

ПОВЫШЕНИЕ КАЧЕСТВА КРАТКОСРОЧНОГО ПРОГНОЗИРОВАНИЯ ЭЛЕКТРОПОТРЕБЛЕНИЯ НА ОСНОВЕ УЧЕТА ДИНАМИКИ ФАКТОРА СЕЗОННОСТИ

IMPROVING OF THE POWER CONSUMPTION SHORT-RANGE FORECASTING ACCURACY BY TAKING INTO CONSIDERATION THE SEASONALITY FACTOR DYNAMICS

*A. Dulesov
A. Shilov*

Annotation

The short-range forecast power consumption forecasting problem with considering of "Day status" (workday or day off) factor (attribute) and seasonality associated with temperature changes is shown. The detailed analysis of outside air temperature factor influence on the process of power consumption is given and 3 seasons depending on the enabled/disabled building heating system: season without heating; heating season; off-season. The clustering procedure based on above-mentioned partition is made. The values obtained in the hourly power consumption) forecasting in the Republic of Khakassia (from the zero to the twenty-third hour) both in general and with a focus on the trends turning points: 8–9; 12–13; 17–18 hours with using of linear regression, moving averages and neural network methods are analyzed. Conclusions about the accuracy improving of the forecast values with applying of clustering procedure based on the "Season depending on heating" factor are made.

Keywords: short-range forecast; forecast error; clustering; season; linear regression; moving averages; neural networks.

*Дулесов Александр Сергеевич
ФГБОУ ВПО "Хакасский государственный
университет им. Н.Ф. Катанова, г. Абакан
Шилов Антон Владимирович
ФГБОУ ВПО "Хакасский государственный
университет им. Н.Ф. Катанова, г. Абакан*

Аннотация

Рассмотрена проблема краткосрочного прогнозирования электропотребления с учётом факторов (признаков), среди которых статус дня (будний или праздничный/выходной) и сезонность, связанная со сменной температурного режима. Более подробно проанализирован фактор воздействия температуры наружного воздуха на процесс электропотребления и представлены 3 сезона в зависимости от включения/отключения системы отопления зданий: сезон без отопления; отопительный сезон; межсезонье. Проведена процедура кластеризации согласно вышеуказанному разбиению. Проанализированы значения, полученные при прогнозировании почасового электропотребления в республике Хакасия с 0-го по 23-й час суток, как в целом по суткам, так и с акцентом на участках переломов трендов, а именно: 8–ми–9–ти; 12–ти–13–ти; 17–18–ти часов с использованием методов линейной регрессии, скользящей средней и нейронных сетей. Сделан вывод о повышении эффективности прогнозирования с использованием процедуры кластеризации по фактору "Сезон в зависимости от отопления".

Ключевые слова:

Краткосрочный прогноз; ошибка прогноза; кластеризация; сезон; линейная регрессия; скользящие средние; нейросетевые методы.

Введение

При краткосрочном прогнозировании электропотребления необходим учёт влияния различных факторов [3–4, 9], среди которых: принадлежность дня к будним (рабочим) или праздничным/выходным дням; сезонность в зависимости от температуры наружного воздуха, и, как следствие, включения/отключения системы отопления зданий. Не уделяя должного внимания к воздействию вышеперечисленных факторов, становится очевидным широкий диапазон разброса данных и неоднородность анализируемых значений. Такая проблема оказывает существенное влияние на величину прогнозной ошибки. Соответственно, необходим учёт факторов (признаков), воздействующих на динамику изменения величины электропотребления. Эффективным методом в данном случае

является кластеризация [1–2, 10], позволяющая разграничить массив используемых для прогноза данных на основе значений признаков так, чтобы значения в пределах кластера стали более близкими по величине, однородными.

Далее в работе проанализировано влияние температуры наружного воздуха на величину электропотребления.

