

ПРОГНОЗИРОВАНИЕ СУТОЧНОГО ПОТРЕБЛЕНИЯ ЭЛЕКТРОЭНЕРГИИ В МОСКОВСКОМ РЕГИОНЕ С ИСПОЛЬЗОВАНИЕМ ИСКУССТВЕННЫХ НЕЙРОННЫХ СЕТЕЙ

В. В. Иванов^{а, б, 1}, А. В. Крянев^{а, б, 2}, Е. С. Осетров^{а, в, 3}

^а Объединенный институт ядерных исследований, Дубна

^б Национальный исследовательский ядерный университет «МИФИ», Москва

^в ФКУ «Ространсmodernизация», Москва

Ранее нами была продемонстрирована принципиальная возможность краткосрочного прогноза суточных объемов пассажирских перевозок в Московском метрополитене с помощью искусственных нейронных сетей. При обучении и прогнозировании на вход нейронной сети подавалась выборка, составленная из факторов, от которых зависит суточный трафик пассажироперевозок в метро. Один из таких факторов — это суточное энергопотребление в Московском регионе. Поэтому для того, чтобы прогнозировать пассажироперевозки в метро, нужно предварительно решить проблему прогноза суточного энергопотребления в Московском регионе.

Earlier we demonstrated a possibility for a short-term forecasting of daily volumes of passenger's traffic in the Moscow metro with the help of artificial neural networks. For training and predicting, a sample of the factors that influence the daily passenger traffic in the subway passes to the input of neural network. One of these factors is a daily power consumption in the Moscow region. Therefore, to predict the volume of the passenger traffic in the subway, we must first solve the problem of forecasting the daily energy consumption in the Moscow region.

PACS: 02.50.-r; 05.45.Tp; 07.05.Mh; 89.60.-k

ВВЕДЕНИЕ

В работе [1] нами была продемонстрирована принципиальная возможность краткосрочного прогноза суточных объемов пассажирских перевозок в Московском метрополитене с помощью искусственных нейронных сетей (ИНС). Однако из-за того, что исходный временной ряд был сильно зашумлен, нам не удалось добиться приемлемых точности и горизонта прогнозирования. Поэтому в следующей работе [2] для исключения шума из анализируемого ряда нами была применена вейвлет-фильтрация исходных данных. Это позволило более чем в четыре раза повысить точность прогноза и, как следствие, увеличить горизонт прогнозирования.

¹E-mail: ivanov@jinr.ru

²E-mail: avkryanev@mephi.ru

³E-mail: esosetrov@gmail.com

В работах [1, 2] для обучения и прогнозирования с помощью ИНС на вход нейронной сети подавалась выборка, составленная из ключевых факторов, от которых зависит суточный трафик пассажироперевозок в метро. Один из таких факторов — суточное энергопотребление в Московском регионе.

В связи с вышеизложенным для прогнозирования суточных объемов пассажирских перевозок в метрополитене с помощью ИНС предварительно нужно выполнить прогноз суточного энергопотребления в Московском регионе. Решению указанной задачи посвящена настоящая работа.

1. ДИНАМИКА ЭНЕРГОПОТРЕБЛЕНИЯ В МОСКОВСКОМ РЕГИОНЕ

На рис. 1 представлен временной ряд, отражающий динамику суточного энергопотребления в Московском регионе за последние 14 лет (всего 5114 наблюдений).

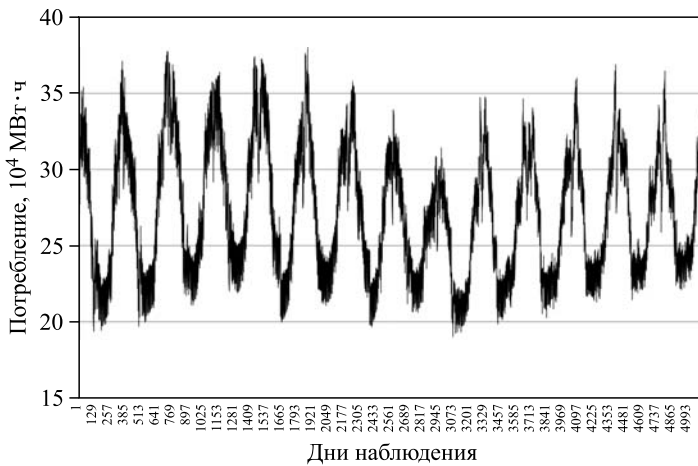


Рис. 1. Динамика суточного энергопотребления в Московском регионе за последние 14 лет

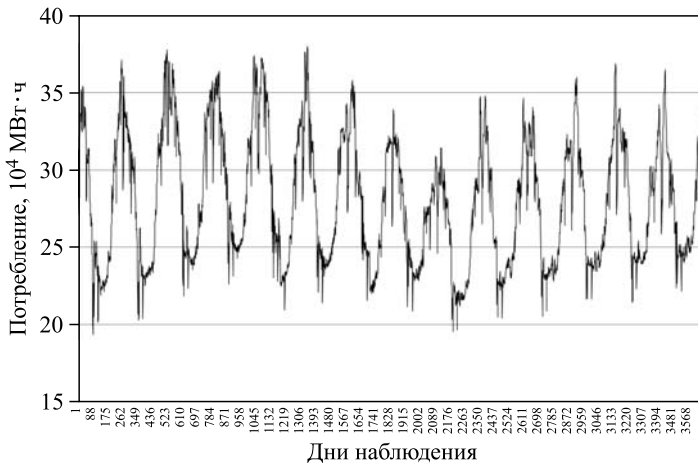


Рис. 2. Динамика суточного энергопотребления в Московском регионе в будние дни

С учетом того, что в работах [1,2] прогнозирование суточных объемов пассажирских перевозок в метрополитене проводилось только в будние дни, мы исключили выходные дни из временного ряда, приведенного на рис. 1.

