

ПРИМЕНЕНИЕ МЕТОДА ЭКСПОНЕНЦИАЛЬНОГО СГЛАЖИВАНИЯ ДЛЯ ВОССТАНОВЛЕНИЯ УТЕРЯННЫХ ДАННЫХ ТЕХНИЧЕСКОГО УЧЕТА ЭЛЕКТРОЭНЕРГИИ НА ПРОМЫШЛЕННЫХ ПРЕДПРИЯТИЯХ

У статті детально досліджена проблема відновлення одиночних даних технічного обліку електроенергії методом експоненційного згладжування. Перевірена гіпотеза про ефективність знаходження оптимального значення постійної згладжування α в діапазоні $0 < \alpha < 2$. Встановлено діапазони зміни значень середньої абсолютної відсоткової помилки прогнозу залежно від значень коефіцієнта варіації часового ряду для конкретних процесів електроспоживання з різним розкидом даних.

В статье детально исследована проблема восстановления одиночных данных технического учета электроэнергии методом экспоненциального сглаживания. Проверена гипотеза об эффективности нахождения оптимального значения постоянной сглаживания α в диапазоне $0 < \alpha < 2$. Установлены диапазоны изменения значений средней абсолютной процентной ошибки прогноза в зависимости от значений коэффициента вариации временного ряда для конкретных процессов электропотребления с различным разбросом данных.

Постановка проблемы

Отсутствие учетных данных энергопотребления на промышленных предприятиях приводит к недоучету энергоресурсов, отсутствию возможности контроля энергоэффективности производственных объектов, невозможности решить разноплановые задачи, стоящие перед электрослужбой промышленных предприятий. Для решения данных задач принято использовать процедуру восстановления данных.

Анализ последних исследований и публикаций

Метод экспоненциального сглаживания описан в работах Дж. Бокса [1], Г. Дженкинса [1], Лукашина Ю.Г. [2], Светунькова С.Г. [3], Четыркина Е.М. [4]. Применение этого метода в электроэнергетике рассматривалось в работах Фармера Е.Д. [5], Бэнн Д.В. [5], Праховника А.В. [6], Гордеева В.И. [7], Розена В.П. [6], Гурского С.К. [8]. В работах [1-5] отмечено, что метод экспоненциального сглаживания может быть использован для краткосрочных прогнозов будущей тенденции на один период вперед. Этот метод автоматически корректирует любой прогноз в свете различий между фактическим и спрогнозированным результатом. Поэтому возможно использовать метод экспоненциального сглаживания для восстановления одиночных пропущенных данных электропотребления. Ведь восстановление данных – это процедура, аналогичная прогнозированию, только не на будущий интервал времени, а на уже прошедший период времени.

В работе д.т.н. Светунькова С.Г. (г. Санкт-Петербург) [9] было предложено расширить границы постоянной сглаживания α при прогнозировании экономических процессов не в пределах $0 < \alpha < 1$, а в пределах $0 < \alpha < 2$.

Цель статьи – проверить целесообразность расширения границ изменения постоянной сглаживания при проведении восстановления пропущенных одиночных данных временного ряда среднечасовых значений электропотребления в цехе аммиака и исследовать, от каких причин зависит точность метода экспоненциального сглаживания.

Основной материал исследований

В ситуации, когда наблюдаются одиночные пропуски данных, целесообразно использовать такие методы, как методы сплайн-интерполяции и экспоненциального сглаживания.

Рабочая формула одного из экстраполяционных методов – метода экспоненциального сглаживания:

$$U_{t+1} = \alpha \cdot y_t + (1 - \alpha) \cdot U_t,$$

где t – период, предшествующий прогнозируемому; $t+1$ – прогнозный период; U_{t+1} – прогнозируемый показатель; α – постоянная сглаживания; y_t – фактическое значение исследуемого показателя за период, предшествующий прогнозируемому; U_t – экспоненциально взвешенная средняя для периода, предшествующего прогнозируемому.

При прогнозировании данным методом возникает две проблемы:

- выбор значения постоянной сглаживания α ;
- определение начального значения U_t .

