

А. В. ИГНАТОВ, О. П. ОСИПОВА, А. С. БАЛЫБИНА

Институт географии им. В. Б. Сочавы СО РАН, 664033, Иркутск, ул. Улан-Баторская, 1, Россия, ignatov@irigs.irk.ru, olga@irigs.irk.ru, balybina@irigs.irk.ru

ПРОСТРАНСТВЕННАЯ СТРУКТУРА ВЗАИМОСВЯЗЕЙ ГОДОВЫХ СУММ ОСАДКОВ НА МЕТЕОСТАНЦИЯХ СИБИРИ И КАЗАХСТАНЕ

Средствами многофакторного стохастического моделирования выполнен анализ совместной изменчивости сумм годовых осадков на территории Сибири и Казахстана. Показано, что между этими переменными существуют статистически значимые взаимосвязи, удовлетворяющие условию устойчивости во времени. Такие связи, отражающие закономерности формирования поля годовых осадков, носят и положительный, и отрицательный характер. Положительные связи наблюдаются между близко расположенными метеостанциями и указывают главным образом на локальную однородность формирования поля годовых осадков. Они надежно выявляются методами как многофакторной, так и однофакторной статистики. Отрицательные взаимосвязи наблюдаются между существенно более отдаленными метеостанциями и выявляются по условию значимости отклонения от нуля отрицательных коэффициентов перед переменными в многофакторных моделях линейной регрессии. Парные отрицательные статистически значимые связи между годовыми осадками на метеостанциях отмечаются крайне редко. В прикладном плане наличие значимых закономерных взаимосвязей между годовыми суммами осадков может быть использовано для построения моделей, базирующихся на таких закономерностях, для приближенного восстановления отсутствующих данных в прошлом или в задачах совместного прогнозирования ожидаемых годовых осадков на рассматриваемой территории в будущем.

Ключевые слова: изменчивость осадков, анализ взаимосвязей, многофакторное моделирование, надежность оценок, стоячие волны.

A. V. IGNATOV, O. P. OSIPOVA, A. S. BALYBINA

V. B. Sochava Institute of Geography, Siberian Branch, Russian Academy of Sciences, ul. Ulan-Batorskaya, 1, Irkutsk, 664033, Russia, ignatov@irigs.irk.ru, olga@irigs.irk.ru, balybina@irigs.irk.ru

SPATIAL STRUCTURE OF THE RELATIONSHIPS OF ANNUAL PRECIPITATION AMOUNTS IN SIBERIA AND KAZAKHSTAN

Joint variability in annual precipitation amounts across the territory of Siberia and Kazakhstan was analyzed using multi-factor stochastic modeling tools. It is shown that there are statistically significant correlations between these variables, which satisfy the temporal stability condition. Such correlations that reflect the regularities in the formation of the annual precipitation field have also both a positive and a negative character. Positive correlations are observed between closely located weather stations and reflect mainly the local homogeneity of the formation of the annual precipitation field. They are reliably identified by methods of both multi-factor and single-factor statistics. Negative correlations are observed between the significantly more remote weather stations. They are revealed according to the condition of the significance of the deviation from zero for the negative coefficients of the variables in multi-factor linear regression models. Paired negative, statistically significant correlations between annual precipitation amounts at weather stations occur extremely rarely. From an applied perspective, the presence of significant regular correlations between annual precipitation amounts can be used to construct models based on such regularities, for an approximate reconstruction of missing data in the past or in problems of a combined forecasting of the expected annual precipitation for the territory under consideration in the future.

Keywords: precipitation variability, analysis of interrelations, multifactor modeling, reliability of estimates, standing waves.

Пространственная связность температурных полей и наличие тенденций их изменения во времени — ныне признанный факт [1]. Что касается другой важной характеристики климата — количества выпадающих осадков, то их пространственная структура более сложная, а временная динамика, особенно в долгосрочном плане, трудно предсказуема. В связи с этим изучение закономерностей пространственного распределения и межгодовой изменчивости осадков востребовано как в научном, так и в практическом плане. Интерес к данной проблеме подтверждается тем, что эти вопросы неоднократно рассматривались различными исследователями и для разных регионов [2–6].

