

and sulfur. The geographical domain of PEEХ covers natural and urban environments of the Northern Eurasian region. China is included due to its crucial importance on its climate – air quality interactions. Siberia and Arctic Ocean are at the core geographical region within the PEEХ domain. The majority of the PEEХ domain is within Russia and China.

---

УДК 004.043+004.67+551.582.2

А.М. Новикова<sup>1</sup>, А.Б. Полонский<sup>2</sup>, А.А. Новиков<sup>3</sup>

## СРАВНЕНИЕ ВОЗМОЖНОСТЕЙ ИНТЕРПОЛЯЦИОННЫХ МОДУЛЕЙ QGIS ДЛЯ МОРСКИХ КЛИМАТИЧЕСКИХ ИССЛЕДОВАНИЙ ПРИ РАБОТЕ С МАССИВОМ ДАННЫХ МАЛОЙ ОБЕСПЕЧЕННОСТИ

***Резюме.** В статье с позиции геоинформационного и геостатистического подходов показана актуальность активного применения современных методов пространственного анализа океанологических данных. Проанализированы возможности некоторых статистических модулей открытой ГИС (QGIS) для решения практических задач оперативной оценки качества данных. Проведена оценка интерполяционных модулей QGIS, использующих методы кригинга и радиальных базисных функций, при работе с массивом данных малой обеспеченности.*

***Ключевые слова:** QGIS, статистические модули, кригинг, регуляризованный сплайн с натяжением, бикубическая сплайн-интерполяция с регуляризацией Тихонова.*

**Введение.** Несмотря на значительный объём океанологических данных, получаемых дистанционными методами зондирования, а также другими современными способами исследования акваторий, существует проблема неравномерности распределения данных для изучения океанологических процессов даже климатического масштаба, которую обычно решают при помощи различных методов интерполяции. В океанологии наиболее широко используется метод оптимальной интерполяции, разновидности которого реализованы в различных геоинформационных системах (ГИС). ГИС могут предложить практичную рабочую среду для интеграции, анализа и визуализации климатических и океанографических данных, несмотря на то, что в этой области ГИС пока еще мало используются (подробнее в [Новикова, 2015]). В настоящей работе приводится опыт использования открытой ГИС QGIS для океанографических приложений на примере построения климатических карт температуры поверхностного слоя Чёрного моря с использованием массива данных малой обеспеченности.

Для достижения поставленной цели в QGIS можно задействовать несколько модулей. Некоторые из них являются приложениями самой QGIS, например, «Модуль интерполяции» и «Геоалгоритмы QGIS», но большая часть задействует ресурсы сторонних ГИС, таких как SAGA и GRASS, а также осуществляет доступ к библиотекам GDAL/OGR через панель «Инструменты анализа». Так как встроенные модули QGIS и геоалгоритмы GDAL/OGR изначально создавались для получения цифровых моделей рельефа, их использование для интерполяции данных наблюдений, характеризующихся интенсивной пространственно-временной изменчивостью, проблематично. Гораздо более широкое применение находят модули GRASS-геоалгоритмы и SAGA-геоалгоритмы, где в качестве интерполяторов представлены различные варианты сплайнов (радиальных базисных функций) и кригинга.

---

<sup>1</sup> ФГБНУ «Институт природно-технических систем», отдел актуальных проблем океанографии, Севастополь, 299011, Россия, старш. инж.-исслед.; e-mail: anelkapi@bk.ru.

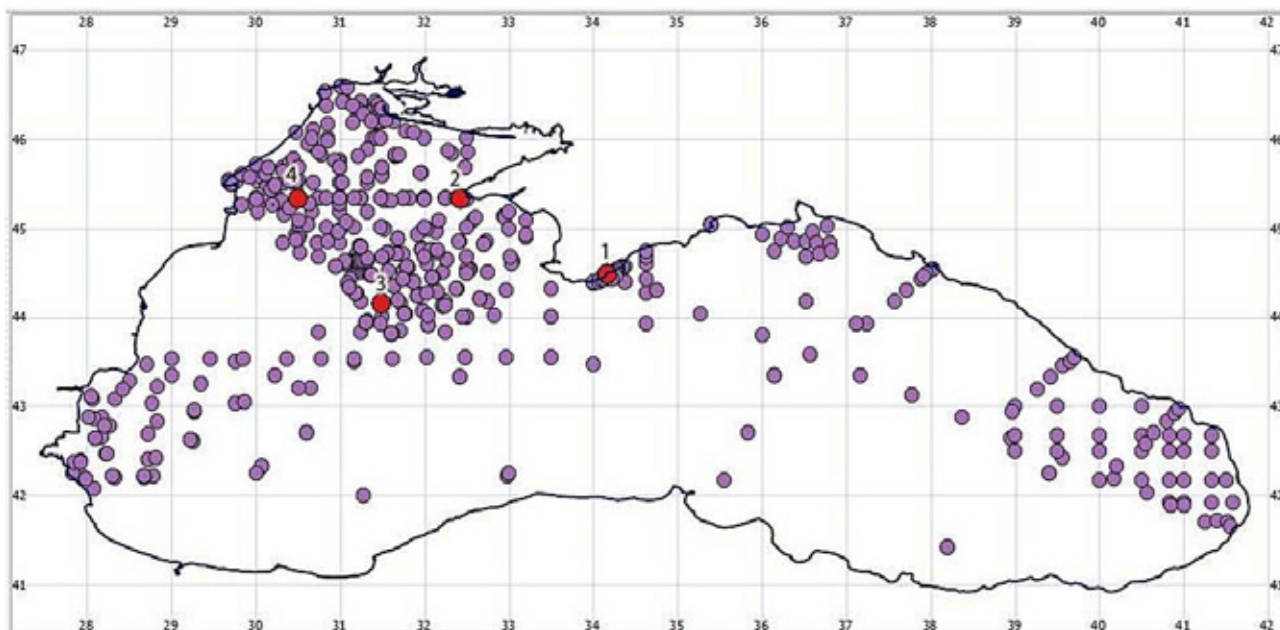
<sup>2</sup> ФГБНУ «Институт природно-технических систем», Океанографический центр, Севастополь, 299011, Россия, руководитель, профессор, докт. геогр. н.; e-mail: apolonsky5@mail.ru.

<sup>3</sup> Филиал Московского государственного университета им. М.В. Ломоносова в г. Севастополе, отделение географии, Севастополь, 299001, Россия, старш. преподаватель; e-mail: a\_novik@bk.ru.

Нами оценивались возможности следующих плагинов QGIS, позволяющих провести пространственную интерполяцию океанологических характеристик при различном разрешении наблюдательной системы: GRASS-геоалгоритмы «Билинейная/бикубическая сплайн-интерполяция с регуляризацией Тихонова» (v.surf.bspline) и «Регуляризованный сплайн с натяжением» (v.surf.rst); SAGA-геоалгоритмы «Обычный кригинг» и «Универсальный кригинг».

