

УДК 681.3

*Великанова Л.И., ludvel@mail.ru**Лыченко Н.М., nlychenko@mail.ru**Институт машиноведения и автоматизации НАН КР, Бишкек, Кыргызстан*

РЕЗУЛЬТАТЫ МОДЕЛИРОВАНИЯ СРЕДНЕСРОЧНОГО ПРОГНОЗА ИНДЕКСА КАЧЕСТВА ВОЗДУХА Г. БИШКЕК

Развита методология прогнозирования индекса качества воздуха AQI г. Бишкек на 24 часа вперед с учетом метеорологических факторов и загрязняющего фактора (количества тонн угля, ежедневно сжигаемого на городской ТЭЦ). Представлены результаты моделирования на базе ИНС прямого распространения сигнала с обратным распространением ошибки FFBNN для 3-часовых сроков прогноза AQI летнего и зимнего периодов для горизонтов прогнозирования до 24 часов. Показано, что точность прогнозирования на 24 часа вперед в среднем по всем срокам прогноза по сравнению с краткосрочным прогнозом (на 3 часа вперед) ухудшилась на 34% для летнего периода и 44% — для зимнего.

Ключевые слова: искусственная нейронная сеть FFBNN, обучающая выборка, индекс качества воздуха, метеофакторы, загрязняющий фактор, горизонт прогнозирования, ошибка прогнозирования.

Введение

Предыдущие работы авторов были связаны с разработкой методологии и краткосрочных моделей прогнозирования концентраций твердых частиц PM_{2.5} и индекса качества воздуха AQI [1] г. Бишкек. При этом в качестве входных параметров прогноза использовались метеофакторы (температура воздуха; температура точки росы, относительная влажность воздуха; скорость ветра, интенсивность осадков) и доступная информация о загрязняющем воздух факторе – ежесуточном потреблении угля городской ТЭЦ. Была развита методология построения прогностических моделей, суть которой заключалась в следующем [2, 3]. Наблюдаемые значения концентраций твердых частиц PM_{2.5} и индекса качества воздуха декомпозировались по сезонам (зима — отопительный/лето — неотопительный) и по срокам измерений: 0 часов, 3 часа, 6 часов, 9 часов, 12 часов, 15 часов, 18 часов, 21 час. Для каждого сезона и каждого срока измерений строились модели прогноза загрязненности воздуха в следующий срок наблюдений на основе искусственных нейронных сетей (ИНС). Были исследованы прогностические свойства моделей, построенных на основе двух типов ИНС: обобщенно-регрессионной нейронной сети GRNN [4] и ИНС прямого распространения с использованием алгоритма обратного распространения ошибки FFBNN [5]. При этом для каждой модели были определены наиболее эффективные, с позиции средних ошибок прогноза, варианты входных векторов сети, учитывающих различную историю метеопараметров (на 3 часа, 6 часов и т.д. ранее). Показано, что прогноз AQI в подавляющем большинстве случаев дает существенно меньшую ошибку по сравнению с прогнозом PM_{2.5} как в отопительный, так и в неотопительный периоды при использовании обоих типов нейросетей.

Однако с практической точки зрения востребован прогноз загрязнения городского воздуха твердыми частицами с упреждением на сутки и более – для того, чтобы было возможным принять необходимые меры по предотвращению влияния недопустимого качества воздуха на здоровье людей. Многие задачи прогноза временных рядов наблюдений за параметрами окружающей среды с горизонтом прогнозирования до суток и более успешно решаются с применением нейронных сетей [6, 7, 8]. В настоящей статье приведены результаты моделирования среднесрочного (до 24 часов) прогноза индекса

качества воздуха г. Бишкек на основе развитой ранее в работах [2, 3] методологии и на базе ИНС прямого распространения с обратным распространением ошибки FFBN.

В качестве исходных данных¹ использованы временные ряды измерения индекса качества воздуха (AQI) г. Бишкек за период с 09.02.2019 по 30.09.2022 с интервалом в 3 часа [9], данные метеослужбы RP5 [10], а также учтены данные о ежедневно потребляемом ТЭЦ г. Бишкек количестве тонн угля и его типе (по месту добычи).

Методология среднесрочного прогноза AQI

Так же, как и для краткосрочного прогноза [2, 3], проведена декомпозиция архивных данных измерений AQI [9] и данных метеослужбы RP5 [10] по сезонным периодам и по срокам измерений: $I = 0$ часов, 3 часа, 6 часов, 9 часов, 12 часов, 15 часов, 18 часов, 21 час. Для каждого i -го срока и обоих периодов (зимний/летний) на базе ИНС с прямым распространением сигнала и обратным распространением ошибки (функция *newff* в среде программирования MatLab) были построены модели для вычисления прогнозных значений AQI с горизонтом прогнозирования на 1, 2, ..., 8 шагов вперед (1 шаг = 3 часа), в зависимости от значений AQI и метеопараметров, предшествующих сроку прогноза ($i-1, i-2$).

