

**СИСТЕМЫ ДЛЯ НАУЧНЫХ  
И ПРОМЫШЛЕННЫХ ПРИМЕНЕНИЙ**

УДК 519.216.8

**МУЛЬТИПЛИКАТИВНАЯ МОДЕЛЬ  
СЕЗОННОГО ЭНЕРГОПОТРЕБЛЕНИЯ ПРЕДПРИЯТИЙ**

**Р. Р. Ахметьянов, Л. А. Делегодина, Н. П. Копылова,  
Б. Н. Луценко, Г. М. Собстель, Г. П. Чейдо**

*Конструкторско-технологический институт вычислительной техники СО РАН,  
г. Новосибирск  
E-mail: lad@kti.nsc.ru*

Рассмотрен комплекс преобразований исходного процесса с нестационарными сезонными колебаниями при его описании мультипликативной моделью типа авторегрессии и проинтегрированного скользящего среднего с целью дальнейшего прогнозирования. Проводимые преобразования завершаются приведением исходного процесса к стационарному виду, необходимому для последующей идентификации модели.

**Постановка задачи.** Ресурсопотребление предприятий обычно обладает суточной, недельной и более продолжительной нестационарной сезонностью. Описание, обработка и прогнозирование этих процессов с помощью полиномиальных и гармонических функций малоэффективны, поскольку потребовалось бы чрезмерное количество параметров, характеризующих данные. Применение для описания этих стохастических процессов нестационарных параметрических моделей авторегрессии (АР) – скользящего среднего (СС) позволяет компактно, сравнительно небольшим числом параметров охарактеризовать их поведение и использовать для последующей пролонгации.

Целью предлагаемой работы является создание математического инструмента для исследования процессов с сезонными колебаниями и их описание с помощью однородных нестационарных стохастических моделей авторегрессии – проинтегрированного скользящего среднего (АРПСС). Описание и исследование моделей АРСС и АРПСС дано в [1–5].

Такие процессы широко представлены не только в производственной и коммерческой деятельности человека, но и в природе как живой, так и неживой.

В данной работе основное внимание уделяется вопросам практического использования модели.

**Мультипликативная сезонная модель АРПСС.** Обозначим сам случайный процесс, возможно подвергшийся предварительным преобразованиям, через  $Z_t$ ,  $t \in (1, N_Z)$ . Отсчеты  $Z_t$  эквидистантны, и в качестве временного дискрета принимается периодичность его регистрации. Если процесс стационарен и его среднее равно  $\mu$ , будем далее работать с центрированным процессом

$$\tilde{Z}_t = Z_t - E(Z_t) = Z_t - \mu.$$

При нестационарности из процесса можно извлечь постоянную константу  $m$ , по-прежнему обозначая результат

$$\tilde{Z}_t = Z_t - m.$$

Оператор сдвига на один такт назад обозначим через В (back):  $BZ_t = Z_{t-1}$ , а вперед – через F (forward):  $FZ_t = Z_{t+1}$ . Оператор взятия разности  $\nabla$  представим в виде

$$\nabla Z_t = (1 - B)Z_t = Z_t - Z_{t-1}.$$

Оператором, обратным  $\nabla$ , будет оператор суммирования (дискретный аналог интегрирования).

Мультипликативная сезонная модель задается своей структурой, в которую входят:  $k$  – количество интервалов сезонности, включая для единообразия и единичный,  $i = 1, \dots, k$ ;  $S_i$  – значения интервалов сезонности;  $p_i$  – количество параметров АР для каждого интервала сезонности;  $q_i$  – количество параметров СС для каждого интервала сезонности;  $d_i$  – количество сезонных разностей для каждого интервала сезонностей.

При этом оператор АР для периода сезонности  $S_i$  представим в виде

$$\Phi_{p_i}(B^{S_i}) = 1 - \sum_{j=1}^{p_i} \phi_{ij} B^{jS_i}, \quad (1)$$

а оператор СС – в виде

$$\Theta_{q_i}(B^{S_i}) = 1 - \sum_{j=1}^{q_i} \theta_{ij} B^{jS_i}. \quad (2)$$

В этих обозначениях сравнительно компактно мультипликативная модель АРПСС будет иметь вид

$$\prod_{i=1}^k \Phi_{p_i}(B^{S_i}) \nabla_{S_i}^{d_i} \tilde{Z}_t = \prod_{i=1}^k \Theta_{q_i}(B^{S_i}) a_t, \quad (3)$$

