



ИЗБРАННЫЕ ТРУДЫ

МЕЖДУНАРОДНАЯ КОНФЕРЕНЦИЯ

и школа молодых ученых по измерениям, моделированию
и информационным системам для изучения окружающей среды

enviromis 2016

INTERNATIONAL CONFERENCE

and Early Career Scientists School
on Environmental Observations, Modeling and Information Systems

SELECTED PAPERS



Опыт поиска оптимальных моделей, соответствующих реальным данным, показывает, что поверхность критерия выбора модели не имеет выраженного экстремума. В первую очередь это касается выбора комбинации предикторов. В меньшей степени, но подобная ситуация наблюдается и при выборе оптимального оператора регрессии. Такая ситуация обуславливает высокую чувствительность процедур выбора наиболее подходящих предикторов и операторов к малым вариациям используемых данных. Для преодоления последствий такой чувствительности приходится применять различные приемы, делающие процесс моделирования более долгим и сложным.

Совокупность перечисленных проблем делает задачу автоматизированного моделирования совместной динамики достаточно большой группы характеристик весьма громоздкой и требующей существенных вычислительных ресурсов. Однако, даже при наличии таких ресурсов, результаты автоматизированного моделирования, как правило, необходимо подвергать экспертному контролю со стороны разработчиков модели, так как использования только статистических критериев для подтверждения качества модели часто оказывается недостаточно. С необходимостью такого контроля связана еще одна проблема – малое количество экспертов, обладающих одновременно интересом к данной задаче, достаточно глубокими знаниями в области технологии моделирования и в обсуждаемой предметной области исследования.

Наличие перечисленных проблем не является фатальным для рассматриваемой задачи. Однако их преодоление требует серьезной и обширной работы по улучшению данных о динамике исследуемых процессов, совершенствованию технологии моделирования, обучению пользователей, в частности, нашей программы, формированию групп заинтересованных экспертов и так далее.

JOINT RESEARCH OF THE NORTHERN HEMISPHERE SURFACE PRESSURE AND AIR TEMPERATURE FIELDS

¹Kataev S.G., ²Kataeva S.S., ^{3,4}Ivanova E.V.

¹*Tomsk State Pedagogical University, Tomsk, Russia*

²*Tomsk State University, Tomsk, Russia*

³*Institute of Monitoring of Climatic and Ecological Systems SB RAS, Tomsk, Russia*

⁴*Tomsk Architectural Construction University, Tomsk, Russia*

E-mail: sgkataev@sibmail.com, sskataeva@sibmail.com, ehllai@rambler.ru

Joint research of different meteorological fields will be more productive if values of the measured parameters exist in the same points. For the solution of this task we offer the approach consisting of consecutive application of procedures of classification and interpolation of the studied data fields. At the same time classification is understood as the procedure allowing to break all set of objects into classes, being guided by a certain rule of reference of concrete object to a certain class. Spatial interpolation represents broadcast of properties of the field on the sites which aren't captured by measurements. Combined use of the specified procedures allows to receive and describe structure of data fields. It should be noted that classification and interpolation well supplement each other and increase efficiency of all approach. So, for example, as it has been shown in [1-2], carrying out a clustering of the initial geophysical field gives the chance with higher quality to carry out interpolation procedure; the average square mistake at mutual interpolation decreases by 20-30%.

Our approach consists of 3 stages.

Classification [3] initial fields and determination of the limits of each class.

Generation uniform for all considered fields of a set of virtual points and carrying out for each of them procedures of optimum interpolation for modified Ganding [4] method, with selection only of those points which belong spatially to the same class, as the interpolated point.

Classification of the new fields. If the set of all points forms the regular grid which is completely covering the studied area, then classes received as a result of secondary classification will cover it completely, forming structure of the field.

The described algorithm has been applied at research of fields of surface temperature and pressure over the northern hemisphere. Introduction to interpolation of procedure of the accounting of a class association and generation of virtual points significantly improve the accuracy of reproduction of signals on basis points. So, in the reviewed example of SKO has decreased from value 3.20MPa till 1.54 MPa, i.e. it is more, than twice. The classification enables to improve the interpolation, and the interpolation, in their turn, permits solve the problem of finding of interclass borders.