Общее число сигналов во входных векторах – 11, их соответствие контролируемым параметрам указано в таблице 1. В этой таблице:

- Ws – скорость ветра в предыдущие сроки, то есть на $i-1$ и $i-2$ шагах прогноза/измерения (м/с);
- Pr – интенсивность осадков в предыдущие сроки, то есть на $i-1$ и $i-2$ шагах прогноза/измерения;
- T – температура воздуха, усредненная за $i-1$ и $i-2$ шага прогноза/измерения;
- RH – относительная влажность воздуха, усредненная за $i-1$ и $i-2$ шага прогноза/измерения;
- $Coal$ – общее количество тонн угля, ежедневно потребляемого городской ТЭЦ, и количество тонн угля определенного типа в зависимости от места добычи, соответственно: $Coal_1$ – Кара-Кече, $Coal_2$ – Кара-Жыра, $Coal_3$ – Шубаркуль.

Таблица 1 – Контролируемые параметры по срокам прогноза с учетом угля и его типа

AQI	Ws	PR	T	RH	$Coal_1$	$Coal_2$	$Coal_3$	Coal	Число сигналов
$i-1$	$i-1$ $i-2$	$i-1$ $i-2$	$i-1$ $i-2$	$i-1$ $i-2$	$i-1$	$i-1$	$i-1$	$i-1$	11

Для оценки адекватности моделей и точности прогнозов, построенных на их основе, так же, как и в [1], использованы: средняя абсолютная процентная ошибка (Mean Absolute Percentage Error, MAPE): $MAPE = \frac{1}{N} \sum_{t=1}^N \frac{|e_t|}{y_t} 100\%$ и средняя абсолютная ошибка MAE, деленная на среднее (MAE/Mean Ratio): $MAEMR = \frac{1}{N} \sum_{t=1}^N |e_t| / Mean(y_t)$ [11].

Здесь $e_t = (y_t - y_{ft})$ – ошибка прогноза или остатки модели, y_t – наблюдаемое (измеренное) значение в срок t , y_{ft} – значение в срок t , рассчитанное по модели, $Mean(y_t)$ – среднее значение y_t , N – длина выборки наблюдений.

