

УДК 551.510.42:502.175

## ПРОСТРАНСТВЕННАЯ ИНТЕРПОЛЯЦИЯ КОНЦЕНТРАЦИИ ВЗВЕШЕННЫХ ЧАСТИЦ $PM_{10}$ МЕТОДОМ КОКРИГИНГА

*С.В. Крюкова, Т.Е. Симакина*

Российский государственный гидрометеорологический университет,  
tatiana.simakina@gmail.com

Определены значения концентрации взвешенных частиц  $PM_{10}$  в точках, для которых результаты измерений на сети экологического мониторинга Санкт-Петербурга отсутствовали. Реализован метод кокригинга с тремя дополнительными переменными — значениями концентрации оксидов азота  $NO$ ,  $NO_2$  и углерода  $CO$ . Выполнена оценка кластерности сети мониторинга путем построения гистограмм расстояния между станциями. Доказана меньшая однородность сети, состоящей только из станций с фактическими измерениями  $PM_{10}$  (10 станций), по сравнению с существующей (20 станций).

*Ключевые слова:* интерполяция, кригинг, кокригинг, взвешенные частицы  $PM_{10}$ , кросс-валидация, кластерность сети.

## SPATIAL INTERPOLATION OF PARTICULATE MATTER ( $PM_{10}$ ) CONCENTRATIONS USING THE COKRIGING METHOD

*S.V. Kryukova, T.E. Simakina*

Russian State Hydrometeorological University

The  $PM_{10}$  concentrations are determined at points where measurements in the St. Petersburg ecological monitoring network were absent. The cokriging method with three additional variables — the concentrations of nitrogen oxides  $NO$ ,  $NO_2$  and carbon monoxide  $CO$  has been implemented. The monitoring network clustering assessment is made by histogramming of the distances between stations, and a smaller uniformity of the network consisting of only stations with actual  $PM_{10}$  measurements (10 stations) compared to the existing (20 stations) is proved.

*Keywords:* interpolation, kriging, cokriging,  $PM_{10}$  particles, cross-validation, clustering of the network.

### Введение

Загрязнение воздуха крупных промышленных регионов остается острой экологической проблемой. Она представляет интерес и для Санкт-Петербурга, где расположено много предприятий, выбрасывающих в атмосферу вредные вещества, и ежегодно увеличивается парк автотранспорта. Мониторинг уровня загрязнения воздуха производится по значениям концентрации загрязняющих веществ (ЗВ) на всей территории, которые определяются с помощью интерполяции нескольких измеренных значений. Точность построенных таким образом карт распределения ЗВ обусловлена в основном количеством фактических данных и расположением пунктов их измерения.

Некоторые вещества измеряются в малом числе пунктов, что приводит к большим погрешностям в определении статистических параметров распределения данного ЗВ по территории. Для повышения точности интерполяции и стати-

стических оценок предлагается оценивать содержание этих веществ не только на основе непосредственного их измерения, но и по данным о других (дополнительных) переменных, коррелированных с искомыми (метод кокригинга) [3].

Целью настоящей работы являлась оценка возможности применения метода кокригинга для построения пространственного распределения значений концентрации взвешенных частиц ( $PM_{10}$ ) на территории Санкт-Петербурга. Массив измерений  $PM_{10}$  содержал значения, полученные лишь на десяти станциях экологического мониторинга из имеющихся двадцати. Однако на всех станциях измерялись другие ЗВ и метеорологические параметры, которые были использованы как дополнительные в модели кокригинга.

### Взвешенные частицы РМ

Термин «взвешенные частицы» в метеорологии и физике атмосферы относится к твердым веществам, диспергированным в воздухе [4, 5]. Используемая для их обозначения аббревиатура РМ — производная от английского *Particulate Matter*.

Взвешенные частицы являются совокупностью разнообразных химических соединений, отличаясь этим от остальных загрязнителей атмосферного воздуха — оксидов азота  $NO_x$ , угарного газа  $CO$ , озона  $O_3$ , оксида серы  $SO_2$ , каждое из которых представляет собой одно химическое вещество.

Среднедисперсные частицы с аэродинамическим диаметром менее 10 мкм ( $PM_{10}$ ) нормируются в РФ следующим образом: предельно допустимая максимальная разовая концентрация (ПДК<sub>мр</sub>) составляет 300 мкг/м<sup>3</sup> (20-минутное осреднение), предельно допустимая среднесуточная концентрация частиц в воздухе — 60 мкг/м<sup>3</sup> (24-часовое осреднение), предельно допустимая среднегодовая концентрация — 40 мкг/м<sup>3</sup> (годовое осреднение) [2].