**Материал и методы исследований.** QGIS обладает широким спектром функций по статистическому анализу, проведению выборок из массивов данных, а также пространственной геостатистической обработке. Представлены такие функции как в собственных модулях программы, так и в плагинах, подключающих возможности сторонних ГИС. В ходе работы с массивом океанологических данных были, в первую очередь, использованы возможности «Таблицы атрибутов» слоя, где с помощью инструмента «Калькулятор полей» осуществлена выборка и выполнены расчёты на основе существующих значений атрибутов или заданных функций. Для оценки качества данных использовались различные статистические плагины: «Group Stats», «Statist» – для векторных слоев, «Зональная статистика» – для растровых. Все они рассчитывают базовые статистические характеристики (среднее, максимальное и минимальное значения, стандартное отклонение и пр.), которые можно получить или для группы параметров, отфильтровав данные по необходимому показателю («Group Stats»), или представив значения параметра на графике его распределения («Statist»), или в пределах области поиска для растрового файла, у которого каждой ячейке соответствует определенное значение анализируемого параметра («Зональная статистика»).

Для оценки возможностей интерполяционных плагинов и влияния различных значений вводимых параметров на результат интерполяции был выбран массив по температуре поверхностных вод Чёрного моря из базы океанографических данных МГИ за 1923–2011 гг. – с малой плотностью данных, неравномерно распределенных по акватории за июль-август 1989 г. (рис. 1). В этот период всего было выполнено 554 измерения температуры поверхности. Массив был подвергнут предварительному анализу. Вначале удалялись данные с ошибочными координатами. После проверки соответствия выборочных гистограмм нормальному закону распределения из массива были удалены экстремальные значения с использованием критерия  $\pm 3\sigma$ . Это дало возможность получить более гладкие климатические поля температуры, характеризующиеся небольшими значениями стандартного отклонения ( $\sigma = 1,72$ ).



**Рис. 1.** Распределение данных по акватории Чёрного моря за июль-август 1989 г.  
Красные точки – станции, использованные для кросс-валидации

В океанологии для интерполяции наиболее часто используется метод кригинга (см. формулу (1) ниже), который можно считать разновидностью метода оптимальной интерполяции. Реже используется метод радиальных базисных функций. К радиальным функциям (см. формулу (6) ниже) относят различные сплайны. В частности, в данной работе использовались регуляризованный сплайн с натяжением (RST) и бикубический сплайн с регуляризацией Тихонова (Bspline). Несмотря на отличие теоретических подходов и формул, методы кригинга и радиальной интерполяции приводят к одной и той же модельной функции при условии использования одинаковой вариограммы и радиальной функции [Кошель, 2004]. Этот факт может быть строго доказан математически, и одно из таких доказательств приведено в работе основателя геостатистики Матерона [Matheron, 1981].

Термин «кригинг» служит для обозначения семейства алгоритмов линейной пространственной регрессии. Он происходит от фамилии инженера Д. Крига, который первым применил интерполятор на основе модели пространственной корреляции данных. Л.С. Гандин независимо от Д. Крига применил аналогичный метод и сформулировал теорию оптимальной интерполяции для объективного анализа метеополей [Демьянов, 2010].

Все интерполяторы семейства кригинга являются различного рода модификациями базового линейного регрессионного оценителя  $Z^*(x)$ , определяемого следующим образом:

$$Z^*(x) - m(x) = \sum_{i=1}^{n(x)} \lambda_i(x) [Z(x_i) - m(x_i)], \quad (1)$$

где  $\lambda_i(x)$  – весовые коэффициенты, относящиеся к данным  $Z(x_i)$ . Величины  $m(x)$  и  $m(x_i)$  являются математическими ожиданиями случайных переменных  $Z(x)$  и  $Z(x_i)$ . Число данных, использующихся при оценке, и значения весовых коэффициентов могут меняться в зависимости от местоположения оцениваемой точки  $x$  [Демьянов, 2010]. При условии, что процесс имеет стационарное математическое ожидание, процедура оценивания называется обычным кригингом. Если же процесс не является стационарным (например, в данных присутствует значимый тренд), используют процедуру универсального кригинга [Кошель и др., 2001].

Кригинг – один из наиболее гибких и часто используемых методов, однако на множествах большого размера он работает медленно. Основная проблема кригинга – необходимость подбора параметров вариограммы, от которых зависит результат интерполяции. На первом этапе вычислений по исходным данным подбирается выборочная (экспериментальная) вариограмма  $\gamma^*(h)$  и строится график функции, где каждому интервалу значений  $h$  (разность расстояний между парами точек; имеет размерность в единицах карты – °) соответствует вариация (квадрат разности значений величин в этих точках);  $h$  откладывается по оси  $x$  или  $y$  в предположении, что поле изотропно. Выборочную вариограмму нельзя напрямую использовать в уравнениях кригинга, её необходимо приблизить некоторой модельной функцией вариограммы (см. формулы (2)-(5) ниже), которая используется на втором этапе вычислений. Наилучшим вариантом подобранной модельной функции можно считать тот, который дает наименьшую дисперсию отклонений эмпирических значений от теоретических [Поротов, 2006].

Для анализируемого набора данных по температуре поверхностных вод Чёрного моря использовались следующие формулы модельных функций вариограмм:

- линейная функция:

$$\gamma(h) = a + b \cdot x, \quad (2)$$

- квадратическая параболическая функция:

$$\gamma(h) = a + b \cdot x + c \cdot x^2, \quad (3)$$

- кубическая функция:

$$\gamma(h) = a + b \cdot x + c \cdot x^2 + d \cdot x^3, \quad (4)$$

- экспоненциальная функция

$$\gamma(h) = a + b \cdot \left(1 - \exp\left(-\left(\frac{x}{b}\right)^2\right)\right), \quad (5)$$

где  $a$  – так называемый «эффект самородка», определяемый случайной составляющей измерений, которая показывает насколько велико различие значений в очень близко расположенных точках; фактически эта величина характеризует степень «зашумленности» параметра [Демьянов, 2010; Поротов, 2006];  $b$ ,  $c$ ,  $d$  – коэффициенты;  $x$  – определен в пределах радиуса влияния вариограммы, где функция имеет максимальный коэффициент достоверности аппроксимации. Значения параметров вариограммы определяют весовые коэффициенты  $a$ ; ближайших соседей оцениваемой точки в (1), а также среднее квадратичное отклонение оценки от истинного значения.

Интерфейс окна кригинга в QGIS позволяет провести его тонкую настройку, в том числе задать формулу вариограммы и её ряд, а также настроить оптимальный охват данных (необходимый и достаточный для получения качественного поля при минимальных временных затратах) и тип показателя качества (стандартное отклонение или вариация).

Кроме вышеупомянутых задаются следующие параметры кригинга и вариограммы: точечный слой и его атрибут для анализа; логарифмическая трансформация и блочный кригинг (с указанием размера блока); количество точек, учитываемых для построения поля (задается максимальное количество точек для расчета или все точки в пределах радиуса поиска); максимальное расстояние поиска (радиус, в единицах измерения карты); предварительное создание качественной сетки, на которую будут наноситься результаты (если в таблице более 300 точек, при отсутствии готовой сетки процесс анализа затягивается во времени или выдается ошибка); размер ячейки сетки (оптимальный размер для каждого набора данных и масштаба анализируемого явления ускоряет процедуру интерполяции); области поиска (узлы или ячейки); количество классов лага расстояния; диапазон поиска (локальный или глобальный).