где  $a_t$  представляет белый шум, т. е. последовательность независимых одинаково распределенных случайных величин с нулевым средним и постоянной дисперсией  $\sigma^2$ :

$$E(a_t) = 0, \quad E(a_t, a_{t+\tau}) = \sigma^2 \rho(\tau) = \begin{cases} \sigma^2 & \text{при } \tau = 0, \\ 0 & \text{при } \tau \neq 0. \end{cases}$$

Структура модели позволяет описывать сезонные колебания, используя как стохастические параметры, так и детерминированные. При отличии параметров сезонной разности  $d_i$  от нуля процесс может содержать детерминированную компоненту произвольного вида с периодом, равным соответствующему интервалу сезонности. При этом наличие  $d_i > 0$  должно сопровождаться надлежащим количеством параметров СС, возникающих в процессе взятия сезонных разностей. Соответствующих данной сезонности параметров АР может и не быть.

Если обозначить результат взятия сезонных разностей процесса  $\tilde{Z}_t$  через  $w_t$ :

$$\prod_{i=1}^k \nabla_{S_i}^{d_i} \tilde{Z}_t = w_t, \quad t \in \left( \sum_{i=1}^k S_i d_i + 1, N_Z \right), \quad (4)$$

то мультипликативная модель уже процесса АРСС примет вид

$$\prod_{i=1}^k \Phi_{p_i}(B^{S_i}) w_t = \prod_{i=1}^k \Theta_{q_i}(B^{S_i}) a_t.$$

Процесс  $w_t$  будет стационарным в том случае, если корни характеристического уравнения

$$\prod_{i=1}^k \Phi_{p_i}(x^{S_i}) = 0,$$

где  $x$  – комплексная переменная, расположены вне единичного круга в комплексной области (характеристическое уравнение получается из оператора АР заменой оператора сдвига  $B$  комплексной переменной  $x$ ) [1, 2]. При этом

$$w_t = \left[ \prod_{i=1}^k \Phi_{p_i}(B^{S_i}) \right]^{-1} \prod_{i=1}^k \Theta_{q_i}(B^{S_i}) a_t = \sum_{i=0}^{\infty} \psi_i B^i a_t, \quad \psi_0 = 1,$$

т. е.  $w_t$  можно представить процессом СС в виде бесконечного сходящегося ряда. Это и есть утверждение теоремы Вольда о декомпозиции стационарного процесса [1, 2].

Удовлетворение условия стационарности процесса АРСС обеспечивает одновременно и эргодичность процесса [6]. На практике это чрезвычайно важно, так как зачастую имеется лишь фрагмент единственной реализации стохастического процесса и параметры его распределения неизвестны. Свойство же эргодичности позволяет и в этой ситуации делать обоснованные выводы о точности получаемых оценок параметров и результатов прогноза.

Важным свойством процесса является также его обратимость [1, 2]. Процесс  $w_t$  будет обратим, если корни характеристического уравнения

$$\prod_{i=1}^k \Theta_{q_i}(x^{S_i}) = 0$$

(как и ранее,  $x$  – комплексная переменная) расположены вне единичного кру-

га в комплексной области. Свойство обратимости заключается в возможности представления

$$a_t = \left[ \prod_{i=1}^k \Theta_{q_i}(B^{S_i}) \right]^{-1} \prod_{i=1}^k \Phi_{p_i}(B^{S_i}) w_t = \sum_{i=0}^{\infty} \pi_i B^i w_t, \quad \pi_0 = 1,$$

т. е. возможности выражения белого шума процессом АР с бесконечным числом членов в виде сходящегося ряда.

Оба свойства – стационарность и обратимость – важны при построении модели для реального процесса.

Мультипликативная модель позволяет экономно описывать достаточно сложные последовательности. Однако ее компактность оплачивается некоторым сужением класса описываемых процессов. При итерационном уточнении параметров модели удобно объединить их все в единый вектор  $\beta$ :

$$\beta = (\varphi_{1,1}, \dots, \varphi_{1,p_1}, \dots, \varphi_{k,1}, \dots, \varphi_{k,p_k}, \theta_{1,1}, \dots, \theta_{1,q_1}, \dots, \theta_{k,1}, \dots, \theta_{k,q_k}), \quad (5)$$