Взвешенные твердые частицы поступают в воздух в основном при эксплуатации автотранспорта. Они представляют серьезную угрозу здоровью населения, поскольку способны адсорбировать на своей поверхности токсичные и канцерогенные вещества, а размеры этих микрочастиц позволяют им проникать в кровеносную систему человека, повышая риск бронхо-легочных, сердечно-сосудистых и онкологических заболеваний [8]. По оценкам, приведенным в работе [1], при каждом увеличении концентрации  $PM_{10}$  на 10 мкг/м<sup>3</sup> суточная смертность возрастает на 0,2—0,6 %. Взвешенные частицы с 2005 г. отнесены Всемирной организацией здравоохранения к приоритетным поллютантам воздуха наравне с  $NO_x$ ,  $O_3$ ,  $SO_2$ , формальдегидом. Однако вопросы загрязнения воздушной среды взвешенными среднедисперсными частицами в нашей стране изучены недостаточно, в отличие от зарубежных стран, где данные проблемы уже давно исследуются [9].

### Материалы и методы

Исходными данными для построения карты распределения взвешенных частиц на территории Санкт-Петербурга служил массив значений концентрации

загрязняющих веществ: оксида углерода (CO), оксида азота (NO), диоксида азота (NO<sub>2</sub>) и твердых примесей (PM<sub>10</sub>). Измерения выполнялись с дискретностью 20 минут двадцатью стационарными станциями, входящими в Автоматизированную систему мониторинга (АСМ) Санкт-Петербурга, схема размещения которых представлена на рис. 1.

Станции расположены в 18 административных районах города, функционируют непрерывно в автоматическом режиме и обеспечивают регулярное получение оперативной информации об уровне загрязнения атмосферного воздуха Санкт-Петербурга основными загрязняющими веществами [11]. 10 станций, на которых измерялась концентрация взвешенных частиц PM<sub>10</sub>, на рис. 1 выделены черным контуром.

Для интерполяции измерений PM<sub>10</sub> на всю исследуемую территорию использовались методы кригинга и кокригинга. Метод кригинга в настоящее время стал фундаментальным инструментом в области геостатистики. В нем учитывается не только удаленность точек, но и их взаимное расположение.

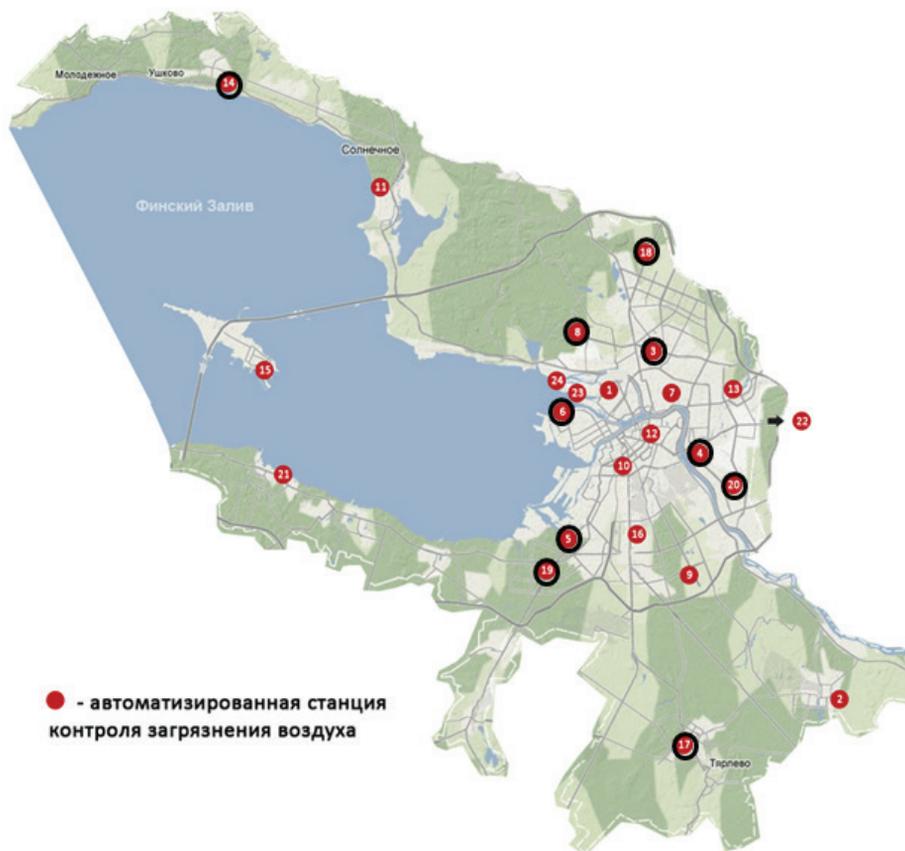


Рис. 1. Схема размещения экологических станций в Санкт-Петербурге.

