

*Russian Academy of Sciences
A.M. Obukhov Institute of atmospheric physics
Laboratory of mathematical ecology*



**Nikolay N. Zavalishin,
Alexander S. Ginzburg,
Gleb G. Alexandrov**

**Методика статистического прогноза загрязнения
атмосферного воздуха в мегаполисах на примере г. Москвы**

**Statistical forecasting method for atmospheric air pollution in
megapolises: Moscow case study**

ENVIROMIS-2018

Томск, 05 - 11 июля 2018

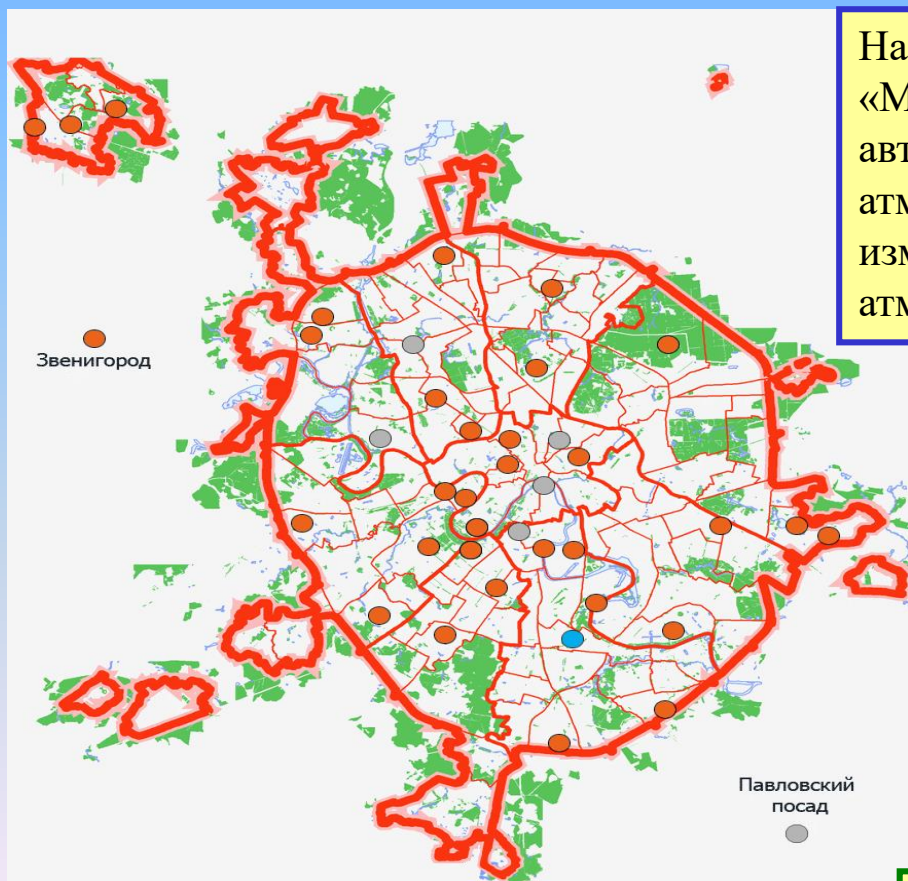


Загрязнение атмосферного воздуха и его измерение в г. Москве

Основными факторами, влияющими на экологическую обстановку в городах, являются:

- промышленные выбросы в атмосферный воздух и поверхностные воды;
- выбросы от автотранспорта;
- загрязнение территории всеми видами твердых отходов;
- шумовое, электромагнитное загрязнение;
- нарушение энергетического, водного, тектонического баланса с окружающей средой.

На середину 2017 г. в Москве под управлением ГУП «Мосэкомониторинг» (mecom.ru) функционируют 62 автоматических станций контроля загрязнения атмосферы (АСКЗА), сгруппированные в 6 групп, измеряющие концентрации максимум 12 примесей в атмосфере города.



Группы:

Вблизи автомагистралей

Смешанные

Жилые

Природные фоновые

Жилые в ТиНАО

Многоуровневые

Оксид углерода

Диоксид азота

Оксид азота

Метан

Диоксид серы

Аммиак

Сероводород

Озон

Формальдегид

Стирол (винилбензол)

ПМ-2.5

ПМ-10

Карта по состоянию на 2012 г.