В результате работы плагина исследователь получает два поля – анализируемого параметра и ошибки процедуры кригинга, что позволяет оценить достоверность результатов как по абсолютным значениям стандартного отклонения, так и по конфигурации поля стандартного отклонения. Стандартное отклонение, как тип показателя качества, в большей мере определяет качество самих данных и частично – процедуры кригинга при различных параметрах. Ошибки кригинга зависят от плотности сети измерений (вариация кригинга меньше в областях с большей плотностью измерений). Зная вариацию кригинга, можно делать выводы о точности получаемого поля.

Метод радиальных функций объединяет целую группу методов, которые изначально развивались самостоятельно и позднее были объединены под одним названием, когда было замечено, что решение многих задач приводит к их математической формулировке в одной и той же форме (6). Модельная интерполяционная функция имеет вид:

$$f(\mathbf{p}) = \sum_{i=1}^n \lambda_i R(r_i) + P_k(\mathbf{p}); \quad r_i = |\mathbf{p} - \mathbf{p}_i| \quad (6)$$

где  $R(r)$  – функция одной переменной, называемая радиальной, а  $P_k(\mathbf{p})$  – полином степени  $k$  [Кошель, 2004].

Регуляризованные сплайны с натяжением (RST) принадлежат к интерполяционной функции, которая сводят к минимуму отклонения от измеренного значения функции в точке. Функция, которую использует RST сплайн описана в [Petrasova et al., 2015], а модуль «v.surf.rst», использующий её, формально эквивалентен универсальному кригингу с выбором

ковариационной функции, определяемой параметром сглаживания [Neteler et al., 2008]. Параметр натяжения в модуле регулирует приближение получаемой поверхности к исходным данным. Чем больше натяжение, тем больше шумов и больше деление изолиний на мелкие контуры; меньшие значения натяжения увеличивают площадь влияния отдельной точки в интерполяционном поле и приводит к большему сглаживанию. Регулируя натяжение, можно получить более гладкую, общую или более подробную шероховатую поверхность. Слишком малые значения натяжения приводят к большим погрешностям в областях с быстрым изменением пространственного градиента характеристики.

Большие массивы данных модуль разбивает на дерево квадрантов (прямоугольных сегментов) по заданному пользователем количеству точек «segmax», входящих в сегмент. Интерполяция проводится по каждому сегменту с учетом точек внутри (параметр «segmax») и снаружи сегмента (параметр «prmin», который должен быть больше, чем «segmax») для достижения неразрывности функции. Анизотропия поля может быть учтена через угол анизотропии «theta» и коэффициент масштабирования «scalex».

Так как модуль проверяет массив на наличие близких по пространству точек (параметр «dmin») и удаляет дублирующие точки, необходимо устанавливать значения «dmin» на 0. Если же между точками, наоборот, слишком большой пространственный разрыв, то параметром «dmax» устанавливает расстояние между двумя точками, при превышении которого будут вычислены дополнительные точки между исходными для дальнейшей интерполяции. Модуль позволяет проводить интерполяцию с использованием фильтра данных (через SQL-запрос) и растровой маски-слоя, а также получать на выходе растры с топографическими характеристиками.

Функция билинейной/бикубической сплайн-интерполяции с регуляризацией Тихонова описана в [Mitasova et al, 2005]. Регуляризирующая функция сплайна предполагает, что существует ошибка в экспериментальных данных, которую необходимо локально сгладить. Эту функцию выполняет параметр регуляризации  $\lambda$  («lambda\_i»). Регуляризация осуществляется путем сведения к минимуму наклона или кривизны интерполяционной функции. Если выбраны слишком малые значения  $\lambda$ , функция сплайна плохо обусловлена в районах с малым количеством или отсутствием данных. Если выбраны большие значения  $\lambda$ , получается более гладкая поверхность.

Этот вид сплайн-интерполяции доступен через плагин QGIS «GRASS-геоалгоритмы», в который включены три вида модуля «v.surf.bspline», один из которых – «v.surf.bspline.lambda» позволяет вычислить несколько значений параметра регуляризации Тихонова «lambda\_i», используя метод кросс-валидации, изымающий поочередно значения из исходных данных и вычисляющий при этом смещение  $\Delta m$  (см. формулу (8) ниже) и среднюю квадратичную ошибку (см. формулу (9) ниже) для каждого значения параметра «lambda\_i».

Для проведения корректной процедуры интерполяции необходимо подобрать значения таких параметров, как длина шага сплайна, охват карты, алгоритм сплайна, параметр регуляризации Тихонова, размер ячеек интерполяционной сетки (задается в единицах карты).

Для оценки погрешности полей, полученных интерполяционными модулями, использовался метод кросс-валидации, позволяющий получить значения невязки (7) на контрольных станциях (рис. 1), а также среднеквадратического отклонения (8) и смещения (9) для всего поля. Кросс-валидация – метод подбора оптимальных параметров модели интерполяции при помощи оценки значения в точке измерения без учета самого измерения в этой точке [Демьянов, 2010]. Невязка – разница между измеренными и оцененными значениями:

$$\Delta Z(x) = Z(x) - Z^*(x). \quad (7)$$

Глобальные характеристики кросс-валидации:  
- смещение:

$$\Delta m = m - m^*, \quad (8)$$

где  $m$  – среднее по исходным данным,  $m^*$  – среднее по полученным результатам;

- средняя квадратичная ошибка (root mean square error – RMSE):

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^n [Z(x_i) - Z^*(x_i)]^2}. \quad (9)$$

Так как в QGIS не во всех интерполяционных модулях есть встроенный инструмент кросс-валидации, были выбраны несколько станций измерения, в каждой из которых производилось более 4 наблюдений за температурой поверхностных вод. По этим станциям проведено осреднение с использованием всех значений. Затем контрольные станции были удалены из массивов данных и проведена контрольная процедура интерполяции для каждого набора параметров интерполяционных функций. Наконец экспериментальные средние значения температуры сравнивались с проинтерполированными значениями температуры в соответствующих точках поля, а также вычислялась средняя температура поверхностных вод каждого интерполяционного поля.