References:

1. Kataev, S.G., Kataeva S.S., Kuskov, A.I. [2013] Use of classification in a problem of spatial interpolation. *Izv.Vus. Fizika. T.56, No. 9/2, 235-238*
2. Kataev, S.G., Kataeva S.S., Ivanova E.V. The description of structure of the meteorological field on the basis of combined use of procedures of classification and interpolation. In the book: XI Siberian meeting on klimato-environmental monitoring. Institute of monitoring of climatic and ecological systems of the Siberian Branch of the Russian Academy of Science. Tomsk, 2015. Page 30-31.
3. Kataev, S.G., Kataev, M.YU., [2012] Mathematical formalism and algorithm of structuring multidimensional data. *Polzunovsky sbornik, No. 2/1, 99-103.*
4. Ganding, L.S., Kagan, R. L. [1963] Statistical methods of interpretation of meteorological data. L., Gidrometeoizdat, 287.

СОВМЕСТНОЕ ИССЛЕДОВАНИЕ ПОЛЕЙ ПРИЗЕМНОГО ДАВЛЕНИЯ И ТЕМПЕРАТУРЫ ВОЗДУХА НАД СЕВЕРНЫМ ПОЛУШАРИЕМ

¹Катаев С.Г., ²Катаева С.С., ^{3,4}Иванова Э.В.

¹Томский государственный педагогический университет, Томск, Россия

²Томский государственный университет, Томск, Россия

³Институт мониторинга климатических и экологических систем СО РАН, Томск, Россия

⁴Томский государственный архитектурно-строительный университет, Томск, Россия

E-mail: sgkataev@sibmail.com, sskataeva@sibmail.com, ehllai@rambler.ru

Введение.

Часто в геофизических исследованиях встречается ситуация, когда на одной территории измеряются различные параметры, но в разных точках. Очевидно, что интерпретация была бы лучше, если бы разные геофизические параметры фиксировались в одних и тех же пространственных точках, поскольку это дает возможность выявлять дополнительные количественные закономерности между изучаемыми полями. Все это говорит о необходимости решать задачу приведения путем интерполяции разнородных геофизических данных, полученных на одной территории, к единой сетке точек.

Для решения этой задачи мы предлагаем подход, состоящий из поэтапного применения процедур классификации и интерполяции поля геофизического параметра (ГП). При этом под классификацией понимается процедура, позволяющая разбить все множество объектов на классы, руководствуясь определенным правилом отнесения конкретного объекта к определенному классу. Пространственная интерполяция представляет собой трансляцию свойств поля на участки, не охваченные измерениями. Совместное применение указанных процедур позволяет получать и описывать структуру полей данных. Надо отметить, что классификация и интерполяция хорошо дополняют друг друга и усиливают эффективность всего подхода. Так, например, как было показано в [1-2], проведение кластеризации исходного геофизического поля дает возможность более качественно проводить процедуру интерполяции; средняя квадратичная ошибка при взаимной интерполяции уменьшается на 20-30%.

Формулировка задачи и основные этапы решения.

Пусть на исследуемой территории произведены измерения сигнала F_l^k для каждого из M_s типов геофизических величин, причем каждая из них фиксировалась на своей сетке точек с координатами (x_l^k, y_l^k) , $l = \overline{1, N_k}$; $k = \overline{1, M_s}$. В общем виде измеренный сигнал геофизического параметра k представим в виде вектора $F_l^k = (f_1^k, f_2^k, \dots, f_{m_k}^k)$, $l = \overline{1, N_k}$. Можно сформулировать следующую задачу: найти структуру всех изучаемых геофизических полей на данной территории. Ранее, для решения этой задачи в случае единственного ГП нами была предложена процедура, состоящая из нескольких этапов.

Классификация поля [3] исходного сигнала и определение пространственных границ между классами. При этом предполагается, что многоугольники, определяющие границы классов должны быть выпуклыми.

Генерация единого для всех рассматриваемых полей множества виртуальных точек и проведение для каждой из них процедуры оптимальной интерполяции по методу Гандина [4], с отбором только тех пунктов, которые принадлежат пространственно тому же классу, что и интерполируемая точка. В результате получается набор дополнительных пространственных точек

$X'_{ij}, i=1, \dots, k'_x, j=1, \dots, k'_y; k'_x \cdot k'_y = N'$, обладающих интерполированными значениями сигнальных функций.