Методы прогноза атмосферного загрязнения

- инерционный - экстраполирует в будущее текущие наблюдаемые значения концентраций ;
- климатический – дает вероятностные характеристики наблюдения в данное время и в данном месте определенных уровней загрязнения согласно многолетней статистике ;
- детерминированные модели качества воздуха (3-D models) – методы прогноза на основе системы численных моделей, способных воспроизводить эмиссию, перенос, диффузию, трансформацию и удаление атмосферных примесей;
- критериальный – дает прогноз опасных (согласно принятым критериям) уровней загрязнений на основе данных о загрязнении воздуха и прогнозируемой метеорологической ситуации;
- метод классификационных карт и регрессионных деревьев (classification chart and regression tree) – статистическая процедура для разбиения данных на определенные группы с целью построения определенной структуры (дерева) для прогноза (решения);
- **метод (не)линейной регрессии – статистический метод, основанный на экстраполяции найденных по длительному ряду наблюдений корреляционных зависимостей между предсказываемой переменной и прошлыми значениями выбранных в качестве предикторов переменных;**
- метод искусственных нейронных сетей – метод основан на применении компьютерных алгоритмов в терминах процессов обучения и распознавания образов, в результате применения которых устанавливаются нелинейные соотношения между входными переменными (предикторами) и прогнозируемыми выходными переменными;
- феноменологические и интуитивные методы– методы экспертной оценки ситуации, основанной на опыте и интуиции экспертов.

(рекомендации EPA - Агентства по охране окружающей среды США)

Основания и цель работы

Демченко П.Ф., Гинзбург А.С., Александров Г.Г., Вересков А.И., Горчаков Г.И., Завалишин Н.Н., Захарова П.В., Лезина Е.А., Юдин Н.И., Статистическое моделирование среднесуточных концентраций загрязняющих веществ атмосфере Московского мегаполиса методом множественной регрессии. // *Метеорология и гидрология*, 2015, № 10, с. 31-43.

Цель работы: разработать регрессионную формулу статистического прогноза среднесуточной концентрации примеси, методику ее калибровки, верифицировать ее на имеющихся данных с АСКЗА и реализовать в виде пользовательского программного обеспечения.

Преимущество **метода статистического моделирования** в том, что он не требует информации об интенсивности источников выбросов, как не требует и сложных вычислений или набора специальных данных, которые необходимы для эйлеровых или лагранжевых трехмерных фотохимических моделей.

Общая регрессионная формула:

$$Y_i = \sum_{l=1}^L a_l Y_{i-l} + \sum_{m=1}^M F_m(X_i^{(m)}, \vec{c}_m)$$

Y_i – прогнозируемая на i -е сутки концентрация, F_m – функция влияния внешних переменных, $\mathbf{a}$, $\mathbf{c}_m$ – векторы неизвестных коэффициентов.

Структура и предикторы формулы статистического прогноза

Формула регрессии-авторегрессии нашего статистического прогноза:

$$Y_{i+1} = a_1 X_1 + a_9 \cos(X_2 + a_2) + \frac{a_3}{b + X_3} + a_4 X_4 + a_5 X_5 + a_6 X_6 + a_7 X_7 + a_8 + a_{10} Y_i + a_{11} Y_{i-1}$$

Экзогенные (внешние) предикторы (X-переменные)

Временные предикторы на день прогноза

X_1	Признак дня	рабочий - 1, выходной - 0	a_1
Метеорологические предикторы на день прогноза			
X_2	Среднесуточное значение направления ветра на уровне 925 мб	градусы	a_2 и a_9
X_3	Среднесуточное значение скорости ветра на уровне 925 мб	м/с	a_3
X_4	Среднесуточное значение относительной влажности приземного воздуха	%	a_4
X_5	Среднесуточное значение скорости ветра у поверхности	м/с	a_5
X_6	Среднесуточное значение температуры воздуха у поверхности	°С	a_6
X_7	Среднесуточное значение атмосферного давления у поверхности	мб	a_7
	Свободный член прогностического уравнения	мг/м ³	a_8

Эндогенные (внутренние) предикторы (Y-переменные)

Y_i	Среднесуточная концентрация атмосферного загрязняющего вещества на день составления прогноза	мг/м ³	a_{10}
Y_{i-1}	Среднесуточная концентрация атмосферного загрязняющего вещества накануне дня составления прогноза	мг/м ³	a_{11}

Для расчета коэффициентов используется итерационный алгоритм Гаусса-Ньютона решения методом наименьших квадратов, в котором на каждой итерации аппроксимация производных, входящих в классический метод Гаусса-Ньютона решения задач нелинейной регрессии, достигается использованием значений регрессионной функции, полученных на предыдущих итерациях. Используется библиотека численного анализа **AlgLib** 3.14 (www.alglib.net).

