

ГОСУДАРСТВЕННЫЙ КОМИТЕТ СССР
ПО ГИДРОМЕТЕОРОЛОГИИ
И КОНТРОЛЮ ПРИРОДНОЙ СРЕДЫ

ТРУДЫ
ОРДЕНА ТРУДОВОГО КРАСНОГО ЗНАМЕНИ
ГЛАВНОЙ ГЕОФИЗИЧЕСКОЙ ОБСЕРВАТОРИИ
им. А. И. ВОЕЙКОВА

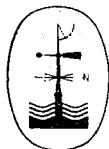
Выпуск

480

ПРИМЕНЕНИЕ
СТАТИСТИЧЕСКИХ МЕТОДОВ
В МЕТЕОРОЛОГИИ

Под редакцией

д-ра физ.-мат. наук Р. Л. КАГАНА,
д-ра физ.-мат. наук, проф. М. И. ЮДИНА
канд. геогр. наук А. В. МЕЩЕРСКОЙ



ЛЕНИНГРАД ГИДРОМЕТЕОИЗДАТ 1985

00

778 УДК 551.509; 551.501

Сборник посвящен вопросам применения статистических методов в метеорологии и климатологии. В сборнике помещен обзор научной деятельности д-ра физ.-мат. наук Р. Л. Кагана и последние подготовленные им статьи, в которых проанализированы и обобщены данные о статистической структуре метеорологических полей применительно к вопросам методики наблюдений и контроля их качества, а также корректной оценке климатообразующих факторов. Рассматривается применение информационного подхода в исследованиях изменений климата, методика долгосрочного прогнозирования и оптимальной стратегии использования прогностической и климатологической информации и вопросы объективного анализа.

Расчитан на метеорологов, климатологов, преподавателей и аспирантов гидрометеорологической специальности.

The publication deals with the problems of using statistical methods in meteorology and climatology. A significant contribution to this direction was made by the late Dr. R. L. Kagan. The survey of research activities of R. L. Kagan is given in the publication as well as his last articles in which there are analyzed and generalized data on the statistical structure of meteorological fields as applied to the problems of techniques for observations and control of their quality, correct estimation of climate-forming factors.

The other articles discuss the use of information approach in climate change studies, techniques for long-range forecasting and optimal strategy of using prognostic and climatological information, the problems of objective analysis.

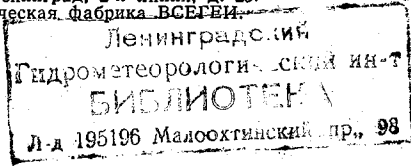
It is mean for meteorologists, climatologists, instructors and post-graduates of hydrometeorological speciality.

V57193

Труды ГГО, вып. 480

ПРИМЕНЕНИЕ СТАТИСТИЧЕСКИХ МЕТОДОВ В МЕТЕОРОЛОГИИ

Редактор Е. А. Соснина. Технический редактор Г. В. Ивкова. Корректор Л. И. Хромова.
 Сдано в набор 29.05.85. Подписано в печать 21.11.85. М-22667. Формат 60×90^{1/16}.
 Бум. тип. № 1. Лит. гарнитура. Печать высокая. Печ. л. 8,5. Кр.-отг. 8,75.
 Уч.-изд. л. 10,03. Тираж 610. Индекс МОЛ-233. Заказ 322. Цена 70 коп. Заказное.
 Гидрометеоздат. 199053. Ленинград, 2-я линия, д. 23.
 Ленинградская картографическая фабрика ВСЕСОИ.



П 1903040000-217 43-82(2)
 069(02)-85

© Главная геофизическая обсерватория
 им. А. И. Воейкова (ГГО), 1985.

Л. В. Руховец, Е. И. Хлебникова

О НАУЧНОЙ ДЕЯТЕЛЬНОСТИ Р. Л. КАГАНА

Р. Л. Каган начал свою научную деятельность еще будучи студентом физического факультета Ленинградского государственного университета. В 1951 г. под руководством М. И. Юдина он выполнил дипломную работу «О рассеянии солнечной радиации в облаках и туманах». В дальнейшем основные результаты этой работы были опубликованы [1]. В работе был предложен новый приближенный способ решения задачи о рассеянии света. Как известно, эта задача сводится к решению системы интегро-дифференциальных уравнений для интенсивности потоков восходящей и нисходящей радиации. Трудности в решении этой задачи связаны с необходимостью учета диффузности радиации. Шварцшильд и Эддингтон предложили сравнительно простые приближенные методы решения указанной задачи, основанные, однако, на довольно грубых предположениях относительно искомых величин, что повлияло на точность получаемых результатов. Более точные методы, предложенные Чандрасекаром и Е. С. Кузнецовым, требовали очень большого объема вычислений.

Метод приближенного решения задачи, изложенный в [1], по объему вычислений мало отличался от методов Шварцшильда и Эддингтона, но давал гораздо лучшие результаты, достаточно близкие к точному решению.

В основе этого метода лежало подтвержденное эмпирическими данными предположение о зависимости осредненных значений секанса зенитного угла для восходящего и нисходящего потоков только от общей толщины слоя. Такое предположение позволило существенно упростить математическую формулировку задачи и получить сравнительно простые расчетные формулы.

Интересно отметить, что много лет спустя Каган использовал опыт этой работы в задаче расчета радиационных потоков в условиях городской застройки. Этот вопрос весьма важен для корректного учета влияния радиации на тепловой режим зданий. Каган разработал теорию и алгоритмы таких расчетов [101, 102, 105]. Этот алгоритм был реализован и использован в работах Л. П. Клягиной.

После окончания ЛГУ Каган в течение нескольких лет работает инженером-синоптиком в Средней Азии. В 1955 году он поступил в аспирантуру Главной геофизической обсерватории им. А. И. Воейкова. Здесь под руководством М. И. Юдина Каган выполнил цикл работ по динамике крупномасштабных процессов в атмосфере. В этих работах решаются задачи численного прогноза полей метеорологических элементов при учете конденсационных притоков тепла и влияния стратификации стратосферы [2, 4, 5, 6, 9, 15]. С математической точки зрения обе эти задачи могут рассматриваться как частные случаи одной общей проблемы учета вертикальной расслоенности атмосферы в задаче гидродинамического прогноза. Для решения этой проблемы необходимо было сформулировать граничные условия на поверхностях раздела слоев. Каган вывел эти условия для поверхностей слабого разрыва в предположении изобаричности этих поверхностей [6]. В частности, им было показано, что на указанных поверхностях не терпят разрыва аналог вертикальной скорости \mathcal{W} в ξ -системе координат и его производная $d\mathcal{W}/d\xi$. Эти два условия непрерывности на поверхностях раздела и были использованы Каганом для решения ряда задач численного прогноза для многослойной атмосферы, что позволило учитывать изменяющуюся с высотой стратификацию. В частности, было оценено влияние стратосферной стратификации на динамику процессов в тропосфере [4]. При решении задачи прогноза с учетом облачных слоев был применен интересный метод, идея которого была предложена Юдиным, обратившим внимание на то, что слагаемые уравнения притока тепла, описывающие вертикальный перенос эквивалентно-потенциальной температуры, в облачных слоях значительно меньше других слагаемых этого уравнения. Пренебрежение указанным слагаемым для отмеченного слоя позволило Кагану упростить решение пространственной задачи прогноза. На основе полученных теоретических результатов он разработал практическую схему расчета тенденций приземного давления, вертикальных токов и конденсации в облачном слое с учетом притоков тепла конденсации [9] и показал, используя фактические данные, что учет конденсационных притоков уточняет расчеты указанных элементов.