На рис. 2 представлен временной ряд, из которого удалены только выходные дни (всего 3654 наблюдений). Для того, чтобы исключить из исходных данных праздничные дни за последние 14 лет, нужно проделать довольно большую рутинную работу. По нашему мнению, праздничные дни, не удаленные из исходных данных, не должны хоть как-то заметно повлиять на окончательный результат. Поэтому мы не стали проводить указанную процедуру.

Заметим, что после исключения из исходных данных выходных дней рассматриваемый временной ряд стал выглядеть менее зашумленным (см. рис. 2).

Так же как и в работах [1,2], прогнозирование суточного энергопотребления будет проводиться с помощью заранее обученной ИНС (см. детали ниже).

2. ПРОГНОЗИРОВАНИЕ СУТОЧНОГО ЭНЕРГОПОТРЕБЛЕНИЯ С ПОМОЩЬЮ ИНС

ИНС широко и успешно применяются для решения многих задач в области науки, техники, экономики и бизнеса, связанных с анализом и обработкой данных наблюдения, таких как анализ изображений, распознавание образов (в частности, отделение сигнала от фона в задачах физики высоких энергий), прогнозирование фондового рынка, предсказание поведения финансовых рядов и т. д. [3,4].

Наиболее подходящей архитектурой ИНС для решения рассматриваемой нами задачи является многослойный перцептрон (МСП) (см., например, [4,5]). МСП содержит несколько слоев нейронов: слой входных нейронов, один или несколько скрытых слоев и слой выходных нейронов. Связи в ИНС такого типа организованы между соседними слоями, а передача информации, как правило, происходит в одном направлении — от входного слоя к выходному.

2.1. Архитектура и основные принципы работы многослойного перцептрона.

На рис. 3 приведена схема МСП с двумя скрытыми слоями и одним выходным нейроном, стрелками указаны направления передачи информации в сети. Здесь x_k обозначают нейроны входного слоя; h_j и h_i — нейроны двух скрытых слоев; y_0 — выходной нейрон; w_{jk} — весовые связи (веса) входных нейронов с первым скрытым слоем; w_{ij} — веса нейронов первого скрытого слоя с нейронами второго скрытого слоя; w_{0i} — веса нейронов второго скрытого слоя с выходным нейроном.

Сигналы $a_j = \sum_k w_{jk} x_k$ и $a_i = \sum_j w_{ij} h_j$ поступают на входы нейронов первого и второго скрытых слоев, а сигнал $a_0 = \sum_i w_{0i} h_i$ подается на выходной нейрон. Величины сигналов с выходов нейронов указанных слоев определяются из следующих соотношений:

$$h_j = g\left(\frac{a_j}{T}\right) + \theta_j, \quad h_i = g\left(\frac{a_i}{T}\right) + \theta_i, \quad y_0 = g\left(\frac{a_0}{T}\right) + \theta_0,$$

где $g(a, T)$ — переходная функция (T — «температура», которая определяет величину наклона переходной функции), а θ — величина порога соответствующего нейрона. Чаще

всего в МСП используется переходная функция сигмоидного вида $g(x) = (1/2)[1 + \tanh(x)]$.

Настройка сети (определение весовых и пороговых коэффициентов) на решаемую задачу проводится путем ее обучения, что обычно реализуется с помощью алгоритма обратного распространения ошибок [6].

В процессе обучения нейронной сети происходит ее настройка на анализируемые образцы, что достигается путем коррекции набора весов: w_{jk} , w_{ij} и w_{0i} (см. детали в [1]). Эта процедура выполняется путем минимизации функционала ошибок (называемого также функционалом энергии сети). Функционал ошибок E представляет собой сумму разностей между выходными сигналами и эталонными значениями:

$$E = \frac{1}{2} \sum_p (y_0^{(p)} - t_0^{(p)})^2, \quad (1)$$

где p — номер входного образца из того набора данных, которые используются для обучения сети; $y_0^{(p)}$ — значение с выходного нейрона в p -м событии, а $t_0^{(p)}$ — величина соответствующего эталона. Обучение МСП продолжается до тех пор, пока сигналы с выхода сети не приблизятся к эталонным значениям.

Следует заметить, что на практике веса обычно корректируются не для каждого отдельного события, а на небольшой выборке событий. Это позволяет ускорить тренировочный процесс. Обычно достаточно сделать несколько десятков/сотен итераций на всей тренировочной выборке для того, чтобы минимизировать функционал ошибок и получить искомый набор весов. После завершения тренировочного процесса веса и пороги фиксируются, а качество обучения сети оценивается путем прогонки через нее тестовой выборки событий.

2.2. Структура сети и формат входных данных. На этапе построения структуры МСП необходимо: 1) выбрать число слоев; 2) задать количество нейронов в каждом слое; 3) определить формат входных данных.

Анализируемые данные подаются на входные нейроны; число таких нейронов определяется размером выборки. В нашем случае анализируемая выборка включает пять переменных: 1) год; 2) месяц; 3) день недели; 4) «подсказку» для МСП, которая берется либо из отфильтрованного ряда анализируемых данных (на этапе обучения сети), либо из прогнозных значений, вычисленных с помощью пакета «Гусеница»-SSA [7] (на этапе тестирования сети); 5) значение, взятое из исходного ряда (на этапе обучения сети), либо предсказанное на текущий день обученной ИНС (на этапе тестирования сети).