**Результаты исследований и их обсуждение.** Кригинг обычный и универсальный. Так как процедура кригинга в QGIS зависит от множества параметров, то в первую очередь для анализируемых наборов данных определялись значения основных параметров, одинаковых для всех моделей вариограмм:

- размер ячейки сетки – в ходе сравнения результатов интерполяции по сетке с разной размерностью установлено, что размер ячейки, равный  $0,01^\circ$ , позволяет получить достаточно подробную карту распределения температур вод Чёрного моря при наименьших временных затратах;

- количество классов лага расстояния – варьирует для каждого набора данных и определяется в зависимости от плотности сети наблюдений и количества данных;

- диапазон поиска – локальный с учетом неоднородности пространственной структуры поля исходных данных и определяемого максимального расстояния поиска;

- направление поиска – полный радиус;

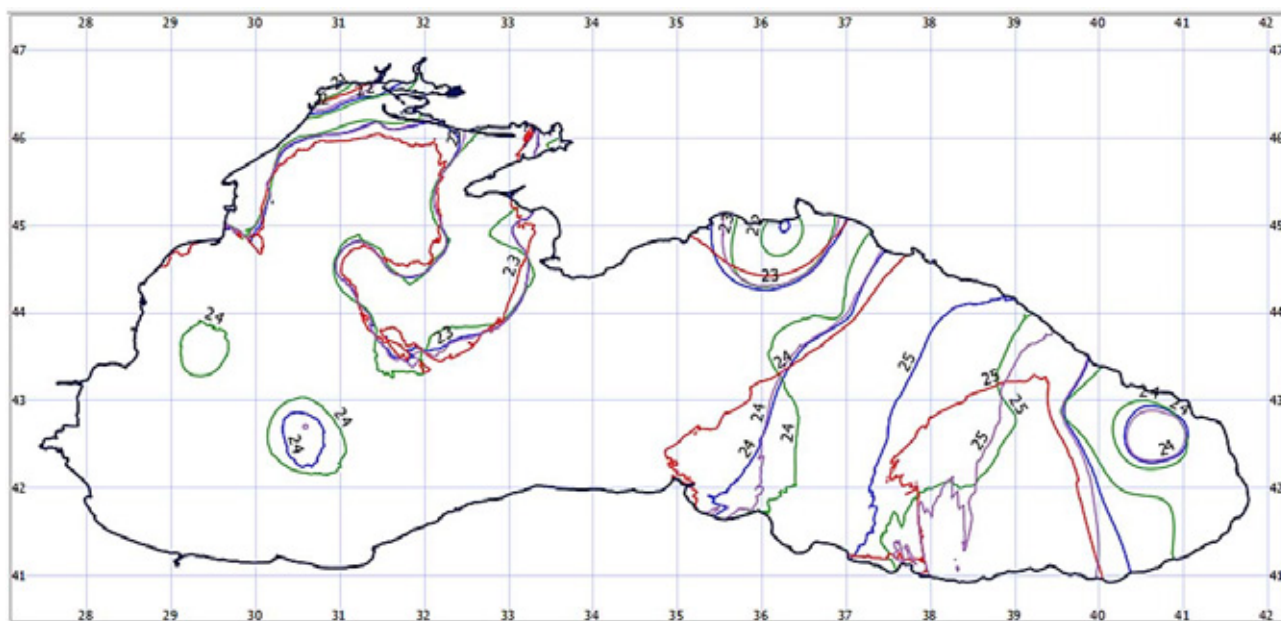
- максимальное расстояние поиска – определяется исходя из распределения точек по акватории согласно необходимости описания крупномасштабной структуры поля и принимается равным  $3^\circ$ . Большие значения радиуса приводят к значительному увеличению времени обработки при том же результате интерполяции, меньшие – к увеличению зашумленности поля;

- максимальное количество точек для расчета – подбор параметра показал, что с увеличением количества точек при одинаковом радиусе охвата уменьшается зашумленность поля и сглаживаются изолинии (сравнивались поля получаемые, при количестве точек для анализа, равном 40, 60, 80 и 100). Вариация ошибки при этом увеличивается в южной части моря, где плотность данных меньше, однако уменьшается общая погрешность восстановления поля. Для имеющихся наборов данных достаточно выбрать максимальное количество точек для расчета, равное 80. Если взять для расчета все точки в заданном радиусе поиска, то проявляется один из недостатков кригинга – на множествах большого размера он работает медленно.

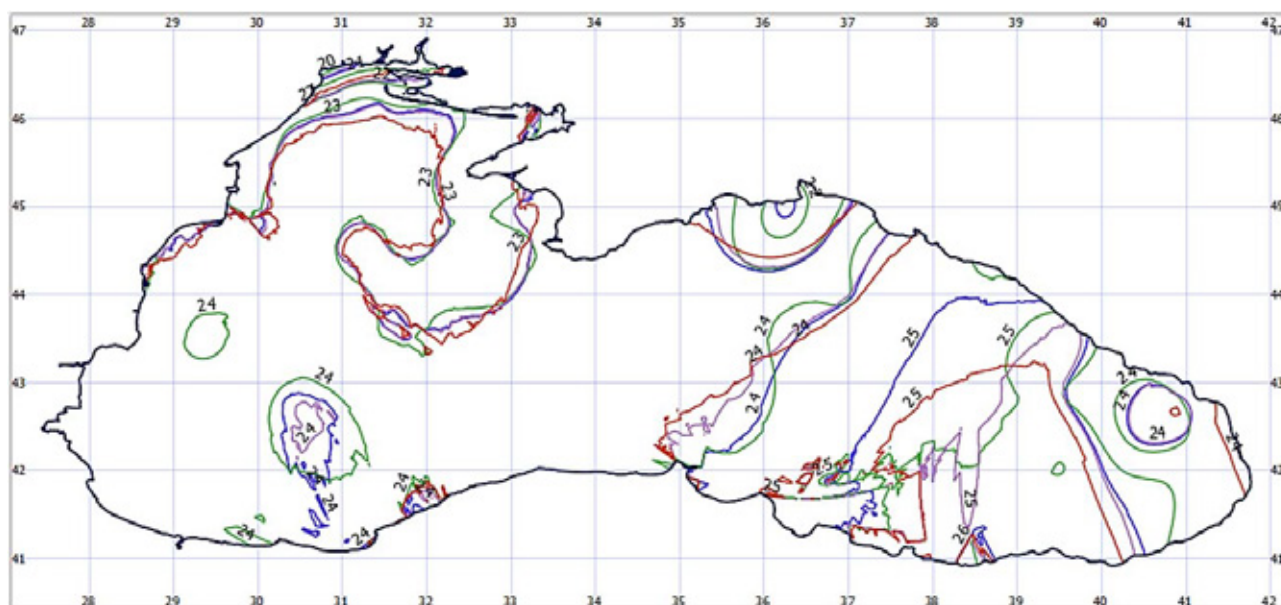
Так как массив неоднороден как по пространственному распределению данных наблюдений, так и по экспериментальной вариограмме, то для процедуры кригинга проводился подбор радиуса влияния вариограммы для каждого вида модельной вариограммы ((2)-(5)) исходя из максимального коэффициента достоверности аппроксимации функции (Кд) и конфигурации получаемого поля.

Для оценки погрешности обычного и универсального кригингов методом кросс-валидации были выбраны поля температуры, наиболее приближенные к реальным условиям, с наименьшими значениями вариации ошибки кригинга при различных функциях вариограмм (рис. 2). Для линейной функции модельной вариограммы значение радиуса влияния принято равным  $2,13^\circ$  (Кд=88%), квадратической –  $2,06^\circ$  (Кд=91%), кубической –  $7,34^\circ$  (Кд=80%), экспоненциальной –  $3^\circ$  (Кд=83%).





