

ПРЕДСКАЗАНИЕ СОБСТВЕННОГО ЭНЕРГОПОТРЕБЛЕНИЯ АЭС С ИСПОЛЬЗОВАНИЕМ МЕТОДОВ ИНТЕЛЛЕКТУАЛЬНОГО АНАЛИЗА ДАННЫХ

М.А. Клещёва, А.В. Нахабов

*Обнинский институт атомной энергетики НИЯУ МИФИ
249040, Обнинск, Калужской обл., Студгородок, 1.*



В процессе работы атомная электростанция потребляет значительный объем электроэнергии, так называемый расход на собственные нужды. При этом существующая в Российской Федерации практика заключается в том, что для АЭС должен быть заранее заказан у оператора энергосистемы необходимый объем закупаемой электроэнергии, причем отклонения фактического энергопотребления от прогнозного в ту или иную сторону влекут за собой определенные штрафные экономические санкции. По этой причине важное значение приобретает точность прогнозирования энергопотребления на ближайший период. Рассматривается применение различных методов анализа данных для прогнозирования собственного энергопотребления АЭС с использованием фактических данных, сравнение этих методов как между собой, так и с используемым в настоящее время на АЭС. В результате предлагается метод прогнозирования собственного энергопотребления АЭС с существенно более высокой точностью.

Ключевые слова: атомная станция, энергопотребление, методы прогнозирования, экспоненциальное предсказание, метод опорных векторов.

ВВЕДЕНИЕ

В состав оборудования атомной станции входит большое число потребителей электроэнергии, расход которой называется расходом на собственные нужды. АЭС в процессе эксплуатации закупает электроэнергию для собственных нужд из единой энергосистемы. Закупка электроэнергии осуществляется на основании прогнозирования. Очень важно как можно точнее планировать собственное потребление: чем точнее прогноз энергопотребления, тем выше финансовый результат. Некачественное планирование потребности в энергоресурсах ведет, в одних случаях, к переплате и, соответственно, к неэффективному отвлечению денежных средств атомной станции, а в других случаях – к риску возможных ограничений поставки энергоресурсов. Поэтому необходимо определить величину объема закупаемой электроэнергии для внесения соответствующей предоплаты.

Атомная станция берет на себя некоторый риск, который связан с невозможностью точного планирования заявки на потребление электроэнергии. Практика показывает, что для прогнозирования необходимого количества энергоресурсов для атомных станций не существует общего единого метода, поскольку огромное коли-

© М.А. Клещёва, А.В. Нахабов, 2015

чество протекающих технологических циклов, суммируясь, образуют уникальный временной процесс.

ПОДХОДЫ К ПЛАНИРОВАНИЮ СОБСТВЕННОГО ПОТРЕБЛЕНИЯ НА АЭС

Насколько известно, на атомных станциях не существует единого разработанного метода прогнозирования энергопотребления. Подход к данному вопросу был рассмотрен на Смоленской АЭС. На сегодня заказ энергоресурсов на САЭС происходит на основе визуального и статистического анализа данных за текущий период. За величину прогнозного значения на будущий период берется некоторая средняя величина (рис. 1). Из рисунка 2 видно, что такой подход дает существенную ошибку в прогнозировании и не учитывает поведение случайной величины.