В теории достаточно одного скрытого слоя МСП для аппроксимации любой непрерывной функции [8]. Однако на практике для аппроксимации таких сложных времен-

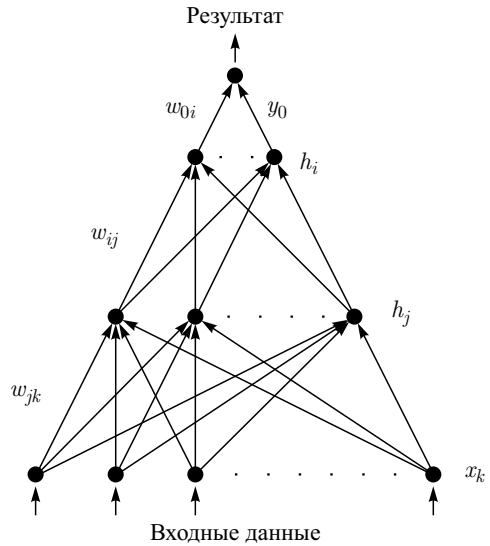


Рис. 3. Схема многослойного перцептрона с двумя скрытыми слоями и одним выходным нейроном. Стрелками указаны направления передачи информации в сети

ных рядов, как в нашем случае, лучший результат дает МСП с двумя скрытыми слоями (см., например, [9] и ссылки в ней). Числа нейронов в скрытых слоях подбирались исходя из результатов обучения сети на исследуемых данных. Качественные и количественные результаты анализа выборок, подаваемых на вход МСП, оценивались по величине сигнала с выходного нейрона.

Перед тем, как приступить к работе с МСП, переменные 1–5 (обозначим их через z) были приведены к диапазону $[-1; +1]$ с помощью следующего преобразования:

$$x_i = \frac{(z_i - A_i) \cdot 2}{B_i - A_i} - 1, \quad i = 0, 1, \dots, 4, \quad (2)$$

где z_i — исходное значение переменной z ; A_i и B_i — минимальное и максимальное значения переменной z_i , а x_i — значение нормированной переменной.

В настоящей работе нами использовался МСП из пакета TMVA 4.2.0 [10] в среде ROOT [11].

Структура сети была взята следующей: 5 нейронов на входе, первый скрытый слой содержал 5 нейронов, второй — также 5 нейронов, один выходной нейрон. Для обучения и тестирования сети использовалась выборка, содержащая 2048 наблюдений (из полного числа в 3654 наблюдения). В результате проведенного нами анализа для обучения был выбран метод BFGS (см., например, [10, 12]), обеспечивший наилучшие результаты как на этапе обучения, так и при тестировании МСП. Процедура обучения обычно состояла из 1000 эпох.

Прежде чем приступить к обучению и тестированию сети, обсудим процедуру формирования файла-«подсказки», играющего ключевую роль в проведении среднесрочного прогнозирования с помощью МСП.

Ранее, в работе [2], нами было показано, что можно существенно повысить точность и увеличить горизонт прогноза, если предварительно, до обучения ИНС, провести фильтрацию анализируемого временного ряда с целью исключения из него высокочастотной (шумовой) компоненты. В связи с этим при подготовке начального участка файла-«подсказки», предназначенного для обучения МСП, нами использовалась вейвлет-фильтрация.

Для подготовки прогнозной части файла-«подсказки», используемой на этапе проведения среднесрочного прогноза с помощью обученной нейронной сети, мы применили метод «Гусеница»-SSA, развитый в работах [7, 24, 25].

3. ВЕЙВЛЕТ-ФИЛЬТРАЦИЯ ВРЕМЕННОГО РЯДА

Временной ряд, отвечающий суточному энергопотреблению, может быть записан в виде

$$f(t_i) = x_i = x[(i - 1)\Delta t], \quad i = 1, 2, \dots, K, \quad (3)$$

где Δt — интервал выборки ($\Delta t = 1$ в нашем случае).

Мы можем представить рассматриваемый временной ряд в виде регулярного процесса и стохастической составляющей, играющей роль высокочастотного шума. Исключение из исходных измерений указанного шума может позволить повысить точность и увеличить перспективу прогноза.

3.1. Общая схема вейвлет-фильтрации. Вкратце изложим общую схему вейвлет-анализа. Детали можно найти, например, в работах [13, 14].

Дискретное вейвлет-преобразование (ДВП) функции $f(t) \in L_2(R)$, заданной в виде одномерного временного ряда (3), может быть представлено в виде следующего разложения:

$$f(t) = \sum_{j,k \in Z} d_{jk} \psi(2^j t - k). \quad (4)$$

Здесь набор базисных функций (вейвлетов) $\{\psi_{jk}(t) = \psi(2^j t - k), j, k \in Z\}$ получается из «родительской» вейвлет-функции $\psi(t) \in L_2(R)$ применением двоичного разложения 2^j и бинарной трансляции $k/2^j$.

Согласно мультиразложению вейвлет-анализа равенство (4) может быть переписано в более удобной форме:

$$f(t) = \sum_k s_k^J \phi(2^J t - k) + \sum_{j \geq J} \sum_{k \in Z} d_k^j \psi(2^j t - k), \quad (5)$$

где $\phi(t)$ — функция масштабирования, соответствующая выбранной вейвлет-функции $\psi(t)$ (см., например, [13]). В (5) первый член описывает гладкую (низкочастотную) составляющую ряда (5), ограниченную уровнем точности J , а второй член связан с высокочастотной составляющей анализируемого ряда.