(а)



(б)

**Рис. 2.** Поля температуры поверхностных вод за июль-август 1989 г., полученные методом (а) обычного, (б) универсального кригинга для разных моделей вариограмм: изолинии синего цвета – линейная модель; красного – квадратическая; фиолетового – кубическая; зеленого – экспоненциальная

Анализ полей показал, что для данной экспериментальной вариограммы кубическая модельная функция (4) при рекомендуемом в методе кригинга радиусе влияния вариограммы (в пределах  $1/2$  максимального расстояния между точками) не приводит к хорошим результатам: густота и кривизна изолиний полей превышает разумные пределы. Поэтому для анализа выбран радиус влияния вариограммы с максимальным Кд, где изолинии наиболее приближены к остальным полученным вариантам интерполяционных полей. Несколько другой результат получен при использовании квадратической модельной функции вариограммы (3), при которой применение радиуса влияния вариограммы с максимальным Кд кардинально улучшило вид интерполяционного поля по сравнению с другими значениями радиуса. В случае с линейной и экспоненциальной функциями ((2) и (5)) вариограммы результаты для ра-

диуса влияния вариограммы с максимальным Кд практически не отличались от результатов, получаемых для иных значений радиуса.

В полях, полученных методом универсального кригинга для всех функций модельной вариограммы, виден резкий скачок функции, обусловленный неравномерностью расположения данных наблюдений, тогда как обычный кригинг показал гладкую функцию для линейной, квадратической и экспоненциальной моделей. Таким образом, универсальный кригинг для такого региона, как Чёрное море, в условиях с неравномерным пространственным распределением данных, следует использовать, внимательно подбирая модельную функцию вариограммы.

Также следует отметить общее равномерное распределение по полю значений ошибки кригинга у всех моделей вариограмм, кроме экспоненциальной, при использовании которой ошибки кригинга резко возрастают по мере удаления от точек с измеренной температурой, хотя в целом не превышают значений  $1,2^\circ$ . Такие поля ошибок, несмотря на неравномерность распределения экспериментальных точек, говорят о достаточной достоверности получаемых полей, а также о широкой возможности применения модуля «Кригинг» ГИС QGIS в морских климатических исследованиях.

Регуляризованный сплайн с натяжением. Для анализируемого массива данных за июль-август 1989 г. были проверены несколько наборов настроек модуля «v.surf.rst». При соответствии выборочной гистограммы нормальному закону распределения данных и достаточном количестве точек с данными, значения параметров модуля, установленные по умолчанию, дают нормальный результат интерполяции. Общими для всех полей были выбраны следующие настройки: охват карты – по слою с границами Чёрного моря (по умолчанию охват меньше); размер ячейки региона GRASS (или узлов регулярной сетки, в которые проводится интерполяция) –  $0,01^\circ$ ; минимальное расстояние между точками «dmin»= $0^\circ$ . Установленные по умолчанию настройки, оставленные неизменными: максимальное расстояние между точками «dmax»= $2,5^\circ$  (изменение параметра практически не влияло на результат интерполяции); угол и масштаб анизотропии – равные 0; коэффициент пересчета «zmult»=1 (используется для конвертации значений, записанных в единицах измерения, отличных от единиц карты, т.е. в данном случае пересчет не использовался).

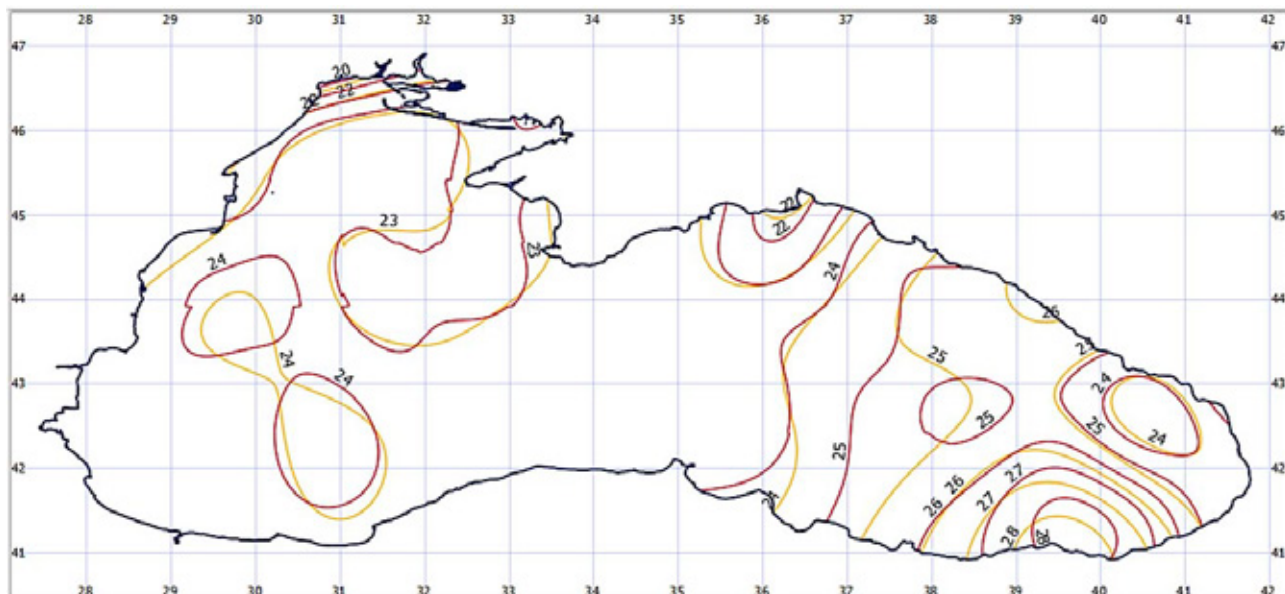
Из рис. 3 и 4 видно, что результат интерполяции со значением охвата точек снаружи сегмента «nprmin»=150 (точки пересечения с соседними сегментами), и натяжением «tension»=10 аналогичен результату с натяжением, равным 15 и «nprmin»=300. Т.е. чем меньше точек взято для анализа (или чем больше натяжение), тем ближе к исходным данным результаты интерполяции. При увеличении количества точек за пределами сегмента (квадранта) или уменьшении натяжения интерполяционная поверхность становится более сглаженной.

С увеличением натяжения растёт количество изолиний, особенно в областях с недостатком данных, что приводит к необходимости тщательного подбора параметра натяжения для каждого набора данных. В анализируемом наборе данных шаг натяжения, при котором происходят значительные изменения результатов интерполяции, очень мал (порядка 2-5 ед.). Слишком малый интервал между количеством точек в квадранте («segmax») и количеством точек за его пределами («nprmin») приводит к разрыву функции в областях с большой плотностью данных (С-З часть Чёрного моря) (рис. 3 и 4).