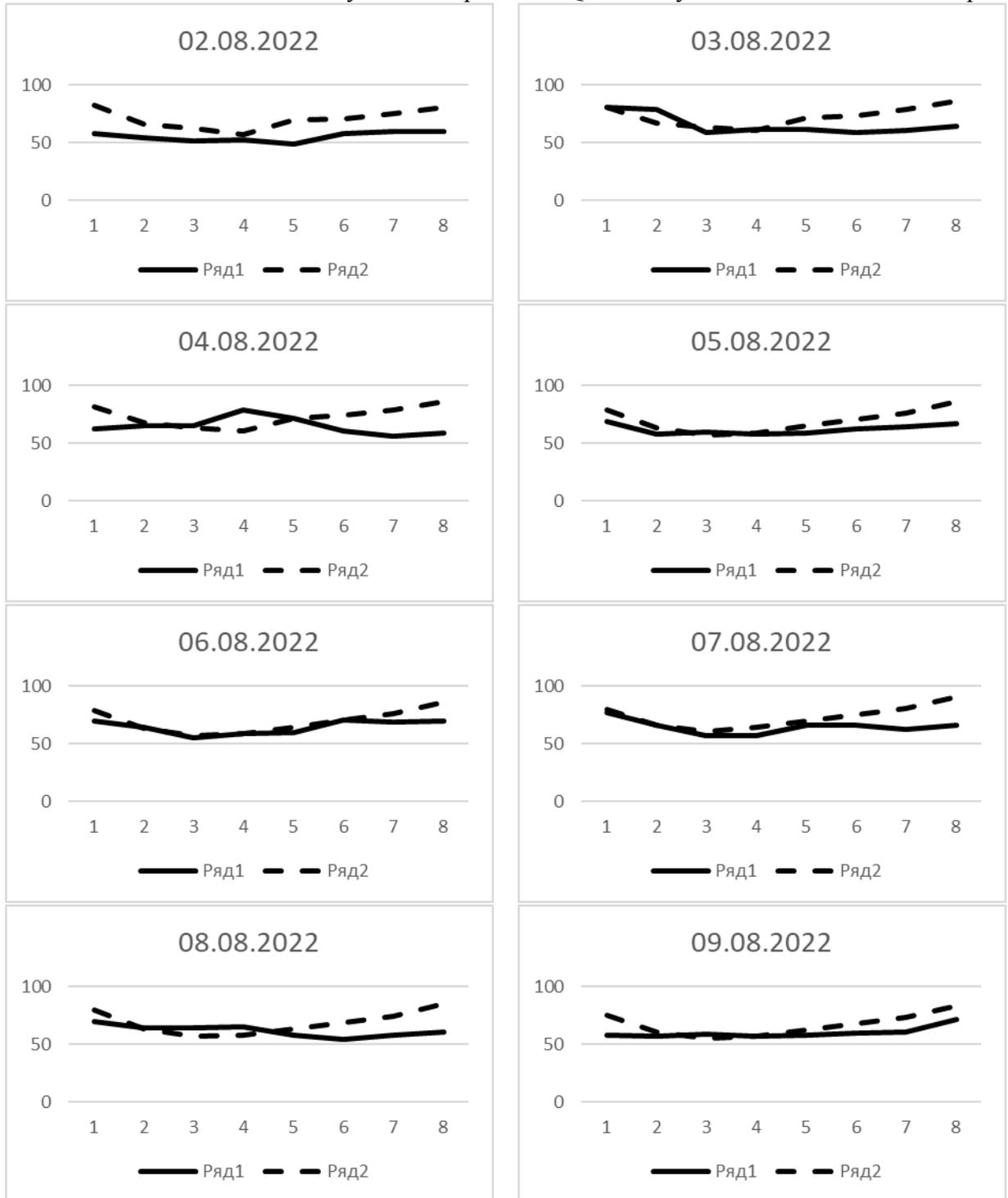
Результаты моделирования

Тестовое моделирование прогноза для летнего периода было проведено на выборке 02.08.2022 – 18.08.2022. Обучающее множество – вся доступная выборка наблюдений, за исключением летнего периода 2020 года [3]: (01.04.2019 – 31.07.2019) + (01.04.2021 – 30.09.2021) + (04.04.2022 – 30.09.2022).

¹ Данные подготовлены Комбаровою Л.А.

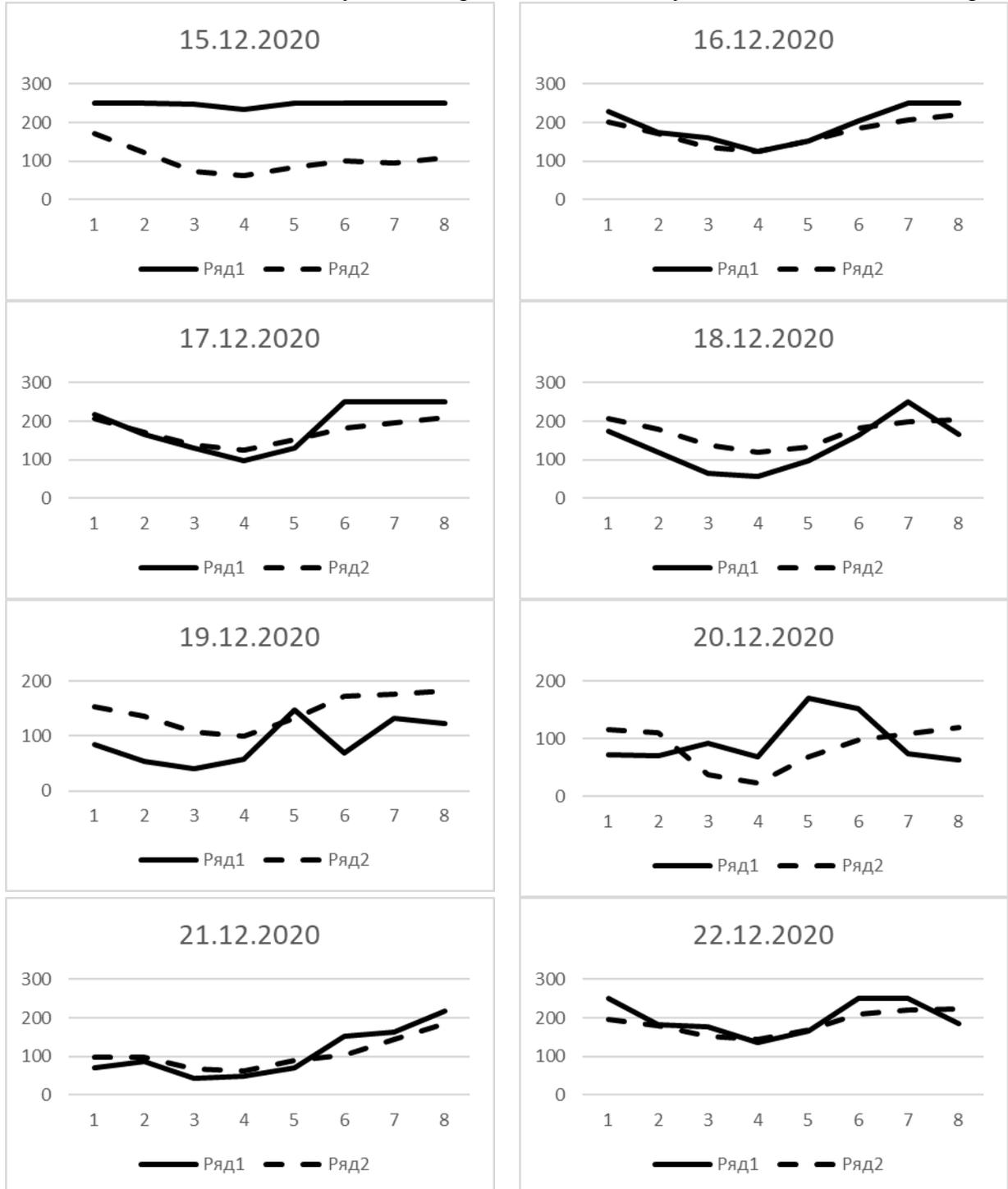
В таблице 2 представлены графики суточного прогноза AQI для 8 суток летнего тестового периода. На этих графиках по оси абсцисс отложены номера i -тых сроков измерений (прогноза): 1 – 0 часов, 2 – 3 часа, 3 – 6 часов, ..., 8 – 21 час; по оси ординат – реальные (измеренные) значения AQI (ряд 1) и прогнозные значения AQI (ряд 2) в соответствующие сроки. При этом для срока 0 часов прогнозное значение рассчитано как прогнозное на 1 шаг вперед от известного реального значения в срок 21 часа предыдущих суток, для срока 3 часа, соответственно на 2 шага вперед, для срока 6 часов – на 3 шага вперед и так далее, для срока 21 час – на 8 шагов вперед.