Этот цикл работ составил основу кандидатской диссертации Кагана [5], успешно им защищенной в 1961 г.

К рассмотренному циклу исследований близко примыкает работа [3], выполненная Каганом совместно с М. Е. Швецом. В этой работе предложен метод решения линейного дифференциального уравнения эллиптического типа с переменными коэффициентами, к которому, в частности, сводится уравнение для вертикальной составляющей скорости в случае переменного по высоте параметра статистической устойчивости. Для решения этого уравнения была применена конечно-разностная аппроксимация по вертикальной координате и решение было получено в явном виде. Хотя это решение едва ли целесообразно применять для практических расче-

тов, однако для качественного анализа такое решение весьма полезно.

Если в начальный период своей научной деятельности Каган занимался главным образом динамическими (детерминистически) задачами, то в дальнейшем он почти полностью переключился на применение стохастических методов. Толчком к этому явилась важная в практическом отношении задача рационализации снегомерных измерений [10—13, 15, 69]. Каган был одним из авторов нового подхода к решению этой задачи. Этот подход базируется на априорных оценках точности определения осредненных вдоль той или иной линии значений метеорологического элемента по данным дискретных измерений. В ряде упомянутых работ [10, 12, 13, 14, 31] Каган развил общий подход к данной задаче. Наряду с точными, но довольно громоздкими формулами, он вывел приближенную формулу, позволяющую весьма просто оценивать дисперсию ошибок осредненной величины на основании данных о статистической структуре рассматриваемого элемента, точности и степени дискретности его измерений. Он выполнил также исследования по оптимальному в стохастическом смысле одномерному осреднению, обобщив известные результаты Р. Л. Соловейчика и др.

Для применения разработанных методов к характеристикам снежного покрова необходимо было изучить их статистическую структуру. Кагану удалось заинтересовать этой проблемой ряд сотрудников ГГО (М. В. Гушину, А. И. Полищук, И. Д. Копанева) и совместно с ними выполнить соответствующую обработку эмпирических данных. Практическим результатом этого цикла работ явилось существенное усовершенствование используемой ранее методики снегомерных наблюдений.

Примерно в это же время [17, 19, 22, 29, 47] Р. Л. Каган разработал оригинальный и плодотворный подход к интерпретации измерений, производимых инерционными приборами, который базируется на применении данных о временной статистической структуре измеряемого элемента. Идея применения стохастического подхода к интерпретации приборных измерений сама по себе не была новой. Важные исследования в этом направлении были выполнены, в частности, А. М. Ягломом. Однако все предыдущие исследования касались точности определения мгновенных значений измеряемых элементов. Каган впервые обратил внимание на тот факт, что основной интерес представляют не мгновенные значения, а осредненные по времени. Ему удалось установить, что эта задача носит экстремальный характер, т. е. что при заданных параметрах прибора существует такой интервал времени, когда осредненные по этому интервалу значения определяются точнее всего. С практической стороны еще интереснее обратная задача — о выборе параметров прибора с целью наиболее точного определения значений, заданным образом осредненных по времени. Для линейного прибора первого порядка Каган получил соотношение, связывающее между собой радиус временной корреляции измеряе-

мого элемента τ , постоянную времени прибора T и оптимальный в указанном смысле интервал T_0 . Он показал, что это соотношение в широком интервале значений τ практически не зависит от τ и имеет простой вид $T_0 = 1,7 T$. В дальнейшем Р. Л. Каган обобщил это исследование в ряде направлений (замена арифметического осреднения оптимальным, использование нескольких отсчетов инерционного прибора и т. п. (см. [97], глава 2)).

Предложенный Каганом подход нашел широкое применение и получил дальнейшее развитие в общей теории проектирования измерительных систем. Применению этого подхода к измерениям температуры воздуха посвящена диссертационная работа Л. Л. Брагинской, выполненная под руководством Кагана.

Впоследствии, в 70-х годах, Каган неоднократно возвращался к вопросу о точности одномерного осреднения: в [126] при изучении третьего и четвертого моментов величин, получаемых в результате осреднения, и отклонений этих величин от истинного среднего, а также в ряде статей, написанных в последние годы [123, 125, 129, 137, 141], где рассматривается вопрос о точности осреднения периодически коррелированных величин. Последний вопрос возник в связи с изучением влияния выбора сроков наблюдений и способов осреднения на точность средних суточных значений метеорологических параметров. В работах этого цикла, выполненных при участии Е. Е. Сибир, убедительно показано, что, для того чтобы в полной мере использовать информацию, содержащуюся в современных детальном измерениях, необходим переход от принятой методики климатологической обработки к более корректным способам осреднения. Такие способы были разработаны Каганом. Внедрение их является задачей ближайшего будущего.

Кагану принадлежит большой цикл работ, посвященных стохастическому подходу к построению методов осреднения метеорологических элементов по площади [21, 25, 32, 33, 38, 41, 42, 44, 67, 84, 85, 90, 94, 110, 114, 123], причем если в задаче об одномерном осреднении у Кагана были предшественники, то теория площадного осреднения по существу целиком им создана и разработана. Он предложил метод оптимального осреднения по площади, являющийся обобщением метода оптимальной двумерной интерполяции. Наряду с оптимальным осреднением он исследовал другие методы осреднения, проанализировал их точность и определил границы применимости методов, а также разработал усовершенствования некоторых из них.

Практическое значение двумерного осреднения особенно велико применительно к проблеме интерпретации осадкомерных данных. Различным аспектам этой проблемы, в разработке которой принимала активное участие А. И. Полищук, посвящены работы [25, 34, 35, 38, 45, 62, 69].

Аналогичным образом обстоит дело и со многими другими метеорологическими элементами, в частности с рядом компонентов водного, радиационного и теплового балансов подстилающей

поверхности. В последние годы актуальность проблемы площадного осреднения еще более возросла благодаря развитию новых методов измерений, в первую очередь спутниковых и радиолокационных, позволяющих непосредственно получить осредненные по площади значения.

Еще одной областью приложения стохастической теории осреднения, связанной с проблемой исследований глобального климата и его изменений, является оценка осредненных по широтным кругам значений температуры. Этот вопрос рассмотрен в ряде работ Кагана, выполненных совместно с К. М. Лугиной и др. [100, 120, 123, 133]. Было установлено, в частности, что применявшиеся многими авторами простое арифметическое осреднение данных станций, расположенных в разнородных по плотности сети районах, является совершенно неудовлетворительным. Вместе с тем оценки показали достаточную точность сравнительно простого метода, заключающегося в предварительном осреднении данных по равновеликим широтно-долготным «квадратам» с последующим арифметическим их осреднением.

Исследования по стохастической теории осреднения были обобщены Каганом в работе «Осреднение метеорологических полей» [127], опубликованной в 1979 г. — единственной монографии на эту тему не только в отечественной, но и в мировой литературе.