От величины α зависит, как быстро снижается вес влияния предшествующих наблюдений. Чем больше α , тем меньше сказывается влияние предшествующих значений временного ряда. Если значение α близко к единице, то это приводит к учету при прогнозе в основном влияния лишь последних наблюдений. Если значение α близко к нулю, то веса, по которым взвешиваются уровни временного ряда, убывают медленно, т.е. при прогнозе учитываются все (или почти все) прошлые наблюдения.

На практике постоянная сглаживания определяется путем поиска на сетке. Возможные значения параметра разбиваются сеткой с определенным шагом. Например, рассматривается сетка значений от $\alpha = 0,1$ до $\alpha = 1,9$ с шагом $0,1$.

Постоянная сглаживания должна изменяться в пределах [9]:

$$0 < \alpha < 2.$$

Модель Брауна имеет право на существование как при нахождении постоянной сглаживания в пределах:

$$0 < \alpha < 1,$$

которые называются "классическими", так и в пределах [9]:

$$1 \leq \alpha < 2,$$

которые называются "запредельным множеством".

Оптимальное значение постоянной сглаживания определяется, используя критерий минимума дисперсии ошибки аппроксимации:

$$\frac{1}{T} \sum_{t=1}^T (Y_t - \hat{Y}_t)^2 \rightarrow \min.$$

Когда оптимальное значение постоянной сглаживания находится в классических пределах, модель адаптивна, а в том случае, когда оно находится в за-предельном множестве, модель не только адаптивна, но и самообучаема [9]. Это говорит о том, что оптимальное значение постоянной сглаживания определяется свойствами исходного ряда.

Если в процессе оптимизации постоянная сглаживания лежит в пределах – от нуля до единицы, то модель Брауна может использоваться для прогнозирования достаточно эффективно. Если же оптимальное значение постоянной сглаживания оказалось находящимся в запредельном множестве, то это диагностирует ситуацию, когда средняя взвешенная в принципе не может использоваться в качестве хорошей оценки прогнозного значения моделируемого процесса. В этом случае возможно два варианта действий [8].

Первый. Процесс вышел за рамки простой динамики. У него появилась некоторая тенденция в развитии.

Второй. Процесс находится на грани между эволюционной и хаотической динамикой, и его математическое описание невозможно с помощью какой-либо модели. Поэтому такой процесс лучше всего прогнозировать с помощью модели Брауна, работающей в запредельном множестве.

Задача выбора U_t (экспоненциально взвешенного среднего начального) решается следующими способами:

- если есть данные о развитии явления в прошлом, то можно воспользоваться средней арифметической и приравнять к ней U_t ;
- если таких сведений нет, то в качестве U_t используют исходное первое значение базы прогноза U_1 .

Для исследований были взяты временные ряды среднечасовых значений электропотребления крупнотоннажного цеха аммиака за 7 подряд идущих суток (табл. 1). Для постоянной сглаживания, принимающей значения от 0,1 до 1,9 с шагом 0,1, рассчитывалось оптимальное значение постоянной сглаживания, используя в расчетах как первое значение базы прогноза, так и среднее почасовое значение электропотребления за предыдущие сутки. Результаты вычислений приведены на рис. 1, 2.

Таблица 1

Значения почасового электропотребления крупнотоннажного цеха аммиака за сутки, МВт и коэффициента вариации, %

Номер часа	Номер суток						
	1	2	3	4	5	6	7
1	31,930	36,609	36,570	37,107	37,146	37,264	37,118
2	31,875	36,641	36,585	37,054	37,091	37,275	37,088
...
21	36,634	36,563	36,965	37,072	37,147	37,472	37,314
22	36,647	36,603	37,031	37,137	37,148	37,267	37,327
23	36,594	36,594	37,035	37,159	37,138	37,074	37,387
24	36,578	36,535	37,050	37,146	37,188	37,133	37,420
Значение коэффициента вариации за сутки, %	6,90	0,24	0,49	0,81	0,23	0,39	0,30