Классификация нового поля, состоящего из $N_k + N'$ векторов сигнала. Если набор всех точек $X_l + X'_{ij}$ образует регулярную сетку, полностью охватывающую исследуемую область, то получаемые в результате вторичной классификации классы будут ее полностью покрывать, формируя структуру поля.

Проводя указанные этапы для каждого в отдельности ГП, мы получим структуры всех исследуемых полей. Причем, в общем наборе виртуальных точек X'_{ij} у нас будут сигналы всех ГП, что позволит проводить разнообразные дополнительные совместные исследования.

Пример.

Исследование полей приземного давления и температуры воздуха над Северным полушарием

Исходные данные: а) временные ряды за 46 лет в 288 пунктах измерения (спутниковые данные, равномерная сетка 16 точек на широту) среднемесячных значений приземного давления P над территорией Северного полушария; б) временные ряды за 56 лет на 818 метеостанций среднемесячных значений приземной температуры T в Северном полушарии.

В соответствии с описанной выше методикой на первом этапе была проведена классификация исходных полей данных. Было выделено 44 класса, как для поля давления, так и для температуры.

Далее была сгенерирована более густая равномерная сеть X'_{ij} , состоящая из 2040 точек, охватывающая все Северное полушарие, и в этих точках получены путем интерполяции значения давления и температуры. Проведена с использованием процедуры взаимной интерполяции проверка качества интерполяции полей T и P для всех точек X'_{ij} . Для этого были рассчитаны с использованием процедуры взаимной интерполяции пространственные распределения среднеквадратичных ошибок (СКО) для P и T . Были получены следующие результаты. СКО исходного поля, (только для базовых точек), для давления в результате применения процедуры взаимной интерполяции без учета классового деления дает величину 3.20 МПа; с учетом классов – 2.87 МПа. Эта же величина при взаимной интерполяции всех базовых и виртуальных точек равна 1.54 МПа. Т.е. точность (качество) интерполяции при этой процедуре возрастет в более чем 2 раза и можно говорить о корректной трансляции поля давления на территории, не охваченные измерениями. Для температуры соответствующее уменьшение СКО равно, примерно 20%, т.е., в целом, улучшение качества также наблюдается, но не такое существенное, как для давления. На рисунке 1 приводятся, в качестве иллюстрации, пространственные распределения СКО для полей давления и температуры. Видно, что и для температуры и для давления существуют области с аномальным, не связанным с окружающими ареалами, поведением параметров. И, наконец, проведена классификация (как для годового хода, так и отдельно за каждый месяц) полей P и T для новой сетки. Результаты классификации для годового хода помещены на рисунке 2. Анализ рис. 2 показывает, что средняя годовая температура воздуха имеет очаговый характер. Очаг минимальных температур расположен в районе Верхоянск-Оймякон и моря Лаптевых.

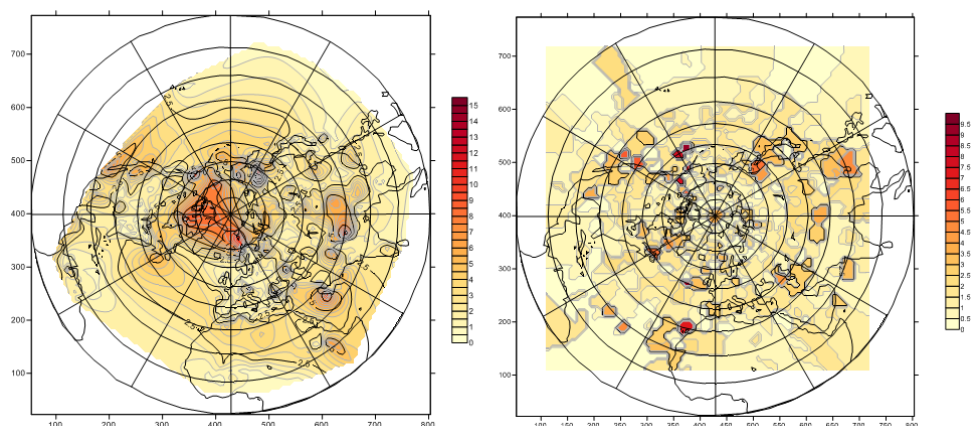


Рисунок 1. Пространственное распределение СКО температуры для только базовых (818) станций (левый фрагмент) и СКО давления для всех базовых и виртуальных точек (правый фрагмент). Северное полушарие.