В известной мере противоположной осреднению является задача определения дифференциальных характеристик метеорологических полей (в первую очередь их градиентов) по данным о самих дифференцируемых элементах. В противоположность осреднению дифференцирование связано не с повышением, а с уменьшением точности. Поэтому в данном случае особенно важны методы, позволяющие свести потерю точности к минимуму.

Этому вопросу посвящен ряд работ Кагана [41, 43, 52, 53]. Им разработан метод оптимального, в стохастическом смысле, дифференцирования, который, как и оптимальное осреднение, является естественным обобщением оптимальной интерполяции. В результате обоих обобщений получилась стройная система стохастически оптимальных методов интерпретации метеорологических данных.

Каган исследовал точность, которая достигается при использовании простых разностных аппроксимаций производных, сопоставил ее с точностью оптимального дифференцирования и тем самым оценил пределы применимости простых разностных формул. Он разработал также методику совместного использования данных о самих элементах и об их производных для восстановления полей этих элементов. Эта задача важна для объективного анализа метеорологических полей, а именно, при совместном анализе данных о ветре и давлении (геопотенциале). В последнее время методика такого совместного анализа (так называемая мультивариантная интерполяция) нашла применение в системах оперативного объективного анализа в прогностических центрах ряда стран. Ведется подготовка к внедрению этой методики в Гидрометцентре СССР.

Вопросы осреднения и дифференцирования метеорологических полей составили ядро докторской диссертации Кагана «Проблемы интерпретации метеорологических наблюдений», защищенной им в 1968 г. Наряду с этими вопросами в диссертации рассмотрен ряд других аспектов стохастической теории метеорологических полей и ее применений. Некоторые из них излагаются ниже.

Наряду с объективным анализом стохастический подход широко используется в задаче контроля метеорологической информации, т. е. выявления и устранения грубых ошибок. Применение этого подхода позволяет объективизировать процедуру контроля и автоматизировать ее на базе ЭВМ.

Каган разработал соответствующую методику для так называемого критического контроля, т. е. для выявления и устранения систематических погрешностей наземной измерительной аппаратуры. Этому вопросу посвящены, в частности, работы [23, 48, 61] (см. также [106], гл. VII, п. 4). В настоящее время данная методика внедрена в практику неоперативного контроля метеорологической информации.

Все перечисленные выше методы, а также ряд других, которые будут рассмотрены ниже, базируются на использовании данных о статистической структуре метеорологических полей, и прежде всего о корреляционных функциях метеорологических элементов. В связи с этим Каган много занимался непосредственным исследованием структуры различных полей [26, 33, 36, 50, 100, 108], причем для ряда полей, например для составляющих теплового баланса океанов, такие исследования были выполнены им впервые. Но еще большее значение имеет разработанная Каганом [60, 66, 80, 89] единая методика определения характеристик статистической структуры метеорологических полей, учитывающая ряд свойств реальной метеорологической информации, в частности наличие пропусков в исходных данных. Им и при его участии эта методика доведена до конкретных алгоритмов и программ. В настоящее время она является общепринятой.

Каганом предложены также методы более детального исследования статистической структуры, позволяющие учесть такие факторы, как анизотропия и сдвиги центров корреляции по высоте и по времени [109, 121, 124, 128]. В реализации этих методов принимали участие К. М. Лугина, В. П. Тараканова, Д. Девени и др.

Для правильного использования данных о статистической структуре необходимо знать, насколько эти данные надежны, т. е. каковы их погрешности, в частности погрешности выборочные. Оценка этих погрешностей является весьма непростой задачей, поскольку к ней в силу внутренней связности исходных данных неприменимы обычные способы оценки выборочных погрешностей коэффициентов корреляции. Ряд работ Кагана [72, 76, 83, 91, 96] был направлен на преодоление этих трудностей. Ему удалось разработать более общий подход к оценке выборочных погрешностей и показать, что реальные выборочные погрешности корреляционных функций существенно выше их оценок по классическим фор-

мулам статистики. Это обстоятельство необходимо должным образом учитывать как при сопоставлении результатов исследования статистической структуры, проведенных по различным выборкам, так и при интерпретации выводов, получаемых при использовании данных о статистической структуре.

Разрабатывая новые области применения данных о статистической структуре метеорологических полей, Каган уделял также много внимания совершенствованию таких ставших уже традиционными направлений применения стохастических методов, как объективный анализ метеорологических полей [18, 27, 52, 74, 75, 112] и рациональное планирование систем метеорологических наблюдений [39, 40, 46, 51, 64, 68, 82]. В частности, он был соавтором экономического подхода к проблеме планирования систем наблюдений — единственного подхода, действительно позволяющего определять оптимальные параметры различных наблюдательных систем, в том числе оптимальную густоту сети станций.

Ряд исследований был выполнен Каганом в связи с проблемой так называемого четырехмерного (пространственно-временного) анализа метеорологических полей, предусматривающего использование асиноптической, относящейся к произвольным моментам времени, информации [74, 78, 98, 101, 121, 128]. Анализируя полученные Лугиной данные о пространственно-временной структуре давления, Каган показал, что применявшаяся рядом отечественных и зарубежных авторов аппроксимация пространственно-временной корреляции в виде произведения функций от расстояния и от интервала времени является неудовлетворительной. Он предложил иную аппроксимацию, базирующуюся на гипотезе о горизонтально-временной изотропии давления, и показал, что эта гипотеза приводит к результатам, хорошо согласующимся с данными о структуре дифференциальных характеристик поля давления. Пользуясь упомянутой аппроксимацией, он выявил далее, что точная привязка асиноптических наблюдений во времени, которой придают большое значение зарубежные исследователи, в действительности не имеет большого значения. Гораздо большую роль играет корректный учет других свойств асиноптической информации, а именно, большой дисперсии ошибок этой информации и, в особенности, значительной корреляции последних. Каган выдвинул идею построения системы четырехмерного анализа на основе оптимальной пространственно-временной интерполяции. Возможности реализации этой идеи еще далеко не исчерпаны в настоящее время.

Работы Кагана, посвященные различным аспектам проблемы интерпретаций метеорологических наблюдений, основывались главным образом на применении данных о корреляционных функциях метеорологических полей. В конце 60-х годов им было начато новое направление исследований. В этот период Каган и его сотрудники приступили к изучению таких характеристик метеорологических элементов, которые зависят не только от их корреля-

ционных функций, но и от распределения метеорологического элемента в целом. Прежде всего предметом исследований стали так называемые выбросы метеорологических элементов, т. е. характеристики непрерывной продолжительности пребывания метеорологических элементов выше или ниже заданного уровня. Такие характеристики имеют большое значение для решения задач прикладной климатологии. Вместе с тем определение их путем непосредственной обработки данных наблюдений требует привлечения весьма большого количества исходного материала и зачастую приводит к малонадежным результатам. Кагана заинтересовала возможность определения характеристик выбросов с помощью стохастических методов. Первые работы [57, 58, 88], посвященные этому направлению, дали обнадеживающие результаты. Они показали, что климатические характеристики выбросов метеорологических элементов, вообще говоря, могут быть получены с достаточной точностью на основе использования дискретных стохастических моделей и положили начало широкому кругу исследований по разработке методики косвенного расчета характеристик выбросов и анализу влияния различных факторов на характеристики выбросов и некоторые другие связанные с ними величины.