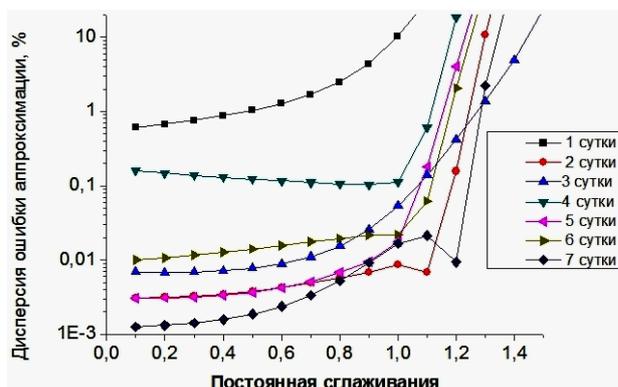


Рис. 1. Зависимость величины дисперсии ошибки аппроксимации от постоянной сглаживания при использовании в качестве экспоненциально взвешенного среднего начального первого значения базы прогноза

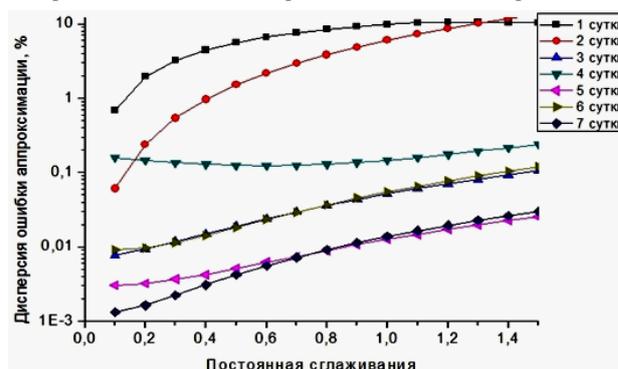


Рис. 2. Зависимость величины дисперсии ошибки аппроксимации от постоянной сглаживания при использовании в качестве экспоненциально взвешенного среднего арифметического значения за предыдущие сутки

Оптимальная величина постоянной сглаживания α из 7 суток 6 раз равнялась 0,1, а 1 раз – 0,9.

Для каждого значения постоянной сглаживания определялась средняя абсолютная процентная ошибка $MAPE$ при обоих способах выбора экспоненциально взвешенного среднего начального значения.

Средняя абсолютная процентная ошибка ($MAPE$) [10]:

$$MAPE = \frac{1}{n} \sum_{i=1}^n \frac{|y_m(i) - y_p(i)|}{y_m(i)},$$

где $y_m(i)$ – измеренное значение энергопотребления, $y_p(i)$ – прогнозное значение, n – горизонт прогнозирования (количество временных интервалов в выборке).

Результаты вычислений приведены на рис. 3, 4.

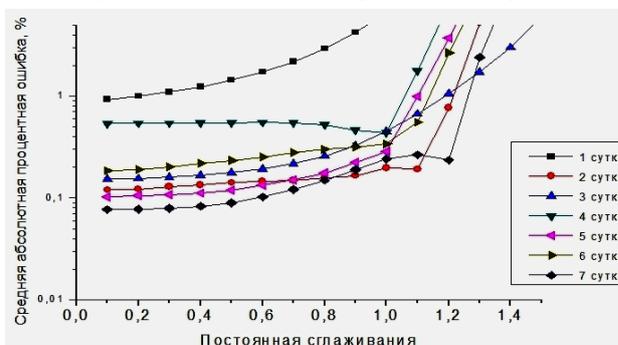


Рис. 3. Зависимость величины средней абсолютной процентной ошибки от постоянной сглаживания при использовании в качестве экспоненциально взвешенного среднего начального первого значения базы прогноза