В данной работе использовался метод ординарного кригинга, в основе которого лежит следующая формула:

$$Z = w_i Z(x_i),$$

где  $Z(x_i)$  — измеренное значение в точке  $i$ :  $Z(i) = \bar{Z} + \varepsilon(i)$ ;  $w_i$  — весовое значение для  $i$ -й точки;  $\bar{Z}$  — постоянное среднее значение;  $\varepsilon(i)$  — стандартная случайная ошибка исходных данных.

Веса определяются из эмпирической вариограммы, вычисляемой по формуле [3, с. 67]:

$$\gamma(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} [Z(x_i) - Z(x_i + h)]^2, \quad (1)$$

где  $Z(x_i)$ ,  $Z(x_i + h)$  — значение параметра, взятое в точке  $i$ , и через интервал  $h$  соответственно.

Вариограмма отражает тот факт, что значения переменной в близко расположенных точках пространственно коррелированы, в то время как в удаленных друг от друга точках они статистически независимы.

Кригинг — многошаговый процесс, состоящий из статистического анализа исходных данных, оценки пространственной структуры на основе пространственной корреляционной или структурной функции и создания интерполирующей поверхности.

Метод кокригинга или совместного кригинга производит интерполяцию переменной величины с учетом нескольких других переменных, коррелирующих между собой. Поскольку из двадцати станций экологического мониторинга, массив данных которых был в нашем распоряжении, только на десяти производились измерения  $PM_{10}$ , первой задачей было восстановление данных  $PM_{10}$  на других станциях с использованием измеренных там параметров. В этом случае речь идет о частичной гетеротопии данных — переменные имеют и общие, и различные точки измерений [3, с. 152].

Оценка обычного кокригинга функции  $Z_{\alpha 0}(x_0)$  — линейная комбинация значений различных переменных из окрестности точки  $x_0$ :

$$Z_{\alpha 0}^*(x_0) = \sum_{\alpha=1}^K \sum_{i=1}^{n_{\alpha}} w_i^{\alpha} Z_{\alpha}(x_i).$$

Для определения весовых значений строится кросс-вариограмма по экспериментальным точкам, разделенным вектором  $h$ , в которых есть данные измерений переменных  $Z_{\alpha}(x)$  и  $Z_{\beta}(x)$  [3, с. 146]:

$$\gamma_{\alpha\beta}(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} [Z_{\alpha}(x_i) - Z_{\alpha}(x_i + h)][Z_{\beta}(x_i) - Z_{\beta}(x_i + h)]. \quad (2)$$

Несмещенность вариации ошибки будет иметь место при следующем условии:

$$\sum_{i=1}^{n_{\alpha}} w_i^{\alpha} = \delta_{\alpha\alpha_0} = \begin{cases} 1, & \alpha = \alpha_0, \\ 0, & \alpha \neq \alpha_0. \end{cases}$$

Таким образом, сумма весов при основной переменной равна единице, а при каждой из дополнительных переменных — нулю. С учетом этого условия после минимизации вариации ошибки получим систему уравнений традиционного обычного кокригинга:

$$\begin{cases} \sum_{\beta=1}^K \sum_{j=1}^{n_{\beta i}} w_j^{\beta} \gamma_{\alpha\beta} (h_{ij}) + \mu_i = \gamma_{\alpha\alpha_0} (h_{i0}), \alpha = 1, \dots, K; i = 1, \dots, n_{\alpha}, \\ \sum_{j=1}^{n_{\alpha}} w_j^{\alpha} = \delta_{\alpha\alpha_0}, \alpha = 1, \dots, K. \end{cases}$$

### Кластерность сети мониторинга

На репрезентативность глобальной статистики — средних, вариаций, гистограмм — большое влияние оказывает наличие кластерных структур или разреженностей в сети мониторинга [6]. Наличие кластерности уменьшает точность и быстроту работы кригинга. Оценка топологии сети экологического мониторинга Санкт-Петербурга с точки зрения ее кластерности была выполнена путем построения гистограммы расстояний между станциями. На рис. 2 а представлена гистограмма для десяти станций, обозначенных кружками на рис. 1, где измеряются концентрации твердых примесей  $PM_{10}$ . Для сравнения на рис. 2 б показана гистограмма для двадцати станций, включающих те, на которых будут восстановлены значения  $PM_{10}$  с помощью кокригинга.