Коэффициенты s_k^j и d_k^j обычно определяются с помощью пирамидальной схемы [14] быстрых вейвлет-преобразований (см., например, [15]) с использованием следующих соотношений:

$$s_k^{j+1} = \sum_m h_m s_{2k+m}^j, \quad d_k^{j+1} = \sum_m g_m s_{2k+m}^j, \quad (6)$$

где h_m и g_m — коэффициенты низко- и высокочастотных фильтров соответственно.

Нами использовались дискретные вейвлеты Добечи [13, 16], так как данные вейвлеты обеспечивают наилучшее описание как высокочастотных, так и низкочастотных составляющих временного ряда [15].

Вейвлет-фильтрация подразумевает отбрасывание или модификацию части коэффициентов разложения с абсолютными значениями, меньшими некоторого наперед заданного порогового значения λ . Существуют различные алгоритмы вейвлет-фильтрации, среди которых наиболее распространенным является жесткий пороговый алгоритм (см., например, [15]). В этом алгоритме все коэффициенты с абсолютными значениями, меньшими λ , отбрасываются, а точнее, приравниваются нулю.

В таких алгоритмах процедура фильтрации воздействует на все коэффициенты без учета соответствующего им разрешающего уровня J . Поэтому такая процедура может приводить к отбрасыванию не только коэффициентов $\{d_k^j\}$, соответствующих высокочастотной составляющей (5), но и коэффициентов $\{s_k^J\}$, которые отвечают за низкочастотную компоненту временного ряда.

Для того, чтобы исправить указанный недостаток, в работе [17] была предложена модификация алгоритма с жесткой пороговой схемой таким образом, чтобы фильтрация коэффициентов вейвлет-разложения проводилась с учетом разрешающего уровня J .

Процедура фильтрации в модифицированном алгоритме выполняется следующим образом. Пусть K — это количество элементов в анализируемом ряду, а M — число коэффициентов, которые должны быть отброшены. Предположим, что $M < K/2$. В этом случае отбрасываются M наименьших коэффициентов из числа $K/2$ коэффициентов, отвечающих высокочастотной составляющей анализируемого ряда (5). Если $K/2 < M < 3K/4$, то отбрасываются все $K/2$ высокочастотных коэффициентов, а также $M - K/2$ наименьших коэффициентов, соответствующих более низкому уровню аппроксимации J (полное число таких коэффициентов составляет $K/4$) и т. д.

По сравнению с пороговым алгоритмом вейвлет-фильтрации модифицированная схема обеспечивает более эффективное удаление высокочастотной компоненты из исходных реализаций анализируемого временного ряда.

После применения ДВП M отобранных коэффициентов приравниваются нулю. Затем с использованием обратного вейвлет-преобразования восстанавливается регулярная составляющая временного ряда. Разница между исходным рядом и отфильтрованным сигналом рассматривается далее как шумовая компонента.

3.2. Вейвлет-фильтрация анализируемого ряда. Нами были проанализированы результаты вейвлет-фильтрации исходного временного ряда, содержащего суточные данные об энергопотреблении за первые 2048 рабочих дня, при отбрасывании шумовой компоненты, включающей разное количество коэффициентов разложения.

Для оценки возможного числа отбрасываемых вейвлет-коэффициентов и горизонта прогноза анализировалось поведение линейной автокорреляционной функции [18]:

$$C(\tau) = \frac{\sum_{i=1}^N (x_{i+\tau} - \bar{x})(x_i - \bar{x})}{\sum_{i=1}^N (x_i - \bar{x})^2}, \quad (7)$$

где N — это число измерений в анализируемом временном ряду $\{x_i\}$, а $\bar{x} = (1/N) \sum_{i=1}^N x_i$. В качестве оценки интервала корреляции τ — горизонта прогноза — принимается величина первого пересечения функцией $C(\tau)$ временной оси в доверительном интервале, отвечающем белому шуму [19, 20].

На рис. 4, *a* приведена зависимость автокорреляционной функции от величины интервала корреляции τ для шумовой компоненты временного ряда, число отбрасываемых коэффициентов равняется 1518 (чтобы минимизировать возможную примесь шума в регулярной компоненте, количество отбрасываемых коэффициентов было взято максимально возможным).

Видно, что для шумовой компоненты при $\tau = 1$ $C(\tau)$ пересекает ось абсцисс. Это означает, что соседние члены рассматриваемого ряда некоррелированы. Абсолютные значения $C(\tau)$ для $\tau > 1$ малы и с ростом τ быстро спадают по абсолютной величине.

На рис. 5 представлены (сверху вниз): рис. *a* — исходный временной ряд, содержащий суточные данные об энергопотреблении за первые 2048 рабочих дня; рис. *b* — указанный ряд после применения к нему вейвлет-фильтрации (оставлено 530 коэффициентов из 2048); рис. *в* — исключенная из исходного ряда высокочастотная (шумовая) компонента.

