

УДК [004.67+004.9]:[550.3+57.045]

## ПРОБЛЕМЫ АНАЛИЗА ВРЕМЕННЫХ РЯДОВ С ПРОПУСКАМИ И МЕТОДЫ ИХ РЕШЕНИЯ В ПРОГРАММЕ WINABD

© 2016 г. А.В. Дещеревский<sup>1</sup>, В.И. Журавлев<sup>1</sup>, А.Н. Никольский<sup>2</sup>, А.Я. Сидорин<sup>1</sup>

<sup>1</sup> Институт физики Земли им. О.Ю. Шмидта РАН, г. Москва, Россия

<sup>2</sup> ООО «КМК Консалтинг», г. Москва, Россия

Рассмотрены технологии, применяемые при анализе временных рядов с пропусками. Обсуждаются некоторые алгоритмы выделения (очистки) сигнала и оценки его характеристик, в частности ритмических составляющих, для рядов с пропущенными наблюдениями. Приведены примеры анализа данных, полученных при долговременных наблюдениях на Гармском геофизическом полигоне и в других регионах. Рассмотрены используемые в программе WinABD технические решения, позволяющие наиболее эффективно организовать работу обсуждаемых алгоритмов при наличии дефектов наблюдений.

*Ключевые слова:* данные геофизического мониторинга, анализ временных рядов с пропусками, алгоритм, программа, обработка в скользящем окне, ритмы.

### Литература

- Абраменкова И.В., Круглов В.В.* Методы восстановления пропусков в массивах данных // Программные продукты и системы. 2005. № 2. С. 18–22.
- Автоматизированная обработка данных на Гармском полигоне / Отв. ред. А.Я. Сидорин. М.: ИФЗ АН СССР, 1991. 216 с.
- Богдасарян Р.А.* Частотно-индивидуальный косайнор анализ. Частота биоритмов – критерий раннего выявления патологии: Методические рекомендации. Ереван, 1980. 55 с.
- Бокс Дж., Дженкинс Г.* Анализ временных рядов. В 2-х т. М.: Мир, 1974. 405 с. 197 с.
- Валеев Р.Т.* Метод взвешенного скользящего среднего и математическая модель «японских свечек» в условиях фондового рынка и их применение для его анализа: Дис. ... канд. техн. наук. Томск: Том. гос. ун-т, 2001. 113 с.
- Гармский геофизический полигон / Отв. ред. А.Я. Сидорин. М.: ИФЗ АН СССР, 1990. 240 с.
- Грачев А.В.* К восстановлению пропусков в экспериментальных данных // Вестн. ННГУ им. Н.И. Лобачевского. Сер. Радиофизика. 2004. Вып. 2. С. 15–23.

