

УДК 004.032.26

ПРОГНОЗИРОВАНИЕ СОЛНЕЧНОЙ АКТИВНОСТИ И ИОНОСФЕРНЫХ ПАРАМЕТРОВ С ИСПОЛЬЗОВАНИЕМ НЕЙРОННЫХ СЕТЕЙ

Ю.С. Масленникова, В.В. Бочкарев

FORECASTING OF SOLAR ACTIVITY AND IONOSPHERE PARAMETERS USING ARTIFICIAL NEURAL NETWORK

Yu.S. Maslennikova, V.V. Bochkarev

Предложен метод долгосрочного прогнозирования временных рядов, основанный на совместном использовании сингулярного спектрального анализа, нейросетевого моделирования и порядковых статистик. Показано, что точность данного метода для 22-го и 23-го циклов солнечной активности более чем на 50 % выше, чем у метода линейного прогнозирования, а дополнительный учет статистически связанных рядов при обучении нейронной сети способствует улучшению качества прогноза. Выполнено прогнозирование временного ряда полного электронного содержания над конкретной областью с учетом различных дополнительных данных.

In this report, a method for long-term prediction of time series, based on singular spectral analysis and neurofuzzy modelling is proposed. The method is used for sunspot number time series of solar cycles 22 and 23, and the accuracy is more than 50 % higher than that of the linear predictor. The complementary, statistically-related time series are used for training neuro networks to improve quality of the forecast. Forecast of the Total Electron Contents (TEC) over a concrete area has been made, taking additional data into account.

Введение

В настоящее время решение задачи прогнозирования солнечной активности и состояния ионосферы имеет большое значение для обеспечения устойчивой радиосвязи и безопасной навигации. Такие прогнозы позволили бы успешно предвидеть нарушения радиосвязи, подготовить альтернативные радиотрассы, перейти на более выгодные длины волн.

Несмотря на то, что солнечная активность имеет определенную периодичность, ее прогнозирование весьма затруднительно. И хотя различные методы статистического анализа и экспериментальные факты показывают наличие долгосрочных и крупномасштабных изменений, разумным кажется долгосрочное прогнозирование не более чем на 10 лет.

Прогнозирование солнечной активности

В последнее время активно развиваются методы прогнозирования, основанные на использовании искусственных нейронных сетей [1]. Но, несмотря на многие преимущества данных систем, имеется ряд сложностей при использовании их на выборках малого объема. Для решения подобных проблем в [2] предложен метод с использованием сингулярного спектрального анализа [3]. Локально-линейная нейросетевая модель оптимизируется для каждой главной компоненты разложения, затем предсказанные на несколько шагов вперед значения используются для восстановления временного ряда. В настоящей работе предлагается несколько иной вариант этой модели, учитывающий также данные меньшего временного масштаба. Предлагаемый алгоритм включает в себя два этапа:

1. Последовательное дообучение нейронной сети по усредненным данным;
2. Инициализация нейронной сети с обучением по исходным данным.

Основная идея метода применительно к рядам солнечной активности состоит в том, что прогнозирование среднегодовых значений числа солнечных пятен производится по среднемесячным данным за период с 1949 по 2008 г. При этом обучение сети

сначала осуществляется на основе ряда, усредненного по методу скользящего среднего (с длиной окна 12 мес.), после чего данная сеть обучается уже с использованием исходных среднемесячных значений.

Поясним подробнее каждый этап предлагаемого метода прогнозирования. Для начала производилось сингулярное разложение усредненного временного ряда среднемесячных значений, при этом число компонент разложения равнялось 400, что соответствовало наибольшему сдвигу на 33 года, чтобы охватить, по крайней мере, три солнечных цикла. Идея алгоритма последовательного дообучения заключается в том, что нейронная сеть постепенно обучается на рядах, восстанавливаемых по различному числу компонент (рис. 1), что позволяет более эффективно учитывать структуру исследуемых данных. При этом начальное состояние нейронной сети при прогнозировании ряда, восстановленного по первым двум компонентам, устанавливалось, исходя из предварительно найденных коэффициентов линейного прогнозирования. На рис. 1 также представлена зависимость ошибки прогнозирования от числа учитываемых при восстановлении ряда компонент в сравнении с линейным методом.

Второй этап прогнозирования включает в себя модификацию полученной нейронной сети, способной прогнозировать средний уровень числа солнечных пятен по усредненным данным. Число входов данной сети увеличивается в 12 раз, для того чтобы получить возможность обучать ее по исходным среднемесячным данным. Соответствующие коэффициенты входного слоя подбираются так, чтобы до обучения сеть была практически идентична ранее обученной. Также осуществляется сортировка по возрастанию входных среднемесячных значений, соответствующих одному и тому же году. Это производится для того, чтобы получить нелинейный аналог метода оптимальных порядковых статистик [4]. Использование подобной сортировки данных способствовало улучшению качества прогнозирования более чем на 30 % (относительная ошибка прогнозирования без сортировки данных составила