Оценка влияния фактора сезонности

Как отмечено выше, на процесс краткосрочного прогнозирования электропотребления влияет фактор сезонности в зависимости от температуры наружного воздуха. В работах [3, 9] выделены 3 сезона, согласно выше перечисленному разбиению:

- ◆ *лето (сезон без отопления)* – зависимость между электропотреблением и температурой воздуха в целом является незначительной;
- ◆ *зима (отопительный сезон)* – температура наружного воздуха существенно влияет на величину электропотребления, связь между данными параметрами отражается простыми линейными моделями;
- ◆ *межсезонье (весеннее и осеннее)* – температура воздуха имеет значительное влияние на электропотребление в межсезонье весной, а также осенью при отключенном отоплении.

Статистические данные об электропотреблении были разбиты на кластеры, согласно выше перечисленным сезонам и введены следующие обозначения: массив значений электропотребления, с учётом влияющих на него факторов времени (Т) и фактора сезонности в зависимости от отопления (Н) $W(T, H)$. С применением процедуры кластеризации: $W_{hs}(T)$ – кластер, содержащий данные в пределах "отопительного сезона"; $W_{ho}(T)$ – сезон "без отопления" и $W_{so}(T)$ – "межсезонье".

Процедура прогнозирования почасового электропотребления, МВт*час, проведена как без учёта кластеризации, так и с использованием выше указанного разбиения. Временной лаг составил 3 месяца (6552 значения). В силу наличия специфических преимуществ использо-

ваны методы, представленные в работах [5–8, 11–12], а именно: линейная регрессия; скользящие средние и методы нейронных сетей. Мерой эффективности полученных прогнозных решений послужила средняя абсолютная ошибка аппроксимации (MAPE). Процедура прогнозирования выполнялась с помощью программного продукта статистического анализа данных IBM SPSS Statistics v. 20 и офисного программного средства MS Excel 2007.

Средние значения абсолютных ошибок прогнозов с учётом и без использования кластеризации по фактору "Сезон в зависимости от отопления" приведены в **табл. 1**.

Согласно результатам прогноза, приведённым в **табл. 1**, после проведения процедуры кластеризации по фактору сезонности с учетом отопления, средняя абсолютная ошибка уменьшилась. Наибольшую эффективность после проведения кластеризации показал метод взвешенной скользящей средней. До проведения процедуры кластеризации средняя ошибка прогноза равнялась 2,3%, с её использованием – 1,5% и 1,6% (с использованием данных в пределах кластеров "Межсезонье" и "Без отопления" соответственно). Детально проанализированы значения ошибок на наиболее проблемных часовых интервалах: 8-й-9-й, 12-й-13-й, 17-й-18-й (**табл. 2**). Данные зоны являются участками переломов трендов, что обусловлено наступлением времени обеденного перерыва,

Таблица 1.

Величины средних абсолютных ошибок прогнозных решений с учётом и без использования кластеризации.

№ п/п	Метод прогнозирования	Вид метода прогнозирования	Средняя абсолютная ошибка (MAPE, %)		
			$W(T, H)$	$W_{so}(T)$	$W_{ho}(T)$
1.	<i>Регрессия</i>	<i>Линейная</i>	2,69	1,86	1,76
2.	<i>Скользящая средняя</i>	<i>Простая</i>	2,07	1,81	1,53
		<i>Взвешенная</i>	2,34	1,54	1,55
		<i>Экспоненциальная</i>	2,6	2,31	2,31
3.	<i>Нейронные сети</i>	<i>Многослойный перцептрон</i>	2,07	1,22	1,73
		<i>Радиальная базисная функция</i>	2,8	1,92	1,62

Таблица 2.

Сравнение величин абсолютной ошибки прогноза, метод взвешенной скользящей средней, до и после проведения кластеризации.

Час	Значение абсолютной ошибки прогноза, % без кластеризации	Значение абсолютной ошибки прогноза, %, кластер "Межсезонье"	Значение абсолютной ошибки прогноза, %, кластер "Без отопления"
8	1,5	1,3	1,4
9	3,42	1,47	1,1
12	3,53	0,6	3,6
13	2,72	0,4	2,6
17	2,67	4,15	1,6
18	0,04	1,92	1,2

прихода и ухода с рабочего места сотрудников в организациях и на предприятиях.