Цель настоящей работы — исследование пространственной структуры и количественных параметров взаимосвязей между годовыми суммами осадков, зарегистрированными в период массовых инструментальных наблюдений на отдельных метеостанциях, расположенных на обширной территории Северной Азии и охватывающих сектор примерно 25° по широте и 70° по долготе.

МАТЕРИАЛЫ И МЕТОДЫ ИССЛЕДОВАНИЯ

Исходные материалы для рассматриваемой задачи взяты из базы данных, созданной ранее в лаборатории гидрологии и климатологии Института географии СО РАН в ходе выполнения плановых проектов института, на основе сведений из различных открытых источников и данных, полученных другими организациями в рамках научного сотрудничества. В сформированную для исследования взаимосвязей таблицу значений переменных входили годовые суммы осадков, зафиксированные на 115 метеостанциях Сибири и Казахстана за 1936–2009 гг., степень заполнения оценками примерно равна 92 %. Отсутствующие данные приходились главным образом на начало рассматриваемого периода. Имеющиеся оценки считались точными. Пространственная привязка данных задавалась географическими координатами метеостанций.

Для исследования пространственной связности полей метеорологических элементов чаще всего используется корреляционный или регрессионный анализ [1–5], основанный на оценке парных взаимозависимостей между временными рядами характеристик, измеряемых на метеостанциях. При выполнении данной работы нами применялся более сложный аппарат, базирующийся на построении многофакторных стохастических моделей взаимосвязи переменных. Для поиска и моделирования таких взаимосвязей между суммами осадков использовалась программа «Стохастическое моделирование» [7], включающая в себя большое число инструментов, специально разработанных для решения подобного рода задач. С описанием методических основ, заложенных в алгоритмы этой программы, можно познакомиться на сайте Института географии им. В. Б. Сочавы СО РАН (<http://www.irigs.irk.ru>) (раздел «Научная работа/Разработки/Пакет программ — стохастическое моделирование»). Возможности этой программы были показаны ранее при решении различных задач [8–11].

В настоящей работе использованы процедуры автоматического подбора оптимального списка предикторов зависимой переменной и аппроксимирующего моделируемую зависимость математического оператора, а также инструменты графического исследования многофакторных взаимосвязей. Определение значений оптимизируемых параметров моделей осуществлялось методом наименьших квадратов в сочетании с численными градиентными алгоритмами поиска экстремума целевой функции [12, 13]. Сравнение результатов моделирования и выбор лучших моделей выполнялись по различным критериям, учитывающим точность и надежность модельных оценок. В частности, использовались: коэффициент детерминации — для оценки силы многофакторной зависимости, критерий Фишера — для оценки значимости связи. Критерий Стьюдента применялся для определения значимости смещения фактического значения переменной от математического ожидания ее модельной оценки при проверке модели на контрольной выборке, определения значимости коэффициентов регрессии и т. д.

Особое внимание уделялось оценке и контролю надежности результатов исследования. Для обеспечения устойчивости подбора оптимального списка аргументов применялись различные способы повышения вероятности правильного выбора предикторов при проверке большого числа пробных гипотез [4]. Надежной признавалась модель, которая выдержала проверку по всем критериям, как на обучающей, так и на контрольной выборке. Содержательная интерпретация выявленных взаимосвязей осуществлялась на основе описаний только надежных моделей.

РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

На основе первичной статистической обработки использованных в работе данных установлено, что для исследуемой территории изменчивость поля годовых сумм осадков как локальной физической переменной характеризуется коэффициентом вариации примерно 20 % по временной координате и примерно 45 % по пространственным координатам.