- Губанов В.А. Анализ воздействия выбросов на результат сезонной корректировки временных рядов // Научные труды: Институт народнохозяйственного прогнозирования РАН. 2004. № 2.
- Децеровская Е.В., Сидорин А.Я. Некоторые результаты изучения суточной периодичности землетрясений Гармского полигона // Сейсмические приборы. 2004. Вып. 40. С. 57–70.
- Децеровская Е.В., Сидорин А.Я. Внутрисезонные колебания сейсмичности Гармского полигона и их связь с атмосферными процессами // Докл. РАН. 2005. Т. 401, № 1. С. 80–83.
- Децеровская Е.В., Сидорин А.Я. Пространственно-временные особенности суточной периодичности слабых землетрясений Гармского полигона // Вопросы инженерной сейсмологии. 2015. Т. 42, № 4. С. 77–84.
- Децеровский А.В., Лукк А.А. Выделение регулярных составляющих во временных вариациях геофизических параметров методом разложения на негармонические компоненты // Вулканология и сейсмология. 2002. № 5. С. 65–78.
- Децеровский А.В., Сидорин А.Я. Алгоритмы и результаты формализованного поиска предвестников землетрясений во временных рядах геоэлектрического мониторинга // Федеральная система сейсмологических наблюдений: Инф.-аналит. бюл. 1996. Т. 3, № 3. С. 11–27.
- Децеровский А.В., Сидорин А.Я. Некоторые вопросы методики оценки среднесезонных функций для геофизических данных. М.: ОИФЗ РАН, 1999. 40 с.
- Децеровский А.В., Сидорин А.Я. База данных биологического мониторинга на Гармском полигоне // Геофизические процессы и биосфера. 2002. Т. 1, № 2. С. 3–15.
- Децеровский А.В., Сидорин А.Я. Проблема фликкер-шума при изучении причинно-следственных связей между природными процессами // Докл. РАН. 2003а. Т. 392, № 3. С. 392–396.
- Децеровский А.В., Сидорин А.Я. Параметризация временных рядов активности животных для геофизических исследований // Моделирование геофизических процессов. М.: ОИФЗ РАН, 2003б. С. 137–155.
- Децеровский А.В., Сидорин А.Я. Исследование значимости корреляции электрической активности рыб и электротеллурического поля // Биофизика. 2004. Т. 49, вып. 4. С. 715–722.
- Децеровский А.В., Сидорин А.Я. Поиск корреляции сейсмичности Гармского полигона с атмосферным давлением и скоростью ветра // Геофизические исследования. 2008. Т. 9, № 1. С. 3–15.
- Децеровский А.В., Сидорин А.Я. Периодограммы наложенных эпох при поиске скрытых ритмов в экспериментальных рядах // Сейсмические приборы. 2011а. Т. 47, № 2. С. 21–43.
- Децеровский А.В., Сидорин А.Я. Сравнение периодограмм наложенных эпох и спектров Фурье экспериментальных рядов // Сейсмические приборы. 2011б. Т. 47, № 3. С. 44–70.
- Децеровский А.В., Сидорин А.Я. Поиск влияния гравитационных приливов на региональную сейсмичность Греции разными методами: 1. Спектральный и периодограммный анализ // Сейсмические приборы. 2012. Т. 48, № 4. С. 5–26.
- Децеровский А.В., Сидорин А.Я. Поиск влияния гравитационных приливов на региональную сейсмичность Греции разными методами: 2. Корреляционный анализ // Сейсмические приборы. 2013а. Т. 49, № 1. С. 35–59.
- Децеровский А.В., Сидорин А.Я. Поиск влияния гравитационных приливов на региональную сейсмичность Греции разными методами: 3. Корреляция с солнечным и лунным компонентами прилива // Сейсмические приборы. 2013б. Т. 49, № 3. С. 41–53.
- Децеровский А.В., Сидорин А.Я. Суточная периодичность землетрясений Южной Калифорнии // Сейсмические приборы. 2014а. Т. 50, № 1. С. 27–50.
- Децеровский А.В., Сидорин А.Я. Оценка корреляции потока землетрясений Аляски с лунно-солнечным приливом // Наука и технологические разработки. 2014б. Т. 93, № 1. С. 29–44.
- Децеровский А.В., Сидорин А.Я. Технические проблемы и ошибки при работе с каталогами землетрясений // Наука и технологические разработки. 2014в. Т. 93, № 4. С. 32–41.

