)_s$ модель, підібрана конструктором моделей пакета SPSS. Сейсмічний процес - це циклічний процес. Результати досліджень були б більш достовірними, якби довжина ряду перевищувала хоча б період одного сейсмічного циклу і була б достатнього обсягу сейсмостатистика.

Труднощі, пов'язані з нестачею відомостей про сейсмічних спостереженнях, можна вирішити створенням штучних каталогів землетрусів за допомогою Монте-Карло моделювання. В межах ЕДЗ просторово-часовий розподіл вогнищ землетрусів стає більш однорідним. Отже, можливе якісне моделювання стохастичних каталогів землетрусів.

Здійснено статистичний аналіз похибок – різниці спостережених і обчислених за рівнянням макросейсмічного поля значень інтенсивності сейсмічних впливів вранчських землетрусів. Досить великий обсяг макросейсмічних даних землетрусів 10.11.1940, 7.04.1977, 31.08.1986, 30.05.1990 та 31.05.1990 рр. вогнища Вранча, дозволяє проводити детальний статистичний аналіз.

Складено алгоритм оцінки міри сейсмічної небезпеки як ймовірності того, що в точці земної поверхні $Q(\varphi, \psi)$ з географічними координатами (φ, ψ) за час t трапиться n сейсмічних поштовхів, m з яких з інтенсивністю не менше I^k балів. Здійснено ймовірнісний аналіз сейсмічної небезпеки територій Болгарії, Румунії, Молдови та Туреччини, як в балах макросейсмічною шкали, так і за нормами ЄС EC8.

За допомогою кластерного аналізу на території Румунії виділено 13 і 50 сейсмічних зон на території Туреччини. Побудована марковська модель сейсмічного режиму виділених зон, і карта сейсмічної небезпеки території Молдови, Румунії, Болгарії та Туреччини. Розроблений метод аналізу сейсмічної небезпеки дозволяє враховувати власний закон загасання інтенсивності струсів для кожної сейсмічної зони і територій. інтенсивності струшувань для кожної сейсмічної зони і територій.

Побудована карта найгіршого сценарію розвитку сейсмічної обстановки на території Молдови, Румунії та Болгарії, на основі максимально можливої магнітуди землетрусів у вогнищі Вранча. При розрахунку сейсмічної небезпеки загасання вплив проміжних землетрусів оцінювався по виведеному автором рівнянню макросейсмічного поля проміжних вранчських землетрусів. Через відсутність макросейсмічних даних корових румунських землетрусів загасання інтенсивності оцінювалося за середніми для Балкан значеннями коефіцієнтів рівняння макросейсмічного поля. Тому результати здійсненого аналізу сейсмічної небезпеки носять попередній характер.

Запропонований метод дозволяє розрахувати сейсмічну небезпеку в населених пунктах, розташованих в радіусі впливу деякого числа сейсмічних зон, і побудувати карту сейсмічного районування великих територій, наприклад, всієї Європи.

СПИСОК ПУБЛІКАЦІЙ ЗДОБУВАЧА ЗА ТЕМОЮ ДИСЕРТАЦІЇ.

Монографія

1. Буртиев Рашид, Методология оценки сейсмической опасности, LAP LAMBERT Academic Publishing, 2017, 281.

Статті

2. Burtiev Rashid, Seismic hazard mapping based of the multi-dimensional markov model of a seismic regime. Gerlands Beitrage zur Geophysik. №95, Leibzig, 1986. 10.
3. Буртиев Рашид, О применении статистики Гумбеля к оценке максимальной магнитуды, Buletinul Institutului de Geologie si Seismologie al ASM. 2007, Nr. 2.5–7.
4. Буртиев Рашид, Статистическая связь между сейсмическими зонами Балкан, Buletinul Institutului de Geologie și Seismologie al ASM.2008, Nr.1. 31–34
5. Буртиев Рашид, О статистическом подходе к картированию сейсмической опасности больших территорий. Buletinul Institutului de Geologie și Seismologie al ASM. 2008, Nr.2. 57–61.
6. Буртиев Рашид, Применение линейной регрессии для преобразования разных магнитуд в шкалу магнитуды mb. Buletinul Institutului de Geologie și Seismologie al ASM. 2009. Nr.1, 62–71.
7. Буртиев Рашид, О возможности применения многомерной линейной регрессии с коллинеарными факторами для преобразования магнитуд. Buletinul Institutului de Geologie și Seismologie al ASM.2009. Nr.2, 35–44
8. Буртиев Рашид, О применении регрессии в условиях гетероскедастичности сейсмических данных. Analele ATIC. I (IV). Chisinau. 2010, 177–180.
9. Буртиев Рашид, О статистической связи сейсмической активности Балканских сейсмических зон. Buletinul Institutului de Geologie si Seismologie al ASM.2010. Nr.1, 5–15