Графики прогнозных и фактических значений на сутки вперёд с использованием и без использования процедуры кластеризации изображены на рис. 1–3.

Анализ графической интерпретации почасового электропотребления (рис. 1–3) позволил выявить эффективность кластеризации. При этом зафиксировано снижение величины прогнозной ошибки на 8–м–9–м, 12–м–13–м часовых интервалах. Согласно табл. 2, средняя ошибка по проблемным часам без применения кластерного подхода равнялась 2,3%. В случае прогнозирования с использованием данных, входящих в кластер $W_{so}(T)$, средняя ошибка снижена до 1,6%, кластер $W_{ho}(T)$ – до 1,9%. Однако на некоторых часовых промежутках значение ошибки остаётся по-прежнему существенным, а именно: 17–й–18–й часы – кластер "Межсезонье", 12–й–13–й часы – кластер "Без отопления" (табл. 2).

Заключение

Выполнен анализ причин низкого качества прогнозных решений почасового электропотребления, получены значения ошибок без применения кластерного подхода. Проведена процедура кластеризации согласно признаку "Сезон в зависимости от отопления", воздействующему на процесс потребления электроэнергии. Значения ошибок, полученных при прогнозировании с использованием ретроспективных данных в выделенных кластерах, указывают на эффективность процедуры кластеризации. Наименьшая величина средней абсолютной ошибки прогноза была достигнута с использованием взвешенной скользящей средней, после применения кластеризации равная 1,5% (кластер "Межсезонье") и 1,6% (кластер "Без отопления"). Тогда как без учёта кластерного подхода величина ошибки прогноза равнялась 2,3%.

Детализированные значения ошибок на проблемных зонах, представленные выше, в целом указывают на сни-

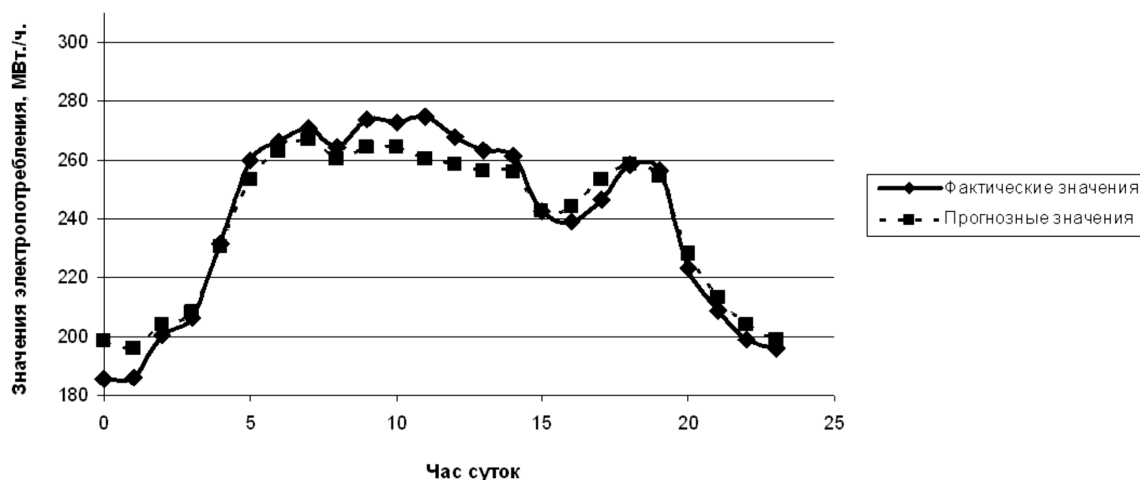


Рисунок 1. Графики фактических и прогнозных значений электропотребления в Хакасии на сутки вперёд, метод взвешенной скользящей средней, без кластеризации.