При равномерном распределении станций в пространстве число пар должно быть одинаково для всех расстояний. Уменьшение числа пар станций с увеличением расстояния вызвано граничным эффектом. Увеличение высоты столбцов с расстоянием на рис. 2 а свидетельствует о большой нерегулярности расположения точек сбора данных. На рис. 2 б пары распределены более равномерно, наблюдаемый пик числа пар на расстоянии от 7 до 14 км говорит о наличии кластера. Таким образом, сеть, состоящая из двадцати станций, обладает меньшей

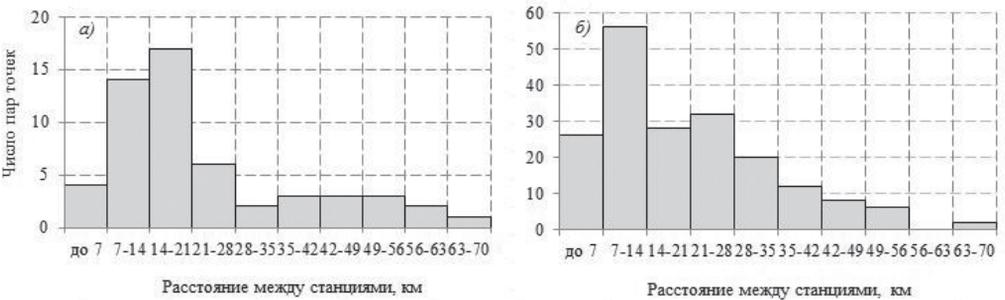


Рис. 2. Гистограмма расстояний между станциями экологического мониторинга в Санкт-Петербурге.

а — для десяти станций, б — для двадцати станций.

кластерностью, и, следовательно, учет восстановленных кокригингом значений  $PM_{10}$  на десяти станциях повысит точность пространственных статистических оценок распределения  $PM_{10}$ .

### Корреляция $PM_{10}$ с другими переменными

Воспользоваться методом кокригинга можно при высокой корреляции между восстанавливаемым параметром и теми параметрами, с помощью которых будет происходить восстановление. В работах авторов [7, 10] исследовалась зависимость концентрации ЗВ и метеорологических параметров в Санкт-Петербурге.

Для сравнения характера связи между всеми переменными, измеряемыми станциями экологического мониторинга, построены диаграммы рассеяния значений концентрации твердых примесей  $PM_{10}$  и остальных измеряемых параметров. На рис. 3 приведены значения концентрации за два дня января 2010 г. на одной

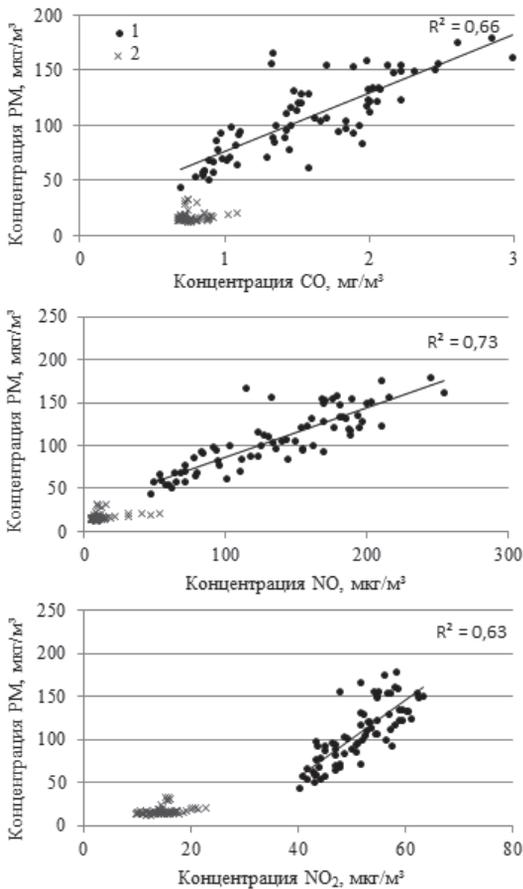


Рис. 3. Диаграммы рассеяния  $PM_{10}$  и других переменных за два дня с разным уровнем ЗВ.

1 — 23 января, 2 — 1 января.

из станций. Для сравнения взяты дни с максимально различающимся уровнем загрязнения: точками обозначены значения в день с наибольшей среднесуточной концентрацией (23 января), крестиками — в «чистый» день с наименьшей среднесуточной концентрацией (1 января). Хорошо прослеживается линейный характер распределения точек, соответствующих дню с высоким уровнем ЗВ. На остальных станциях картина аналогичная.

Для количественного исследования зависимостей между переменными были найдены коэффициенты парной корреляции между  $PM_{10}$  и другими переменными. Результаты расчета приведены на рис. 4. Столбцы гистограммы соответствуют ежедневным значениям коэффициента линейной корреляции концентрации  $PM_{10}$