Для прогнозирования среднесуточных значений концентраций используются выборки данных по **двум интервалам времени: обучающему и тестовому**. По обучающей выборке находятся коэффициенты модели, а тестовая применяется для сопоставления прогнозируемых и измеренных значений концентрации по независимым наблюдениям.

Программная реализация методики статистического прогноза концентраций загрязняющих примесей в атмосфере г. Москвы

Характеристики программного комплекса **AtmoPoll** v. 1.0:

- исходная информация включает метеоданные и значения концентрации выбранной примеси на выбранной группе АСКЗА г. Москвы за тестовый период (не менее 60 суток) для расчета неизвестных коэффициентов прогностической формулы;
- вычисление коэффициентов прогностической формулы методом нелинейной регрессии-авторегрессии и тестирование концентрации примеси на полном интервале с измеренными ее значениями с вычислением корреляций и средних ошибок;
- получение прогноза значений концентраций на прогнозном интервале времени без измеренных значений;
- сохранение и экспорт результатов для повторного использования и сравнений.

Формирование исходных данных для алгоритма статистического прогноза

AtmoPoll - Статистический прогноз атмосферного загрязнения г. Москвы

Файлы Наборы База данных О программе...

Калибровка

28.12.2015

29.12.2015

30.12.2015

31.12.2015

01.01.2016

02.01.2016

03.01.2016

04.01.2016

05.01.2016

06.01.2016

07.01.2016

08.01.2016

09.01.2016

10.01.2016

11.01.2016

12.01.2016

13.01.2016

Дата	1	2	3	4	5
	CO avg	NO2 avg	NO avg		
01.11.2015	2	0.39	0.0315	0.005	
02.11.2015	3	0.46	0.0309	0.007	
03.11.2015	4	0.4	0.0283	0.005	
04.11.2015	5	0.23	0.0367	0.01	
05.11.2015	6	0.13	0.0255	0.003	
06.11.2015	7	0.43	0.0512	0.043	
07.11.2015	8	0.65	0.05	0.058	
08.11.2015	9	0.23	0.034	0.01	
09.11.2015	10	0.43	0.0499	0.052	
10.11.2015	11	0.43	0.0313	0.036	
11.11.2015	12	0.48	0.0258	0.037	
12.11.2015	13	0.36	0.0256	0.029	
13.11.2015	14	0.39	0.0429	0.028	
14.11.2015	15	0.39	0.0381	0.028	
15.11.2015	16	0.27	0.0279	0.014	

	1	0 м-давл...	CO Конц. тек...	NO2 Конц. тек...	NO Конц...
		8	9	10	11
01.11.2015	1		0.39	0.0315	0.005
02.11.2015	2	1015	0.46	0.0309	0.007
03.11.2015	3	1013	0.4	0.0283	0.005
04.11.2015	4	1013	0.23	0.0367	0.01
05.11.2015	5	1015	0.13	0.0255	0.003
06.11.2015	6	1022	0.43	0.0512	0.043
07.11.2015	7	1028	0.65	0.05	0.058
08.11.2015	8	1020	0.23	0.034	0.01
09.11.2015	9	1011	0.43	0.0499	0.052
10.11.2015	10	1001	0.43	0.0313	0.036
11.11.2015	11	994	0.48	0.0258	0.037
12.11.2015	12	998	0.36	0.0256	0.029
13.11.2015	13	1009	0.39	0.0429	0.028
14.11.2015	14	1019	0.39	0.0381	0.028
15.11.2015	15	1004	0.27	0.0279	0.014