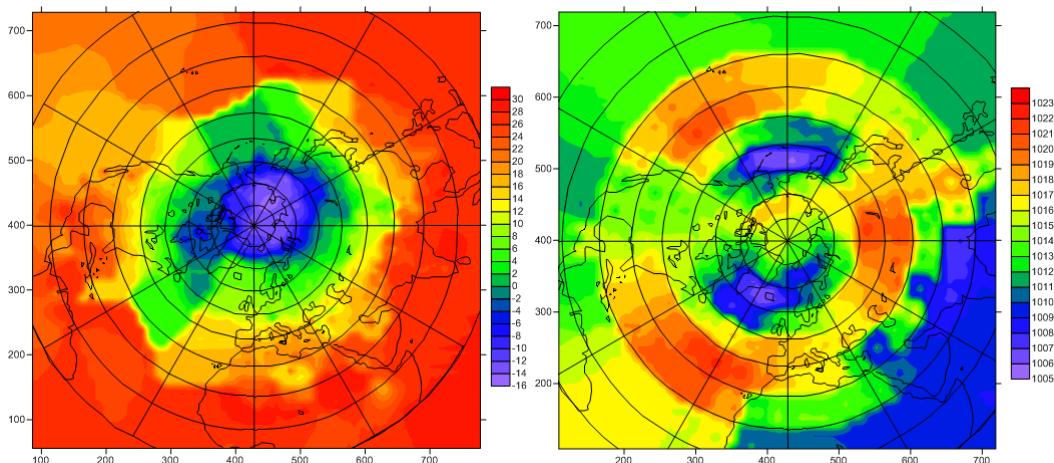


Рисунок 2. Результат классификации поля температуры воздуха (левый фрагмент) и давления (правый фрагмент). Цветом выделены среднеклассовые значения температуры и давления.

Гребень тепла в Атлантическом океане (выделен зеленым цветом) образует отдельный класс, который расположен практически меридионально и связан с влиянием Гольфстрима. Наибольшую территорию Северного полушария занимает один класс (выделен красным цветом), который располагается в южных широтах Северного полушария и имеет зональный характер распределения.

Классификация поля давления показывает и подтверждает тот факт, что в структуре циркуляции атмосферы над Северным полушарием имеют место две хорошо выраженные в течение всего года циркуляционные системы – Северо-Атлантическая, представленная Исландским минимумом (выделен синим цветом) и Азорским максимумом давления (красный цвет), и Северо-Тихоокеанская, включающая в себя Алеутский циклон (выделен синим цветом) и Гавайский антициклон (выделен красным цветом). Знание структуры поля (расположение и характеристики классов) позволяет не только получить крупномасштабную картину, но и достаточно просто проводить анализ и получать нетривиальные выводы, оперируя среднеклассовыми значениями. В качестве иллюстрации такого анализа на рисунке 3 приводится динамика размаха полей температуры и давления. Под размахом понимается разность среднеклассовых параметров между классами с максимальным его значением и минимальным. Величина размаха характеризует контрастность поля метеопараметра, чем выше размах, тем сильнее степень контрастности. Так, из рисунка 3 можно видеть, что наблюдается устойчивость классовой структуры и для Т и для Р за все три периода и устойчивое поведение минимальных и максимальных среднемесячных значений. Появляющиеся отличия в структуре классов не носят устойчивого трендового характера. Отсюда можно сделать вывод: изменения климатической системы, происходящие на планете, на уровне среднемесячного усреднения не проявляются.

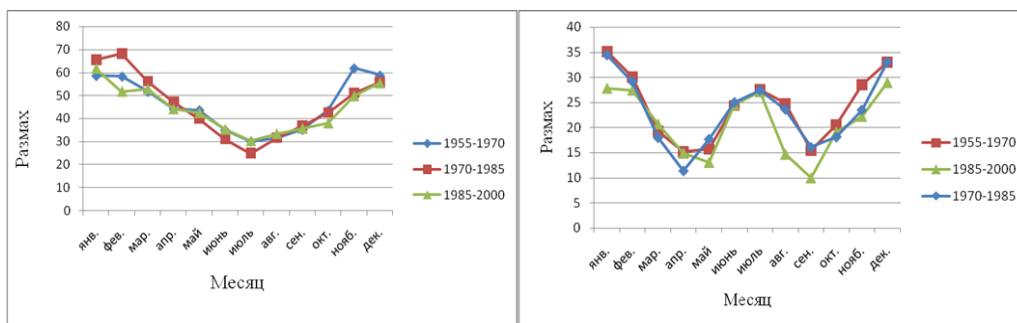


Рисунок 3. Временная динамика «размаха» классов для температуры (слева) и давления (справа) за 3 периода времени