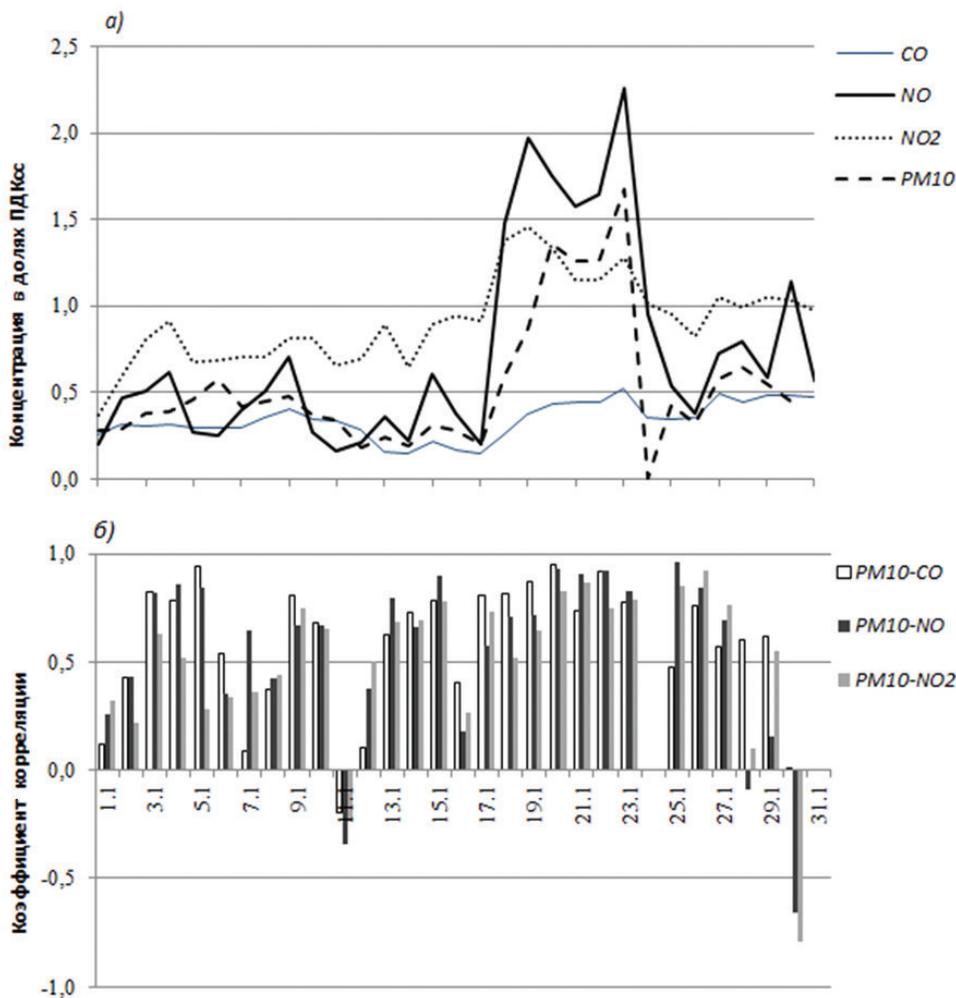


Рис. 4. Изменение среднесуточных значений концентрации ЗВ в долях ПДК<sub>сс</sub> (а) и коэффициента корреляции концентрации  $PM_{10}$  и ЗВ (б) в январе.

и оксидов углерода CO и азота NO, NO<sub>2</sub> за январь. Для сопоставления на график также нанесены среднесуточные значения концентрации ЗВ в долях ПДК<sub>сс</sub>. Заметно увеличение значений коэффициентов выше 0,8 в периоды роста загрязненности воздуха. Данные расчеты подтверждают, что более точное восстановление  $PM_{10}$  возможно в дни с высокой загрязненностью воздуха.

### Выполнение интерполяции и ее результаты

Пространственный характер данных отражают вариограммы переменных, представленные на рис. 5. По оси абсцисс отложено расстояние между отсчетами, на оси ординат — значения  $\gamma(h)$ , вычисленные по формуле (1), для всех пар точек, разделенных расстоянием  $h$ .

Экспериментальная вариограмм  $PM_{10}$  и аппроксимирующая ее сферическая модель приведены на рис 5 а, вариограммы дополнительных переменных — на рис. 5 б — г и кросс-вариограмма  $PM_{10} — NO — NO_2 — CO$ , вычисленная по формуле (2), — на рис. 5 д. Наиболее подходящей моделью экспериментальных вариограмм  $PM_{10}$ , NO и NO<sub>2</sub> оказалась сферическая, вариограмм CO и  $PM_{10} — NO — NO_2 — CO$  — модель Гаусса.

Вариограммы оксидов азота заметно отличаются от остальных характером поведения в нуле. Скачок в нуле, называемый наггет-эффектом, происходит в результате ошибок измерений и/или наличия неоднородностей меньшего масштаба, чем масштаб рассматриваемых данных. Вариограмма CO и кросс-вариограмма  $PM_{10} — NO — NO_2 — CO$  демонстрируют параболическое поведение вблизи нуля, что характеризует довольно регулярную структуру переменных.