Отмеченные данные: Концентрация текущая, г/куб. м

Границы выделения: верхнее левое: 2 3 Очистить

нижнее правое: 79 3 Направить

Очистить Установка

Исходный файл в таблице: **Гурьянова ноябрь 2015-январь 2016**

Одиночная обработка Пакетный прогноз Без производных

Интервал времени

Дата начала интервала: 13.01.2016

Дата конца интервала: 17.01.2016

Вид операции: ☒ Добавить ☐ Заменить

Вид интервала: ☐ Тестовый ☒ Прогнозный

Группировка АСКЗА и примесей

Станции АСКЗА: Гурьянова

Виды примесей: выборочно

☒ CO ☒ NO2 ☒ NO ☐ SO2 ☐ H2S ☐ CH4 ☐ PM2.5 ☐ PM-10

Добавить/изменить

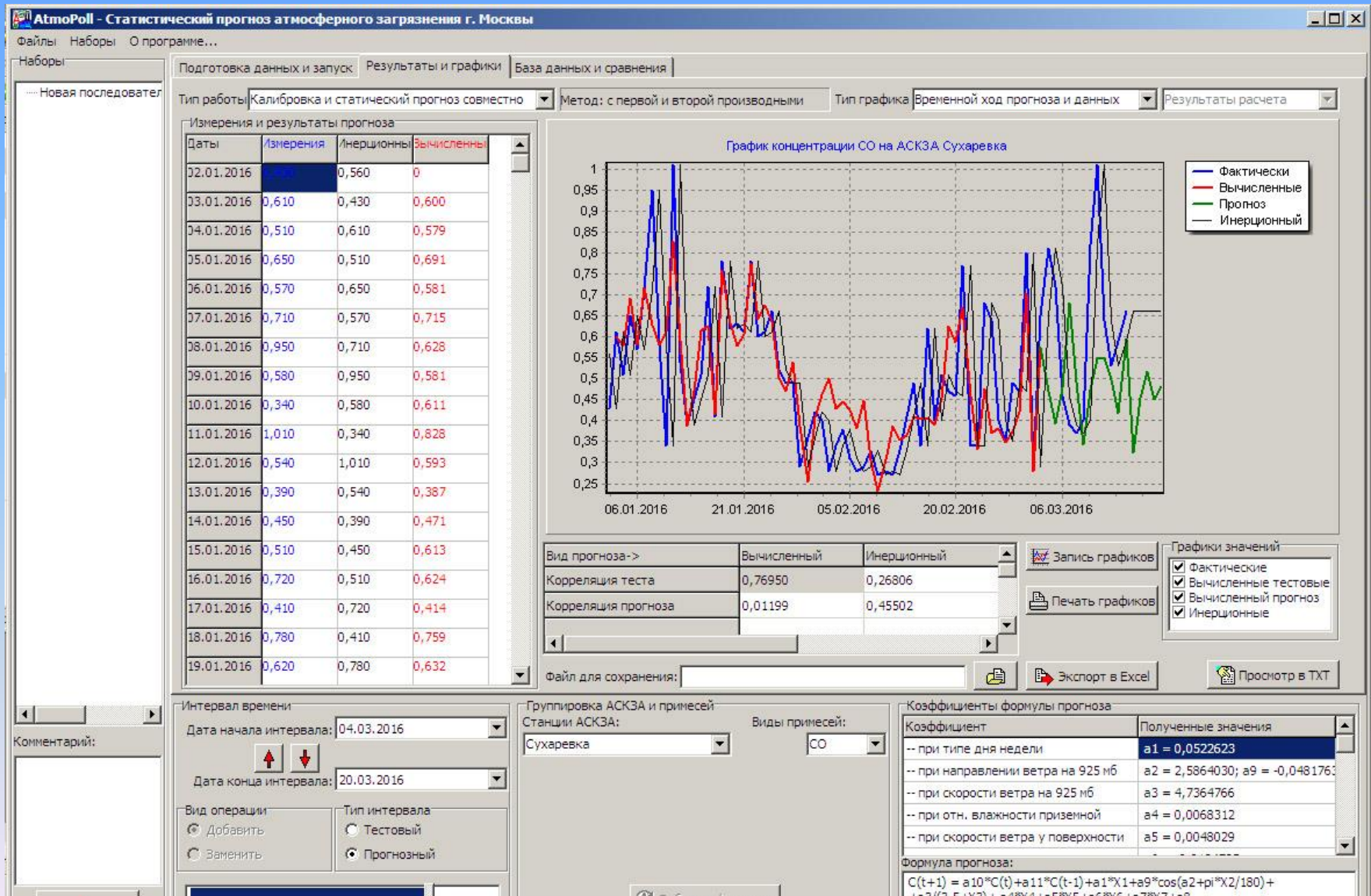
Коэффициенты формулы прогноза

Коэффициент	Полученные значения
-- при типе дня недели	a1=
-- при направлении ветра на 925 мб	a2= a9=
-- при скорости ветра на 925 мб	a3=
-- при отн. влажности приземной	a4=
-- при скорости ветра у поверхности	a5=