Поскольку применение стохастического подхода к определению характеристик выбросов производится на основе анализа фактических данных, Каган уделял большое внимание созданию единых алгоритмов эмпирической обработки, дающих информацию, необходимую для формулировки моделей, и обеспечивающих возможно меньше искажение характеристик выбросов [99, 111, 115].

Под руководством Кагана была разработана методика расчета характеристик выбросов средней суточной температуры воздуха по данным, опубликованным в климатических справочниках. Первые результаты были опубликованы в [57, 58, 88]. В дальнейшем, в работах Е. И. Федорченко, выполненных по инициативе и под руководством Кагана, было продолжено изучение возможностей стохастического подхода для определения характеристик выбросов температурных рядов, а методика косвенного расчета характеристик выбросов средней суточной температуры воздуха была доведена до реализации для территории СССР. Эта методика включает в себя как формулировку принятой стохастической модели, так и способы задания параметров этой модели. Для определения необходимых параметров с учетом нестационарности годового хода очень полезной оказалась методика восстановления [92] моментов распределения метеорологических элементов на каждый день месяца по данным о повторяемости значений в месячных совокупностях, основанная на сплайн-аппроксимациях.

Предложенные в [87] приближенные формулы для расчета вероятности выброса гауссовской (в [104] релеевской) последовательности позволили не только значительно упростить вычисления при реализации методики косвенного расчета, но и явились основой для изучения влияния различных факторов на характеристики выбросов.

К числу таких факторов, отражающих реальные свойства метеорологической информации, в первую очередь следует отнести наличие ошибок наблюдений и дискретность измерений метеорологических процессов. В [91] показано, что при наличии ошибок наблюдений существует оптимальная дискретность измерений, и приведены оценки оптимальной, в среднем и в среднем квадратическом, дискретности измерений при определении числа выбросов процесса. В [127] аналогичная постановка задачи рассматривается применительно к определению характеристик экстремумов метеорологических процессов.

В [103] Каган обобщил полученные им результаты о влиянии ошибок наблюдений на дифференциальные характеристики процессов и показал, что задача выбора дискретности отсчетов, обеспечивающей отсутствие систематической или, как минимум, случайной ошибки в определении числа выбросов, может быть сведена к описанной выше задаче выбора оптимального шага дифференцирования [41].

Каган был одним из пионеров применения метода статистического моделирования в метеорологических задачах. Впервые он обратил внимание на возможность использования этого метода в связи с задачей оценки таких характеристик выбросов, аналитический расчет которых представляет значительные вычислительные трудности [65]. Дальнейшие проработки показали целесообразность применения метода статистического моделирования не только для получения оценок климатических характеристик, но главным образом для решения различных методических вопросов. Каган широко использовал в своих работах метод статистического моделирования. Особую роль этот метод сыграл при оценке выборочных погрешностей различных характеристик с учетом реальных свойств метеорологических данных [117, 134].

Ряд исследований Кагана [79, 116, 134, 135, 136] посвящен разработке алгоритмов статистического моделирования коррелированных случайных процессов, являющихся достаточно экономичными и учитывающими специфические свойства рядов различных метеорологических элементов. Наибольшие трудности при моделировании возникают при рассмотрении существенно негауссовских метеорологических процессов. Каган уделял таким процессам много внимания [116, 134—136] главным образом в связи с изучением характеристик поля ветра. Основываясь на некоторых применяемых стохастических моделях ветра, с помощью метода статистического моделирования Каган получил интересные результаты о точности статистической экстраполяции вероятных скоростей ветра малой обеспеченности. Он, в частности, показал [131], что для оценки экстремальных значений скорости ветра по данным выборок ограниченного объема целесообразно использовать метод максимального правдоподобия, который обеспечивает существенно большую точность, чем обычно применяемый для этой цели метод наименьших квадратов.

В последние годы Каган провел исследования двух важных вопросов, связанных с актуальной проблемой изменений глобального климата [100, 120, 123, 131, 133, 137]. Первый из них — это оценка возможностей определения средних по большим площадям концентраций так называемых малых примесей в атмосфере, которые могут существенно влиять на глобальный климат. На примере наиболее «благополучного» из такого рода элементов — общего содержания озона в вертикальном столбе атмосферы, который измеряется на сравнительно густой сети станций, Каган убедительно показал, что определение изменений глобальных концентраций озона по таким измерениям практически невозможно. Учитывая малую густоту и неравномерность существующей наземной озонметрической сети, можно сделать вывод, что для получения сколько-нибудь достоверных суждений о тенденциях изменения содержания озона необходимо значительное пространственное и временное осреднение на основе сведений о статистической структуре общего содержания озона, а также привлечение данных спутниковых озонметрических наблюдений.

Этот вывод имеет важное значение для правильной интерпретации существующей информации о якобы имевших место изменениях глобального содержания некоторых малых примесей, но особенно для правильного планирования системы наблюдений для прослеживания глобальных изменений во времени — так называемого мониторинга малых примесей.

Второй вопрос связан с интерпретацией данных нестабильной, постоянно совершенствующейся системы метеорологических наблюдений. Этому вопросу уделялось незаслуженно малое внимание, хотя он имеет кардинальное значение при изучении изменений климата. Каган показал [131, 139], что имеющиеся данные об увеличении в последние годы изменчивости температуры в некоторых районах арктического бассейна могут быть полностью объяснены происшедшим расширением сети станций в этих районах и, следовательно, отнюдь не свидетельствуют о каких-либо реальных изменениях климата.

Для того чтобы избежать подобных ложных эффектов, Каган предложил видоизменение метода оптимальной интерполяции. Оставаясь во всех прочих отношениях оптимальной, такая интерполяция обеспечивает сохранение (а не уменьшение) дисперсии интерполируемого элемента по сравнению с ее фактическим значением.

Изучая конкретные вопросы исследования характеристик структуры метеорологических полей и их применения, Каган вместе с тем уделял большое внимание обобщению работ данного направления, их систематическому рассмотрению. Выше уже упоминались монографии [106] и [122]. К этому же направлению относится ряд статей [24, 27, 29, 41, 43, 44, 81, 93, 107, 118]. Большую роль в пропаганде достижений отечественной науки за рубежом сыграли и продолжают играть написанные Каганом по заданию Всемирной метеорологической организации ключевые разделы Посо-

бия по проектированию сети гидрологических станций [68—70], изданные ВМО в 1972 г. Каган принимал также активное участие в ряде международных симпозиумов по вопросам статистической структуры и ее приложений [78—85], не только выступая с докладами, но и активно руководя подготовкой докладов других сотрудников.

Несмотря на тяжелую болезнь, Каган буквально до последнего дня своей жизни активно занимался научными исследованиями. Он вынашивал планы применения разработанного им стохастического подхода к изучению широкого круга характеристик общей циркуляции атмосферы и глобального климата, описывающих перенос энергии, тепла, влаги и моментов количества движения. Он намечал переход от чисто стохастических исследований к работам стохастико-динамического направления.

Безвременная кончина помешала Кагану осуществить эти планы. Реализация их является долгом его многочисленных учеников и последователей.