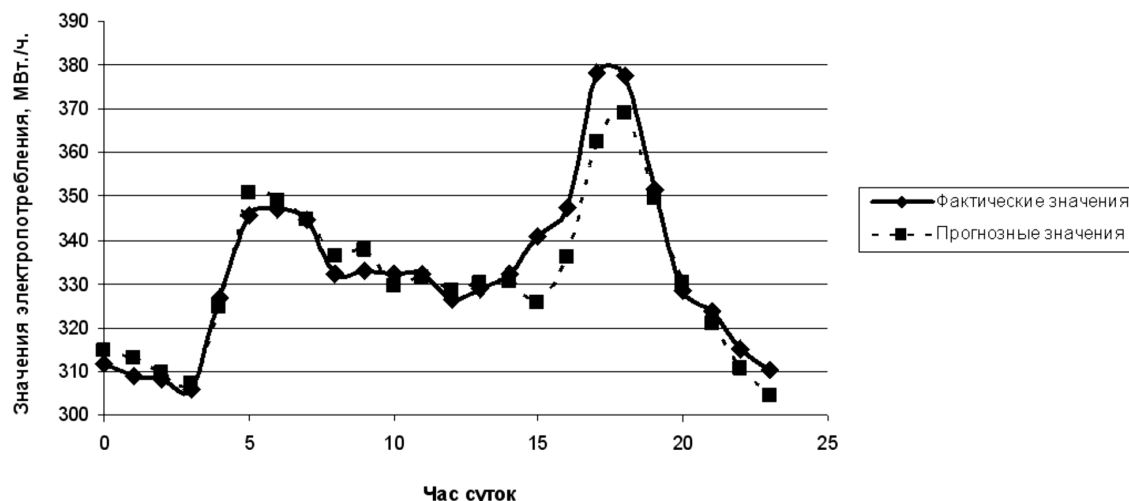


Рисунок 2. Графики фактических и прогнозных значений электропотребления в Хакасии на сутки вперёд, метод взвешенной скользящей средней, кластер "Межсезонье".

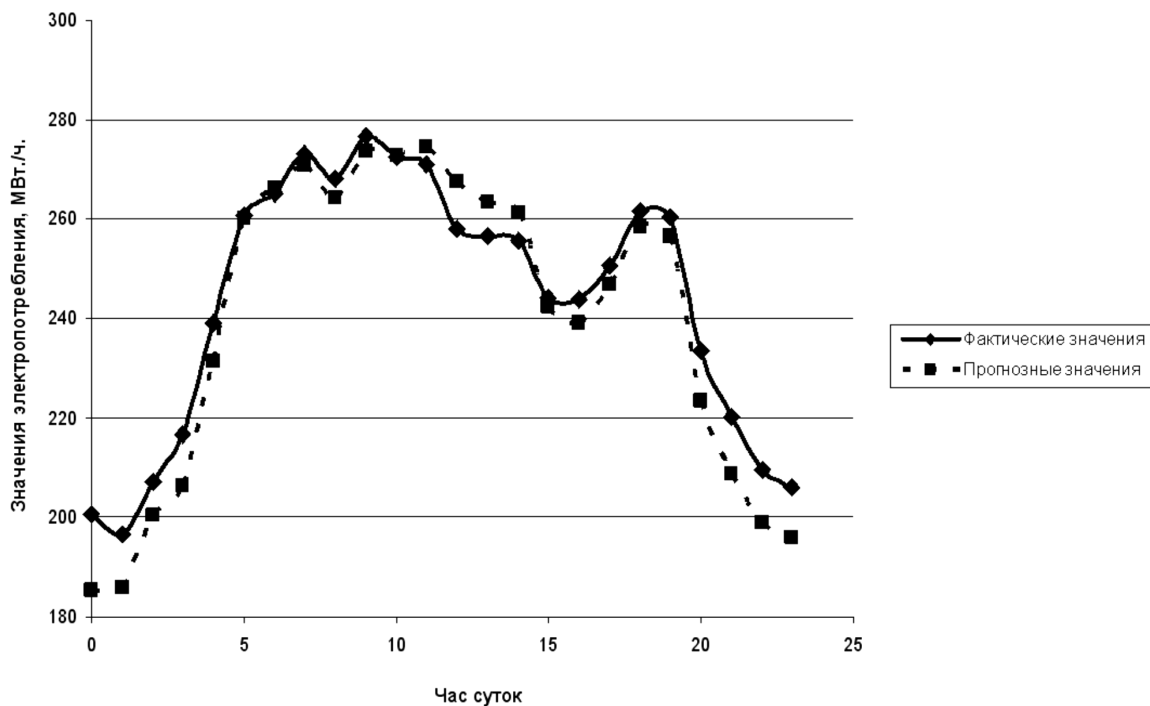


Рисунок 3. Графики фактических и прогнозных значений электропотребления в Хакасии на сутки вперед, метод взвешенной скользящей средней, кластер "Без отопления".