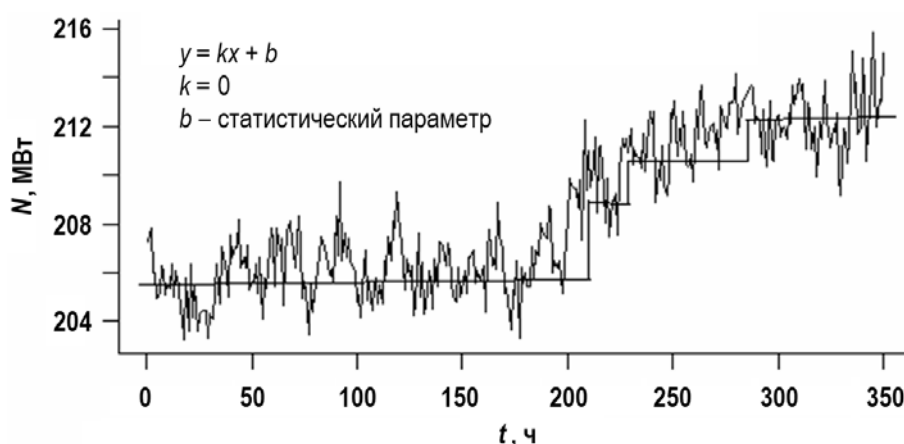


Рис. 1. Существующий подход прогнозирования собственного энергопотребления на САЭС

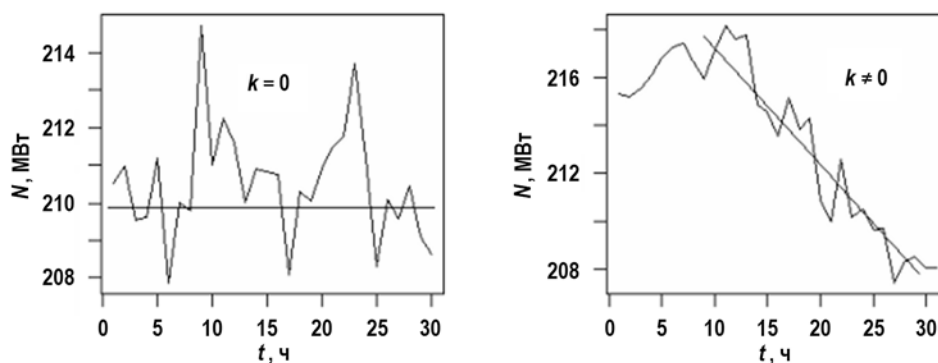


Рис. 2. Недостатки подхода прогнозирования, принятого на САЭС

МЕТОДЫ ПРОГНОЗИРОВАНИЯ

Временной ряд представляет собой набор или последовательность наблюдаемых данных, расположенных в хронологическом порядке в равноотстоящие промежутки времени. Прогнозирование временных рядов является предсказанием будущих событий на основе уже известных прошлых событий с использованием подходящей модели. Существует классификация методов, которые могут быть применены для прогнозирования энергопотребления [1].

В качестве исходных данных используются значения фактического потребления электроэнергии Смоленской АЭС – филиалом ОАО «Концерн Росэнергоатом».

Прогноз по максимумам основан на разбиении временного ряда на участки. Приняв, что величина энергопотребления за некий предшествующий временной период считается постоянной и равняется его максимальному значению, можно построить простейшую модель прогнозирования. Метод медиан по тройкам основан на предыдущем методе. Отличие в том, что за величину прогнозного значения берется не максимальное значение, а медиана по тройкам, т.е. отсекаются самые большие и самые маленькие значения. На рисунке 3 показаны примеры таких прогнозов.

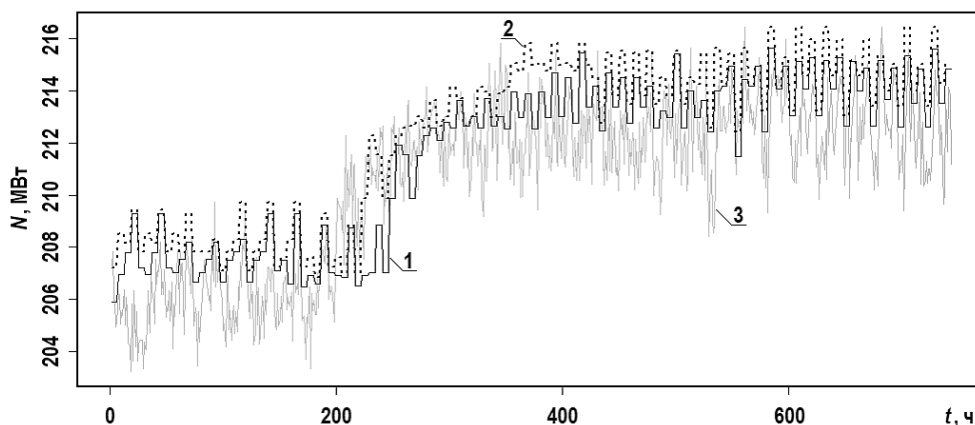


Рис. 3. Прогноз по максимумам и предсказание с помощью медиан по тройкам: 1 – прогнозирование с использованием медиан по тройкам; 2 – прогноз по максимумам; 3 – фактическое энергопотребление