Таблица 2– Суточный прогноз AQI для 8 суток летнего тестового периода



Тестовое моделирование для зимнего периода было проведено на выборке 15.12.2020 – 30.12.2020. Обучающее множество – вся доступная зимняя выборка наблюдений: (09.02.2019 – 23.03.2019) + (01.11.2019 – 31.03.2020) + (01.11.2020 – 09.12.2020) + (01.01.2021 – 31.03.2021) +(08.10.2021 – 03.04.2022). В таблице 3 представлены графики прогноза AQI для восьми суток зимнего тестового периода.

Таблица 3 – Суточный прогноз AQI для 8 суток зимнего тестового периода



В таблицах 4 – 5 приведены ошибки прогноза AQI для всех сроков прогноза, усредненные за 17 суток для летнего и за 16 суток для зимнего периодов. При этом для

сравнения прогнозные значения вычислялись для различных горизонтов прогнозирования: на 3, 6, 9, 12, 15, 18, 21 и 24 часа вперед.

Таблица 4 – Усредненные ошибки моделированных значений AQI по срокам прогноза и для различных горизонтов прогнозирования (летняя тестовая выборка)

Сроки прогноза	Горизонт прогнозирования							
	3 часа		6 часов		9 часов		12 часов	
	<i>MAPE</i>	<i>MAEMR</i>	<i>MAPE</i>	<i>MAEMR</i>	<i>MAPE</i>	<i>MAEMR</i>	<i>MAPE</i>	<i>MAEMR</i>
00	16.95	16.24	24.12	23.18	26.99	25.80	28.22	26.91
03	6.98	7.10	7.32	7.35	8.55	8.53	9.72	9.61
06	7.76	7.62	9.23	9.20	7.76	7.63	6.51	6.49
09	6.05	6.35	7.19	7.52	8.64	8.99	7.73	8.07
12	10.26	9.63	10.70	10.06	10.91	10.32	10.66	10.10
15	12.85	12.39	16.88	16.19	17.05	16.42	16.65	16.05
18	17.21	16.92	22.80	22.45	24.67	24.26	24.63	24.23
21	20.91	20.10	28.64	27.73	31.43	30.42	32.34	31.29
<i>MEAN</i>	12.37	12.04	15.86	15.46	17.00	16.54	17.06	16.59
Сроки прогноза	Горизонт прогнозирования							
	15 часов		18 часов		21 час		24 часа	
	<i>MAPE</i>	<i>MAEMR</i>	<i>MAPE</i>	<i>MAEMR</i>	<i>MAPE</i>	<i>MAEMR</i>	<i>MAPE</i>	<i>MAEMR</i>
00	28.64	27.31	28.68	27.37	28.63	27.33	28.60	27.31
03	10.26	10.12	10.46	10.32	10.49	10.35	10.47	10.33
06	6.29	6.26	6.21	6.17	6.22	6.18	6.20	6.17
09	7.96	8.31	7.90	8.23	7.86	8.19	7.87	8.19
12	11.15	10.57	9.02	9.01	9.24	9.21	9.31	9.28
15	16.42	15.84	16.80	16.22	16.55	16.03	16.69	16.16
18	24.51	24.11	24.32	23.91	24.48	24.07	24.57	24.06
21	32.33	31.29	32.24	31.21	32.14	31.11	32.23	31.19
<i>MEAN</i>	17.19	16.73	16.96	16.56	16.95	16.56	16.99	16.59

Как видно, для большинства сроков прогноза ошибка возрастает с увеличением горизонта прогнозирования, что объясняется ее накоплением: при вычислении прогноза для следующего горизонта в вычисления включаются значения AQI, рассчитанные для предыдущего горизонта прогноза. Увеличение средней ошибки для всех сроков прогноза на 24 часа вперед по сравнению с краткосрочным прогнозом (на 3 часа вперед) составило 34% для летнего периода и 44% – для зимнего (значения *MAEMR* 16.59/12.04 в таблице 4 и 27.27/18.90 в таблице 5 – соответственно).