Расстояние, на котором вариограмма достигает ограничивающего значения (плато), называется радиусом влияния. Данные, удаленные друг от друга на расстояние, превышающее радиус влияния, перестают быть коррелированными. Наличие плато у вариограммы означает, что процесс удовлетворяет стационарности второго порядка. Для рассматриваемых переменных  $PM_{10}$ , NO, NO<sub>2</sub> и CO плато приблизительно соответствует дисперсии концентраций.

Параметры всех вариограмм — наггет, плато и радиус влияния — приведены в табл. 1.

Таблица 1

Параметры авто- и кросс-вариограммы для  $PM_{10}$ , NO, NO<sub>2</sub>, CO

Переменная	Наггет	Плато	Радиус
$PM_{10}$ , (мкг/м <sup>3</sup> ) <sup>2</sup>	0	5900	330
NO, (мкг/м <sup>3</sup> ) <sup>2</sup>	3100	14140	270
NO <sub>2</sub> , (мкг/м <sup>3</sup> ) <sup>2</sup>	290	520	80
CO, (мг/м <sup>3</sup> ) <sup>2</sup>	0	0,5	180
$PM_{10} — NO — NO_2 — CO$ , 10 <sup>-6</sup> (мг/м <sup>3</sup> ) <sup>2</sup>	0	152	260

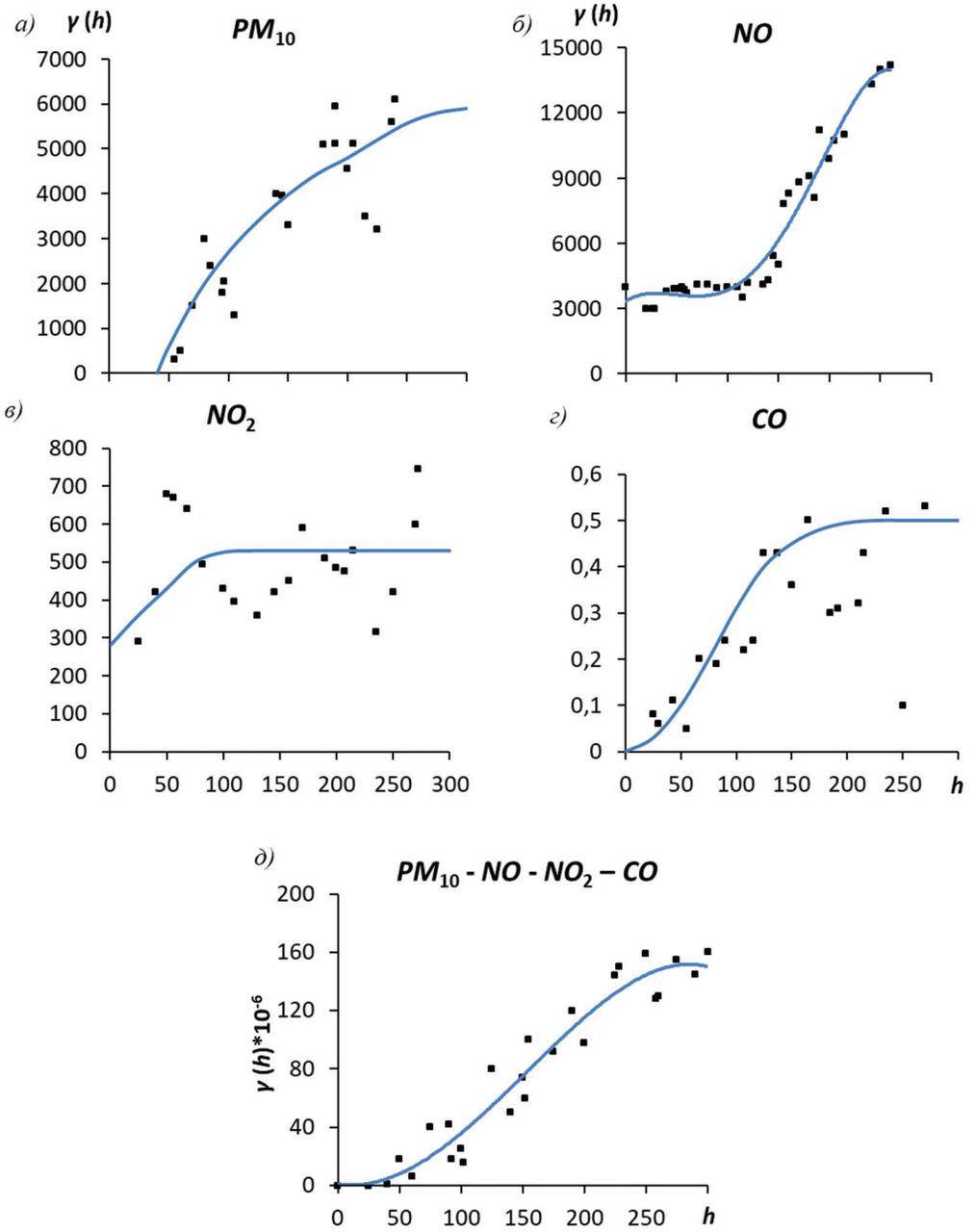


Рис. 5. Экспериментальные и модельные вариограммы.