Первый шаг в решении задачи поиска вероятных закономерностей, которым может подчиняться переменная составляющая годовых сумм осадков на метеостанциях, состоял в выборе наиболее устойчивого способа поиска оптимальных моделей взаимосвязи между этими характеристиками. Результатом такого поиска являются модели, каждая из которых характеризуется конкретной комбинацией

Таблица 1

Сравнение средних характеристик качества моделей, построенных с использованием различных способов подбора аргументов и операторов регрессии

Параметр	Значение				
	1	2	3	4	5
Номер способа построения оптимальной модели					
Отношение средней квадратичной ошибки аппроксимации обучающей выборки к стандартному отклонению функции (S_0)	0,713	0,716	0,664	0,629	0,644
Нормированное смещение модельной оценки на контрольной выборке (критерий Стьюдента)	0,198	0,195	0,171	0,201	0,164
Отношение фактической ошибки контрольного расчета (S_{Φ}) к ожидаемой ошибке	1,15	1,13	1,2	1,23	1,21
Отношение S_{Φ} к стандартному отклонению функции на обучающей выборке (S_0)	0,936	0,933	0,979	1	0,947
Доля оправдавшихся с точностью ($S_{\Phi} < 0,67S_0$) расчетов по модели на контрольной выборке	0,54	0,55	0,52	0,54	0,55
Мера качества модельного расчета, определяемая по знаку его корреляции с фактическими значениями функции на контрольной выборке	0,42	0,44	0,41	0,38	0,44

Примечание. Жирным шрифтом выделены параметры, отражающие более высокое качество моделей.

наиболее репрезентативных предикторов зависимой переменной и оператором регрессии, аппроксимирующим эту зависимость. Для сравнительной оценки способов поиска для каждой из 115 переменных в исходных данных мы построили по пять моделей, используя различные алгоритмы выбора оптимального решения.

Для одной и той же зависимой переменной результаты подбора разными способами ее предикторов и оператора описания зависимости от них в определенной мере могут отличаться. Для выбора наиболее устойчивого способа сравнивались средние параметры построенных с использованием разных алгоритмов моделей, характеризующие их способность аппроксимировать обучающую и контрольную выборки (табл. 1). Лучшую среднюю устойчивость всех моделей при их проверке с помощью независимых данных продемонстрировали способы 2 и 5. Способ 4, показавший лучшую точность аппроксимации обучающей выборки, оказался худшим при проверке построенных моделей на контрольной выборке. Этот результат сравнения различных методов выбора оптимальных моделей коррелируется с результатами, полученными ранее при использовании программы «Стохастическое моделирование» в процессе разработки методик для прогнозирования притока воды в енисейские водохранилища [11]. Из двух лучших по контрольной проверке способов построения моделей нами выбран способ 2. Для него ожидаемая ошибка расчета в большей степени соответствовала фактической ошибке, что является важным элементом подтверждения надежности модели.

Второй шаг решения задачи — это выбор наиболее надежных моделей. Из всех 115 моделей, отобранных по способу 2, мы оставили только те зависимости, для которых коэффициент детерминации превышал 0,5. Таких моделей оказалось 58. Формально на обучающей выборке по критерию Фишера все зависимости, описываемые этими моделями, оказались статистически значимыми на уровне $<0,05$.

Таблица 2

Средние параметры качества надежных (на уровне значимости $<0,05$) моделей, определенные по обучающей и контрольной выборке

Параметр	Значение
Мера точности аппроксимации функции на обучающей выборке (коэффициент детерминации)	0,626
Отношение средней квадратичной ошибки аппроксимации к стандартному отклонению функции (S_0)	0,611
Нормированное смещение модельной оценки на контрольной выборке (критерий Стьюдента)	0,12
Отношение фактической ошибки контрольного расчета (S_{Φ}) к ожидаемой ошибке	1,01
Отношение S_{Φ} к стандартному отклонению функции на обучающей выборке (S_0)	0,665
Доля оправдавшихся с точностью ($S_{\Phi} < 0,67S_0$) расчетов по модели на контрольной выборке	0,664
Мера качества контрольного модельного расчета по знаку его корреляции с фактическими значениями функции	0,656

Из отобранных 58 моделей исключены те, надежность которых с необходимым уровнем значимости не подтвердилась на контрольной выборке. Таких исключенных моделей оказалось 21. Описания оставшихся 37 моделей, подтвердивших свою надежность по всем критериям, использовались для анализа пространственной структуры взаимосвязей между годовыми суммами осадков на различных метеостанциях. Средние характеристики качества этих 37 моделей приведены в табл. 2.