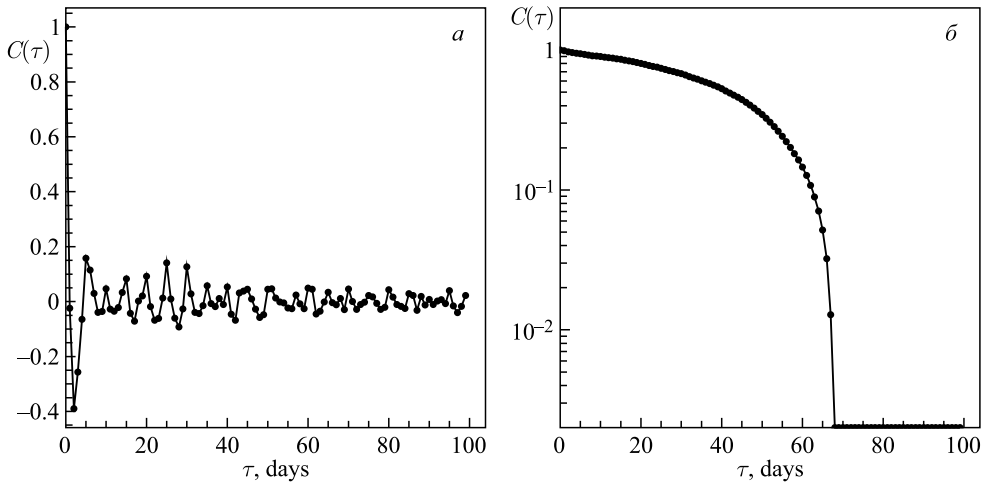


Рис. 4. Поведение автокорреляционной функции $C(\tau)$ в зависимости от интервала корреляции τ для шумовой компоненты (а) и регулярной составляющей временного ряда (б)

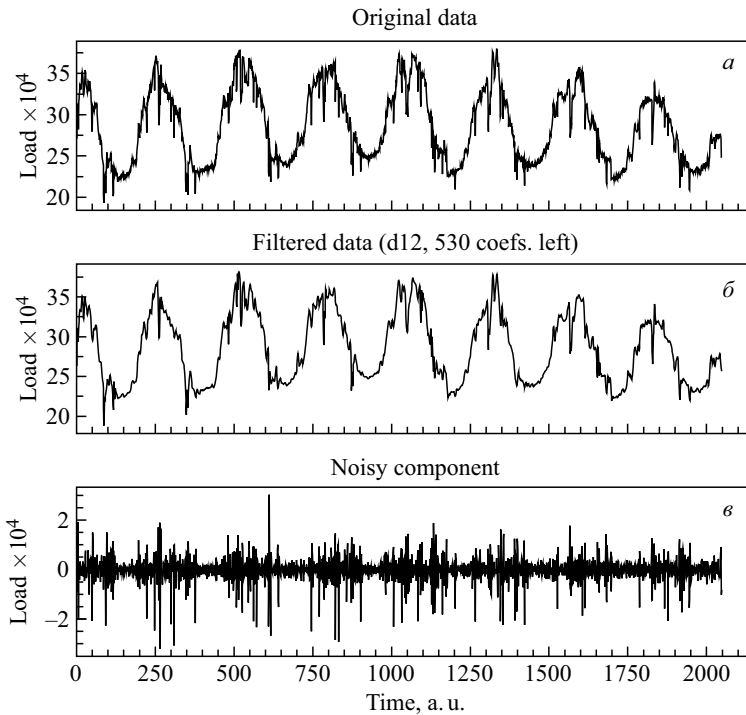


Рис. 5. а) Исходный временной ряд, содержащий данные об энергопотреблении за 2048 рабочих дней; б) реконструированная (на основе 530 коэффициентов) регулярная составляющая временного ряда; в) исключенная из исходного ряда высокочастотная (шумовая) компонента

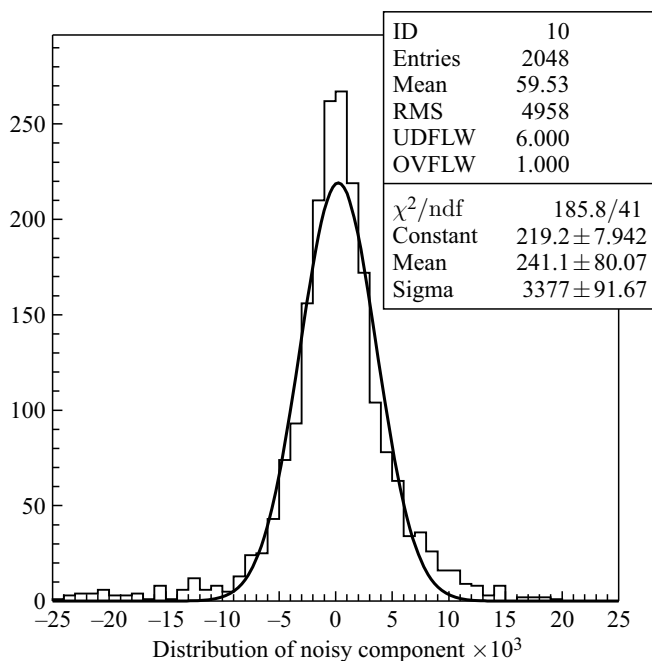


Рис. 6. Распределение шумовой компоненты (рис. 5, в)

На рис. 6 приведено распределение шумовой компоненты исследуемого нами временного ряда. На этом же рисунке представлен результат проверки нулевой гипотезы о соответствии наблюдаемой случайной величины нормальному распределению с помощью критерия согласия χ^2 [21]. Процедура фитирования была реализована с помощью пакета MINUIT [22] в среде пакета PAW (Physical Analysis Workstation) [23].