Экспоненциальное предсказание – метод математического преобразования, используемый при прогнозировании временных рядов. В его основе лежит расчет экспоненциальных средних. Экспоненциальное предсказание ряда осуществляется по рекуррентной формуле:

$$S_t = \alpha \cdot X_t + (1 - \alpha) \cdot S_{t-1}, \quad (1)$$

где S_t – значение экспоненциальной средней в момент времени t ; α – параметр сглаживания, $0 < \alpha < 1$. Как видно из выражения (1), каждое новое сглаженное значение (которое является также прогнозом) вычисляется как взвешенное среднее текущего наблюдения и сглаженного ряда. Более старым наблюдениям приписываются экспоненциально убывающие веса, учитывающие все предшествующие наблюдения ряда. Результат сглаживания зависит от параметра α . Если $\alpha = 1$, то предыдущие наблюдения полностью игнорируются. Если $\alpha = 0$, то игнорируются текущие наблюдения. Экспоненциальное предсказание является простейшим вариантом самообучающейся модели. Вычисления просты и выполняются итеративно, массив для инициализации вычислений может быть уменьшен до одного элемента S_{t-1} [2, 6].

Параметр α следует выбирать так, чтобы значение суммы квадратов отклонений фактического потребления электроэнергии от прогнозного было минимальным. Оказалось, что наименьшая среднеквадратичная ошибка достигается при $\alpha = 0.3$. На рисунке 4 представлены графики фактического энергопотребления и прогноз с помощью экспоненциального предсказания. Как видно, данный метод наилучшим образом повторяет распределение имеющихся данных.

Метод опорных векторов SVM (англ. – Support Vector Machines) решает задачи классификации и регрессии путем построения нелинейной плоскости, разделяющей решения. В представленной работе метод опорных векторов используется в решении задачи регрессии SVR (англ. – Support Vector Regression). В исходном пространстве SVM строит оптимальную гиперплоскость, применяя итерационный алгоритм обучения, использующийся для минимизации функции ошибок.

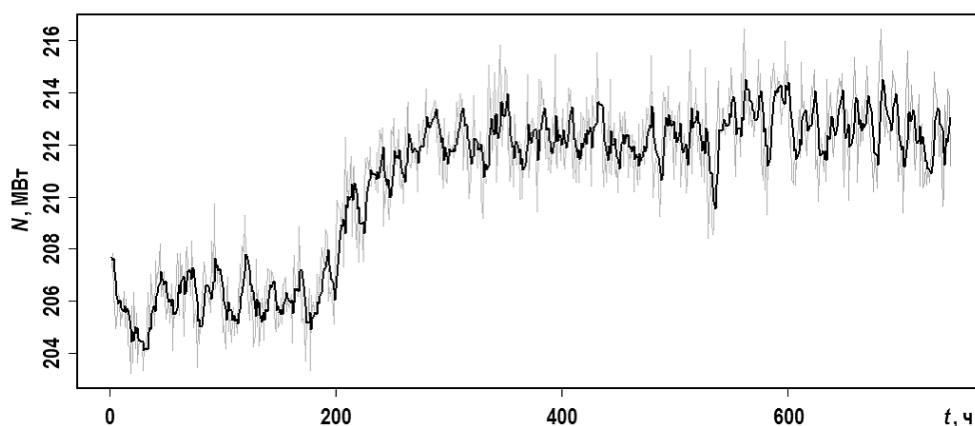


Рис. 4. Экспоненциальное предсказание энергопотребления

Опишем случайную линейную функцию f , имеющую форму

$$f(x) = \langle \omega, x \rangle + b, \quad \omega \in X, \quad b \in R. \quad (2)$$