Таблица 5 – Усредненные ошибки моделированных значений AQI по срокам прогноза и для различных горизонтов прогнозирования (зимняя тестовая выборка)

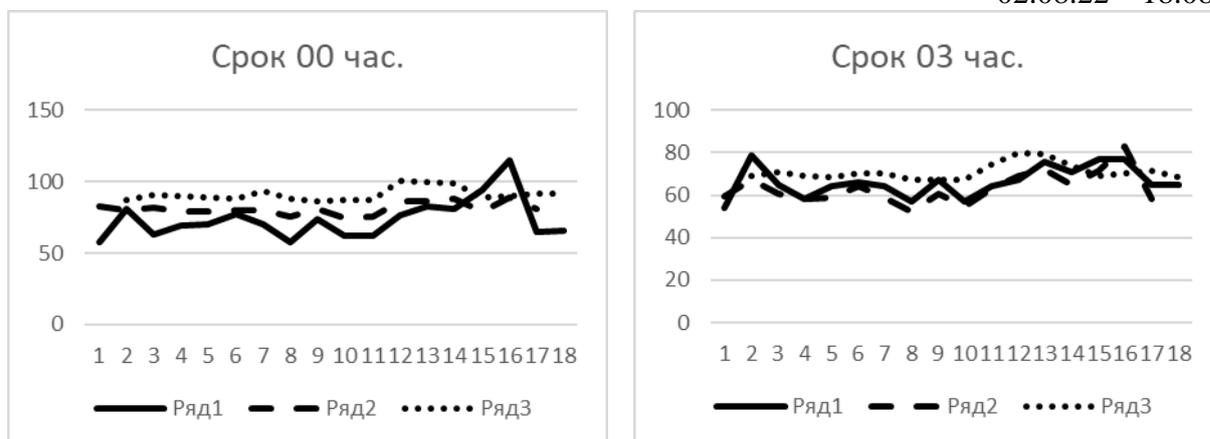
Сроки прогноза	Горизонт прогнозирования							
	3 часа		6 часов		9 часов		12 часов	
	<i>MAPE</i>	<i>MAEMR</i>	<i>MAPE</i>	<i>MAEMR</i>	<i>MAPE</i>	<i>MAEMR</i>	<i>MAPE</i>	<i>MAEMR</i>
00	25.74	20.49	35.69	26.93	34.61	26.52	36.53	27.65
03	18.71	15.85	36.71	25.81	43.29	27.39	41.94	27.05
06	22.53	24.87	29.96	33.07	52.48	37.31	56.10	34.73
09	22.24	22.00	29.69	28.31	39.39	34.80	50.48	39.20
12	18.98	19.99	27.79	27.74	31.03	31.60	30.59	31.26
15	21.06	17.92	2.65	2.26	27.27	22.72	29.68	25.17

18	16.07	14.21	22.87	19.97	19.96	18.87	21.65	20.42
21	18.82	15.87	21.00	17.07	23.67	19.09	23.34	19.72
<i>MEAN</i>	20.52	18.90	25.80	22.65	33.96	27.29	36.29	28.15
Горизонт прогнозирования								
Сроки прогноза	15 часов		18 часов		21 час		24 часа	
	<i>MAPE</i>	<i>MAEMR</i>	<i>MAPE</i>	<i>MAEMR</i>	<i>MAPE</i>	<i>MAEMR</i>	<i>MAPE</i>	<i>MAEMR</i>
00	36.26	28.10	36.18	28.18	36.74	28.79	37.14	28.89
03	42.59	27.35	42.22	27.56	41.75	27.24	42.36	27.77
06	55.67	35.43	53.64	34.21	53.82	34.75	53.84	34.81
09	50.59	34.87	50.65	35.58	48.37	34.14	48.70	34.60
12	28.24	28.53	24.90	24.06	25.69	24.89	25.01	24.30
15	28.92	24.24	32.03	25.98	31.25	24.15	31.28	24.42
18	20.54	20.22	20.86	20.33	22.06	21.02	24.66	22.86
21	23.88	20.15	23.49	20.36	23.75	20.48	24.05	20.48
<i>MEAN</i>	35.84	27.36	35.50	27.03	35.43	26.93	35.88	27.27

Также стоит отметить, что, как и в предыдущих работах [2, 3], полученные ошибки прогноза для летнего периода в полтора – два раза меньше, чем для зимнего периода. Это можно объяснить более стационарным характером процессов загрязнения городского атмосферного воздуха в летний период и влиянием неучтенных факторов, связанных с источниками загрязнений, для зимнего периода (прежде всего – печное отопление частного сектора). При этом для летнего периода отмечается увеличение ошибок прогнозирования для сроков прогноза второй половины суток, чего, в общем, не наблюдается для зимнего периода (см. таблицы 4 и 5). Это можно объяснить более выраженным влиянием на загрязненность воздуха твердыми частицами городского автотранспорта в летнее время.

В таблицах 6-7 представлены графики измеренных (ряд 1) и моделированных значений AQI с горизонтами прогнозирования на 3 часа вперед (ряд 2) и 24 часа вперед (ряд 3) для всех сроков прогноза в летний (02.08.2022 – 18.08.2022) и зимний (15.12.2022 – 30.12.2022) тестовые периоды. Здесь также иллюстрируется увеличение ошибки прогноза для горизонта прогнозирования в 24 часа в сравнении с краткосрочным прогнозом на 3 часа вперед и отмеченная выше особенность процессов загрязнения в летний период.