10. Буртиев Рашид, Применение кластерного анализа для сейсмического районирования Балканского региона. *Buletinul Institutului de Geologie si Seismologie al ASM*.2010. Nr.2, 55–62.
11. Буртиев Рашид, Статистическая связь между коровой и подкоровой сейсмическими активностями очага Вранча. *Buletinul Institutului de Geologie si Seismologie al ASM*.2011. Nr. 2, 133–148
12. Burtiev Rashid, Evaluation of seismic hazards from several seismic zones. *Environmental Engineering and Management Journal*. N12, 2012, 2141–2150.
13. Буртиев Рашид, Вероятностный метод оценки сейсмической опасности, *Buletinul INCERCOM*, Nr. 4, 2013, 41–47.
14. Буртиев Рашид, Павлов Петр, Исследование азимутальной неоднородности геологической среды по макросейсмическим данным. *Buletinul Institutului de Geologie si Seismologie al ASM*, 2013, Nr.2, 38–46.
15. Буртиев Рашид, Определение расстояния от точки наблюдения до элементарного источника землетрясений. *Buletinul Institutului de Geologie si Seismologie al ASM*, 2014, Nr.1, 23–29.
16. Burtiev Rashid, Fitting an Ellipse to the set of intensity data points of Vrancea earthquakes. *Journal of Engineering Studies and Research – Volume 20 (2014), No, 4 Baceu, Romania*, 2014, 15–24.
17. Burtiev Rashid, Time Series in the Study of Seismic Regime of Vrancea (Romania) Seismic Zone, *The Global Environmental Engineers*, 2014, Volume 1, N2, Karachi, Pakistan, 54–63.
18. Буртиев Рашид, Вероятностный анализ экстремальной сейсмической опасности в Молдове и Румынии. *Buletinul INCERCOM*, Nr. 4, 2015, 41–47.
19. Буртиев Рашид, Вероятностный анализ сейсмической опасности территории Молдовы и сопредельных районов, *Геофизический журнал № 6, Т. 37, 2015, Киев*, 124–130.

20. Burtiev Rashid, Alcaz Vasile, Cardanets Vladlen, Probabilistic Seismic Hazard Analysis on the Base of the Stochastic Models of Seismicity, INDIAN JOURNAL OF APPLIED RESEARCH, V.6, Issue 8, 2016, India. 454–466.
21. Буртиев Рашид, Методы расчета сейсмического риска и составления синтетического каталога землетрясений с использованием Монте-Карло моделирования, Buletinul Institutului de Geologie si Seismologie al ASM, 2016, Nr.2. 87–96.
22. Burtiev Rashid, Seismic Hazard Assessment Method Based on the Stochastic Models of Seismicity, Bulletin of the International Institute of Seismology and Earthquake Engineering, Vol.51, 2017, Tsukuba, Japan, 22–38.

Тези конференцій

23. Burtiev Rashid, Cluster analysis to select the seismic zones of Romania. First International Conference on Moldavian Risks-from Global to local scale. Book of abstracts, 2012, Bacau, Romania. p. 25
24. Burtiev Rashid, Probabilistic Seismic Hazard Analysis for the Territory of Moldova and Romania. Second International Conference on Moldavian Risks-from Global to local scale, Book of abstracts, 2014, Bacau, Romania, p. 24.
25. Burtiev Rashid, A new approach to probabilistic seismic hazard analysis, Georisks in the Mediterranean and Their Mitigation. University of Malta, 2015, p. 154.
26. Burtiev Rashid, Evaluation of seismic hazard on base of probabilistic models of Vrancea zones, The National Symposium “75 Years from November 10th 1940 Vrancea Earthquake”, Book of Abstracts, Bucarest, 2015, p. 14.