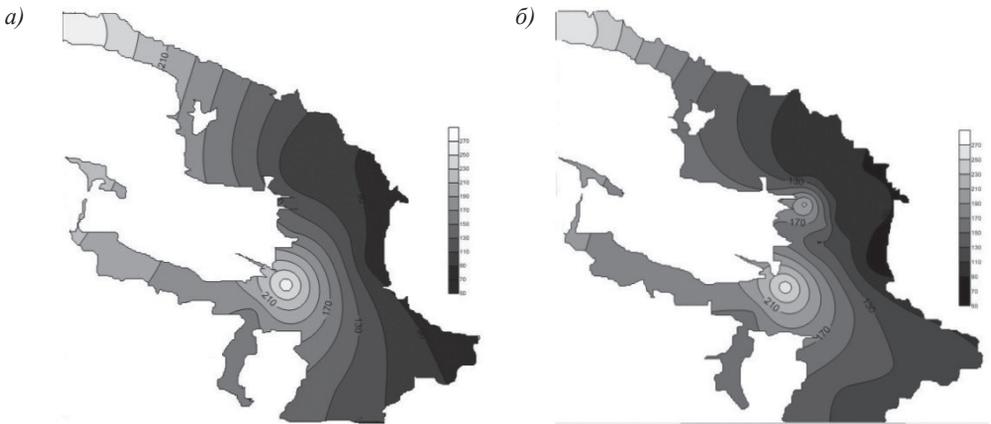


Рис. 6. Карты результатов оценки загрязнения  $PM_{10}$  территории Санкт-Петербурга, полученные с помощью кригинга (а) и с помощью кокригинга с тремя дополнительными переменными —  $CO$ ,  $NO$ ,  $NO_2$  (б).

Полученные модели с указанными в табл. 1 параметрами определяют веса в вычислениях кригинга и кокригинга. Сферическая модель вариограммы  $PM_{10}$  была использована при интерполяции методом кригинга, модель Гаусса кросс-вариограммы  $PM_{10} — NO — NO_2 — CO$  — методом кокригинга.

Результирующие карты оценки концентрации взвешенных частиц, построенные на прямоугольной сетке  $80 \times 90$  ячеек с размером ячейки  $0,9 \times 0,9$  км, приведены на рис. 6. На рис. 6 а представлена оценка загрязнения  $PM_{10}$  с использованием кригинга, на рис. 6 б — результаты кокригинга при совместном моделировании пространственной структуры  $PM_{10}$  и  $NO$ ,  $NO_2$ ,  $CO$ . При сравнении карт видно, что в случае использования кригинга произошло усреднение в центре города и остались не оцененными области по краям, что связано с отсутствием в этих областях данных измерений  $PM_{10}$ . Большая однородность и меньшая вариабельность оценок при использовании кригинга обусловлены завышением полученных значений в точках с низким загрязнением и его занижением в точках с высоким загрязнением. Метод кокригинга с использованием дополнительных данных в областях, где измерения  $PM_{10}$  отсутствуют, позволил получить существенно менее сглаженную оценку. На рис. 6 б видна локальная область значительного загрязнения ближе к северной границе города, которой не было на рис. 6 а.

Сравнение итоговых статистик оценки кокригинга и данных исходных измерений, приведенных в табл. 2, также подтверждает его превосходство над оценкой обычного кригинга. Стандартное отклонение по кригингу больше стандартного отклонения по кокригингу.

Статистические параметры и результаты кросс-валидации значений  $PM_{10}$ 

Параметр, мкг/м <sup>3</sup>	Кригинг	Кокригинг CO, NO, NO <sub>2</sub>
Среднее	143,6	145,8
СКО	64,8	52,7
Медиана	124,9	144,9
Минимум	75,5	75,5
Максимум	266,1	266,1
Средняя квадратическая ошибка	9,2	4,1

### Кросс-валидация для оценки результатов интерполяции

Оценка результатов трех процедур интерполяции выполнена с помощью метода кросс-валидации [3], или перекрестной проверки, который заключается в следующем:

1) из начального массива данных  $PM_{10}$  на десяти станциях случайным образом изымаем одну точку и проводим оценку в ней с использованием всех оставшихся точек;

2) полученное значение в изъятой точке сравниваем с известным истинным значением в этой точке и фиксируем ошибку интерполяции (невязку);

3) возвращаем изъятую точку в число данных, изымаем другую и фиксируем ошибку интерполяции уже в этой второй точке; выполнив поочередно данную процедуру для всех точек данных, получаем число невязок, равное числу точек.