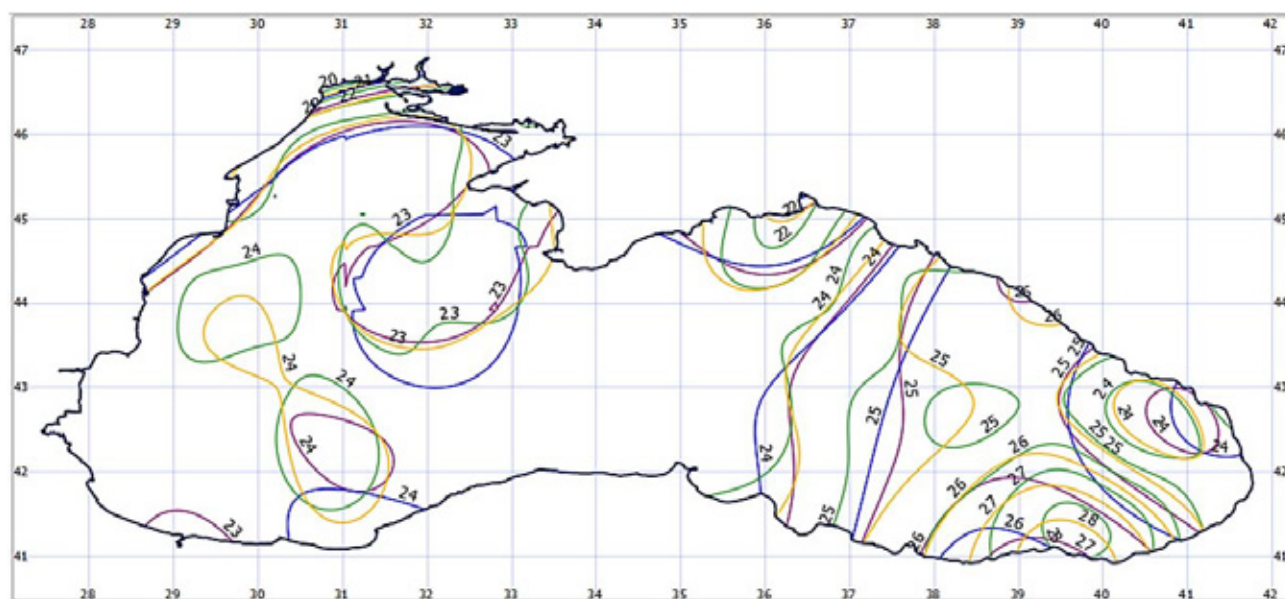
Бикубическая или билинейная интерполяция с регуляризацией Тихонова. Размер ячейки региона GRASS был задан таким же, как и в предыдущих интерполяционных модулях, –  $0,01^\circ$ . Так как данные распределены неравномерно, требуется статистическая регуляризация. В таких случаях модуль «v.surf.bspline» будет пытаться минимизировать кривизну бикубических сплайнов в областях, не имеющих наблюдений. Основное правило при подборе длины шага сплайна: она должна быть больше, чем среднее расстояние между точками наблюдения. Небольшая длина шага сплайна – меньше, чем расстояние между точками наблюдения – может значительно увеличить время обработки. При помощи модуля «v.surf.bspline.lambda» были оценены значения параметра регуляризации Тихонова «lambda\_i» для разных шагов сплайна, равных  $1^\circ$ ,  $2^\circ$  и  $3^\circ$ , по рассчитанным для них смещениям и среднеквадратическим отклонениям. Билинейный и бикубический интерполяционные алгоритмы показали одинаковые значения смещения и средне-



квадратического отклонения для соответствующих шагов сплайна. Бикубическая интерполяция дает в результате поверхность, которая является более гладкой функцией по сравнению с билинейной. Поэтому в данной работе билинейный алгоритм не использовался.



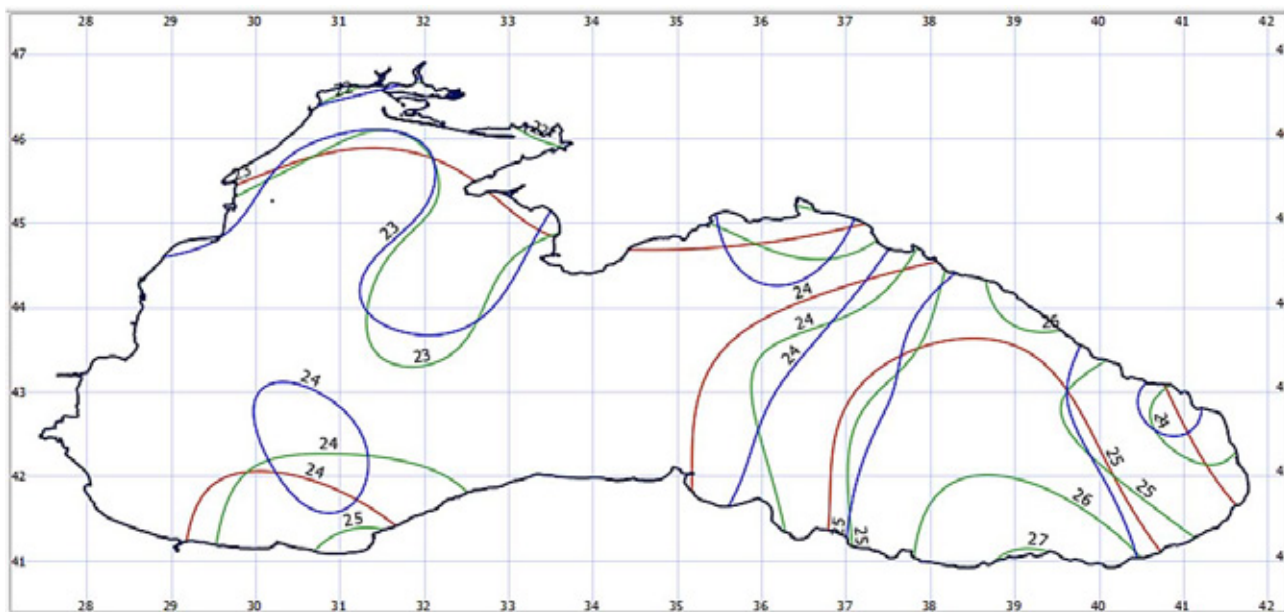
**Рис. 3.** Поля температуры поверхностных вод Чёрного моря за июль-август 1989 г., полученные методом регуляризованного сплайна с натяжением. Значение натяжения «tension»=10. Изолинии красного цвета – «segmax»=50, «npmin»=150 (rst1); жёлтого цвета – «segmax»=40, «npmin»=300 (rst2)



**Рис. 4.** Поля температуры поверхностных вод Чёрного моря за июль-август 1989 г., полученные методом регуляризованного сплайна с натяжением. Значения «segmax»=40, «npmin»=300. Изолинии синего цвета – «tension»=5 (rst3); фиолетового цвета – «tension»=7 (rst5); жёлтого цвета – «tension»=10 (rst2); зелёного цвета – «tension»=15 (rst4)

Для дальнейшего анализа и уменьшения погрешности процедуры интерполяции сплайном взяты значения «lambda\_i» с минимальной среднеквадратической ошибкой при наименьшем смещении: шаг сплайна=1°,  $\lambda=0,2$ ; шаг=2°,  $\lambda=0,01$ ; шаг=3°,  $\lambda=0,01$  (рис. 5). Поле поверхностной температуры, полученное при наибольшем шаге сплайна – 3°, показало зна-