Таблица 6 – Графики измеренных (ряд 1) и моделированных значений AQI на 3 (ряд 2) и 24 (ряд 3) часа вперед для всех сроков прогноза в летний тестовый период 02.08.22 – 18.08. 22



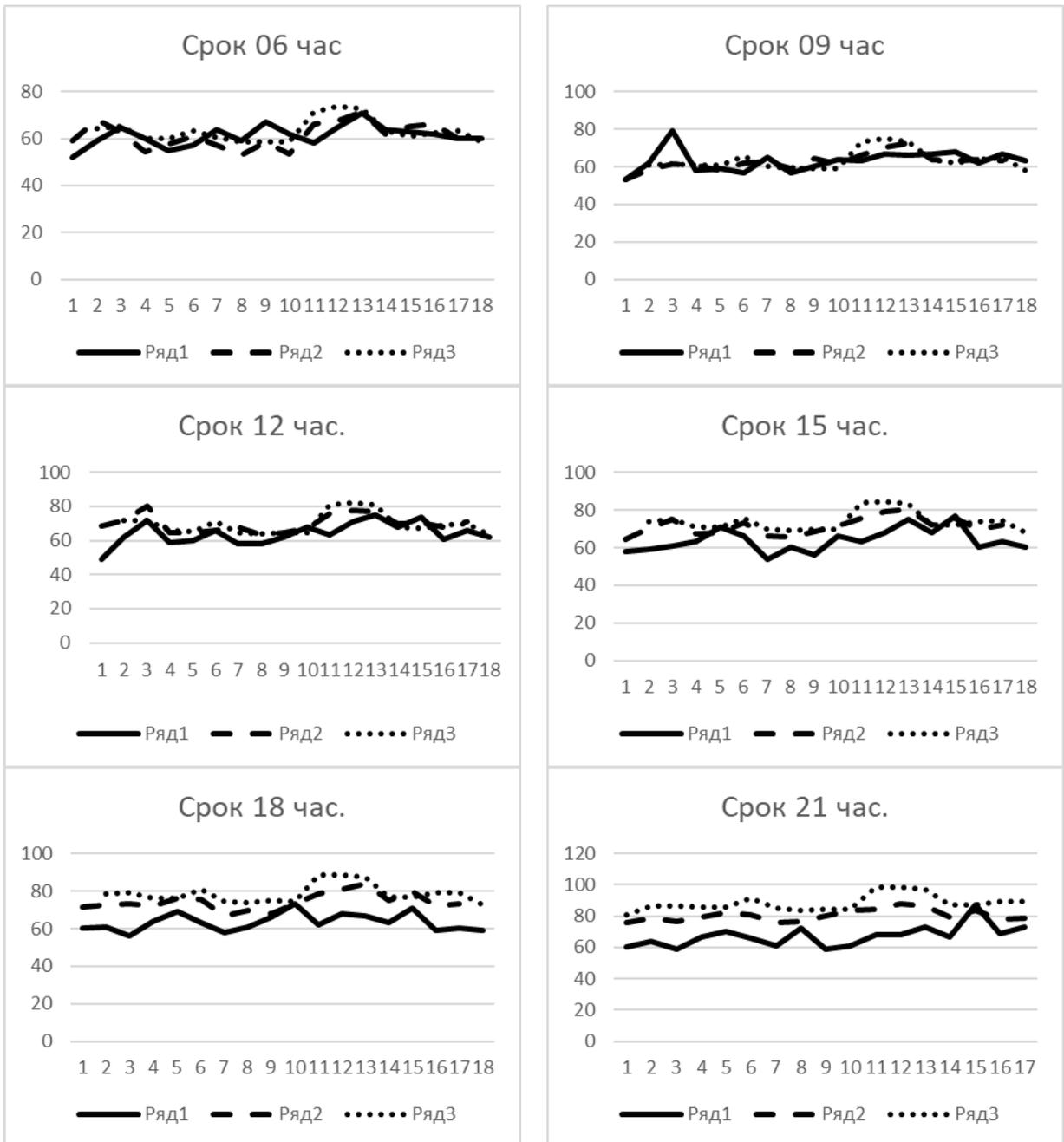
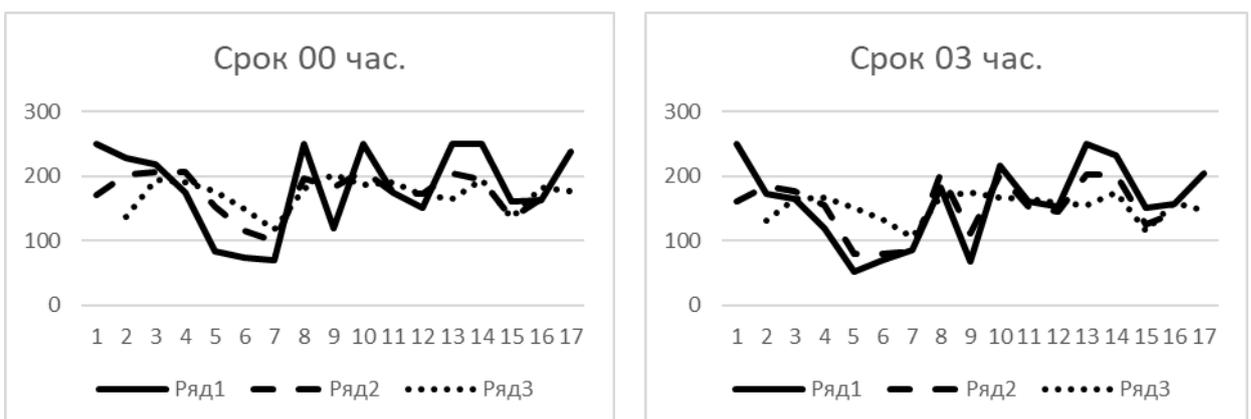
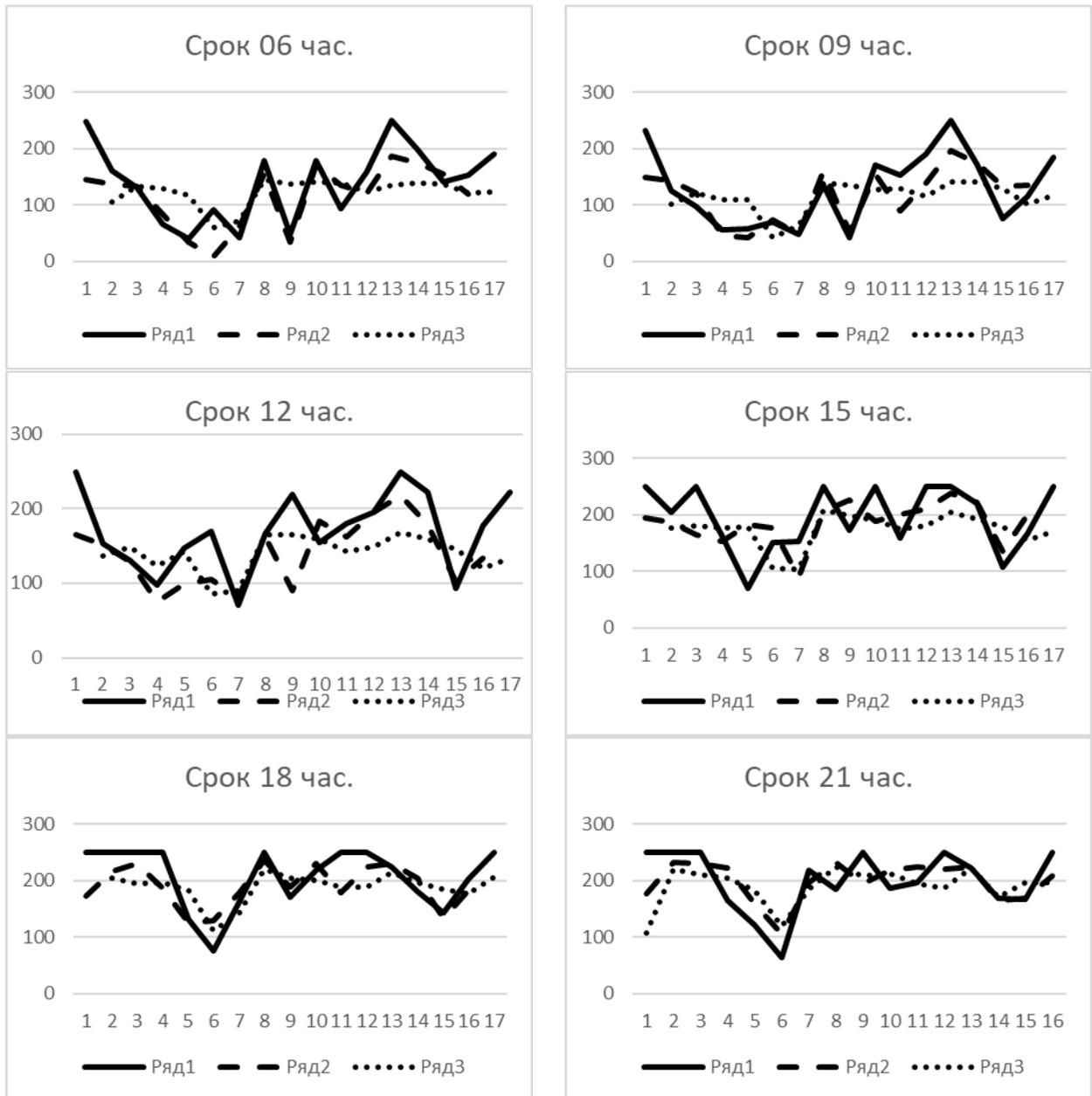


Таблица 7 – Графики измеренных (ряд 1) и моделированных значений AQI на 3 (ряд 2) и 24 (ряд 3) часа вперед для всех сроков прогноза в зимний тестовый период 15.12.22 – 30.12. 22





Заклучение

Таким образом, в работе представлено развитие методологии для среднесрочного (на 8 шагов вперед) прогноза AQI на основе метеоданных и данных о ежедневно потребляемом ТЭЦ г. Бишкек количестве тонн угля и его типе (по месту добычи). Использoваны модели, построенные на основе нейронной сети с прямым распространением сигнала и обратным распространением ошибки FFBN. Показано, что точность прогнозирования для всех сроков прогноза на 24 часа вперед по сравнению с краткосрочным прогнозом (на 3 часа вперед) ухудшилась на 34% для летнего периода и 44% – для зимнего. Это объясняется накоплением ошибок при вычислении прогноза на следующий временной шаг.