Задача нахождения параметров формулируется в виде минимизации функционала

$$\min ||\omega||^2 / 2 + C \sum_{i=1}^l (\xi_i + \xi_i^*) \quad (3)$$

при ограничениях

$$\begin{aligned} y_i - \langle \omega, x_i \rangle - b &\leq \varepsilon + \xi_i, \\ \langle \omega, x_i \rangle + b - y_i &\leq \varepsilon + \xi_i^*, \\ \xi_i, \xi_i^* &\geq 0. \end{aligned} \quad (4)$$

В функционале (3) первое слагаемое $||\omega||^2 / 2$ неявно обеспечивает некую степень гладкости функции, а второе – штрафует любые отклонения $f(x_i)$ от y_i , большие, чем ε для всех обучающих данных, где ε – так называемая функция потерь. В оптимизационную задачу вводятся параметры ξ , с помощью которых контролируются отклонение остатков регрессии в большую или меньшую стороны от заданной границы ε .

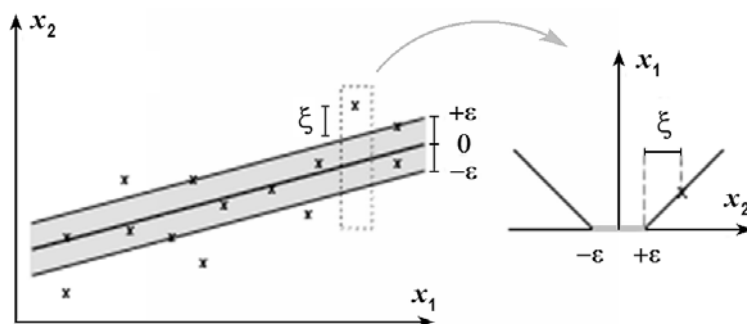


Рис.5. Установление оптимальной функции потерь для линейного SVM

Оказывается, что в большинстве случаев задача оптимизации (3) может быть решена более легко в своей двойной формулировке:

$$f(x) = \sum_{i=1}^l (\alpha_i - \alpha_i^*) K(x_i, x) + b. \quad (5)$$

Элементы данных, которым соответствуют ненулевые значения α_i , называют опорными векторами, задающими границы построенной гиперплоскости. Ядерная функция K заменяет скалярное произведение в преобразованном пространстве. Существует ряд ядер, которые могут быть использованы в моделях метода опорных векторов. Выбор подходящей ядерной функции – ключевая задача при получении рабочей модели регрессии [3, 5].

Данный алгоритм реализуется в пакете LibSVM в системе статистических вычислений R [4]. Использовался вариант алгоритма с полиномиальным ядром. Подбор параметров регрессии осуществлялся на обучающей выборке. Точность метода и прогнозируемая способность оценивались на проверочной выборке. Таким образом, была выбрана модель с оптимальными параметрами, которые обеспечивают наименьшую ошибку предсказания и удовлетворительное число опорных векторов.

Таблица 1

Результаты прогнозирования энергопотребления

Методы	Сумма квадратов отклонений	Сумма положительных ошибок	Сумма отрицательных ошибок
Существующий метод на САЭС	3532.2	728.2	–599.9
Прогнозирование по максимумам	4374	86.9	–1444.4
Медиана по тройкам	2893.2	232.2	–963.5
Экспоненциальное предсказание	1253	391.8	–365.5
SVR	0.19	6.97	–3.64

Сравнение результатов показало, что метод SVR является наилучшим с точки зрения точности прогнозирования энергопотребления. Вектором ошибок для каждого метода является разница между фактическим потреблением электроэнергии и предсказанным. В таблице 1 эти результаты занесены во вторую колонку. Прогнозирование осуществлялось на один месяц (744 значения).