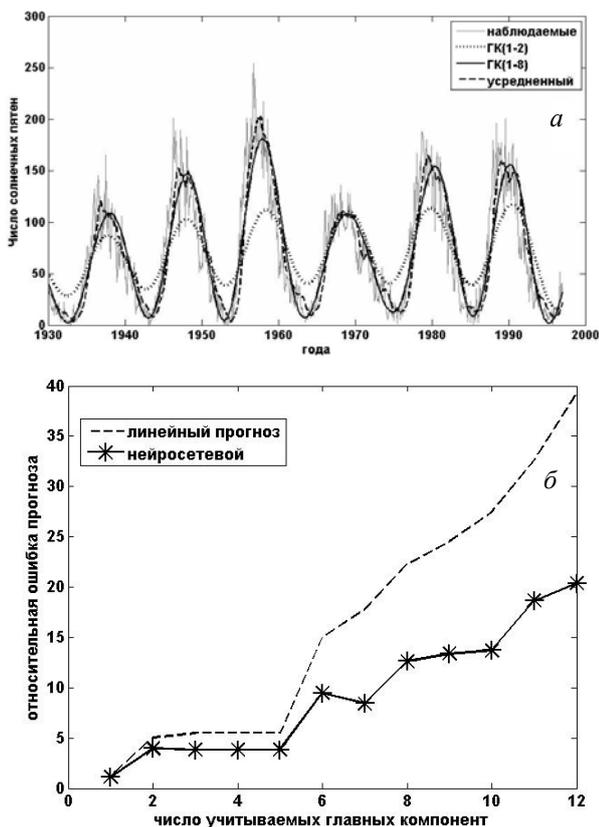


Рис. 1. Временные ряды среднемесячных значений числа солнечных пятен: наблюдаемый, восстановленный по различному числу компонент разложения (ГК) и усредненный по 12 месяцам (а); зависимость ошибки прогнозирования от числа используемых при восстановлении ряда компонент разложения (б).

10.4 %, с сортировкой – 8.13 %).

Изложенный выше алгоритм был использован для прогнозирования среднего уровня солнечных пятен на пять и девять лет вперед для 22- и 23-го циклов солнечной активности (рис. 2). Конечная нейронная сеть содержала 120 нейронов во входном слое, 10 в скрытом и 1 в выходном, обучение производилось по методу Левенберга–Марквардта в течение 1100 эпох [1]. В табл. 1 представлены среднеквадратичное отклонение (СКО) и максимум абсолютной ошибки прогнозирования данным методом и методом, предложенным в статье [2] (отмечено *), для различных интервалов времени.

Таким образом, удалось показать эффективность предлагаемого метода начальной инициализации и последовательного дообучения нейронной сети, который позволил спрогнозировать последнюю тенденцию затянувшегося минимума. Дополнительно к улучшению качества прогноза также способствовала сортировка данных.

Так как временной ряд для индекса $F_{10.7}$ содержит значения не более чем для шести циклов солнечной активности (1947–2008 г.), для обучения прогностической сети были задействованы также данные по числу солнечных пятен за тот же период времени. Среднегодовые значения для этих рядов были найдены путем усреднения данных по среднемесячным значениям. Начальное состояние нейронной сети устанавливалось, исходя из найденных коэффи-

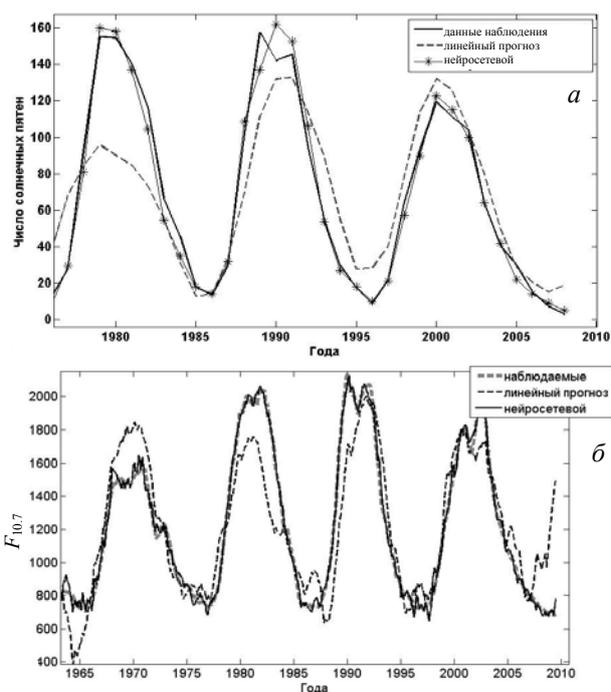


Рис. 2. Прогнозирование среднего уровня числа солнечных пятен на пять лет вперед для 22-го и 23-го циклов солнечной активности (а); прогнозирование среднего уровня индекса $F_{10.7}$ на пять лет вперед (б).

циентов линейного прогнозирования. Относительная ошибка прогнозирования индекса $F_{10.7}$ на 5 лет вперед при совместном использовании данных составила 3.3929 % (рис. 2), а без учета статистически связанного ряда – 6.1897 %, что свидетельствует об эффективности данного алгоритма прогнозирования.