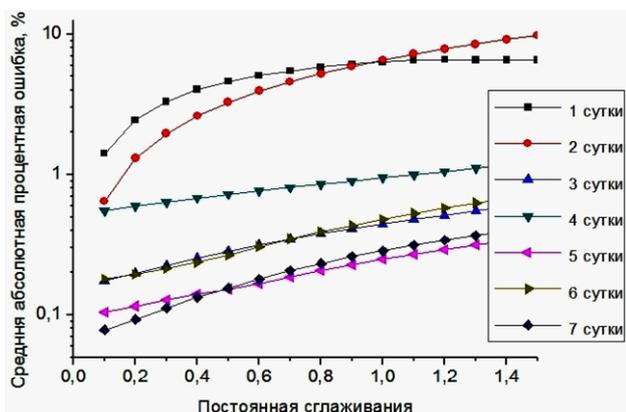


Рис. 4. Зависимость величины средней абсолютной процентной ошибки от постоянной сглаживания при использовании в качестве экспоненциально взвешенного среднего начального среднего арифметического значения за предыдущие сутки

Величина средней абсолютной процентной ошибки для обоих случаев определения экспоненциально взвешенного среднего начального значения не превышает 2%, а в среднем равна 0,5 %, что свидетельствует о высокой точности восстановления оригинальных пропущенных данных методом экспоненциального сглаживания. Исследования показали, что величина средней абсолютной процентной ошибки *MAPE* зависит от значения коэффициента вариации (*KB*) 24 значений временного ряда почасовых значений электропотребления цеха аммиака (табл. 2).

Таблица 2

Значения средней абсолютной процентной ошибки *MAPE* за сутки, полученные при использовании в качестве экспоненциально взвешенного среднего начального среднего арифметического значения за предыдущие сутки и первого значения базы прогноза, в цехе аммиака

Номер суток	1	2	3	4	5	6	7
Значение <i>MAPE</i> при использовании первого значения базы прогноза, %	0,92	0,12	0,15	0,44	0,10	0,19	0,08
Значение <i>MAPE</i> при использовании среднего арифметического значения за предыдущие сутки, %	1,41	0,65	0,17	0,55	0,10	0,18	0,08
Значение коэффициента вариации <i>KB</i> , %	6,90	0,24	0,49	0,81	0,23	0,39	0,30

Значениям *KB*, выраженным в процентах, от 0 до 0,5 %, соответствует величина средней абсолютной процентной ошибки *MAPE* около 0,2 %. Значениям *KB*, выраженным в процентах, от 0,5 до 1 % соответствует величина средней абсолютной процентной ошибки *MAPE* около 0,5 %. Значениям *KB*, выраженным в процентах, от 1 до 7 %, соответствует величина средней абсолютной процентной ошибки *MAPE* около 1 %. Но эти зависимости верны только для конкретного процесса электропотребления.

Зависимость средней абсолютной процентной ошибки *MAPE* от *KB* проверена также на временном ряду (почасовых значениях электропотребления за 7 суток) на насосной станции, обеспечивающей водой химические производства. В программе Microsoft Excel

создан шаблон, который может рассчитывать и оптимальное значение постоянной сглаживания, и величину *MAPE* по почасовым значениям электропотребления.

Получены следующие значения *MAPE* при определенных значениях *KB* (табл. 3).

Таблица 3

Значения средней абсолютной процентной ошибки *MAPE* и коэффициента вариации *KB* за сутки, полученные при использовании в качестве экспоненциально взвешенного среднего начального среднего арифметического значения за предыдущие сутки, на насосной станции

Номер суток	1	2	3	4	5	6	7
Значение <i>MAPE</i> при использовании среднего арифметического значения за предыдущие сутки, %	0,22	0,18	0,36	0,86	0,33	0,29	0,2
Значение коэффициента вариации <i>KB</i> , %	0,29	0,37	0,76	1,59	0,8	0,56	0,4

С табл. 2 и 3 следует, что при коэффициенте вариации меньше 1 % величина *MAPE* в среднем в два раза меньше величины *KB*.

Зависимость средней абсолютной процентной ошибки *MAPE* от *KB* проверена также на временном ряду (почасовых значениях электропотребления за 7 суток) в цехе слабоазотной кислоты, где наблюдается более высокий разброс данных временного ряда. Получены следующие значения *MAPE* при определенных значениях *KB* (табл. 4).