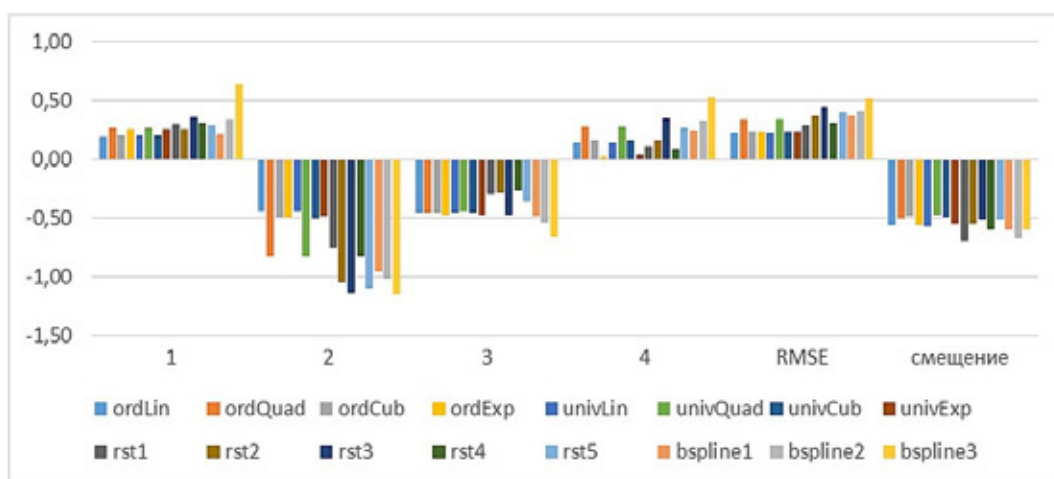
чительное осреднение анализируемого параметра. Остальные поля показали результаты, близкие к результатам модулей «v.surf.rst» и «Кригинг».



**Рис. 5.** Поля температуры поверхностных вод Чёрного моря за июль-август 1989 г., полученные методом бикубического сплайна с регуляризацией Тихонова.

Изолинии синего цвета – шаг сплайна = 1,  $\lambda=0,2$  (bspline1); зелёного цвета – шаг сплайна = 2,  $\lambda=0,01$  (bspline2); красного цвета – шаг сплайна = 3,  $\lambda=0,01$  (bspline3)

Сравнение ошибок кросс-валидации для полей поверхностной температуры вод Чёрного моря, полученных методом кригинга и сплайнов. При сравнении невязок и среднеквадратических погрешностей обычного и универсального кригингов выявляется, что они аналогичны в обоих видах кригинга для соответствующих функций вариограмм (рис. 6).



**Рис. 6.** Графики распределения невязок температур по станциям (1–4), использованным в процедуре кросс-валидации, а также среднеквадратическая ошибка (RMSE) и смещение полей, полученных методом кригинга (обычного «ordLin», «ordQuad», «ordCub», «ordExp» и универсального «univLin», «univQuad», «univCub», «univExp» с использованием линейной, квадратической, кубической и экспоненциальной моделей вариограмм), регуляризованного сплайна с натяжением («rst1» – «rst5», соответствующие им параметры см. на рис. 3 и 4) и бикубического сплайна с регуляризацией Тихонова («bspline1» – «bspline3», соответствующие им параметры см. на рис. 5) для массива за июль-август 1989 г.

Среднеквадратическая погрешность полей, полученных методом сплайн-интерполяции, немного выше, чем у полей, полученных методом кригинга, и не превышает  $0,5^\circ$ . Смещение (разница средней температуры между измеренным и полученными в результате интерполяции значениями) во всех методах практически идентична (около  $-0,5^\circ$ ), что также говорит о приближенных друг к другу интерполяционных функциях.

**Выводы.** Сравнительный анализ полей, полученных различными методами интерполяции в QGIS показал, что все виды анализируемых функций позволяют получать поля с небольшой ошибкой, близкой к минимальной. Бикубический сплайн с регуляризацией Тихонова и регуляризованный сплайн с натяжением при подобранных параметрах регуляризации и натяжения дают результаты, близкие к результатам, получаемым методом кригинга.

Все сплайны обладают интерполяционными свойствами, однако могут использоваться и в сглаживающем варианте. Параметр сглаживания может быть задан индивидуально для каждой опорной точки, что реализовано в модулях ГИС QGIS «GRASS-геоалгоритмы» для регуляризованных сплайнов с натяжением и бикубических/билинейных сплайнов с регуляризацией Тихонова.

Возможность дополнительной настройки модулей кригинга и сплайнов с помощью множества параметров делают данные методы наиболее мощным инструментом интерполяции. Практические эксперименты показывают, что для ограниченного набора данных с неравномерным распределением опорных точек, наилучшие результаты при соответствующей настройке параметров дают обычный кригинг и сглаживающие варианты сплайнов, с соответствующим образом подобранным параметром сглаживания (параметр регуляризации Тихонова для бикубических сплайнов) или натяжения (в регуляризованных сплайнах с натяжением).

Из-за того, что очень часто анализируемые массивы данных характеризуются плохой пространственной структурой, в океанологии остро стоит вопрос о достоверности получаемых результатов. Используя для интерполяции несколько независимых модулей, QGIS позволяет провести как качественную (визуальную), так и количественную (методом кросс-валидации) оценку получаемых полей, используемых для описания пространственно-временной изменчивости параметров морской среды. Проанализированные в статье методы интерполяции в сочетании с возможностями статистических модулей и модулей управления базами данных в QGIS открывают широкие возможности для морских климатических исследований.

## СПИСОК ЛИТЕРАТУРЫ

1. Демьянов В.В., Савельева Е.А. Геостатистика: теория и практика, под ред. Р.В. Артюняна. М.: Наука, 2010. 327 с.
2. Кошель С.М. Теоретическое обоснование структуры и функций блока моделирования рельефа в ГИС: диссертация. М., 2004. 119 с.
3. Кошель С.М., Мусин О.Р. Методы цифрового моделирования: кригинг и радиальная интерполяция. – Информационный бюллетень ГИС-Ассоциации, 2001. № 2(29)–3(30). С. 23–24.
4. Новикова А.М., Новиков А.А. ГИС как инструмент хранения, обработки и визуализации климатических данных. – Материалы Научной конференции «Ломоносовские чтения» 2015 г. и Международной научной конференции студентов, аспирантов и молодых ученых «Ломоносов-2015», под ред. М.Э. Соколова, В.А. Иванова, Н.Н. Миленко, В.В. Хапаева, Н.В. Величко. Севастополь: ООО Экспресс-печать, 2015. С. 34–35.
5. Поротов Г.С. Математические методы моделирования в геологии. СПб., 2006. 223 с.
6. Matheron G. Splines and Kriging: their formal equivalence // Down to the earth statistics, D.F. Merriam (ed.), Academic Press, N.Y., Geology Contributions 8, 1981. Pp. 77–95.