СПИСОК НАУЧНЫХ ТРУДОВ Р. Л. КАГАНА

- 1*. Приближенное решение уравнения рассеяния света. — Изв. АН СССР, сер. геофиз., 1956, № 8, с. 968—975.
2. Влияние конденсационных притоков тепла на изменения давления во времени. — Метеорология и гидрология, 1959, № 2, с. 7—14.
- 3*. К вопросу о предвычислении вертикальной скорости в атмосфере. — Труды ГГО, 1959, вып. 99, с. 123—130.
4. К учету влияния конденсационных притоков тепла и стратосферы на изменение поля давления. — Труды ГГО, 1959, вып. 99, с. 15—36.
5. Влияние стратификации на динамику крупномасштабных движений в атмосфере. Автореф. дисс. на соискание учен. степени канд. физ.-мат. наук. — Л.: Изд-во ЛГУ, 1960. — 7 с.
6. К расчету вертикальных токов и конденсации в облаках. — Труды ГГО, 1960, вып. 114, с. 9—28.
- 7*. Некоторые вопросы рационализации снегосъемок. — Труды ГГО, 1960, вып. 108, с. 3—18.
8. О влиянии конденсационных притоков тепла на вертикальные токи в атмосфере. — Метеорология и гидрология, 1960, № 5, с. 3—10.
9. Опыт вычисления некоторых метеорологических элементов с учетом конденсационных притоков тепла в облаках. — Труды ГГО, 1961, вып. 121, с. 67—79.
- 10*. О точности определения средней высоты снежного покрова по дискретным данным. — Труды ГГО, 1962, вып. 130, с. 3—10.
11. О точности экстраполяции высоты снежного покрова во времени. — Труды ГГО, 1962, вып. 130, с. 38—50.
- 12*. К выбору оптимальной методики снегосъемок. — Труды ГГО, 1963, вып. 112, с. 78—86.
- 13*. К точности определения средней величины по дискретным данным. — Труды ГГО, 1963, вып. 112, с. 87—99.
- 14*. О рационализации снегомерных наблюдений на метеорологической сети. — В кн.: Труды всесоюзного научного метеорологического совещания, 1963, т. 9, с. 353—358.
15. Предвычисление давления, вертикальных токов и осадков с учетом тепла конденсации. — В кн.: Труды всесоюзного научного метеорологического совещания, 1963, т. 2, с. 144—152.

* Статьи, отмеченные звездочкой, выполнены совместно с другими авторами.

- 16*. Вопросы картографирования радиационных измерений метеорологических спутников. — Труды ГГО, 1964, вып. 166, с. 227—234.
17. К учету инерции прибора при метеорологических измерениях. — Изв. АН СССР, сер. геофиз., 1964, № 2, с. 302—308.
18. Об одном частном случае интерполяции. — Труды ГГО, 1965, вып. 168, с. 94—100.
19. Об учете инерции прибора при метеорологических измерениях. — Труды ГГО, 1965, вып. 174, с. 21—34.
20. О расчете потоков теплового излучения в безоблачной атмосфере. — Труды ГГО, 1965, вып. 174, с. 158—174.
21. О точности определения средней по площади по данным точечных измерений. — Труды ГГО, 1965, вып. 175, с. 117—131.
- 22*. О точности определения средних величин по мгновенным отсчетам инерционных приборов. — Труды ГГО, 1965, вып. 174, с. 3—20.
- 23*. К вопросу об автоматизации контроля метеорологических данных. — Труды ГГО, 1966, вып. 194, с. 11—15.
- 24*. К вопросу об аппроксимации статистической структуры. — Труды ГГО, 1966, вып. 191, с. 18—21.
25. К оценке репрезентативности осадкомерных данных. — Труды ГГО, 1966, вып. 191, с. 28—34.
- 26*. Об аномалиях членов теплового баланса океана. — Метеорология и гидрология, 1966, № 1, с. 24—28.
- 27*. Об интерполяции метеорологических элементов во времени. — Труды ГГО, 1966, вып. 194, с. 3—10.
- 28*. Об использовании методов объективного анализа метеорологических полей для исследования полей физических характеристик океана и атмосферы над ним. — В кн.: Проблемы получения и обработки информации о физическом состоянии океана и атмосферы над ним. Киев: Наукова думка, 1966, с. 68—78.
- 29*. О влиянии инерции прибора на точность метеорологических измерений. — Труды ГГО, 1966, вып. 191, с. 72—76.
30. О некоторых приложениях статистической структуры гидрометеорологических полей. — В кн.: Проблемы получения и обработки информации о физическом состоянии океана и атмосферы над ним. Киев: Наукова думка, 1966, с. 47—58.
31. Оптимальные формулы численных квадратур для случайных функций специального вида. Труды ГГО, 1966, вып. 191, с. 77—88.
32. О редукции метеорологических элементов на площади. — Труды ГГО, 1963, вып. 191, с. 122—132.
- 33*. О репрезентативности осадкомерных данных. — Метеорология и гидрология, 1967, № 3, с. 101—106.
- 34*. О статистической структуре осредненных значений метеорологических элементов. — Труды ГГО, 1966, вып. 191, с. 92—121.
- 35*. О статистической структуре поля осадков. — Труды ГГО, 1966, вып. 191, с. 35—46.
36. О статистической структуре средних месячных значений составляющих теплового баланса в Северной Атлантике. — В кн.: Проблемы получения и обработки информации о физическом состоянии океана и атмосферы над ним. Киев: Наукова думка, 1966, с. 27—34.
37. К оценке эффективности активных воздействий по данным наземной осадкомерной сети. — В кн.: Исследования по физике облаков и активным воздействиям на погоду. М.: Гидрометеиздат, 1967, с. 107—111.
38. Некоторые вопросы интерпретации осадкомерных данных. — Труды ГГО, 1967, вып. 208, с. 64—75.
- 39*. Об экономическом подходе к планированию сети метеорологических станций. — Труды ГГО, 1967, вып. 208, с. 120—131.
- 40*. Об экономическом подходе к размещению сети станций. — В кн.: Проблемы изучения и обработки информации о физическом состоянии океана. Севастополь, 1967, с. 167—174.
41. О построении оптимальных конечно-разностных и квадратурных формул для однородных и изотропных случайных полей. — Изв. АН СССР. Физика атмосферы и океана, 1967, т. 3, № 6, с. 592—601.