- Децеровский А.В., Сидорин А.Я.* Повышение робастности и устойчивости оценок параметров годографов Рэлея–Шустера с помощью различных способов нормировки векторов // Сейсмические приборы. 2015а. Т. 51, № 2. С. 56–80.
- Децеровский А.В., Сидорин А.Я.* Тестирование метода годографов Рэлея–Шустера на модельных временных рядах и потоках землетрясений // Сейсмические приборы. 2015б. Т. 51, № 3. С. 59–84.
- Децеровский А.В., Сидорин А.Я.* Анализ суточной периодичности в Душанбино-Вахшском каталоге землетрясений методом годографов Рэлея–Шустера // Вопросы инженерной сейсмологии. 2015в. Т. 42, № 3. С. 71–92.
- Децеровский А.В., Журавлев В.И., Сидорин А.Я.* Некоторые алгоритмы фильтрации для геофизических временных рядов // Физика Земли. 1996. № 2. С. 56–67.
- Децеровский А.В., Лукк А.А., Сидорин А.Я.* Признаки фликкер-шумовой структуры во временных реализациях геофизических полей // Физика Земли. 1997. № 7. С. 3–19.
- Децеровский А.В., Сидорин А.Я., Харин Е.П.* Исследование влияния гелиогеофизических факторов на активность животных в лабораторных условиях // Докл. РАН. 2005. Т. 401, № 6. С. 837–841.
- Децеровский А.В., Сидорин А.Я., Харин Е.П.* Геомагнитные возмущения и активность животных в лабораторных условиях // Биофизика. 2009. Т. 54, вып. 3. С. 554–562.
- Децеровский А.В., Журавлев В.И., Никольский А.Н., Сидорин А.Я.* Технологии анализа геофизических временных рядов. Ч. 1. Требования к программе обработки // Сейсмические приборы. 2016а. Т. 52, № 1. С. 61–82
- Децеровский А.В., Журавлев В.И., Никольский А.Н., Сидорин А.Я.* Технологии анализа геофизических временных рядов. Ч. 2. WinABD – пакет программ для сопровождения и анализа данных геофизического мониторинга // Сейсмические приборы. 2016б. Т. 52, № 3. С. 51–80.
- Децеровский А.В., Мирзоев К.М., Лукк А.А.* Критерии группирования землетрясений с учетом пространственной неоднородности сейсмичности // Физика Земли. 2016в. № 1. С. 79–97.
- Емельянов И.П.* Формы колебаний в биоритмологии. Новосибирск: Наука, 1976. 128 с.
- Журавлев В.И.* Моделирование сейсмического режима уравнением авторегрессии // Экспериментальная сейсмология. М.: Наука, 1983. С. 99–108.
- Журавлев В.И., Лукк А.А.* Полуденная активизация сейсмичности в Турции и ряде других регионов мира // Геофизические исследования. 2011. Т. 12, № 4. С. 31–57.
- Журавлев В.И., Сидорин А.Я.* Спектральные исследования суточной периодичности землетрясений Гармского полигона // Геофизические исследования. 2005. Вып. 1. С. 48–57.
- Злоба Е., Яцкив И.* Статистические методы восстановления пропущенных данных // Computer Modelling & New Technologies. 2002. V. 6, N 1. С. 51–61.
- Инструкция к программе «Cosinor Ellipse 2006». Омск: ООО Научно-методический центр «Аналитик», 2015. 13 с.
- Карлов И.А., Проворова О.Г.* Новый подход к исследованию устойчивости алюминиевого электролизера // Вестн. Краснояр. гос. ун-та. Физико-математические науки. 2002. № 1. С. 116–120.
- Карп В.П., Катинас Г.С.* Математические методы исследования биоритмов // Хронобиология и хрономедицина / Под ред. Ф.И. Комарова. М.: Медицина, 1989. С. 29–45.
- Концевая Н.В.* Оптимизация процедур сглаживания показателей финансовых рынков // Аудит и финансовый анализ. 2011. № 1. С. 122–127.
- Концевая Н.В.* Анализ методов заполнения пропусков во временных рядах показателей финансовых рынков // Вестн. Воронеж. гос. техн. ун-та. 2012. Т. 8, № 8. С. 18–20.
- Концевая Н.В.* Скользящий  $\beta$ -коэффициент как инструмент оптимизации торговых стратегий на примере валютного рынка // Вестн. Финансового ун-та. 2013. № 1. С. 73–81.