7. *Mitasova, H., Mitas, L., and Harmon, R.* Simultaneous spline approximation and topographic analysis for lidar elevation data in open-source GIS. *IEEE Geoscience and Remote Sensing Letters*, 2005. № 2. Pp. 375–379.

8. *Neteler M., Mitasova H.* Open Source GIS: A GRASS GIS Approach: Third edition. New York: Springer, 2008. 420 p.

9. *Petrasova A., Harmon B., Petras V., Mitasova H.* Tangible Modeling with Open Source GIS. Springer International Publishing, 2015. 135 p.

---

**A.M. Novikova<sup>1</sup>, A.B. Polonskij<sup>2</sup>, A.A. Novikov<sup>3</sup>**

## **COMPARISON OF QGIS INTERPOLATION MODULES CAPABILITIES FOR MARINE CLIMATE RESEARCH ON DUTY WITH AN ARRAY OF LOW-SECURITY DATA**

**Abstract.** *In the article the urgency of modern methods' active use in oceanographic data spatial analysis from the perspective of geo-information and geostatistical approaches is approved. There are analyzed the possibilities of some open GIS QGIS statistical modules for practical problems of data quality rapid assessment solution. The quality of QGIS interpolation modules is estimated representing methods of kriging and radial basis functions (splines) by using an array of low-security data.*

**Key words:** *QGIS, statistics modules, kriging, regularized spline with tension, bicubic spline interpolation with Tikhonov regularization.*

### REFERENCES

1. *Dem'yanov V.V., Savel'eva E.A.* Geostatistika: teoriya i praktika [Geostatistics: Theory and Practice], pod red. R.V. Arutyunyana. Moscow: Nauka, 2010. 327 p.

2. *Koshel' S.M.* Teoreticheskoe obosnovanie struktury i funkcij bloka modelirovaniya rel'efa v GIS: dissertacija [Theoretical Study of the Structure and Functions of the Relief Modeling Unit in GIS: the dissertation]. Moscow, 2004. 119 p.

3. *Koshel' S.M., Musin O.R.* Metody cifrovogo modelirovaniya: kriging i radial'naya interpol'yaciya [Digital Modeling Methods: Kriging and Radial Interpolation]. – Informacionnyj byulleten' GIS-Associacii, 2001, № 2(29)–3(30), pp. 23–24 (in Russian).

4. *Novikova A.M., Novikov A.A.* GIS kak instrument hraneniya, obrabotki i vizualizacii klimaticheskikh dannyh [GIS as a tool for storage, processing and visualization of climate data]. – Materialy Nauchnoj konferencii «Lomonosovskie chteniya» 2015 g. i Mezhdunarodnoj nauchnoj konferencii studentov, aspirantov i molodyh uchenyh «Lomonosov-2015», pod red. M.EH. Sokolova, V.A. Ivanova, N.N. Milenko, V.V. Hapaeva, N.V. Velichko. Sevastopol: OOO EHkspress-pechat', 2015. Pp. 34–35.

5. *Porotov G.S.* Matematicheskie metody modelirovaniya v geologii [Mathematical methods of modeling in geology]. St. Petersburg, 2006, 223 p.

6. *Mitasova, H., Mitas, L., and Harmon, R.* Simultaneous spline approximation and topographic analysis for lidar elevation data in open-source GIS. *IEEE Geoscience and Remote Sensing Letters*, 2005. № 2. Pp. 375–379.

---

<sup>1</sup> FSBSI «Institute of natural and technical systems», Department of Oceanography Actual Problems, Sevastopol, 299011, Russia, researcher; e-mail: anelkapi@bk.ru.

<sup>2</sup> FSBSI «Institute of natural and technical systems», the Oceanographic Center, Sevastopol, 299011, Russia, head, professor; e-mail: apolonsky5@mail.ru.

<sup>3</sup> Branch of M.V. Lomonosov Moscow State University in Sevastopol, Geography Department, Sevastopol, 299001, Russia, senior teacher; e-mail: a\_novik@bk.ru.

7. Neteler M., Mitasova H. Open Source GIS: A GRASS GIS Approach: Third edition. New York: Springer, 2008. 420 p.

8. Petrasova A., Harmon B., Petras V., Mitasova H. Tangible Modeling with Open Source GIS. Springer International Publishing, 2015. 135 p.

---

УДК 528.852, 551.584.4

М.Ю. Грищенко<sup>1</sup>, К.К. Чернулич<sup>2</sup>

## СОПОСТАВЛЕНИЕ НАЗЕМНЫХ И КОСМИЧЕСКИХ ТЕМПЕРАТУРНЫХ ДАННЫХ НА ПРИМЕРЕ ТЕРРИТОРИИ ОСТРОВА ВРАНГЕЛЯ

***Резюме.** В современных условиях в нашей стране многие отдалённые районы характеризуются малой плотностью метеорологических станций, что снижает точность синоптических прогнозов на территории, удалённые от метеостанций. В этой связи использование тепловых космических снимков для моделирования некоторых климатических параметров рассматривается авторами как перспективное направление науки. В статье представлены результаты сопоставления рассчитанных по тепловым космическим снимкам со спутников серии Landsat значений температуры земной поверхности и измеренных с помощью метеоприборов значений температуры приземного слоя атмосферного воздуха. Для рассмотренных сезонов показатели характеризуются выраженной линейной связью с высоким коэффициентом корреляции.*

***Ключевые слова:** Тепловые космические снимки, температура земной поверхности, температура атмосферного воздуха, остров Врангеля.*

**Введение.** Современный способ составления синоптических карт основывается на проведении климатического моделирования и интерполяции данных сети метеостанций. Минусами данного способа являются крайняя трудоёмкость и разная точность получаемых результатов: для областей, расположенных вблизи метеостанций, точность достаточно высокая, но при удалении от станций она снижается.

Таким образом, получение дистанционных и пространственно непрерывных данных о температуре воздуха было бы полезно для нашей страны, учитывая её географические особенности. Такую возможность даёт использование космических снимков в совокупности с разработкой метеорологических моделей.

Тепловые космические снимки имеют огромный потенциал для использования в географической науке, в том числе, в климатологии. На настоящий момент при применении к ним специальных методов обработки возможно получение значений температуры земной поверхности. Между температурой земной поверхности и температурой приземного слоя атмосферного воздуха существует сложная связь, суть которой ещё до конца не изучена, однако она может быть выражена математически с использованием эмпирических коэффициентов. Поиск таких зависимостей является в настоящее время актуальной задачей аэрокосмических методов географических исследований.

Целью работы является определение связи между температурой земной поверхности, рассчитанной по тепловым космическим снимкам, и температурой приземного слоя атмосферного воздуха, полученной по результатам наземных измерений.

**Материалы и методы исследования.** На территорию исследуемых островов имеется ряд наземных данных о температуре приземного слоя воздуха. На острове Врангеля распо-

---

<sup>1</sup> МГУ им. М.В. Ломоносова, географический факультет; e-mail: m.gri@geogr.msu.ru.

<sup>2</sup> МГУ им. М.В. Ломоносова, географический факультет; e-mail: chernulich.kk@gmail.com.