В представленном на рис. 6 примере при числе степеней свободы $\text{ndf} = 41$ уровню значимости $\alpha = 5\%$ отвечает критическая граница, равная 55,8. Это означает, что только в 5% случаев величина критерия χ^2 может превысить указанную границу. Полученное в нашем случае значение $\chi^2 = 185,8$ (см. рис. 6) существенно больше указанной критической границы. Из чего нельзя сделать вывод о том, что исследуемая случайная величина следует нормальному закону. Возможно, что из-за доминирующих в исходном ряду гармонических компонент в результате вейвлет-фильтрации наряду с шумовой компонентой затрагивается и регулярная составляющая.

Однако, учитывая то, что реализации рассматриваемого ряда довольно близки к гауссовскому закону и симметричны относительно нуля (см. рис. 6), можно предположить, что указанный ряд представляет собой белый гауссовский шум (см., например, [7]).

Прогнозирование временного ряда возможно только тогда, когда существует связь последующих значений ряда от предыдущих. На рис. 4, б приведена зависимость автокорреляционной функции от величины интервала корреляции τ для регулярной составляющей временного ряда. Исходя из поведения зависимости автокорреляционной функции от величины интервала корреляции τ для регулярной компоненты временного ряда, можно сделать предварительный вывод о том, что указанный ряд предоставляет неплохие возможности для среднесрочного прогнозирования.

4. МЕТОД «ГУСЕНИЦА»-SSA

В этом разделе мы вкратце изложим основные принципы анализа одномерного временного ряда, развитого в рамках подхода «Гусеница»-SSA [7, 24].

4.1. Основные этапы применения метода «Гусеница»-SSA. Метод «Гусеница»-SSA может быть использован для анализа временного ряда (3), отвечающего произвольной функции $f(t)$, определенной на равномерной сетке.

Стандартная схема «Гусеницы»-SSA включает четыре основных этапа:

- 1) преобразование одномерного временного ряда к многомерному виду;
- 2) разложение многомерного ряда в особых точках;
- 3) анализ этого разложения с помощью метода главных компонент и отбор основных компонент;
- 4) реконструкцию одномерного временного ряда на основе отобранных компонент.

Преобразование одномерного временного ряда (3) к многомерной форме реализуется путем представления (3) в матричной форме:

$$\mathbf{X} = (x_{ij})_{i,j=1}^{k,L} = \begin{pmatrix} x_1 & x_2 & x_3 & \dots & x_L \\ x_2 & x_3 & x_4 & \dots & x_{L+1} \\ x_3 & x_4 & x_5 & \dots & x_{L+2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_k & x_{k+1} & x_{k+2} & \dots & x_K \end{pmatrix}, \quad (8)$$

где $L < K$ называется длиной гусеницы, а $k = K - L + 1$.

Затем находятся собственные значения λ_i , $i = 1, 2, \dots, L$, и собственные векторы $\mathbf{V} = (v_1, v_2, \dots, v_i)$ ковариационной матрицы $C = (1/k)XX^T$. Матрицу \mathbf{V} можно рассматривать как матрицу перехода к главным компонентам:

$$\mathbf{Y} = \mathbf{V}^T \mathbf{X} = (y_1, y_2, \dots, y_L), \quad (9)$$

где y_i ($i = 1, 2, \dots, L$) — столбцы матрицы, состоящие из k элементов.

Соотношение

$$\sum_{i=1}^L \frac{\lambda_i}{L} = \sum_{i=1}^L \alpha_i = 1$$

позволяет оценить вклад α_i в порядке возрастания i -й компоненты в анализируемый ряд.

Этот вклад может быть интерпретирован как доля информации, отвечающая конкретной компоненте, и совместно с аналитическим и визуальным анализом собственных векторов и главных компонент позволяет отобрать характерные компоненты для реконструкции одномерного временного ряда. Обычно отбор характерных компонент зависит от целей, которые мы преследуем, и от информативного содержания отдельных компонент (см. детали в [7, 24]).

4.2. Вычисление прогнозной части файла-«подсказки». Для вычисления прогнозной части файла-«подсказки» мы использовали программу CaterpillarSSA (version 3.40, Professional M Edition, см. детали на сайте [26]). Данная версия программы предназначена для анализа и прогноза одномерных и многомерных временных рядов. Она также содержит модуль обнаружения разладки в структуре одномерных временных рядов.



Рис. 7. Вычисление прогнозной части файла-«подсказки» с помощью программы CaterpillarSSA

В настоящей работе мы опускаем детали вычисления прогнозной части файла-«подсказки» с помощью CaterpillarSSA, так как планируем подробно описать используемую процедуру в готовящейся к публикации новой статье. Остановимся только на отдельных важных моментах:

1) в качестве входной информации для программы CaterpillarSSA использовался отфильтрованный с помощью вейвлет-фильтрации временной ряд, т.е. ряд, из которого была удалена шумовая компонента;

2) перед тем, как перейти к многомерному виду, этот ряд был отнормирован (средствами CaterpillarSSA);

3) при переходе к многомерному виду (8) длина гусеницы была взята равной 600, так как она: а) должна быть кратной 5 (число дней в неделе, исключая выходные) и 12 (число месяцев в году); б) включать несколько годовых циклов (с тем, чтобы можно было выделить характерные особенности анализируемого ряда);

4) на этапе реконструкции одномерного ряда была оставлена 31 компонента; их общий вклад составил чуть более 95 %;

5) на этапе прогнозирования доверительный интервал был задан равным 0,25.