- 42*. О точности определения среднего слоя осадков на площади. — Труды ГГО, 1967, вып. 208, с. 49—63.
43. О расчете интегральных и дифференциальных характеристик случайных полей. — В кн.: Проблемы получения и обработки информации о физическом состоянии океана. Севастополь, 1967, с. 192—204.
44. Проблемы интерпретации метеорологических наблюдений. — Автореф. дисс. на соискание учен. степени доктора физ.-мат. наук. Л., 1967. — 17 с.
45. Интерпретация осадкомерных данных и оценка густоты сети станций. — *Idöjárás*, 1968, том 72, № 4, с. 197—203.
- 46*. К вопросу о рациональном планировании сети наблюдений за температурой воздуха. — Труды ГГО, 1968, вып. 228, с. 30—40.
47. О сравнении показаний инерционных приборов. — Труды ГГО, 1968, вып. 228, с. 131—136.
48. О точности интерполяции при критическом анализе метеорологических данных. — Труды ГГО, 1968, вып. 230, с. 48—52.
- 49*. О точности конечно-разностного представления производных. — Труды ГГО, 1968, вып. 228, с. 49—86.
- 50*. Статистическая структура аномалий аэрологических полей над североамериканским континентом. — Труды ГГО, 1968, вып. 228, с. 98—120.
- 51*. Zur Frage der rationalen Planung der Beobachtungsnetzes der Lufttemperatur. — *Idöjárás*, 1968, t. 72, N 5, с. 269—278.
52. К вопросу о согласовании полей геопотенциала и ветра. — Метеорология и гидрология, 1969, № 1, с. 42—48.
- 53*. О точности расчета производных по данным в регулярной сетке точек. — В кн.: Объективный анализ метеорологических полей. Ташкент, 1969, с. 66—72.
- 54*. Некоторые численные эксперименты по формированию термического режима атмосферы. — Труды ГГО, 1970, вып. 256, с. 130—146.
- 55*. О влиянии сжимаемости атмосферы на эволюцию планетарных волн. — Труды ГГО, 1970, вып. 256, с. 121—129.
- 56*. О задании притоков тепла при численных экспериментах с помощью термотропной модели. — Труды ГГО, 1970, вып. 256, с. 98—107.
- 57*. О применении теории выбросов к исследованию температурных рядов. — Труды ГГО, 1970, вып. 267, с. 86—99.
- 58*. О расчете статистических характеристик выбросов случайной функции. — Труды ГГО, 1970, вып. 268, с. 146—172.
- 59*. Упрощенная модель для численных экспериментов по термическому режиму атмосферы. — Труды ГГО, 1970, вып. 256, с. 72—97.
- 60*. Вычисление автокорреляционных и взаимных корреляционных функций по нескольким реализациям случайного процесса. — Труды ГГО, 1971, вып. 289, с. 20—28.
- 61*. К вопросу об автоматической обработке данных наблюдений над осадками. — В кн.: Автоматическая проверка и исправление данных метеорологических наблюдений. Обнинск, 1971, вып. 6(II), с. 79—86.
- 62*. Объективный анализ поля осадков. — В кн.: Применение статистических методов метеорологии. Л.: Гидрометеониздат, 1971, с. 334—335.
- 63*. Методика анализа полей метеорологических элементов по данным сети станций. — Труды V всесоюзного метеорологического съезда. — Л.: Гидрометеониздат, 1972, т. III, с. 220—233.
- 64*. Об оценке информативности систем метеорологических измерений. — Труды ГГО, 1972, вып. 286, с. 120—140.
- 65*. О расчете характеристик временных рядов методом статистического моделирования. — Труды ГГО, 1972, вып. 286, с. 71—82.
- 66*. О расчете характеристик пространственной корреляции метеорологических полей. — Труды ГГО, 1972, вып. 286, с. 18—25.
- 67*. О точности определения средней величины по дискретным данным. — Труды ГГО, 1972, вып. 286, с. 95—119.
68. Planning the spatial distribution of hydrometeorological stations to meet an error criterion. — In: Casebook on hydrological network design. — WMO, N 324, Geneva, IV.1, 1972, p. 1—8.

69. Precipitation-Statistical principles. — In: Casebook on hydrological network design. — WMO, N 324, Geneva, 1972, vol. 1, p. 1—11.

70. Snow cover-Statistical principles. — In: Casebook on hydrological network design. — WMO, N 324, Geneva, 1972, vol. 13, p. 1—9.

71*. К вопросу о статистическом моделировании двумерных метеорологических полей. — Труды ГГО, 1973, вып. 308, с. 20—26.

72. О точности расчета пространственных корреляционных функций. — Труды ГГО, 1973, вып. 308, с. 3—19.

73. К вопросу об осреднении методом полигонов. — Труды ГГО, 1974, вып. 336, с. 120—138.

74*. К вопросу о пространственно-временном анализе барического поля. — Труды ГГО, 1974, вып. 336, с. 75—94.

75*. О построении системы объективного анализа разнородных данных на основе метода оптимальной интерполяции и согласования. — Метеорология и гидрология, 1974, № 5, с. 3—10.

76. О точности расчета пространственных корреляционных функций. — Труды ГГО, 1974, вып. 336, с. 3—19.

77*. Симпозиум по проблемам изучения статистической структуры, объективного анализа и автоматического контроля метеорологических полей. — Метеорология и гидрология, 1973, № 5, с. 120—124.

78*. Die räumlich-zeitliche Struktur des bodennahen Luftdrucks. — In: „Untersuchungen der statistischen Struktur meteorologischer Felder. Band I“. Obninsk, 1974, p. 212—226.

79*. Einige Angaben über die Zeitliche Structur und Auswertung von Charakteristiken der Auswürfe meteorologischer Elemente. — In: Untersuchungen der statistischen Struktur meteorologischer Felder. Band II. Obninsk, 1974, p. 5—19.

80*. Über die Algorithmen der Berechnung von Korrelationsfunktionen. — In: Untersuchungen der statistischen Struktur meteorologischer Felder, Band I. Obninsk, 1974, p. 65—92.

81*. Über die Arbeiten zur Untersuchung der statistischen Struktur und Ausnutzung deren Angaben, die in der UdSSR durchgeführt werden. — In: Untersuchungen der statistischen Struktur meteorologischer Felder, Band I. Obninsk, 1974, p. 4—28.

82*. Über die Auswertung des Informationsgehaltes der Systeme meteorologischer Beobachtungen. — In: Untersuchungen der statistischen Struktur meteorologischer Felder, Band I. Obninsk, 1974, p. 29—64.

83. Über die Genauigkeit der Berechnung von Charakteristiken der räumlichen Korrelation. — In: Untersuchungen der statistischen Struktur meteorologischer Felder. Band I. Obninsk, 1974, p. 92—116.

84*. Über die Auswertung von Genauigkeit der räumlichen Mittelung meteorologischer Felder. — In: Untersuchungen der statistischen Struktur meteorologischer Felder. Band II. Obninsk, 1974, p. 47—69.

85. Über die Realisierung der Methode der räumlichen Mittelung mit EDVM. — In: Untersuchungen der statistischen Struktur meteorologischer Felder, Band II. Obninsk, 1974, p. 70—93.

86*. К методике расчета потоков коротковолновой радиации на отдельно стоящее здание. — Труды ГГО, 1975, вып. 338, с. 104—118.

87*. К расчету вероятности выброса нормальной последовательности. — Труды ГГО, 1975, вып. 348, с. 69—77.

88*. Некоторые данные по временной структуре и оценки характеристик выбросов метеорологических элементов. — В кн.: Исследования статистической структуры метеорологических полей. М.: Гидрометеиздат, 1975, т. 2, с. 3—12.

89*. Об алгоритмах расчета корреляционных функций. — В кн.: Исследования статистической структуры метеорологических полей. М.: Гидрометеиздат, 1975, т. 1, с. 18—33.

90*. Об оценке точности пространственного осреднения метеорологических полей. — В кн.: Исследования статистической структуры метеорологических полей. М.: Гидрометеиздат, 1975, т. 2, с. 28—42.