Прогнозирование полного электронного содержания (ПЭС)

Основной характеристикой ионизированной атмосферы является величина полного электронного содержания (ПЭС) в единичном атмосферном столбе [5]. В работе осуществляется прогнозирование значений ПЭС на 2 ч вперед над о города Казань. В качестве исходных данных были использованы карты вертикального ПЭС ионосферы, представленные лабораторией JPL (Jet Propulsion Laboratory) за 2008 г. [6]. Наиболее очевидным вариантом прогнозирования подобных данных является использование в качестве обучающей выборки для нейронной сети мгновенных значений ПЭС во всех точках на определенной широте (табл. 2, обучающая выборка 1). Однако данный способ не дал хороших результатов прогнозирования. Возможным путем улучшения качества прогнозирования может являться учет при обучении статистически связанных рядов. Для этого необходимо было выяснить степень корреляционной связи между ними. Чтобы избежать влияния доминирующих периодичностей, из исследуемых временных рядов вычитались данные, соответствующие первым двум компонентам сингулярного разложения. Таким образом, эквивалентный коэффициент корреляции (полученный из меры взаимной информации как $R_0 = \sqrt{1 - 2^{-2.1}}$) [7] между вариациями среднедневных

Таблица 1

Номер солнечного цикла	Интервал прогноза	СКО	Максимум абсолютной ошибки	СКО*	Максимум абсолютной ошибки *
22	5 (1991–1996)	5.5022	13.9756	6.6250	17.5
22	9 (1987–1996)	6.0542	14.5683	7.0424	19.8
23	5 (2003–2008)	5.4404	7.7859	7.185	22.3
23	9 (1999–2008)	6.2292	8.0251	7.8059	21.5

Таблица 2

Метод прогнозирования	Обучающая выборка 1		Обучающая выборка 2		Обучающая выборка 2 с дообучением нейросети	
	СКО	%	СКО	%	СКО	%
тривиальный	18.3	24.3	17.6	24.1	16.2	21.9
линейный	7.45	10.1	6.94	9.49	5.48	7.51
нейросетевой	6.21	8.26	5.43	7.46	3.59	4.92

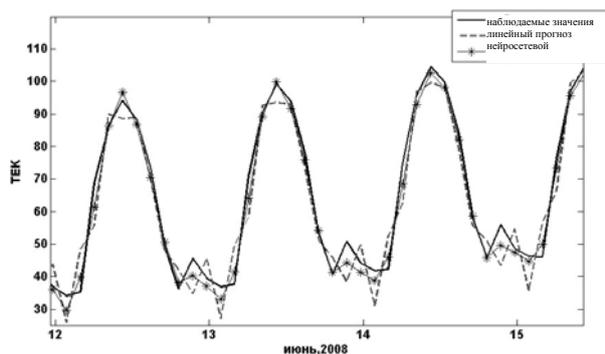


Рис. 3. Прогнозирование ПЭС на 2 ч вперед над окрестностями города Казань.

значений ПЭС и вариациями индекса $F_{10.7}$ составил 0.58, а между ПЭС и числом солнечных пятен – 0.05 (что объясняется низким уровнем солнечной активности в 2008 г.). Итоговая обучающая выборка содержала значения ПЭС над исследуемой областью за предыдущие 12 ч, значения в предыдущие моменты времени над областями, отстающими по долготе на 2, 4 и 6 ч соответственно, восстановленные по 13 компонентам разложения, и ряд среднесуточных значений индекса $F_{10.7}$ (табл. 2, обучающая выборка 2). Также был использован описанный выше метод последовательного дообучения нейронной сети. Итоговые результаты прогнозирования (рис. 3) представлены в табл. 2, в которой также представлены резуль-

таты линейного и тривиального (по предыдущим отсчетам) методов прогнозирования.

Таким образом, учет дополнительной информации позволил повысить качество прогноза более чем на 10 %, а алгоритм последовательного дообучения нейронной сети способствовал улучшению еще на 30 %.

СПИСОК ЛИТЕРАТУРЫ

1. Круглов В.В., Дли М.И., Голунов Р.Ю. Нечеткая логика и искусственные нейронные сети. М.: Изд-во физико-математической литературы, 2001. С. 48–72.
2. Gholipour A., Lucasa C., Arabia B.N., Shafiee M. Solar activity forecast: Spectral analysis and neurofuzzy prediction // J. Atm. Sol.-Terr. Phys. V. 67. 2005. P. 595–603.
3. Голяндина Н.Э. Метод «Гусеница»-SSA: анализ временных рядов: Учеб. пособие. СПб., 2004. 76 с.
4. Боярский Э.А. Порядковые статистики. М.: Статистика, 1972. С. 7–26.
5. Куницын В.Е., Терещенко Е.Д., Андреева Е.С. Радиотомография ионосферы. М.: Физматлит, 2007. С. 336.
6. <http://www.jpl.nasa.gov/>.
7. Масленникова Ю.С., Бочкарев В.В. Анализ и прогнозирование рядов ионосферных параметров // Труды Волжской региональной молодежной конференции «Радиофизические исследования природных сред и информационные системы». Казань: Изд-во Казан. гос. ун-та, 2008. С. 126–129.

Казанский государственный университет, Казань