На рис.7 приведены результаты прогнозирования с помощью программы CaterpillarSSA на 120 дней вперед: сплошной линией изображен отфильтрованный ряд; штриховой кривой представлена прогнозная кривая; пунктирными линиями показаны кривые, отвечающие доверительному интервалу.

5. ОБУЧЕНИЕ И ТЕСТИРОВАНИЕ НЕЙРОННОЙ СЕТИ

На рис. 8 приведен график изменения ошибки при обучении ИНС в зависимости от номера эпохи.

После завершения процедуры обучения МСП и фиксации весов оценивалось качество прогноза. Прогноз оценивался по последним 120 наблюдениям. На рис. 9 приведено

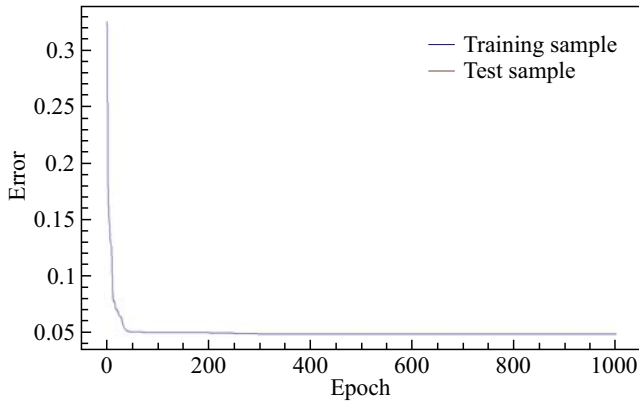


Рис. 8. График изменения ошибки при обучении ИНС в зависимости от номера эпохи

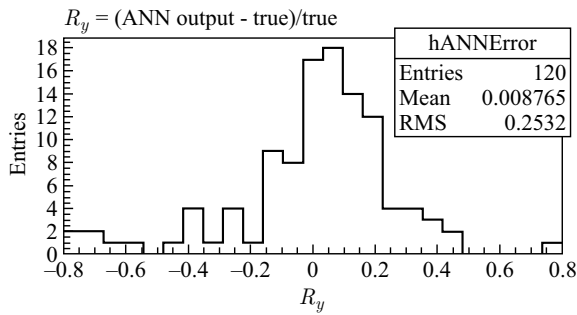


Рис. 9. Распределение относительной ошибки при тестировании обученной сети



Рис. 10. Фактические и предсказанные значения объемов энергопотребления

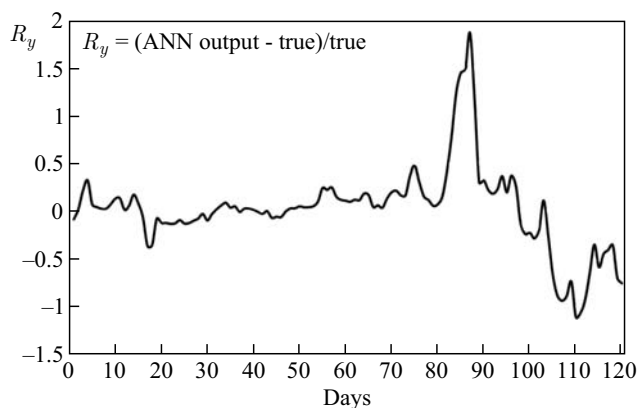


Рис. 11. Зависимость относительной ошибки от порядкового номера дня прогнозирования

распределение относительной ошибки

$$R_y = \frac{y_{\text{real}} - y_{\text{prog}}}{y_{\text{real}}}, \quad (10)$$

где y_{real} — наблюдаемое (реальное) значение суточного энергопотребления, а y_{prog} — прогноз, полученный с помощью МСП. Из приведенного на рис. 9 распределения видно, что, несмотря на довольно большой промежуток прогнозирования, обученная ИНС обеспечивает неплохое прогнозирование суточного энергопотребления в Московском регионе.

На рис. 10 представлены фактические объемы энергопотребления (сплошная линия) вместе с прогнозными значениями (штриховая), предсказанными ИНС на период в 120 дней.

На рис. 11 приведена кривая зависимости относительной ошибки от порядкового номера дня прогнозирования. Следует отметить, что полученная зависимость неплохо согласуется с поведением автокорреляционной функции для регулярной составляющей анализируемого ряда (см. рис. 4, б).

6. ОБСУЖДЕНИЕ РЕЗУЛЬТАТОВ И ВЫВОДЫ

В настоящей работе нами показано, что задача прогнозирования суточного энергопотребления для Московского региона может быть принципиально решена с помощью ИНС.

Важную роль в успешном решении рассматриваемой задачи сыграли следующие факторы: 1) оптимально подобранная архитектура ИНС; 2) адекватная структура выборки, подаваемой на вход сети как на этапе обучения ИНС, так и при ее тестировании; 3) оригинально построенные процедуры обучения и прогнозирования сети.

Первые три переменные, подаваемые на вход ИНС, отвечают за сезонные и периодические колебания энергопотребления в Московском регионе. Особо следует отметить четвертую переменную, играющую роль своеобразной «подсказки» для ИНС, которая берется либо из отфильтрованных данных (на этапе обучения сети), либо из прогнозных

значений, вычисленных с помощью пакета «Гусеница»-SSA [7] (на этапе тестирования сети). Последняя (пятая) переменная представляла собой величину, взятую из исходного ряда (на этапе обучения сети), либо то значение, которое предсказывалось на текущий день обученной ИНС (на этапе тестирования сети).