Следует отметить, что в реальных прогнозах загрязненности воздуха в качестве входных векторов метеопараметров можно использовать их почасовые прогнозные значения на сутки вперед, которые, стоит заметить, достаточно точно соответствуют реальным. Что касается количества тонн угля, то здесь также хорошо бы использовать планируемые показатели загрузки городской ТЭЦ на сутки вперед. При отсутствии такой возможности можно использовать данные за прошедшие сутки, ошибка прогноза при этом увеличится незначительно, поскольку, как показал корреляционный анализ [3]

среднесуточных концентраций загрязнения воздуха частицами PM_{2.5} и количества тонн угля, ежедневно потребляемого городской ТЭЦ, их корреляции в соседние сутки практически не изменяются.

Благодарности

Авторы выражают благодарность Отделу по связям с общественностью ОАО «Электрические станции» КР за предоставленную информацию о количестве тонн ежедневно сжигаемого на ТЭЦ г. Бишкек угля и его типу по месту добычи за период с 9 февраля 2019 года по 30 сентября 2022 года.

Литература

1. [Air Quality Index \(AQI\) - A Guide to Air Quality and Your Health](#). US EPA. 9 December 2011.
2. Великанова, Л. И. Мультирегрессионные и обобщенно-регрессионные нейросетевые модели краткосрочного прогноза загрязнения PM_{2.5} в г. Бишкек с учетом метеорологических параметров / Л. И. Великанова, Н. М. Лыченко // Проблемы автоматизации и управления. – 2019. – № 2(37). – С. 16–25. – DOI 10.5281/zenodo.3594777. – EDN MNNDYU.
3. Великанова, Л. И. Моделирование краткосрочного прогноза загрязненности воздуха твердыми частицами на основе искусственных нейронных сетей с учетом фактора загрязнения / Л. И. Великанова, Н. М. Лыченко // Проблемы автоматизации и управления. – 2022. – № 3(45). – С. 110–118. – EDN TGNBLC.
4. Donald F. Specht A general regression neural network // IEEE Transactions on neural networks. — November 1991. — Vol. 2, № 6. 568–576. DOI: [10.1109/72.97934](#)
5. Widi Aribowo. Optimizing Feed Forward Backpropagation Neural Network Based on Teaching-Learning-Based Optimization Algorithm for Long-Term Electricity Forecasting // International Journal of Intelligent Engineering and Systems, Vol.15, No.1, 2022. DOI: [10.22266/ijies2022.0228.02](#)
6. Верзунов, С. Н. Применение глубоких нейронных сетей для краткосрочного прогноза дальности видимости / С. Н. Верзунов // Проблемы автоматизации и управления. – 2019. – № 1(36). – С. 118–130. – DOI 10.5281/zenodo.3253019. – EDN IFNJHG.
7. Karimian H, Li Q, Wu C, et al. Evaluation of different machine learning approaches to forecasting PM_{2.5} mass concentrations // Aerosol and Air Quality Research.- 2019.- 19(6). –p. 1400-1410. DOI: [10.4209/aaqr.2018.12.0450](#)
8. Tong W, Li L, Zhou X, et al. Deep learning PM_{2.5} concentrations with bidirectional LSTM RNN // Air Quality, Atmosphere & Health.- 2019. - 12(4). –p. 411–423. DOI: [10.1007/s11869-018-3030-4](#)
9. AirNow Department of State // [https://airnow.gov/index.cfm?action=airnow.global_summary#U.S. Department of State Bishkek](https://airnow.gov/index.cfm?action=airnow.global_summary#U.S._Department_of_State_Bishkek) (дата обращения: 05.10.2022)
10. Сайт «Расписание погоды rp5.ru» [Архив погоды в Бишкеке https://rp5.ru/%D0%90%D1%80%D1%85%D0%B8%D0%B2_%D0%BF%D0%BE%D0%B4%D1%8B_%D0%B2_%D0%91%D0%B8%D1%88%D0%BA%D0%B5%D0%BA%D0%B5](https://rp5.ru/%D0%90%D1%80%D1%85%D0%B8%D0%B2_%D0%BF%D0%BE%D0%B4%D1%8B_%D0%B2_%D0%91%D0%B8%D1%88%D0%BA%D0%B5%D0%BA%D0%B5) (дата обращения: 05.10.2022)
11. Современное прогнозирование. URL: <https://forecasting.svetunkov.ru/etextbook/> (дата обращения: 30.09.2022)