Результаты кросс-валидации значений  $PM_{10}$ , полученные методами кригинга и кокригинга с использованием трех дополнительных переменных, также представлены в табл. 2. Добавление новой информации — значений концентрации ЗВ — привело к улучшению оценки.

### Заключение

В настоящей работе выполнено реконструирование значений концентрации взвешенных частиц в точках, для которых есть данные измерений концентрации ЗВ, коррелированных с  $PM_{10}$ . К зонам, где отсутствуют данные измерений  $PM_{10}$ , но имеются данные измерений концентрации CO, NO и NO<sub>2</sub>, относятся области на западе, востоке и в центре Санкт-Петербурга. Восстановление значений  $PM_{10}$  на всех станциях экологического контроля привело к увеличению однородности сети мониторинга и уменьшению ее кластерности, что позволило повысить репрезентативность статистических оценок, а также точность и быстроту работы кригинга.

Экологическая карта пространственного распределения  $PM_{10}$  для территории Санкт-Петербурга была построена методом кокригинга. При ее построении в расчет принимались не только данные измерений картируемой характеристики на станциях экологического мониторинга, но учитывалось влияние других ЗВ. В ре-

зультате применение методики кокригинга позволило преобразовать граничные районы из областей экстраполяции в области интерполяции и повысить точность интерполяции в центральной зоне.

Применение кокригинга целесообразно в случае частичной гетеротопии, когда есть достаточный поднабор изотопных данных для построения пространственных кросскорреляционных моделей. При этом использование дополнительных переменных увеличивает область интерполяции и уменьшает неопределенность оценки. С помощью кросс-валидации проведена оценка точности интерполяции методами кригинга и кокригинга, показано преимущество метода кокригинга с тремя дополнительными переменными ( $\text{CO}$ ,  $\text{NO}$ ,  $\text{NO}_2$ ).

Результаты кокригинга свидетельствуют о более высоком уровне загрязнения воздушного бассейна Санкт-Петербурга взвешенными частицами по сравнению с результатами, полученными методом кригинга.

### Список литературы

1. Воздействие взвешенных частиц на здоровье. Рекомендации в отношении политики для стран Восточной Европы, Кавказа и Центральной Азии / Всемирная организация здравоохранения. Европейское региональное бюро. 2013. [Электронный ресурс]. — Режим доступа [www.euro.who.int](http://www.euro.who.int)
2. Гигиенические нормативы ГН 2.1.6.2604-10. Предельно допустимые концентрации (ПДК) загрязняющих веществ в атмосферном воздухе населенных мест. Дополнение № 8 к ГН 2.1.6.1338-03. М., издание Минздрава России, 2010.
3. Демьянов В.В., Савельева Е.А. Геостатистика: теория и практика / Под ред. Р.В. Арутюняна. М.: Наука, 2010. 327 с.
4. Ивлев Л.С. Аэрозоли и облачность как экологический фактор // Региональная экология. 2013. № 1—2 (34). С. 130—137.
5. Ивлев Л.С. Аэрозоли и глобальные изменения климата // Общество. Среда. Развитие. 2013. № 1 (26). С. 225—231.
6. Каневский М.Ф., Демьянов В.В., Савельева Е.А., Чернов С.Ю., Тимонин В.А. Элементарное введение в геостатистику // Проблемы окружающей среды и природных ресурсов. 1999. № 11. 136 с.
7. Крюкова С.В., Симакина Т.Е. Оценка влияния метеорологических параметров на концентрации загрязняющих веществ в атмосфере Санкт-Петербурга // Актуальные проблемы гуманитарных и естественных наук. 2015. № 5—2. С. 299—305.
8. Лим Т.Е., Недре А.Ю., Недре Ю.А. Анализ заболеваемости болезнями органов дыхания взрослого населения Санкт-Петербурга в зависимости от качества атмосферного воздуха // Здоровье населения и среда обитания. 2014. № 1 (250). С. 26—28.
9. Ложкин В.Н., Невмержицкий Н.В., Ложкина О.В. О физических особенностях моделирования автомагистрали как источника чрезвычайно опасного воздействия частиц  $\text{PM}_{2.5}$  —  $\text{PM}_{10}$  // Проблемы управления рисками в техносфере. 2015. № 3 (35). С. 107—113.
10. Симакина Т.Е., Крюкова С.В. К вопросу о загрязнении воздуха г. Санкт-Петербурга в дни с инверсиями // Актуальные вопросы современной науки. Раздел «Естественные науки (Науки о Земле)». 2015. № 2 (6). С. 45—48.
11. Экологический портал Санкт-Петербурга [Электронный ресурс]. Режим доступа <http://www.infoeco.ru/>