размерность которого

$$n_\beta = \sum_{i=1}^k (p_i + q_i),$$

а при использовании одношагового прогнозирования, основанного на разностной модели, удобнее ввести ее избыточные параметры, перемножив операторы в левой и соответственно в правой части выражения (3). Количество этих избыточных параметров для приведенных оператора АР

$$n_p = \prod_{i=1}^k (p_i + 1) - 1$$

и оператора СС

$$n_q = \prod_{i=1}^k (q_i + 1) - 1.$$

Приведенные параметры ранжируют по степеням В. При этом, если их обозначить соответственно через  $\varphi_i^*$  и  $\theta_i^*$ , получим

$$\prod_{i=1}^k \Phi_{p_i}(B^{S_i}) = \sum_{i=0}^{n_p} \varphi_i^* B^{l_i}, \quad \varphi_0^* = 1, \quad l_0 = 0, \quad (6)$$

$$\prod_{i=1}^k \Theta_{q_i}(B^{S_i}) = \sum_{i=0}^{n_q} \theta_i^* B^{m_i}, \quad \theta_0^* = 1, \quad m_0 = 0, \quad (7)$$

и исходная модель с учетом (6), (7) примет вид

$$\sum_{i=0}^{n_p} \varphi_i^* B^{l_i} w_t = \sum_{i=0}^{n_q} \theta_i^* B^{m_i} a_t, \quad (8)$$

похожий на модель АРСС, только теперь количество параметров избыточно, а степени операторов сдвига  $l_i$  и  $m_i$  следуют с неравномерным шагом.

Естественно, можно использовать и аддитивную сезонную модель. Однако при тех же возможностях описания процесса потребуется большее количество параметров.

К дополнительным достоинствам мультипликативной модели (3) относится упрощение проверки условий стационарности и обратимости, так как соответствующие характеристические уравнения распадаются на сомножители, корни которых можно искать независимо.

Конечно, реально-значимые  $p_i$ ,  $q_i$  и  $d_i$  будут невелики.

Пример использования сезонной модели АРСС с несколькими интервалами сезонности и алгоритмом бутстреп (методом «складного ножа») приведен в [7, гл. 15].

**Схема решения задачи прогнозирования.** Рассматриваемая модель стохастического процесса хотя и обладает определенной представительностью, все же занимает достаточно узкий класс в обширном поле нестационарных процессов. И при визуальном просмотре исходных данных прежде всего следует оценить целесообразность ее привлечения. В случае положительного решения исходные данные необходимо привести к виду, соответствующему требованиям модели АРПСС.

После выбора прогнозной базы может потребоваться функциональное преобразование и, возможно, извлечение тренда.

Предварительная обработка включает: заполнение лакун и по необходимости аномалий с использованием непараметрических методов, опираясь на предшествующие данные; выявление и исключение выбросов путем анализа разностей между фактическими и медианно сглаженными данными; возможно, корректировку данных праздничных дней и приведение их к рабочим.

Применение оператора взятия сезонных разностей к преобразованным таким образом данным переводит процесс АРПСС в АРСС. Следующим этапом является идентификация модели по стационарному процессу АРСС. Задается структура модели, для которой осуществляются идентификация, уточнение параметров и выбор единственного представителя. Для нее проводится проверка адекватности, т. е. соответствия описываемому процессу по различным критериям. При неудовлетворительном результате проверки изменяется структура модели и вычисления повторяются. При успешном результате проверки строится прогноз вместе с доверительными интервалами на заданный интервал пролонгации с помощью уже нестационарной модели АРПСС. Если исходный процесс был функционально преобразован и из него извлекался тренд, результаты прогноза стохастического процесса и тренда объединяются и вместе с доверительными интервалами подвергаются обратному функциональному преобразованию.

**Подготовка исходных данных для использования модели АРПСС** заключается в приведении их к виду, соответствующему требованиям модели, и состоит из:

- 1) визуального просмотра всего массива,
- 2) заполнения пробелов (лакун),
- 3) возможной корректировки дополнительных аномалий,
- 4) выявления и отбраковки выбросов,
- 5) обработки праздничных дней,
- 6) возможно, ступенчатой корректировки недельных данных.

Рассмотрим подробнее каждый из этих этапов.

*Визуальный просмотр* должен выявить пригодность анализируемых данных для описания моделью АРПСС. Это предварительное решение, отвергающее заведомо неприемлемые для данной модели процессы. Кроме того, должен быть оценен объем и характер требуемой предварительной обработки. Возможно необходимое функциональное преобразование или извлечение тренда. На этом этапе осуществляется выбор прогнозной базы, по которой будет строиться модель. Интервал пролонгации задается заранее. Вообще, целесообразно каждый этап предварительной обработки завершать визуальным просмотром результатов, выделяя на графике коррективы другим цветом (на фоне исходных данных), возможно, корректируя последующие действия.