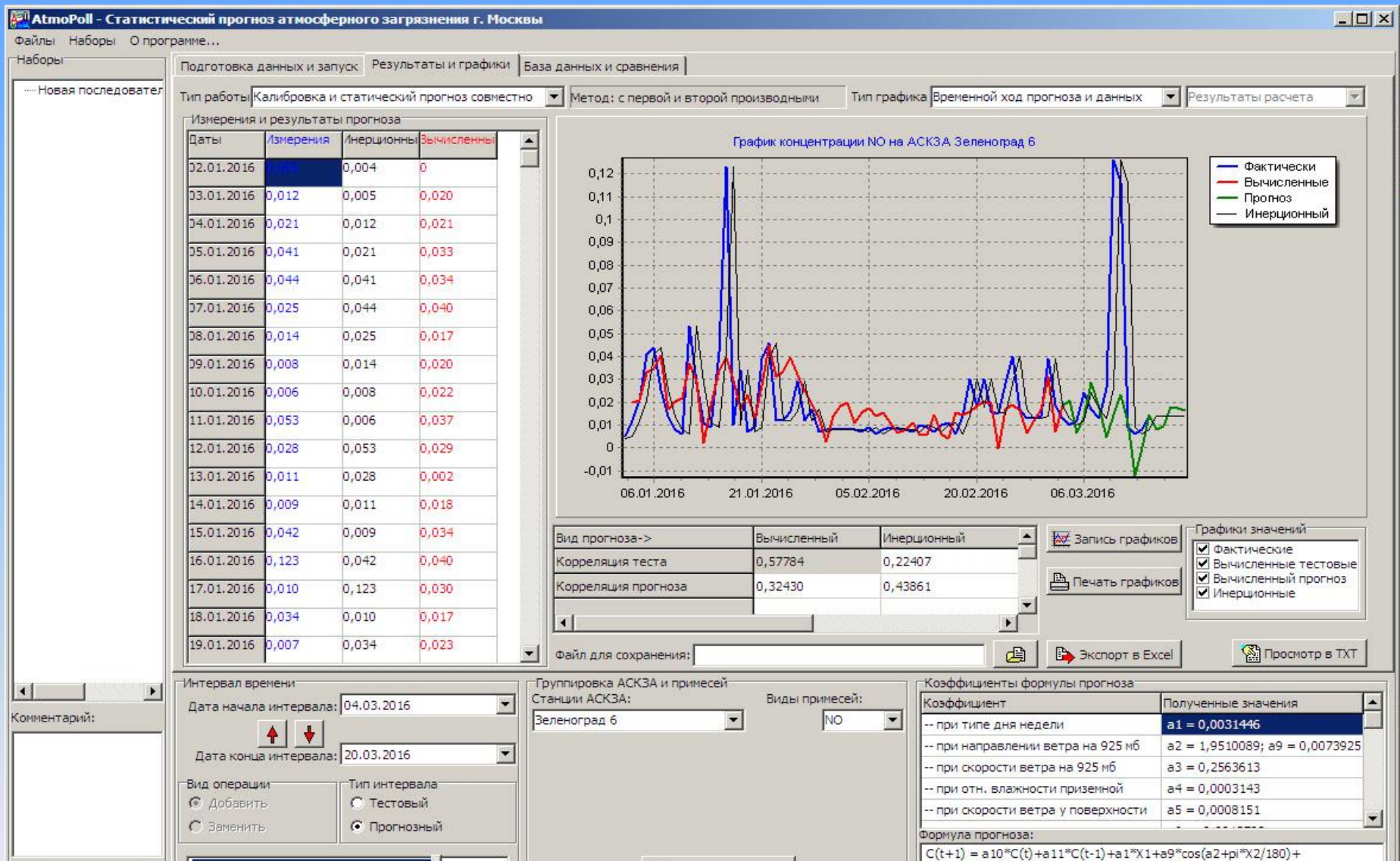
Формула прогноза:

$$C(t+1) = a10 * C(t) + a11 * C(t-1) + a1 * X1 + a9 * \cos(a2 + \pi * X2 / 180) + a3 / (3.5 + X3) + a4 * X4 + a5 * X5 + a6 * X6 + a7 * X7 + a8$$