ПОДХОДЫ К ПРОГНОЗИРОВАНИЮ

При построении модели встает вопрос, какой объем предшествующих данных оптимален для будущего предсказания, поскольку устаревшие данные при моделировании часто бывают бесполезными и оказывают негативное влияние. Также нужно определить, на какой интервал времени осуществлять прогноз. Так как процесс потребления электроэнергии имеет нестационарный характер, то чем больше период прогноза, тем более вероятно изменение количества энергопотребления. Кроме того, возможно возникновение непредсказуемых ситуаций, которые оказывают существенное влияние на изучаемый процесс.

Поскольку SVR дает наилучший результат предсказания энергопотребления, для увеличения точности прогноза необходимо подобрать оптимальные размеры обучающей и проверочной выборок, при которых модель способна отражать изменяю-

щиеся условия во времени и адаптироваться к временному ряду.

Были рассмотрены два подхода к прогнозированию по методологии SVR. Первый вариант – на основе увеличивающейся обучающей выборки. Задается начальный размер обучающей выборки, производится прогноз на некий период (положим, на одни сутки), затем обучающая выборка увеличивается ровно на период прогнозирования. Второй – на основе скользящей обучающей выборки. Задается начальный размер обучающей выборки, производится прогноз на некий период (например, на одни сутки), затем обучающая выборка сдвигается ровно на период прогнозирования.

Таблица 2

Результаты прогнозирования с использованием различных подходов

Подходы	Сумма квадратов отклонений	Сумма положительных ошибок	Сумма отрицательных ошибок
SVR (увеличивающаяся обучающая выборка)	132.4	26.0	–454.6
SVR (скользящая обучающая выборка)	132.0	64.1	–244.5

Для рассматриваемых методик прогнозирования проводились численные эксперименты, результаты которых занесены в табл. 3, где введены обозначения: **train** – обучающая выборка, **test** – проверочная выборка. Как видно из табл. 2, 3, преимущество скользящей обучающей выборки состоит в том, что, имея небольшое количество предшествующих данных, можно добиться меньшей ошибки в прогнозировании и уравновесить баланс ошибок в одну и другую стороны.

Таблица 3

Результаты прогнозирования с различными размерами обучающей и проверочной выборок

		Сумма квадратов отклонений	Сумма положительных ошибок	Сумма отрицательных ошибок
Увеличивающаяся обучающая выборка	train = 7 сут, test = 36 ч	132.4	27.0	–454.5
	train = 7 сут, test = 24 ч	132.4	26.0	–454.6
	train = 14 сут, test = 24 ч	132.3	23.7	–454.6
Скользящая обучающая выборка	train = 7 сут, test = 36 ч	124.1	64.3	–236.5
	train = 7 сут, test = 24 ч	132.0	64.1	–244.5
	train = 14 сут, test = 24 ч	133.7	57.7	–286.3

АЭС выплачивает штраф в том случае, когда электроэнергия заказана с избытком, и, наоборот, при недостатке электроэнергии. Поскольку оплата штрафа различна в этих двух случаях, то нашей целью является сведение экономических затрат АЭС к минимуму. SVR позволяет использовать различные подходы к прогнозированию и регулировать ошибки предсказания в положительную и отрицательную

стороны. Поскольку ошибки прямо пропорционально связаны с ценой на электроэнергию, можно менять соотношение ошибок и тем самым стремиться к уменьшению выплаты штрафа. Таким образом, построенная модель в дальнейшем может быть использована атомными станциями для прогнозирования собственного энергопотребления, что позволит достичь существенного экономического эффекта.

ЗАКЛЮЧЕНИЕ

Выполнен численный анализ пяти методов прогноза собственного энергопотребления АЭС. Сравнение результатов показало, что метод опорных векторов, основанный на интеллектуальной обработке данных, является наилучшим из них и дает наименьшую ошибку. Построенный на базе этого метода вычислительный алгоритм позволяет регулировать предсказательную способность и тем самым сокращать экономические затраты атомной станции. В дальнейшем на его основе возможно создание полноценного программного обеспечения, которое позволит не только сэкономить на затратах, но и оптимизировать процессы поставки, учета, потребления и управления энергоресурсами.