Не вдаваясь в детали, можно сделать вывод о том, что сформированная таким образом входная выборка позволила: 1) достичь быстрого и эффективного обучения нейронной сети, 2) обеспечить приемлемое среднесрочное прогнозирование суточного энергопотребления для Московского региона.

Учитывая то, что подход, реализованный в пакете «Гусеница»-SSA [24, 25], предоставляет дополнительные возможности для детального анализа исследуемых временных процессов, в нашей следующей работе мы планируем провести сравнение результатов по прогнозированию суточного энергопотребления для Московского региона на основе ИНС с тем, чего можно достичь, используя только методику, развитую в подходе «Гусеница»-SSA.

СПИСОК ЛИТЕРАТУРЫ

1. *Иванов В. В., Осетров Е. С.* Прогнозирование объемов пассажирских перевозок в Московском метрополитене с помощью искусственных нейронных сетей // *Вестн. Нац. исслед. ядерного ун-та «МИФИ»*. 2016. Т. 5, № 1. С. 65–74.
2. *Иванов В. В., Осетров Е. С.* Прогнозирование пассажиропотока в Московском метрополитене на основе нейронных сетей с предварительной фильтрацией анализируемых данных // *Там же*. № 2. С. 162–169.
3. *Haykin S.* *Neural Networks: A Comprehensive Foundation*. 2nd Ed. Prentice Hall, 1999.
4. *Denby B.* Tutorial on Neural Networks Applications in High Energy Physics: 1982 Perspective // *Proc. of the Second Intern. Workshop on “Software Engineering, Artificial Intelligence and Expert System in High Energy Physics”, L’Agelaude France-Telecom La Londe-les-Maures, France, Jan. 13–18, 1992. New Comp. Techniques in Phys. Res. II / Ed. by D. Perret-Gallix. World Sci., 1992. P. 287.*
5. *Fogelman Soulie F.* *Neural Networks for Patterns Recognition: Introduction and Comparison to Other Techniques* // *Ibid.* P. 277.
6. *Rumelhart D. E., Hinton G. E., Williams R. J.* Learning Internal Representations by Error Propagation // *Parallel Distributed Processing: Explorations in the Microstructure of Cognition / Eds.: D. E. Rumelhart, J. L. McClelland. V. 1: Foundations. MIT Press, 1986.*
7. *Голяндина Н. Э., Некруткин В. В., Браулов К. А.* Метод «Гусеница»-SSA: анализ временных рядов. <http://www.gistatgroup.com/gus/>.
8. *Cybenko G.* Approximation by Superposition of a Sigmoidal Function // *Math. Control Signals Systems*. 1989. V. 2. P. 303.
9. *Peterson C., Rögnvaldsson Th., Lönnblad L.* JETNET 3.0 — A Versatile Artificial Neural Network Package // *Comp. Phys. Commun.* 1994. V. 81. P. 185.
10. *Hoecker A. et al.* TMVA 4.2.0 — Toolkit for Multivariate Data Analysis with ROOT. arXiv:physics/0703039; Data Analysis, Statistics and Probability. CERN-OPEN-2007-007, TMVA Version 4.2.0. 2013; <http://tmva.sourceforge.net>.
11. *Brun R., Rademakers F.* ROOT — An Object Oriented Data Analysis Framework // *Nucl. Instr. Meth. A*. 1997. V. 389. P. 81.

12. *Lahmiri Salim.* A Comparative Study of Backpropagation Algorithms in Financial Prediction // Intern. J. Comp. Sci. Engin. Appl. 2011. V. 1, No. 4. P. 15–21.
13. *Chui C. K.* An Introduction to Wavelets. New York: Acad. Press, 1992. P. 1–18.
14. *Mallat S. A.* Wavelet Tour of Signal Processing. Acad. Press, 1999.
15. *Press W. H. et al.* Numerical Recipes in C: The Art of Scientific Computing. 2nd Ed. Cambridge Univ. Press, 1988; 1992.
16. *Daubechies I.* Wavelets. Philadelphia: S.I.A.M., 1992.
17. *Antoniu I. et al.* Wavelet Filtering of Network Traffic Measurements. JINR Commun. E11-2002-223. Dubna, 2002; Physica A. 2003. V. 324. P. 733–753.
18. *Broomhead D. S., King G. P.* Time-Series Analysis // Proc. Roy. Soc. London. 1989. V. 423. P. 103–110.
19. *Broomhead D. S., King G. P.* Extracting Qualitative Dynamics from Experimental Data // Physica D. 1986. V. 20. P. 217.
20. *Albano A. M. et al.* Singular Value Decomposition and the Grassberger Procaccia Algorithm // Phys. Rev. A. 1988. V. 38. P. 3017.
21. *Eadie W. T. et al.* Statistical Methods in Experimental Physics. Amsterdam; London: North-Holland Publ. Comp., 1971.
22. *James F., Roos M.* MINUIT — Function Minimization and Error Analysis. CERN Program Library D506. 1988.
23. *Brun R. et al.* PAW — Physics Analysis Workstation. CERN Program Library Q121. 1989.
24. Главные компоненты временных рядов: метод «Гусеница» / Ред.: Д. Л. Данилов, А. А. Жиглявский. Изд-во СПбГУ, 1997.
25. *Golyandina N., Nekrutkin V., Zhigljavsky A.* Analysis of Time Series Structure: SSA and Related Techniques. Chapman & Hall/CRC, 2001.
26. <http://www.gistatgroup.com/cat/>

Получено 21 июня 2016 г.