91*. О влиянии дискретности измерений на точность определения числа выбросов случайного процесса. — Труды ГГО, 1975, вып. 348, с. 78—98.

92*. О восстановлении годового хода моментов метеорологических рядов. — Труды ГГО, 1975, вып. 348, с. 99—111.

93*. О выполненных в СССР работах по исследованию статистической структуры. — В кн.: Исследования статистической структуры метеорологических полей. М.: Гидрометеоздат, 1975, т. 1, с. 5—17.

94. О реализации методов пространственного осреднения на ЭВМ. — В кн.: Исследования статистической структуры метеорологических полей. М.: Гидрометеоздат, 1975, т. 2, с. 43—57.

95. О точности выборочных структурных функций. — Труды ГГО, 1975, вып. 348, с. 58—68.

96. О точности расчета характеристик пространственной корреляции. — В кн.: Исследования статистической структуры метеорологических полей. М.: Гидрометеоздат, 1975, т. 1, с. 34—48.

97*. Пространственно-временная структура приземного давления. — В кн.: Исследования статистической структуры метеорологических полей. М.: Гидрометеоздат, 1975, т. 1, с. 64—72.

98*. Расчет характеристик выбросов и временной структуры случайных последовательностей. — Труды ГГО, 1975, вып. 364, с. 66—76.

99*. Упрощенный способ вычисления потоков суммарной радиации на отдельно стоящее здание. — Труды ГГО, 1975, вып. 338, с. 119—132.

100*. К вопросу о пространственной корреляции аномалий температуры воздуха. — Труды ГГО, вып. 374, 1976, с. 186—209.

101. Некоторые возможности использования данных о пространственной структуре барического поля. — Труды ГГО, 1976, вып. 374, с. 3—21.

102*. О возможности приближенного расчета потоков суммарной радиации на элементы городской застройки. — Труды ГГО, 1976, вып. 365, с. 76—88.

103. О выборе оптимальной дискретности при обработке данных наблюдений. — Труды ГГО, 1976, вып. 374, с. 121—131.

104. О расчете вероятности выброса релейской последовательности. — Труды ГГО, 1976, вып. 374, с. 136—153.

105*. О расчете потоков коротковолновой радиации в условиях городской застройки. — Труды ГГО, 1976, вып. 365, с. 61—75.

106*. Статистические методы интерпретации метеорологических данных. — Л.: Гидрометеоздат, 1976, 357 с.

107. Характеристики статистической структуры метеорологических полей и методы их вычисления. — В кн.: Статистическая структура метеорологических полей. — Будапешт, 1976, с. 31—75.

108*. Временная структура скорости ветра по данным наблюдений над океаном. — Труды ГГО, 1977, вып. 397, с. 39—50.

109*. К вопросу об использовании наземных данных при анализе барического поля на высотах. — Метеорология и гидрология, 1977, № 2, с. 32—38.

110*. К вопросу об оценке точности осредненных по площади величин. — Труды ГГИ, 1977, вып. 247, с. 88—98.

111*. Некоторые вопросы обработки рядов наблюдений за ветром. — Труды ГГО, 1977, вып. 397, с. 51—59.

112*. Об использовании прогностической информации при объективном анализе барического поля. — Труды ГГО, 1977, вып. 397, с. 107—115.

113*. Об оценке статистической значимости некоторых климатических параметров. — В кн.: Труды Второго всесоюзного симпозиума по применению статистических методов метеорологии. Л.: Гидрометеоздат, 1977, с. 157—160.

114. Об учете пространственной связности мелкомасштабных флюктуаций при оценке точности осреднения. — Труды ГГИ, 1977, вып. 247, с. 80—87.

115*. Об учете взаимной связи статистических характеристик при расчете числа выбросов временных рядов. — Труды ГГО, 1977, вып. 397, с. 13—20.

116*. О применении статистического моделирования при решении метеорологических задач. — В кн.: Труды Второго всесоюзного симпозиума по применению статистических методов в метеорологии. Л.: Гидрометеоздат, 1977, с. 36—44.

117. О точности расчета характеристик выбросов средней суточной температуры. — Труды ГГО, 1977, вып. 397, с. 3—12.

118*. Статистическая структура метеорологических полей и ее применение. — В кн.: Современные фундаментальные и прикладные исследования Главной

геофизической обсерватории им. А. И. Воейкова. Л.: Гидрометеиздат, 1977, с. 142—150.

119*. Некоторые вопросы статистической обработки временных рядов. — В кн.: Труды третьей всесоюзной конференции по применению статистических методов в метеорологии. — Обнинск, 1978, с. 237—245.

120*. О точности зонального осреднения метеорологических полей. — В кн.: Труды третьей всесоюзной конференции по применению статистических методов в метеорологии. — Обнинск, 1978, с. 206—211.

121. Об учете особенностей пространственно-временной структуры метеорологических полей при их статистическом анализе. — В кн.: Труды третьей всесоюзной конференции по применению статистических методов в метеорологии. — Обнинск, 1978, с. 148—158.

122. Осреднение метеорологических полей. — Л.: Гидрометеиздат, 1979, с. 212.

123*. О точности пространственного осреднения поля температуры воздуха. — Труды ГГИ, 1979, вып. 257, с. 50—67.

124*. К вопросу об анизотропии пространственно-временной структуры барического поля. — Труды ГГО, 1980, вып. 412, с. 20—29.

125*. К вопросу о выборе интервала осреднения периодически коррелированных величин. — Труды ГГО, 1980, вып. 431, с. 74—90.

120*. К вопросу о сходимости распределения к нормальному для сумм связанных случайных величин. — Труды ГГО, 1980, вып. 431, с. 106—117.

127*. О влиянии ошибок наблюдений на характеристики экстремумов метеорологических процессов. — Труды ГГО, 1980, вып. 412, с. 59—69.

128. О пространственно-временной структуре поля ветра. — Труды ГГО, 1980, вып. 412, с. 29—41.

129*. О точности осреднения периодически коррелированных величин. — Труды ГГО, 1980, вып. 431, с. 90—106.

130*. Некоторые актуальные вопросы методики климатической обработки данных. — В кн.: Исследование взаимодействия мезо- и макропроцессов в атмосфере и применение статистических методов в метеорологии. Алма-Ата, 1981, (Тезисы докладов Всесоюзной конференции).

131*. О влиянии густоты сети станций на характеристики изменчивости интерполированных значений. — Метеорология и гидрология, 1981, № 5, с. 39—47.

132*. О выборе сроков наблюдений при обработке метеорологических данных. — В кн.: Исследование взаимодействия мезо- и макропроцессов в атмосфере и применение статистических методов в метеорологии. Алма-Ата, 1981 (Тезисы докладов Всесоюзной конференции).

133*. О зональном осреднении метеорологических полей. — Труды ГГИ, 1981, вып. 271, с. 55—61.

134*. О моделировании временных рядов некоторых метеорологических элементов. — В кн.: Исследование взаимодействия мезо- и макропроцессов в атмосфере и применение статистических методов в метеорологии. Алма-Ата, 1981 (Тезисы докладов Всесоюзной конференции).

135*. О моделировании связанных случайных последовательностей с равномерным распределением. — Труды ГГО, 1981, вып. 437, с. 95—104.