Таблица 4

Значения средней абсолютной процентной ошибки *MAPE* и коэффициента вариации *KB* за сутки, полученные при использовании в качестве экспоненциально взвешенного среднего начального среднего арифметического значения за предыдущие сутки, в цехе слабоазотной кислоты

Номер суток	1	2	3	4	5	6	7
Значение <i>MAPE</i> при использовании среднего арифметического значения за предыдущие сутки, %	4,06	3,58	2,50	1,66	2,11	0,98	1,13
Значение коэффициента вариации <i>KB</i> , %	16,62	11,73	6,73	3,24	3,53	2,99	2,34

Доверительные границы для прогностической оценки равны [4]:

$$y_{i+L} = \bar{y} \pm t_{\alpha} s \sqrt{\frac{2}{2-\alpha}}$$

где y_{i+L} – экстраполируемое значение уровня, L – период упреждения ($L = 1$), t_{α} – табличное значение t – статистики Стьюдента с $n-1$ степенями свободы и уровнем вероятности p , s – среднее квадратическое отклонение (СКО), \bar{y} – среднее значение временного ряда.

Для 24 значений временного ряда значений электропотребления за первые сутки в цехе аммиака с средним значением 37,048 МВт и СКО, равным 0,087 МВт, значением $t_{\alpha} = 2,07$ (для 23 степеней свободы и уровня доверительной вероятности 95 %), постоянной сглаживания $\alpha = 0,1$ доверительный интервал для цеха аммиака равен $37,048 \pm 0,184$ МВт.

Согласно [4] минимальное число наблюдений для линейного, параболического и кубического

трендов при периоде упреждения $L = 16, 13, 23$ соответственно. Поэтому оправдано использование для экстраполяции временного ряда 24 значений суточного электропотребления.

ВЫВОДЫ

1. Если начальные условия, на основании которых разрабатывается прогноз, верны, то возможно использовать в качестве начального значения первое значение базы прогноза. Если уверенности в начальных данных нет, то лучше использовать в качестве начального значения среднее арифметическое значения предыдущей базы прогноза.

2. Проверена гипотеза целесообразности использования в качестве постоянной сглаживания чисел в интервале $0 < \alpha < 2$. Результаты вычислений свидетельствуют о том, что оптимальное значение α для временных рядов электропотребления химических цехов и других объектов химического предприятия находится в интервале $0,1 < \alpha < 0,2$. Хотя проверка всех значений постоянной сглаживания на интервале $0 < \alpha < 2$ гарантирует правильность определения постоянной сглаживания.

3. Показано практически и доказано теоретически, что средняя абсолютная процентная ошибка прямо пропорционально зависит от коэффициента вариации значений временного ряда электропотребления. Указаны диапазоны изменения значений средней абсолютной процентной ошибки $MAPE$ в зависимости от значений коэффициента вариации для конкретных процессов электропотребления с различным разбросом данных.

4. Установлено, что метод экспоненциального сглаживания обеспечивает высокую точность восстановления пропущенных одиночных данных электропотребления на химических производствах.

5. Оправдано использование для экстраполяции временного ряда почасового электропотребления 24 почасовых значений электропотребления (24 значения в сутки).

СПИСОК ЛИТЕРАТУРЫ

1. Бокс Дж., Дженкинс Г. Анализ временных рядов. Прогноз и управление. – М.: Мир, 1974. – 520 с.
2. Лукашин Ю.Г. Адаптивные методы краткосрочного прогнозирования. – М.: Экономика, 1989. – 214 с.
3. Светушков С.Г., Светушков И.С. Методы социально-экономического прогнозирования: учебник для вузов. Том II. – СПб.: Изд-во СПбГУЭФ, 2010. – 103 с.
4. Четыркин Е.М. Статистические методы прогнозирования. – М.: Финансы и статистика, 1979. – 199 с.
5. Бэнн Д.В., Фармер Е.Д. Сравнительные модели прогнозирования электрической нагрузки. – М.: Энергоатомиздат, 1987. – 200 с.
6. Праховник А.В., Розен В.П., Дегтярев В.В. Энергосберегающие режимы электроснабжения горнодобывающих предприятий. – М.: Недра, 1985. – 232 с.
7. Гордеев В.И., Васильев И.Е., Шуцкий В.И. Управление электропотреблением и его прогнозирование. – Ростов-на-Дону: Изд-во Ростовского ун-та, 1991. – 104 с.