*Пробелы в данных* могут занимать от нескольких часов (при почасовой регистрации) до нескольких суток. Однако при этом обычно производится непрерывное накопление потребляемого ресурса, что позволяет, по крайней мере, в целом не утрачивать итоговой информации и согласовывать с ней результаты заполнения лакун.

Для заполнения пробелов можно прибегнуть к непараметрическому медианному прогнозированию по предшествующим данным недельного или, возможно, трехнедельного периода. Обозначим символом  $MD(Z_t)$  медиану массива из трех соседних точек:

$$MD(Z_t) = \text{med}(Z_{t-1}, Z_t, Z_{t+1}).$$

Тогда медианный прогноз для момента  $t$  по данным недельной давности будет

$$Z_{p1}(t) = MD(Z_{t-S}),$$

где  $S$  – интервал недельной сезонности, выраженный в дискретах регистрации  $Z_t$ . При трехнедельном медианном прогнозе

$$Z_{p3}(t) = \text{med}[MD(Z_{t-3S}), MD(Z_{t-2S}), MD(Z_{t-S})].$$

Заполнив этими прогнозными данными все лакуны, можно таким же образом заменить и *аномальные фрагменты данных*. Тогда замена рассматривается как предварительная. После построения для этой прогнозной базы модели АРПСС можно осуществить пошаговое прогнозирование данных для лакун и аномалий и заменить ими, возможно, менее точные предшествующие медианные прогнозы и затем скорректировать параметры модели по уточненным данным.

*Выявление и отбраковка выбросов* также осуществляется с использованием скользящего медианного сглаживания по трем точкам по разности исходных и сглаженных значений:

$$\Delta Z_t = Z_t - MD(Z_t), \quad t \in (1, N_Z).$$

Если

$$|\Delta Z_t| \geq \text{lev},$$

данное  $Z_t$  полагается выбросом и заменяется медианным сглаженным значением  $MD(Z_t)$ . Для оценки порога отбраковки выбросов  $\text{lev}$  необходимо оце-

нить дисперсию  $S_{\Delta}^2$  массива разностей  $\Delta Z_t$ , отличных от нуля, и задать относительный уровень выбросов  $k$  в единицах СКО для  $\Delta Z_t$ :

$$\text{lev} = kS_{\Delta}.$$

Оценка  $S_{\Delta}$  производится по следующему алгоритму. Из массива  $\Delta Z_t$ ,  $t \in (2, N_Z - 1)$ , отбираются  $n_{\Delta}$ , отличные от нуля, так как большинство (или, по крайней мере, значительная часть)  $\Delta Z_t$  будут нулевыми. Полученный таким образом массив  $\Delta Z_t^*$  цензурируется, т. е. из него исключается определенный процент ( $\text{cenz}/2$ ) максимальных и ( $\text{cenz}/2$ ) минимальных значений  $\Delta Z_t^*$ . Выбор уровня цензурирования  $\text{cenz}$  осуществляется таким образом, чтобы включить по возможности в число удаляемых все выбросы и, естественно, часть достоверных  $\Delta Z_t^*$ . После цензурирования останется

$$n_{\Delta 1} = n_{\Delta} - n_{\Delta} \text{cenz}/100$$

данных, если  $\text{cenz}$  выражен в процентах. Обозначим эти  $n_{\Delta 1}$  значений через  $\Delta Z_i^{**}$ , перегруппировав их в слитный массив, сохранив информацию об их положении в исходном массиве  $\Delta Z_t$  и оценив его среднее значение и дисперсию соответственно

$$m_1 = \frac{1}{n_{\Delta 1}} \sum_{i=1}^{n_{\Delta 1}} \Delta Z_i^{**}, \quad S_1^2 = \frac{1}{n_{\Delta 1}} \sum_{i=1}^{n_{\Delta 1}} (\Delta Z_i^{**})^2 - m_1^2.$$

Поскольку при цензурировании может быть исключена часть достоверных данных, производится попытка их реабилитации и возврата в основную выборку. Здесь приходится постулировать некоторые свойства распределения вероятностей достоверных значений  $\Delta Z_t^*$  и выбросов. Под выбросами будем понимать резко выделяющиеся точки на фоне окружения наблюдения. В данной работе не рассматривается проблема возможной вероятностной или детерминированной природы выбросов. Анализируя выборочное распределение  $\Delta Z_t^*$ , можно наблюдать характер его «хвостов». Если они отделены от основной массы распределения, то это вероятный признак выбросов. Рассматриваемый алгоритм отбраковки ориентирован в первую очередь на такого рода выбросы. Если распределения  $\Delta Z_t^*$  и выбросов перекрываются, проникая друг в друга, потребуется иной, более сложный, алгоритм отбраковки.