В общем случае многофакторные регрессионные зависимости могут быть нелинейными, но если их аппроксимировать линейными операторами, то можно, используя знак коэффициентов регрессии, разделить частные взаимосвязи в многофакторных моделях на положительные и отрицательные. Понятно, что вклад каждого предиктора в значимую многофакторную зависимость может быть разным. Кроме того, предикторы в той или иной мере могут быть связаны между собой, и это обстоятельство необходимо учитывать при оценке значимости и интерпретации коэффициентов регрессии. Тем не менее аккуратный анализ структуры и количественных параметров стохастических моделей позволяет установить определенные закономерности, наблюдающиеся в изменчивости годовых сумм осадков на рассматриваемой территории.

Если выделить зависимости, в которых частные коэффициенты регрессии для всех предикторов положительные, то становится очевидно, что таких моделей большинство. В табл. 3 в качестве примера приведены фрагменты описания нескольких таких моделей. Анализируя значения координат метеостанций для функций и аргументов, можно констатировать, что соответствующие различия между ними много меньше характерных размеров рассматриваемой географической области исследования. Это означает, что такие модели отражают наличие некоторой локальной однородности поля годовых осадков, обусловленной синоптическим механизмом формирования его пространственно-временной изменчивости. Этот достаточно тривиальный результат показывает, что анализ коэффициентов регрессии, с учетом значимости их отклонения от нуля, позволяет выделять закономерные эффекты в пространственной структуре полей осадков. Многофакторная зависимость (табл. 4) сильнее ($K1 > K2$ и $K1 > K3$), чем однофакторная зависимость от главного (наиболее существенного) предиктора, связывает в рамках одной модели годовые суммы осадков на сравнительно близко расположенных метеостанциях. Естественно, этот эффект более выражен при слабых парных связях ($K2 < 0,5$,

Таблица 3

Метеостанции с положительными взаимосвязями сумм осадков

Метеостанция	Координаты, град.	
	с. ш.	в. д.
Модель 1		
Талдыкорган	45	78
Алма-Ата	43	78
Анархай	44	75
Баянаул	51	75
Модель 2		
Багдарин	54	114
Троицкий Прииск	55	113
Романовка	53	113
Модель 3		
Иркутск	52	104
Хомутово	52	104
Балаганск	54	103
Хамар-Дабан	52	104
Байкальск	52	104
Усть-Ордынское	54	106
Модель 4		
Карам	55	108
Казачинское	56	108
Качуг	54	106
Нижнеангарск	56	110
Модель 5		
Байкит	61	96
Бор	61	90
Тура	64	100
Тутончаны	64	94

Примечание. Жирным шрифтом выделены метеостанции, данные о суммах осадков на которых используются в модели в качестве функции. Далее следуют аргументы этой функции.

Таблица 4

Параметры, характеризующие положительную связность переменных в много- и однофакторных моделях

Метеостанция для зависимой переменной (функции) в модели	K1	Метеостанция для переменной главного предиктора в модели	K2	K3	K4
Талдыкорган	0,65	Алма-Ата	0,53	0,58	0,62
Иркутск	0,78	Хомутово	0,69	0,72	0,31
Карам	0,56	Казачинское	0,31	0,39	0,40
Багдарин	0,80	Троицкий Прииск	0,71	0,74	0,56
Байкит	0,56	Бор	0,29	0,40	0,13

Примечание. Здесь и в табл. 5: K1 — коэффициент детерминации многофакторной модели; K2, K3 — коэффициенты детерминации для однофакторных моделей зависимости функции от главного предиктора с двумя разными аппроксимирующими операторами; K4 — коэффициент регрессии для главного предиктора в многофакторной модели с нормированными на стандартное отклонение переменными.