- Корягина Ю.В., Нопин С.В. Cosinor Ellipse 2006. № 2006611345 // Программы для ЭВМ... (офиц. бюл.). 2006. № 3 (56). С. 42.
- Кучеров И.С., Ткачук В.Г., Волков А.В. Длительные биологические ритмы в динамике мышечной работоспособности человека // Кибернетика и вычислительная техника. М.: Наука, 1970. Вып. 7. С. 71–77.
- Лагутин М.Б. Наглядная математическая статистика. 2-е изд., испр. М.: БИНОМ, 2009. 472 с.
- Литтл Р.Дж.А., Рубин Д.Б. Статистический анализ данных с пропусками. М.: Финансы и статистика, 1990. 336 с.
- Любушин А.А. Анализ данных систем геофизического и экологического мониторинга. М.: Наука, 2007. 228 с.
- Макс Ж. Методы и техника обработки сигналов при физических измерениях. В 2-х т. М.: Мир, 1983. Т. 1. 312 с.
- Маркелов О.А. Система информационной поддержки принятия решений врача при лечении вегетативных расстройств: Дис. ... канд. техн. наук. СПб. гос. электротехн. ун-т «ЛЭТИ» им. В.И. Ульянова (Ленина), 2007.
- Маркин А.В., Щербачев М.В. Метод автоматического восстановления значений в потоках данных на основе взвешенной модели // Прикаспийский журнал: управление и высокие технологии. 2013. № 3. С. 49–54.
- Моисеев С.Н. Заполнение пропусков в случайно-цензурированных временных рядах // Автометрия. 1998. № 1. С. 61–66.
- Оранский И.Е., Царфис П.Г. Биоритмология и хронотерапия. М.: Высш. шк., 1989. 159 с.
- Разумихин Д.В. Использование нейронных сетей на уровне семантики в системе распознавания речи // IV Всерос. конф. «Нейрокомпьютеры и их применение», г. Москва, 6–18 февраля 2000 г. М.: ИПУ РАН, 2000. С. 208–210.
- Россиев А.А. Итерационное моделирование неполных данных с помощью многообразий малой размерности. Красноярск: КГТУ, 2000. 84 с.
- Рыженкова К.В. Методы восстановления пропуска данных при проведении статистических исследований // Интеллект. Инновации. Инвестиции. 2011. № 3. С. 127–133.
- Сидорин А.Я. Суточная периодичность сильных землетрясений Гармского полигона // Сейсмические приборы. 2008. Т. 44, № 3. С. 70–76.
- Сидорин А.Я. О применении метода Рэлея–Шустера в исследованиях периодичности землетрясений // Сейсмические приборы. 2009. Т. 45, № 3. С. 29–40.
- Сидорин А.Я. Различия внутрисуточных фазовых диаграмм потоков землетрясений разной энергии // Сейсмические приборы. 2013. Т. 49, № 2. С. 71–84.
- Сидорин А.Я. Техногенная суточная периодичность сейсмических событий в районе Нурекского водохранилища // Наука и технологические разработки. 2015. Т. 94, № 2. С. 28–44.
- Скрипник В.М., Назин А.Е., Приходько Ю.Г., Благовещенский Ю.Н. Анализ надежности технических систем по цензурированным выборкам. М.: Радио и связь, 1988. 183 с.
- Снитюк В.Е. Эволюционный метод восстановления пропусков в данных // Тр. VI Междунар. конф. «Интеллектуальный анализ информации». Киев, 2006. С. 262–271.
- Теребиж В.Ю. Анализ временных рядов в астрофизике. М.: Наука, 1992. 392 с.
- Урбах В.Ю. Математическая статистика для биологов и медиков. М.: Изд-во АН СССР, 1963. 322 с.
- Хампель Ф., Рончетти Э., Рауссеу П., Штаэль В. Робастность в статистике: Подход на основе функций влияния. М.: Мир, 1989. 512 с.
- Хардле В. Прикладная непараметрическая регрессия. М.: Мир, 1993. 349 с.
- Хемминг Р.В. Цифровые фильтры. 2-е изд. М.: Недра, 1987. 221 с.
- Хемминг Р.В. Численные методы для научных работников и инженеров. М.: Наука, 1972. 400 с.