Будем полагать, что основная часть распределения  $\Delta Z_t^*$  лежит в пределах  $\pm 3\sigma$ . Истинное значение дисперсии  $\sigma^2$  неизвестно, в нашем распоряжении пока лишь ее предварительная оценка по усеченной выборке  $S_1^2$ . Проверим все отброшенные при цензурировании точки на возможность их реабилитации. Если

$$|\Delta Z_i^* - m_1| < 3S_1,$$

то эти точки возвращаются к массиву точек  $\Delta Z_i^{**}$ , после чего вычисляются оценки  $m_2$  и  $S_2^2$  по расширенному таким образом массиву  $\Delta Z_i^{**}$ . Такая последовательная процедура реабилитации отброшенных вначале  $\Delta Z_t^*$  и корректи-

ровка  $m_i$  и  $S_i^2$  продолжают до тех пор, пока они не перестанут изменяться. В данном подходе есть опасность чрезмерно завышенного уровня цензурирования  $cen_z$ , при котором отвергнутые достоверные точки уже не будут возвращены обратно и квалифицируются как выбросы. Выбор же слишком малого значения  $cen_z$  таит опасность оставить выбросы в выборке, полагаемой достоверной, что повлечет завышение и дальнейший чрезмерный рост оценок  $S_i^2$ , что, в свою очередь, приведет в конечном счете к признанию всех выбросов достоверными данными. Необходимо пройти между Сциллой и Харбдой. Лучше попробовать несколько вариантов выбора  $cen_z$  и проконтролировать отбраковываемые данные визуально на графике. Процедура выбора уровня отбраковки выбросов иллюстрирует рис. 1, а выявленные выбросы, заполненные в дальнейшем медианно сглаженными значениями процесса, представлены на рис. 2. Массив  $Z_i$  из прогнозной базы не особенно показателен в плане наличия выбросов (их всего четыре с относительными величинами 6,27, 5,88, 4,09, 4,35, выраженными в СКО при заданном пороге отбраковки 4 СКО), и ввиду их малости выбросы можно было бы и не отбраковывать.

*Обработка праздничных дней.* Описание процесса моделью АРПСС предполагает регулярную сезонную периодичность без нарушения ритма. Праздничные дни и сопутствующие им возможные перенесения рабочих дней или введение компенсационных нерабочих нарушает циклическую ритмичность. Чтобы подготовить прогнозную базу процесса для адекватного описания моделью АРПСС, желательно восстановить эту ритмичность. Речь идет только о подготовке исходных данных для прогноза, а не об изменении самих исходных данных. Процесс обработки праздничных дней сводится к искусственному заполнению данных праздничных суток медианными прогнозными данными по предшествующей неделе наблюдений и возможному перемещению данных рабочих и нерабочих дней на их исконные места.

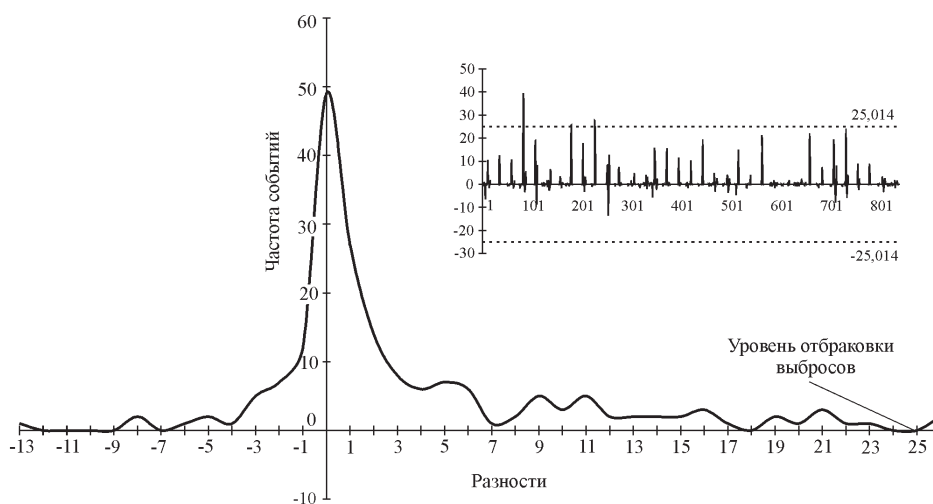


Рис. 1. Гистограмма распределения отличных от нуля разностей между исходными и медианно сглаженными данными. На фрагменте показаны сами разности с порогами отбраковки выбросов  $\Delta Z_i$  (сплошная линия) и  $lev$  (пунктирная линия)