8. Гурский С.К. Адаптивное прогнозирование временных рядов в электроэнергетике. – Мн.: Наука и техника, 1983. – 271 с.

9. Светушков С.Г. Запредельные случаи метода Брауна // Экономические науки: Ученые записки УлГУ. – Ульяновск: Изд-во СВНЦ, 1997. – Вып. 2. – Часть 1.

10. Лук'яненко И.Г., Красникова Л.И. Эконометрика: Пособие. – К.: Товариство "Знання", КОО, 1998. – 494 с.

REFERENCES: 1. Boks Dzh., Dzhenkins G. *Analiz vremennykh riadov. Prognoz i upravlenie* [Time series analysis. Forecast and Control]. Moscow, Mir Publ., 1974. 520 p. 2. Lukashin Yu.G. *Adaptivnye metody kratkosrochnogo prognozirovaniia* [Adaptive methods of short-term forecasting]. Moscow., Ekonomika Publ., 1989. 214 p. 3. Svetun'kov S.G., Svetun'kov I.S. *Metody sotsial'no-ekonomicheskogo prognozirovaniia: uchebnik dlia vuzov. Tom II* [Methods of social and economic forecasting: a textbook for high schools. Vol. II]. St. Petersburg, St. Petersburg State University of Economics and Finance Publ., 2010. 103 p. 4. Chetyrkin E.M. *Statisticheskie metody prognozirovaniia* [Statistical methods for forecasting]. Moscow, Finansy i statistika Publ., 1979. 199 p. 5. Benn D.V., Farmer E.D. *Sravnitel'nye modeli prognozirovaniia elektricheskoi nagruzki* [Comparative models of electric load forecasting]. Moscow, Energoatomizdat Publ., 1987. 200 p. 6. Prakhovnik A.V., Rozen V.P., Degtiarev V.V. *Energoberegaiushchie rezhimy elektrosnabzheniia gornodobyvaushchikh predpriatii* [Power-saving modes of electricity for mining companies]. Moscow, Nedra Publ., 1985. 232 p. 7. Gordeev V.I., Vasil'ev I.E., Shutskii V.I. *Upravlenie elektropotrebleniem i ego prognozirovanie* [Energy management and forecasting]. Rostov-on-Don, Rostov State University Publ., 1991. 104 p. 8. Gurskii S.K. *Adaptivnoe prognozirovanie vremennykh riadov v elektroenergetike* [Adaptive prediction of time series in electric power industry]. Minsk, Nauka i tekhnika Publ., 1983. 271 p. 9. Svetun'kov S.G. *Zapredel'nye sluchai metoda Brauna* [Outrageous cases of the method of Brown]. *Ekonomicheskie nauki: Uchenye zapiski UIGU – Economic sciences: Scientific notes of Ulyanovsk State University*, 1997, no.2, part 1. 10. Luk'ianenko I.G., Krasnikova L.I. *Ekonometrika: Pidruchnik* [Econometrics: Textbook]. Kyiv, Tovarisstvo "Znannia", KOO Publ., 1998. 494 p.

Поступила (received) 11.04.2014

*Бедерак Ярослав Семенович, инженер,
Публичное акционерное общество "АЗОТ",
18014, Черкассы, ул. Первомайская, 72,
тел/phone +38 047 2392979, e-mail: ei@uch.net*

*Ya.S. Bederak
PJSC "AZOT"
72, Pervomayskaya Str., Cherkassy, 18014, Ukraine*

An exponential smoothing method application to restoring lost data on electric power technical record-keeping in industrial enterprises.

The paper studies in detail a problem of restoring single data of electric power technical record-keeping systems by means of exponential smoothing. A hypothesis of efficient finding the optimal value of smoothing constant α in the range of $0 < \alpha < 2$ is tested. Variation ranges for the mean absolute percentage error of the forecast are specified versus values of the time series variation coefficient for specific electricity consumption processes with different data spread.

Key words – exponential smoothing, lost data restore, mean absolute percentage error, variation.