жение прогнозной ошибки. Однако на некоторых участках, а именно 12-ти–13-ти и 17-ти–18-ти часов, отмечена недостаточная точность прогнозных значений, в связи с этим необходимо подключение дальнейшего механизма их корректировки.

Предлагаемая процедура кластеризации может быть полезна и при прогнозировании на длительных интервалах времени.

ЛИТЕРАТУРА

1. Кластеризация [Электронный ресурс] // BaseGroupLabs: технологии анализа данных. – URL: <https://basegroup.ru/deductor/function/algorithm/clustering> (дата обращения: 18.03.2016).
2. Кластерный анализ. Научный подход при изучении сложных явлений [Электронный ресурс] // ФБ. – URL: <http://arcticaoy.ru/fb.ru/article/22911/klasternyiy-analiz-nauchnyiy-podhod-pri-izuchenii-slojnyih-yavleniy.html> (дата обращения: 13.02.2016).
3. Макоклюев Б.И. Анализ и планирование электропотребления. – М. Энергоатомиздат, 2008. – 296 с.
4. Методы принятия решений в электроэнергетике [Электронный ресурс] // ИД "Библио-Глобус". – URL: <http://bgscience.ru/lib/7261/> (дата обр.: 24.03.2016).
5. Нейронные сети [Электронный ресурс] // StatSoft: Электронный учебник по статистике. – URL: <http://www.statsoft.ru/home/textbook/modules/stneunet.html#multilayera> (дата обращения: 19.03.2016).
6. Разработка прогноза с помощью метода скользящей средней [Электронный ресурс] // Отдельные вопросы экономики. – URL: <http://www.ekonomika-st.ru/drugie/metodi/metodi-prognoz-1-3.html> (дата обращения: 11.02.2016).
7. Сглаживание скользящих средних – достоинства и недостатки [Электронный ресурс] // ВебМастерМаксим.ру. – URL: ВебМастерМаксим.ру. – URL: <http://webmastermaksim.ru/foreks/sglzhivanie-skolzyashhix-srednix-dostoinstva-nedostatki.html> (дата обращения: 21.11.2015).
8. Хайкин, С. Нейронные сети [Текст] / С. Хайкин. – М.: Вильямс, 2008. – 1104 с.
9. Чувилкин, А.В., Гордеев, А.С. Влияние метеорологических и производственных факторов на потребление электроэнергии предприятий АПК [Электронный ресурс] / А.В. Чувилкин, А.С. Гордеев // Профессиональный образовательный портал по Тамбовской области. – URL: <http://www.edu68.ru/nauka/290-vliyanie-meteorologicheskix-i-proizvodstvennyx-faktorov-na-potreblenie.html> (дата обращения: 17.03.2016).
10. Чучуева, И.А. Модель прогнозирования временных рядов по выборке максимального правдоподобия / И.А. Чучуева // Математическое бюро. – URL: <http://www.mbureau.ru/articles/dissertaciya-model-prognozirovaniya-vremennyh-ryadov-glava-1> (дата обращения: 11.03.2016).
11. Шилов, А.В. Анализ методов в системе выработки и принятия прогнозных решений [Текст] / А.В. Шилов // С.С. Чернов. – Новосибирск: Издательство ЦРНС, 2014. – С. 27–34.
12. Экспоненциальное скользящее среднее [Электронный ресурс] // FreshForex. – URL: <https://freshforex.org/encyclopedia-forex/exponential-moving-average/> (дата обращения: 21.03.2016).