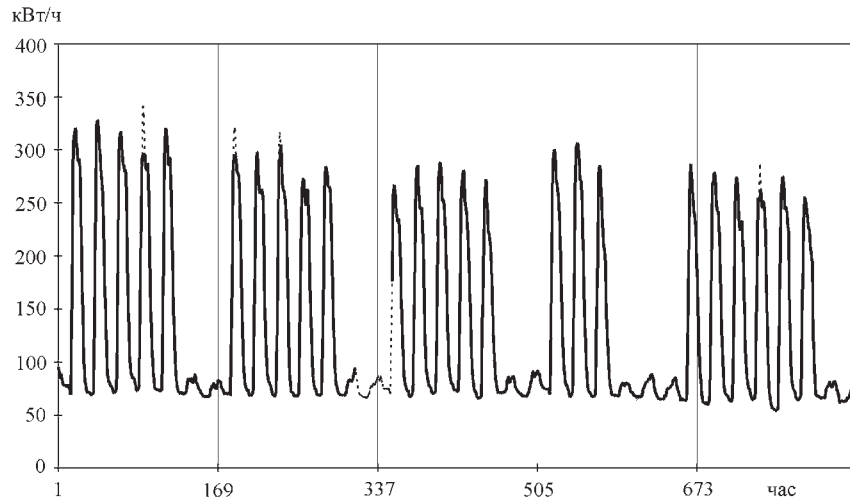


Рис. 2. Заполнение лакун и устранение выбросов: исходные данные (сплошная линия), выбросы и скорректированная лакуна (пунктирная линия)

Праздник может прийти на любой из семи дней недели, что порождает семь возможных ситуаций. Если он приходится на понедельник, среду или пятницу, соответствующие ему почасовые значения  $Z$  праздничных суток заменяются медианно сглаженными по данным предшествующей недели:

$$\hat{Z}_t = \text{MD}(Z_{t-s}).$$

Если праздник приходится на вторник или четверг, он, как правило, сопровождается переносом рабочих понедельника и пятницы соответственно на субботу и воскресенье. А праздник, приходящийся на субботу или воскресенье, обычно сопровождается компенсационными нерабочими днями соответственно в пятницу или понедельник. Только этот алгоритм выдерживается не всегда.

Чтобы идентифицировать эти семь ситуаций, располагая лишь массивом  $Z_t$  без календарной привязки, осуществляется вначале отождествление суточных интервалов  $Z_t$  с рабочими или нерабочими сутками. Для этого предварительно строится массив суточных потреблений ресурса  $\text{St}(mz)$ , где  $mz = N_Z/S_d$ , а  $S_d$  – суточная сезонность, выраженная в дискретах регистрации  $Z_t$ . Если начало массива  $Z_t$  совмещено с началом суток, то

$$\text{St}(i) = \sum_{j=1}^{S_d} Z[(i-1)S_d + j].$$

Далее вводится индикаторная функция, принимающая значение 1 для рабочих и 0 для нерабочих дней. Аргументом индикаторной функции  $\text{ind}_7$  является скользящий по  $\text{St}$  вектор  $V_i(7)$ :

$$V_i(j) = \text{St}(i-4+j), \quad j=1, \dots, 7.$$

Индикаторная функция

$$\text{ind7}(i) = \begin{cases} 1 & \text{при } \text{St}(i) > \text{por}, \\ 0 & \text{иначе,} \end{cases} \quad \text{por} = \frac{1}{2}[\min V_i(j) + \max V_i(j)], \quad j = 1, \dots, 7,$$

т. е. порог регистрации характера суток адаптивно меняется при скольжении индикаторной функции вдоль массива  $\text{St}(mz)$ . С ее помощью строится индикаторный массив для прогнозной базы

$$\text{ind}(i) = \text{ind7}(i), \quad i = 4, \dots, mz - 3.$$

При семикомпонентном скользящем интервале возникает возможность ложной идентификации рабочих суток как нерабочих на стыке недель при резком изменении уровня сигнала  $Z_t$ . Чтобы исправить эту возможную ошибочную идентификацию, в индикаторном массиве  $\text{ind}(mz)$  производится просмотр триад из последовательно идентифицированных нерабочих дней. Их анализ осуществляется с помощью индикаторной функции, аналогичной  $\text{ind7}(i)$ , но со скользящим интервалом из пяти суток, т. е.  $\text{ind5}(i)$ . Теперь при анализе характера пятницы или понедельника в скользящий интервал не будут попадать данные рабочих дней соседних недель. С помощью  $\text{ind5}(i)$  индикаторный массив  $\text{ind}(mz)$  корректируется в районе выделенных триад. Если