Выводы

Классификация дает возможность улучшить интерполяцию, а интерполяция, в свою очередь, позволяет решить проблему нахождения межклассовых границ.

Введение в интерполяцию процедуры учета классовой принадлежности и генерация виртуальных точек существенно улучшают точность воспроизводства сигналов на базисных пунктах. Так, в рассмотренном примере СКО уменьшилась со значения 3.20 МПа до 1.54 МПа, т.е. больше, чем в 2 раза.

Анализ пространственного распределения СКО и изменения СКО при включении в процедуру взаимной интерполяции виртуальных станций позволяют судить о корректности трансляции сигнала по территории.

Предлагаемый в работе подход не привязан к физической природе измеряемого сигнала и исследуемого геофизического поля, и в этом плане он универсален. Можно надеяться, что данный подход будет полезным в разнообразных исследованиях, связанных с анализом пространственного распределения сигналов.

Литература:

1. Катаев, С.Г., Катаева С.С., Кусков, А.И. [2013] Использование классификации в задаче пространственной интерполяции. *Известия Вузов «Физика»*, Т.56, №9/2, 235-238
2. Катаев С.Г., Катаева С.С., Иванова Э.В. Описание структуры метеорологического поля на основе совместного применения процедур классификации и интерполяции. В книге: XI Сибирское совещание по климато-экологическому мониторингу Институт мониторинга климатических и экологических систем СО РАН. Томск, 2015. С. 30-31.
3. Катаев, С.Г., Катаев, М.Ю., [2012] Математический формализм и алгоритм структурирования многомерных данных. *Ползуновский вестник*, № 2/1, 99-103.
4. Гандин, Л.С., Каган, Р.Л.[1963] Статистические методы интерпретации метеорологических данных. Л., Гидрометеиздат, 287.

АНАЛИЗ ВАРИАЦИЙ ТЕМПЕРАТУРЫ ПОВЕРХНОСТИ ТИХОГО ОКЕАНА

Кузин В.И., Лобанов А.С.

*Институт вычислительной математики и математической геофизики СО РАН, Новосибирск,
Россия
E-mail: kuzin@sscc.ru*

В работе представлены исследования по анализу температуры поверхности Тихого океана (ТПО) на основе данных реанализа NCEP/NCAR для периода 1948 -2002 гг. [1].

В работе используются два метода. Первый является классическим анализом естественных ортогональных функций (ЕОФ). Анализ данных на основе ЕОФ позволяет выделить наиболее сильный сигнал межгодовой изменчивости ТПО в тропиках Тихого океана такой как Эль-Ниньо и Ля-Нинья и реконструировать его на основе первых гармоник. Это указывает на то, что аномалии в периоды экстремальных событий в тропиках определяются первыми гармониками и в промежуточные периоды существуют более сложные процессы формирования аномалий. Изменчивость аномалий ТПО в субтропиках и в субполярных регионах на временных масштабах квазидесятилетних колебаний ТПО (PDO) [2, 3] заглушается сильным тропическим сигналом. Эти процессы влияют не только на атмосферные условия над Тихим океаном, но осуществляют климатическое влияние на материки [4, 5, 6].

Для выявления этих процессов в изменчивости ТПО Тихого океана использовался метод кластерного анализа. Кластерный анализ представляет собой способ объединения данных в классы (кластеры) по критерию корреляции между пространственными или временными точками. Анализ пространственного распределения кластеров позволил выделить типичные структуры в изменчивости ТПО Тихого океана между восточными и западными зонами океана, а также между субтропическими и субполярными районами в продолжении Курошио. При этом выделяется сигнал квазидесятилетней изменчивости ТПО, которые может идентифицироваться как проявление сигнала PDO.