Литература

1. *Alfares H.K., Nazeeruddin M.* Electric Load Forecasting: literature survey and classification of methods // *International Journal of Systems Science*. – 2002. – Vol. 33. – № 1. – PP. 23-34.
2. *Лукашин Ю.П.* Адаптивные методы краткосрочного прогнозирования временных рядов: Учеб. пособие. – М.: Финансы и статистика, 2003. – 416 с.
3. *Smola A.J., Schölkopf B.A.* Tutorial on Support Vector Regression // *Statistics and computing*. – 2004. – Vol. 14. – № 3. – PP. 199-222.
4. *Хайкин С.* Нейронные сети: полный курс – М.: Издательский дом «Вильямс», 2006. – 1104 с.
5. *Барсегян А.А., Куприянов М.С., Степаненко В.В., Холод И.И.* Методы и модели анализа данных: OLAP и DataMining. – СПб.: БХВ-Петербург, 2004. – 336 с.
6. *Андерсон Т.* Статистический анализ временных рядов – М.: Мир, 1976. – 755 с.

Поступила в редакцию 26.01.2015 г.

Авторы

Клещёва Марина Александровна, магистр 1 курса ИАТЭ НИЯУ МИФИ

E-mail: kleshcheva_ma@mail.ru

Нахабов Александр Владимирович, доцент, кандидат техн. наук

E-mail: AVNakhabov@mephi.ru

UDC 621.317

PREDICTION OWN ENERGY CONSUMPTION NUCLEAR POWER PLANTS USING DATA MINING METHODS

Kleshchyova M.A., Nakhabov A.V.

Obninsk Institute for Nuclear Power Engineering, National Nuclear Research University «MEPhI». 1, Studgorodok, Obninsk, Kaluga reg., 249040 Russia

ABSTRACT

While in operation, nuclear power plant consumes a significant amount of electricity, the so-called consumption for own needs. The existing practice in the Russian Federation is that nuclear power plants must order in advance from the grid operator required volume of electricity so that any deviation of the actual energy consumption of the predicted value results in some penalties. It is quite difficult to predict required value of energy consumption in advance, so prediction of energy consumption with high accuracy using available operational information is very important.

This article discusses the use of various methods of data analysis and data mining to predict the energy consumption for own needs of nuclear power plant using actual operational data. Among these methods are smoothing by medians, exponential smoothing and support vector regression. First of all the simple method of prediction currently used at the NPP was considered. Then the comparison of proposed methods is performed, both among themselves and with the currently used one at the plant. The special emphasis is placed on support vector regression. As a final result, the method of forecasting of nuclear power plant energy consumption for own needs using support vector regression with significantly higher accuracy is proposed.

Key words: nuclear power plants, energy consumption, forecasting methods, exponential smoothing, Support Vector Machines.

REFERENCES

1. Alfares H.K., Nazeeruddin M. Electric Load Forecasting: literature survey and classification of methods. *International Journal of Systems Science*. 2002, v. 33, no. 1, pp. 23-34.
2. Lukashin Yu.P. Adaptive Methods of Short-term Time Series Prediction – Moscow, Finansy i Statistika Publ., 2003, 416 p. (in Russian).
3. Smola A.J., Schölkopf B.A. Tutorial on Support Vector Regression. *Statistics and computing*. 2004, v. 14, no.3, pp. 199–222.
4. Haykin S. Neural Networks: A Comprehensive Foundation. N.Y., Prentice Hall, 1999.
5. Barsegyan A.A., Kupriyanov M.S., Stepanenko V.V., Holod I.I. Methods and Models of Data Analysis: OLAP and Data Mining. Sankt-Petersburg. BHV-Peterburg Publ., 2004, 336 p. (in Russian).
6. Anderson T.W. The Statistical Analysis of Time Series. John Wiley & Sons, 1971.

Authors

Kleshchyova Marina Aleksandrovna, Master Science Degree Student of OINPE NRNU MEPhI
E-mail: kleshcheva_ma@mail.ru

Nakhabov Alexandr Vladimirovich, Assistant Professor, Cand. Sci. (Engineering),
E-mail: AVNakhabov@mephi.ru