$$\text{ind}(k) = 0, \quad k = i - 1, i, i + 1,$$

то, возможно, скорректированные значения будут

$$\text{ind}(k) = \text{ind5}(k), \quad k = i - 1, i, i + 1.$$

Идентификация перечисленных выше семи ситуаций, связанных с праздничными днями, осуществляется сопоставлением с данными предшествующей недели, которые уже предполагаются прошедшими предварительную обработку с восстановлением сезонной периодичности. После идентификации событий праздничные или, возможно, компенсационные сутки заполняются медианным прогнозом по предшествующей неделе, и если имели место перемещения рабочих и нерабочих дней, они возвращаются на свои места. Результаты обработки предпраздничного дня, приходящегося на четверг четвертой недели прогнозной базы, приведены на рис. 3. При этом рабочий день пятницы был перенесен на воскресенье, а нерабочими были четверг–суббота. Обработка восстанавливает регулярную цикличность прогнозируемого процесса путем заполнения суточных данных предпраздничного четверга результатами медианного прогноза по предыдущей неделе и взаимным обменом данных пятницы и воскресенья.

*Ступенчатая корректировка недельных данных.* После заполнения лакун, устранения аномалий и обработки праздничных дней может оказаться целесообразным проведение ступенчатой коррекции недельных уровней данных прогнозной базы с привязкой к последней предпрогнозной неделе по методу наименьших квадратов (МНК). Коррекция осуществляется недельным изменением масштабных коэффициентов  $\gamma_i$ ,  $i \in (1, k_z), k_z = N_z / S$ .

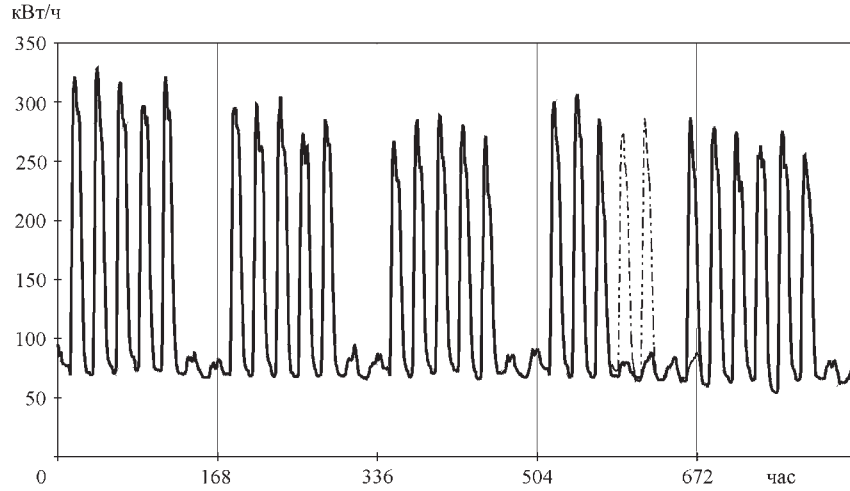


Рис. 3. Обработка праздничных дней с восстановлением на прогнозной базе регулярной цикличности: исходные данные (сплошная линия), восстановленная информация по праздничному и выходному дням (штрихпунктирная линия)

Для нахождения коэффициентов  $\gamma_i$  предварительно вычисляются элементы массива суточных потреблений ресурса  $St(i)$ ,  $i \in (1, m_Z)$ ,  $m_Z = N_Z / S_d$ :

$$St(i) = \sum_{j=1}^{S_d} Z[(i-1)S_d + j], \quad i \in (1, m_Z).$$

Массив корректирующих множителей

$$\gamma_i = \frac{\sum_{j=1}^{S_d} St[(i-1)S_d + j] St[(k_Z - 1)S_d + j]}{\sum_{j=1}^{S_d} (St[(i-1)S_d + j])^2}, \quad i \in (1, k_Z - 1).$$

Скорректированные данные

$$Z_i^*[(i-1)S_d + j] = \gamma_i Z[(i-1)S_d + j], \quad i \in (1, k_Z - 1), \quad j \in (1, S_d),$$

при этом данные последней недели прогнозной базы остаются неизменными:

$$Z_i^*[(k_Z - 1)S_d + j] = \gamma_i Z[(k_Z - 1)S_d + j], \quad j \in (1, S_d).$$

Результаты ступенчатой корректировки уровней недельных данных приведены на рис. 4.

