

ПРОГНОЗИРОВАНИЕ ТОКА НА ОСНОВАНИИ ДАННЫХ ЭНЕРГОМОНИТОРА

М. И. Гальченко¹, Н. В. Васильев²

¹HiQo Solutions, г. Санкт-Петербург, Российская Федерация

*²Санкт-Петербургский государственный аграрный университет,
Российская Федерация*

В рамках проводимых работ по улучшению показателей качества электроэнергии как подзадача рассматривается возможность применения умных устройств с возможностью учета прогноза показателей в кратковременном периоде на основе исторических данных. С целью определения возможности построения кратковременных прогнозов токов по фазам А-С и нейтрали были предоставлены данные одной из трансформаторных подстанций Санкт-Петербурга, снятые энергомонитором в

период с 16 по 22 марта 2017 г. В соответствии с регламентом данные снимались с частотой один раз в 10 мин.

В настоящей статье мы покажем основные этапы проведенного описательного и разведочного анализа данных, а также результаты тестирования различных подходов к моделированию процессов. Для анализа данных применялся язык статистического программирования R как наиболее отвечающий поставленной задаче.

Первичный анализ набора данных показывает, что:

- ряд не содержит пропусков значений в исследуемом интервале, частота снятия показателей не нарушается;
- предоставленный ряд короток, невозможно определить показатели, относящиеся к сезонности, большей дневной;
- данные относятся только к весеннему периоду, невозможно определить, как изменится картина в прочие сезоны;
- отсутствуют погодные данные (температура, влажность, освещенность), что делает прогноз неустойчивым к локальным по времени (несезонное похолодание, дождливая погода) изменениям окружающей среды.

В рассматриваемом наборе данных токи по фазам и нейтрали различны по своим показателям и поведению. Наиболее полно отобразить ситуацию позволяет скрипичный график (рис. 1).

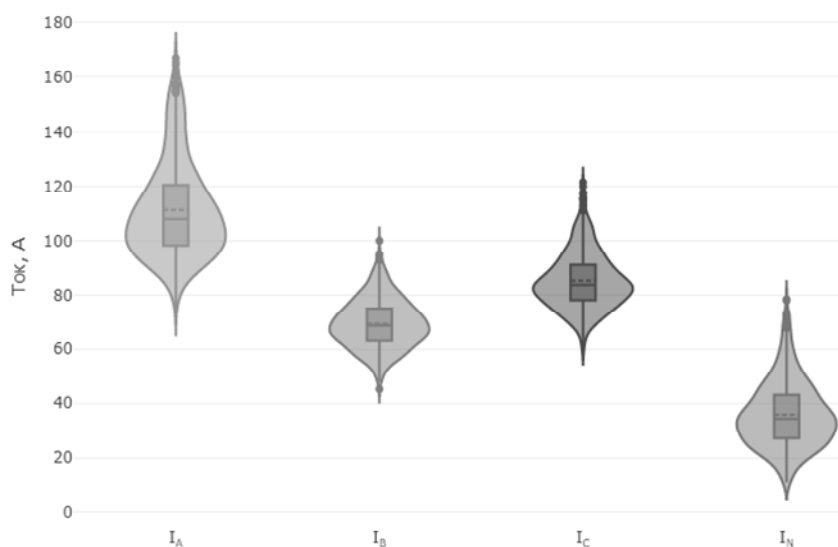


Рис. 1. Распределение тока по фазам и нейтрали

Ток фазы А имеет скошенное распределение, также можно говорить о превосходстве значений фазы А над остальными. Ток нейтрали, наоборот, значительно уступает по значениям току по фазам А-С. Распределения унимодальные, с относительно небольшим количеством выбросов. Ни одно из распределений нельзя отнести к нормальным, что подтверждает тест Шапиро–Уилка.

Для определения схожести временных рядов использовали расстояние DTW. Расстояние для всех рядов достаточно велико (более 7000), наиболее близки между собой ряды для фаз А и В (7539) и А и С (11650). Для определения схожести рядов в части периодичности использовали периодограмму Ломба–Скаргла. Для всех рядов выявляется свой набор статистически значимых мощностей и соответствующих им

периодов. В почасовом разрезе отмечается сходство в поведении серий для токов по А и N, В и С. Наблюдается различие в характере серий для пятницы, субботы и прочих дней.

В сумме все это говорит о существенных различиях в рядах и невозможности применения одной модели для прогнозирования. Кроме того, возникает вопрос о разумности прогнозирования данного ряда как многомерного.

Для оценки возможности прогнозирования токов по фазам и нейтралю применялись алгоритмы SSA [1] (библиотека Rssa), ARIMA (библиотека forecast), NNAR [2] (библиотека forecast), а также искусственные нейронные сети (библиотека keras). Результаты моделирования для ARIMA и SSA (наилучшие модели по PME, тестовый набор – 6 последних измерений) приведены ниже (таблица). Модели подбирались по количеству используемых для прогноза лагов (NNAR) и количеству компонент разложения для автогруппировки (SSA).

Результаты моделирования значений тока с применением различных алгоритмов

Фаза	ARIMA				NNAR				SSA			
	Спецификация	RMS E	PME (CP)	DTW	Спецификация	RMSE	PME (CP)	DTW	Компонент	RMSE	PME (CP)	DTW
A	(4,1,2)	8,86	-0,20	47,92	(13, 50)	7,49	0,14	35,33	6	5,88	0,43	35,23
B	(8,1,3)	6,75	0,33	43,99	(4, 50)	4,56	0,62	33,42	13	4,84	0,60	28,94
C	(6,1,4)	6,13	-0,55	28,61	(2, 50)	4,88	0,04	23,39	6	4,12	0,32	20,50
N	(6,1,4)	5,08	0,39	30,95	(9, 50)	3,22	0,73	26,42	5	3,42	0,66	16,13

По использованным критериям качества наилучший результат показывают NNAR и SSA. Стоит отметить, что модели, построенные с применением алгоритма SSA отличаются компактностью (используется 5-6 компонент для автогруппировки, максимум – 13), а сам расчет не занимает много времени. Применение нейронной сети с частично подобранными параметрами (три слоя, активация selu, в первом слое 100 нейронов, с последовательным уменьшением количества по слоям) к данным по фазе А дало результат, близкий к показанному NNAR.

Таким образом, наилучшие результаты по полученным моделям показывают NNAR и SSA. Показываемые результаты можно оценить как удовлетворительные, однако есть необходимость в продолжении исследований. Дополнительно будут рассмотрены рекуррентные нейронные сети и Гауссова регрессия. Также рассматривается возможность получения более длинных рядов измерений, в том числе на других трансформаторных подстанциях, погодных данных.

Литература

1. Golyandina, N. Basic Singular Spectrum Analysis and Forecasting with R. / N. Golyandina, A. Korobeynikov // Computational Statistics and Data Analysis.– 2014. – Vol. 71. – С. 934–954.
2. Hyndman, Rob J. Forecasting : principles and practice / Rob J. Hyndman and George Athanasopoulos. – OTexts.com [Heathmont, Victoria]. – 2014.