- Щербина С.В.* Экспериментальное исследование динамического хаоса в сейсмогенной среде // Геофиз. журн. 2004. Т. 26, № 3. С. 125–131.
- Baghi Q., Gilles M., Berge J., Christophe B., Touboul P., Rodrigues M.* Regression analysis with missing data and unknown colored noise: Application to the MICROSCOPE space mission // Physical Rev. D. 2015. V. 91, N 6.
- Cornelissen G.* Cosinor-based rhythmometry // Theoretical Biology and Medical Modelling. 2014. V. 11:16.
- Dergachev V.A., Makarenko N.G., Karimova L.N., Danilkina E.B.* Nonlinear methods of analysis of data with gaps // Geochronometria. 2001. V. 20. P. 45–50.
- Filling data gaps in a periodic timeseries in MATLAB. 2014. URL: [http://www.mathworks.com/matlabcentral/newsreader/view\\_thread/337892](http://www.mathworks.com/matlabcentral/newsreader/view_thread/337892)
- Filling gaps in time series with Nan. 2013. URL: <http://www.mathworks.com/matlabcentral/answers/76164-filling-gaps-in-time-series-with-nan>
- Giles D.E.A.* The underground economy: Minimizing the size of government // How to spend the fiscal dividend: The optimal size of government / Ed. by H. Grubel. Vancouver: Fraser Institute, 1998. P. 93–110.
- Gorban A., Rossiev A., Makarenko N., Kuandykov Y., Dergachev V.* Recovering data gaps through neural network methods // Intern. J. of Geomagnetism and Aeronomy. 2002. V. 3, N 2. P. 191–197.
- Guiles M.D.* Effect of diurnal data gaps on regression and FFT analysis. SOEST, University of Hawaii at Manoa, 2007. 9 p.
- Halberg F., Cornélissen G., Tarquini B., Grafe A., Syutkina E.V., Otsuka K., Watanabe Y., Siegelova J., Sanchez de la Pena S., Carandente F., Schwartzkopff O.* Pineals, cancers and feedsideswards in the biosphere and the cosmos (BIOCOS) // Cancer Biotherapy and Radiopharmaceuticals. 1997. V. 12. P. 421–422.
- Halberg F.* Chronobiology // Ann. Rev. Physiol. 1969. V. 31. P. 675–725.
- Halberg F.* Chronobiology: Methodological problems // Acta Med. Rom. 1980. V. 18. P. 399–440.
- Kanasewich E.R.* Time sequence analysis in geophysics. Edmonton [Alta.]: Univ. of Alberta Press, 1981. 480 p.
- Katinas G.S.* Logistic informative complex of time series analysis // Здоровье и образование в XXI. Сер. Медицина. 2012. Т. 14, № 2. С. 128–133.
- Klingenberg B.* Regression models for binary time series with gaps // Computational Statistics and Data Analysis. 2008. V. 52. P. 4076–4090.
- Lomb N.R.* Least-squares frequency analysis of unequally spaced data // Astrophys. and Space Sci. 1976. V. 39. P. 447–462.
- Love J.J.* Missing data and the accuracy of magnetic-observatory hour means // Ann. Geophys. 2009. V. 27. P. 3601–3610.
- Moody G.B.* Spectral analysis of heart rate without resampling // Proc. of «Computers in cardiology» conf., London, 5–8 Sept. 1993. IEEE Computer Society, 1993. P. 71–718.
- Pashova L., Koprinkova-Hristova P., Popova S.* Gap filling of daily sea levels by artificial neural networks // TransNav. 2013. V. 7, N 2. P. 225–232.
- Press W.H., Rybicki G.B.* Fast algorithm for spectral analysis of unevenly sampled data // Astrophysical J. 1989. V. 338. P. 277–280.
- Press W.H., Teukolsky S.A., Vetterling W.T., Flannery B.P.* Numerical recipes in C: The art of scientific computing. 2-nd ed. Cambridge–New York–Port Chester–Melbourne–Sydney: Cambridge Univ. Press, 2002. 925 p.
- Refinetti R., Cornelissen G., Halberg F.* Procedures for numerical analysis of circadian rhythms // Biological Rhythm Res. 2007. V. 38, N 4. P. 275–325.

- Ryan K.F., Giles D.E.A.* Testing for unit roots in economic time-series with missing observations // *Advances in Econometrics* / Ed. by T.B. Fomby & R. Carter Hill. Emerald Group Publ. Ltd., 1998. V. 13 P. 203–242.
- Sandip V. George, Ambika G.* Effect of data gaps on correlation dimension computed from light curves of variable stars // *Astrophys. Space Sci.* 2015. V. 360:5.
- SAS. Knowledge base. Usage Note 22921. URL: <http://support.sas.com/kb/22/921.html>
- Scargle J.D., Norris J.P., Jackson B., Chiang J.* Studies in astronomical time series analysis. VI. Bayesian block representations // *The Astrophys. J.* V. 764, N 2:167.
- Schluter T.* Knowledge discovery from time series: Inaugural-dissertation zur Erlangung des doktorgrades der mathematisch-naturwissenschaftlichen fakultat der Heinrich-Heine-Universität Dusseldorf. 2012. 169 p.
- Seelam M.R.* How do I fill gap in time series data? 2015. URL: [https://www.researchgate.net/post/How\\_do\\_I\\_fill\\_gap\\_in\\_time\\_series\\_data](https://www.researchgate.net/post/How_do_I_fill_gap_in_time_series_data)
- Smith J.* How to fill data gaps in a time series with NaN. 2014. URL: <http://www.mathworks.com/matlabcentral/answers/160053-how-to-fill-data-gaps-in-a-time-series-with-nan>
- Time series analysis and data gaps. URL: <http://epchan.blogspot.ru/2015/07/time-series-analysis-and-data-gaps.html>
- Torres-Reyna O.* Data analysis notes: Links and general guidelines: Online Stata Tutorial. 2014. URL: <https://www.princeton.edu/~otorres/Stata/statnotes>
- Working with missing data. 2015. URL: [http://pandas.pydata.org/pandas-docs/stable/missing\\_data.html](http://pandas.pydata.org/pandas-docs/stable/missing_data.html).
- Yuan G.-C., Lozier M.S., Pratt L.J., Jones C.K.R.T., Helfrich K.R.* Estimating the predictability of an oceanic time series using linear and nonlinear methods // *J. Geophys. Res.* 2004. V. 109, C08002. doi: 10.1029/2003JC002148.