Все представленные выше преобразования исходных данных имеют целью создание благоприятных условий для описания их моделью АРПСС. И завершаются они применением оператора взятия сезонных разностей,

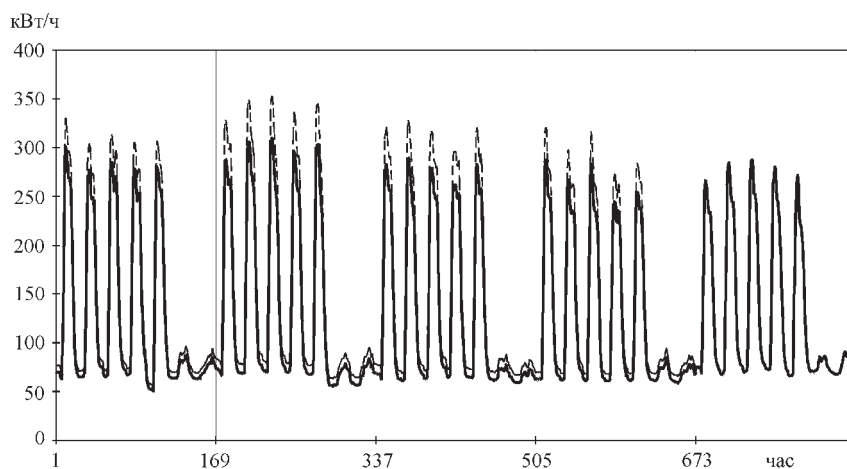


Рис. 4. Ступенчатое выравнивание недельных уровней процесса с привязкой к последней неделе прогнозной базы по МНК: исходные данные (штриховая линия), скорректированная по последней неделе база (сплошная линия)

если таковой предусмотрен структурой модели. При этом следует иметь в виду, что оператор должен сопровождаться наличием соответствующих параметров СС, ибо если даже исходный сигнал их не содержал, то при трансформации данных с применением оператора взятия сезонных разностей они будут порождены. Кроме того, сезонные разности освобождают процесс от детерминированных компонент с соответствующими сезонными периодичностями. Эти детерминированные компоненты вновь проявятся при прогнозе и будут включены в прогнозируемый стохастический процесс.

Оператор взятия сезонных разностей предположительно переводит процесс в стационарное состояние, описываемое уже моделью АРСС, идентификация и оценка параметров которой будут рассмотрены в последующих работах.

**Заключение.** Энергопотреблению предприятий присуща цикличность. Для компактного описания таких процессов с циклическими сезонными колебаниями предлагается использовать мультипликативную модель АРПСС. В данной работе приведено описание структуры и двух форм представления этой модели: первая – мультипликативная с минимальным набором базовых параметров, вторая – аддитивная с расширенным множеством избыточных параметров. Именно аддитивная форма модели будет в основном использоваться при идентификации и прогнозировании процесса с помощью разностного уравнения. Избыточность ее параметров обуславливает расширение количества вовлекаемых в прогнозирование данных и соответственно повышает точность прогноза. В этом достоинство рассматриваемой модели.

Результатом предварительной обработки является приведение исходного процесса к стационарному виду, позволяющему описывать его моделью АРСС.

#### СПИСОК ЛИТЕРАТУРЫ

1. Андерсон Т. Статистический анализ временных рядов. М.: Мир, 1976.

2. **Бокс Дж., Дженкинс Г.** Анализ временных рядов (Прогноз и управление). М.: Мир, 1974. Вып. 1.
3. **Кендалл М. Дж., Стьюарт А.** Многомерный статистический анализ и временные ряды. М.: Наука, 1976.
4. **Кендэл М.** Временные ряды. М.: Финансы и статистика, 1981.
5. **Справочник** по прикладной статистике /Под ред. Э. Ллойда, У. Ледермана. М.: Финансы и статистика, 1990. Т. 2.
6. **Канторович Г. Г.** Анализ временных рядов // Эконом. журн. Высш. шк. экономики. 2002. № 1–4; 2003, № 1.
7. **Бэкон Д. У., Брозкховен Л. Х.** Предсказание временных рядов // Статистические методы для ЭВМ /Под ред. К. Энслейна, Э. Релстона, Г. С. Уилфа. М.: Наука, 1986. С. 434.

*Поступила в редакцию 6 июня 2007 г.*

---