Параметры моделей с наличием отрицательных связей

Метеостанция и ее координаты для зависимой переменной (функции) в модели	K1	Метеостанция и ее координаты для предиктора с отрицательным влиянием на изменение функции	K2	K3	K4
Байкальск (52° с. ш., 104° в. д.)	0,83	Киселёвск (54° с. ш., 87° в. д.)	0	0,04	-0,08
Култук (52° с. ш., 104° в. д.)	0,55	Кемерово (55° с. ш., 86° в. д.)	0	0,04	-0,21
Алыгджер (53° с. ш., 98° в. д.)	0,52	Мамакан (58° с. ш., 114° в. д.)	0,03	0,10	-0,26
		Тайшет (57° с. ш., 98° в. д.)	0	0	-0,24
Барнаул (54° с. ш., 84° в. д.)	0,51	Витим (59° с. ш., 112° в. д.)	0,14	0,19	-0,30
Хабары (54° с. ш., 80° в. д.)	0,56	Черемхово (53° с. ш., 103° в. д.)	0	0	-0,18
Балаганск (54° с. ш., 103° в. д.)	0,72	Жесказган (48° с. ш., 67° в. д.)	0,13	0,19	-0,18
		Кемерово (55° с. ш., 86° в. д.)	0,03	0,10	-0,21
Нижнеудинск (55° с. ш., 99° в. д.)	0,61	Сытомино (61° с. ш., 71° в. д.)	0	0,11	-0,24
		Березово (64° с. ш., 65° в. д.)	0	0,07	-0,12
Большое Голоустное (52° с. ш., 105° в. д.)	0,64	Аральск (47° с. ш., 62° в. д.)	0	0,04	-0,19
Большой Ушканий остров (53° с. ш., 108° в. д.)	0,70	Киселёвск (54° с. ш., 87° в. д.)	0	0	-0,19
		Тара (57° с. ш., 74° в. д.)	0	0	-0,28
Хомутово (52° с. ш., 104° в. д.)	0,74	Адыгджер (53° с. ш., 98° в. д.)	0	0	-0,28
Ванавара (60° с. ш., 102° в. д.)	0,59	Сосновка (51° с. ш., 94° в. д.)	0,33	0,30	-0,47
		Усть-Каменогорск (50° с. ш., 82° в. д.)	0,03	0,08	-0,25

Примечание. Жирным шрифтом выделены значимые на уровне 0,05 по критерию Фишера парные связи между переменными. Обозн. см. табл. 4.

$K3 < 0,5$). При необходимости многофакторную модель с высоким коэффициентом детерминации $K1$ можно использовать в прикладных целях, например для приближенного восстановления отсутствующих значений функции, но при имеющихся оценках значений ее аргументов.

Наряду с положительными связями, во многих моделях также были обнаружены отрицательные частные зависимости, характеризующиеся значимыми по критерию Стьюдента коэффициентами в уравнениях множественной регрессии. Анализ пространственной привязки переменных в таких моделях показал, что отрицательная взаимосвязанность годовых сумм осадков характерна для метеостанций, существенно более отдаленных друг от друга (табл. 5), чем для метеостанций с положительными связями. Соответствующие средние модули разности координат метеостанций составляют $1,5^\circ$ по широте и $1,7^\circ$ по долготе для положительных зависимостей и $4,5^\circ$ по широте и $21,8^\circ$ по долготе — для отрицательных.

Следует отметить, что отрицательные зависимости не всегда выявляются при анализе парных взаимосвязей между переменными. Среди представленных в табл. 5 оценок силы связи однофакторных зависимостей только коэффициенты детерминации для пары метеостанций Ванавара—Сосновка можно признать значимыми. Более сильные положительные связи выявляются методами как многофакторного, так и однофакторного анализа зависимостей.

ЗАКЛЮЧЕНИЕ

Использование статистических методов оценки взаимосвязей между годовыми осадками по материалам реальных метеорологических наблюдений требует тщательного и разностороннего контроля доверия к их результатам на всех этапах анализа данных. Формального применения классических приемов оценки значимости не всегда достаточно для обеспечения выводов с необходимым уровнем надежности.