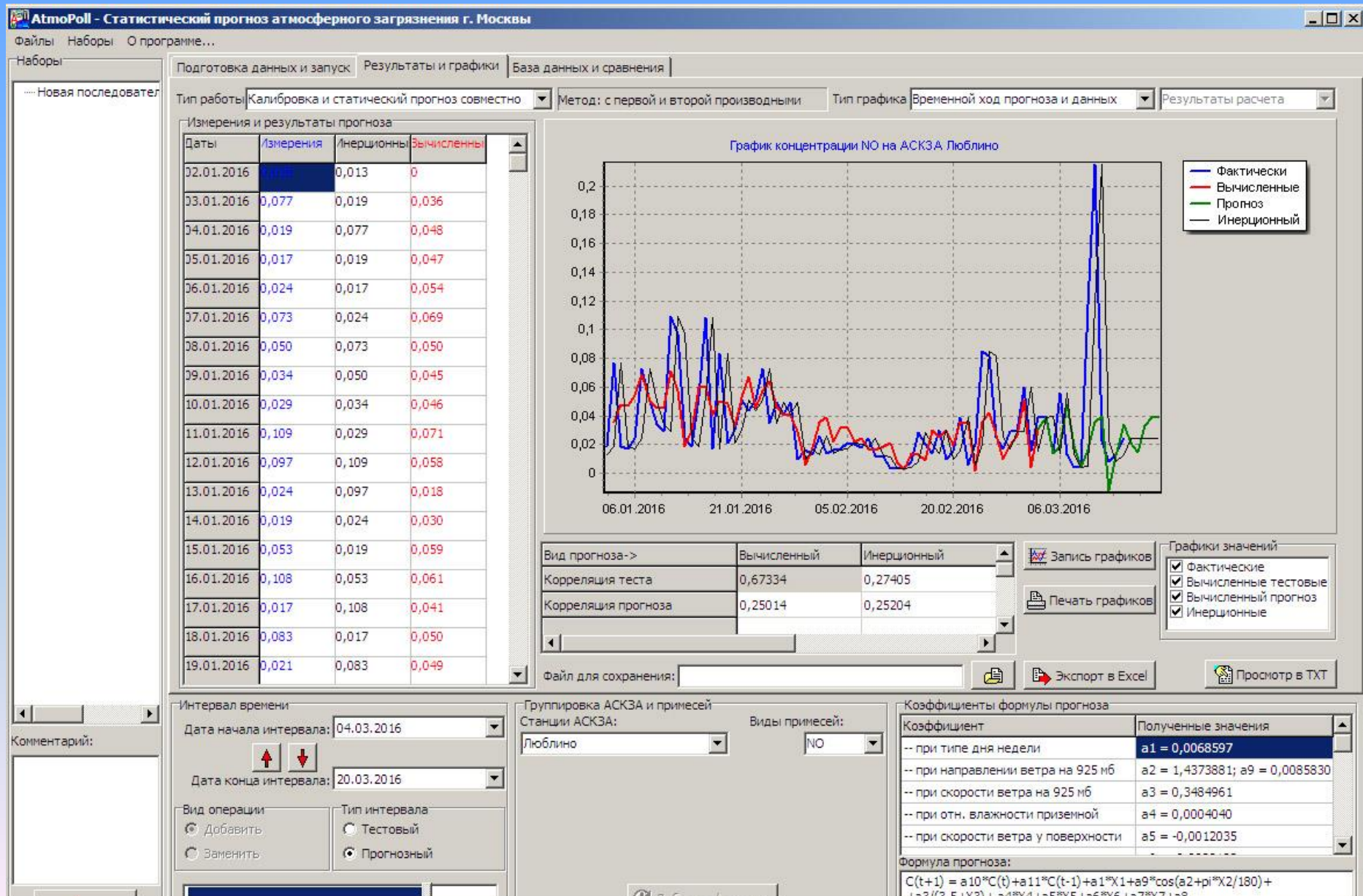
Результаты расчета коэффициентов прогностической формулы и прогноза концентраций



Результаты расчета коэффициентов прогностической формулы и прогноза концентраций



Результаты расчета коэффициентов прогностической формулы и прогноза концентраций





Выводы



- 1) разработана нелинейная формула статистического прогноза среднесуточных концентраций основных загрязнителей атмосферного воздуха в г. Москве;**
- 2) создан и верифицирован алгоритм вычисления коэффициентов формулы по данным с АСКЗА г. Москвы и значениям метеорологических параметров в течение обучающего периода;**
- 3) алгоритм может быть применен для прогноза среднесуточных концентраций загрязнителей атмосферы на период после обучающего, причем корреляция рассчитанных значений с данными наблюдений выше по сравнению с инерционным прогнозом;**
- 4) создано ПО AtmoPoll, реализующее как сам алгоритм прогноза, так и все операции по подготовке данных и интерпретации результатов.**



Спасибо за внимание !



**Детальные имитационные на
малых масштабах времени
(минуты, часы, сутки, декады)**

**Качественные «минимальные»
на больших масштабах
времени (месяц, год и более)**