Во временной изменчивости поля годовых сумм осадков на территории Сибири и Казахстана можно выделить следующие закономерные эффекты. Во всех районах рассмотренной географической области имеет место локальная однородность поля осадков, отражающаяся в положительной корреляции временных рядов их годовых сумм, зарегистрированных на близко расположенных метеостанциях. Это свойство обнаруживается как однофакторными, так и многофакторными методами анализа взаи-

мосвязей между переменными. Значимые отрицательные связи между суммами осадков наблюдаются реже, и обнаруживаются они в основном методами многофакторного моделирования как частные зависимости от отдельных переменных. Эта закономерность проявляется для метеостанций, удаленных друг от друга на значительно большее расстояние, в отличие от метеостанций, на которых осадки положительно коррелированы.

Наличие в моделях взаимосвязи одновременно близкой положительной и дальней отрицательной связности годовых сумм осадков на метеостанциях может указывать на присутствие в их пространственно-временной изменчивости эффектов, напоминающих стоячие волны. Однако проверка этой гипотезы требует более детальных исследований.

Статистически значимая связность переменных в построенных моделях может также использоваться для приближенного восстановления отсутствующих данных в прошлом или для коррекции совместных прогнозов годовых осадков на метеостанциях с заблаговременностью в 1–2 годовых шага.

Работа выполнена в рамках НИР (0347-2016-003).

СПИСОК ЛИТЕРАТУРЫ

1. Анисимов О. А., Кокорев В. А. Об оптимальном выборе гидродинамических моделей для оценки влияния изменений климата на криосферу // Лед и снег. — 2013. — № 1 (121). — С. 83–92.
2. Литвинова О. С., Гуляева Н. В. Анализ временных рядов осадков Обь-Иртышского междуречья в XX–начале XXI вв. // Динамика окружающей среды и глобальные изменения климата. — Изд-во Югор. ун-та, 2010. — С. 38–45.
3. Савич-Шемет О. Г. Пространственно-временная структура многолетних колебаний жидких атмосферных осадков на территории Беловежской пуши // Природопользование. — 2011. — Вып. 20. — С. 35–40.
4. Попова В. В. Пространственные связи в поле осадков на ЕТ СССР и их зависимость от типа циркуляции атмосферы // Изв. АН СССР. Сер. геогр. — 1989. — № 1. — С. 98–104.
5. Максютова Е. В., Кичигина Н. В., Воропай Н. Н., Бальбина А. С., Осипова О. П. Тенденции гидроклиматических изменений на Байкальской природной территории // География и природ. ресурсы. — 2012. — № 4. — С. 72–80.
6. Беженару Г., Мельничук О. Характеристика и оценка статистических параметров годовых атмосферных осадков на территории Республики Молдова [Электронный ресурс]. — http://www.meteo.md/metodf_pdf/harakteristika_godovih_osadkov.pdf (дата обращения 14.04.2017).
7. Игнатов А. В., Кравченко В. В., Чекмарёв А. А. Стохастическое моделирование: версия «Эксперт». Свидетельство о государственной регистрации программы для ЭВМ № 2016617121 от 27.06.2016.
8. Игнатов А. В. Сравнение различных моделей годового гидрологического цикла и расчетных оценок месячного стока с водосбора // География и природ. ресурсы. — 2011. — № 4. — С. 118–124.
9. Чекмарёв А. А. Модели трендов климатических характеристик и их прогностические возможности // Вестн. Иркут. техн. ун-та. — 2015. — № 7. — С. 32–38.
10. Чекмарёв А. А., Игнатов А. В. Апробация программного пакета «Стохастическое моделирование» на примере разработки моделей для прогнозирования годового числа Вольфа // Вестн. Воронеж. ун-та. Сер. География. Геоэкология. — 2015. — № 1. — С. 5–13.
11. Бураков Д. А., Гордеев И. Н., Игнатов А. В., Петкун О. Э., Путинцев Л. А., Чекмарёв А. А. Прогнозирование притока воды в Красноярское и Саяно-Шушенское водохранилища во втором квартале года // География и природ. ресурсы. — 2016. — № 2. — С. 175–183.
12. Христофоров А. В. Особенности задачи прогноза гидрологических характеристик по уравнениям регрессии // Метеорология и гидрология. — 1975. — № 11. — С. 72–80.
13. Христофоров А. В. Надежность расчетов речного стока. — М.: Изд-во Моск. ун-та, 1993. — 166 с.

Поступила в редакцию 21 июня 2017 г.