136*. О статистических характеристиках некоторых моделей связанных вейбулловских последовательностей. — Труды ГГО, 1981, вып. 437, с. 104—119.

137*. О точности климатического мониторинга общего содержания озона в атмосфере. — В кн.: Исследование взаимодействия мезо- и макропроцессов в атмосфере и применение статистических методов в метеорологии. Алма-Ата, 1981 (Тезисы докладов Всесоюзной конференции).

138. О точности экстраполяции статистических характеристик вейбулловских последовательностей по выборочным данным. — Труды ГГО, 1981, вып. 437, с. 119—134.

139*. К вопросу о стационарности средних зональных значений метеорологических полей. — Труды ГГО, вып. 480, (см. наст. сб.).

140. Об аппроксимации характеристик пространственной структуры осредненных величин. — Труды ГГО, вып. 480 (см. наст. сб.).

141*. О корректной оценке средних суточных значений метеорологических параметров. — Труды ГГО, вып. 480 (см. наст. сб.).

НЕКОТОРЫЕ АКТУАЛЬНЫЕ ВОПРОСЫ МЕТОДИКИ КЛИМАТОЛОГИЧЕСКОЙ ОБРАБОТКИ ДАННЫХ

В связи с увеличением масштабов хозяйственной деятельности человека и влияния этой деятельности на окружающую среду значение климатологической информации постоянно возрастает. Качество этой информации, которая используется в различных целях — как для изучения структуры и изменчивости наблюдаемого климата, так и для решения прикладных задач, — зависит от положенных в основу данных наблюдений и от применяемых методов климатологической обработки. Высокая стоимость метеорологических наблюдений и потребность получения возможно более надежной информации о климате делают настоятельно необходимым совершенствование методики климатологической обработки данных.

Традиционно сложившиеся методы не были рассчитаны на использование ЭВМ, и применение этих методов при автоматизированной обработке зачастую либо невозможно в силу трудностей формализации либо нецелесообразно (из-за простоты и малой точности вычислительных процедур). Массовая обработка данных с широким использованием ЭВМ ставит перед климатологами новые задачи. Некоторые вопросы, связанные с их решением, и обсуждаются в настоящей статье.

Ряд трудностей, с которыми приходится сталкиваться при анализе и интерпретации метеорологических данных, вызван постоянным совершенствованием системы метеорологических наблюдений (изменением числа сроков наблюдений, методики измерений, количества станции и т. д.), что неизбежно приводит к нарушению однородности рядов.

В связи с этим возникает задача разработки таких методов климатологической обработки, которые давали бы наибольшую точность рассчитываемых характеристик и вместе с тем обеспечивали сравнимость параметров, относящихся к разным периодам времени. Рассмотрим простейшие аспекты этой проблемы.

При смене сроков наблюдений и переходе от трех- к четырех- и к восьмисрочным наблюдениям довольно детально оценивались возможные систематические погрешности средних суточных значений, вызванные этими изменениями, и практически не учитывалось различие случайных ошибок.

Между тем, как показано в [3—5], внутрисуточная нестационарность вторых моментов значений метеорологических элементов приводит к различию средних квадратических ошибок осреднения, которое существенно зависит не только от числа сроков наблюдений, но и от расположения сроков наблюдений по отношению к началу периода осреднения и способа получения средних суточных величин. О величине этих различий можно судить по значе-

ниями средних квадратических погрешностей осреднения температуры воздуха за сутки, полученным по данным наблюдений на ГМС Ленинград, ГМО, которые представлены в табл. 1. В этой таблице $\epsilon_0(4)$ — ошибка арифметического осреднения по четырем срокам при оптимальном расположении их на периоде осреднения, $\epsilon_1(4)$ и $\epsilon_1(8)$ — ошибки арифметического осреднения соответственно по четырем и восьми неблагоприятно расположенным срокам наблюдений.

Таблица 1

Средняя квадратическая погрешность осреднения температуры воздуха за сутки

Характеристика	Месяц			
	I	IV	VII	X
$\epsilon_1(4)$	0,63	0,46	0,39	0,40
$\epsilon_1(8)$	0,31	0,21	0,19	0,20
$\epsilon_0(4)$	0,16	0,22	0,24	0,15

Из табл. 1 хорошо видно, что простое увеличение числа сроков наблюдений необязательно приведет к повышению точности климатических параметров. Например, для января при благоприятном расположении четырех сроков на периоде можно добиться в два раза большей точности осреднения, чем при неудачном расположении восьми сроков. Заметим, что действующими сейчас наставлениями предусмотрено арифметическое осреднение именно по неблагоприятно расположенным срокам наблюдений. Следовательно, рекомендуемая в настоящее время методика обработки не позволяет в полной мере использовать информацию, содержащуюся в современных детальных измерениях.

При заданном числе и расположении сроков наблюдений (именно такова существующая ситуация) уменьшения погрешности осреднения можно добиться за счет применения более корректных способов осреднения. Переход к таким способам получения средних представляется весьма актуальным, особенно в связи с подготовкой Официального свода климатических параметров (Кадастра по климату СССР).

Представление в Официальном своде информации о параметрах должно, по возможности, сопровождаться оценкой их точности. Наличие сведений о погрешностях позволяет не только выбрать оптимальный способ расчета изучаемых характеристик (как в рассмотренном выше примере), но и является неременным условием правильной интерпретации результатов, полученных с использованием климатических параметров. Очевидность этого факта говорит о необходимости уделять возможно большее внимание вопросам оценки точности характеристик, получаемых в результате климатологической обработки. Необходимо при этом

подчеркнуть важность учета специфики метеорологических данных при получении соответствующих оценок. Это легко проиллюстрировать на примере учета связности при опеределении точности основных статистических параметров метеорологических элементов [3, 8]. В табл. 2 приведены некоторые оценки погрешностей средних \bar{f} , стандартов σ , коэффициентов асимметрии A и эксцесса E , а также межсуточного коэффициента корреляции $r(24)$ для наблюдений за температурой воздуха в январе на ГМС Ленинград, ГМО. Эти

Таблица 2

Отношение оценок ошибок определения статистических параметров с учетом связности ряда к ошибке, оцененной без учета связности (σ_1/σ_0)

Число сроков	Статистические параметры				
	\bar{f}	σ	A	E	$r(24)$
24	12,4	9,6	7,0	5,8	7,5
8	7,1	5,5	4,1	3,4	4,3
4	5,0	3,9	2,9	2,4	3,1
1	2,5	2,0	1,4	1,4	1,6

Примечание. σ_0 — средняя квадратическая погрешность определения статистической характеристики по заданному объему данных без учета связности исходных рядов, σ_1 — с учетом реальной связности.

оценки были получены путем многократного моделирования месячных совокупностей при задании фактической корреляции.

Представленные результаты показывают, что неучет связности завышает кажущуюся точность выводок статистических параметров и может привести к ложным выводам при анализе метеорологических данных.

Естественной тенденцией развития климатологии является переход от эмпирических способов получения различных величин, интересующих потребителей, к построению на основе анализа фактических данных вероятностных моделей метеорологических рядов и последующему расчету требуемых характеристик. Вопрос о том, насколько удачно принята вероятностная модель описывает поведение того или иного метеорологического элемента, является принципиальным и его решение требует тщательного изучения реальных данных наблюдений. В связи с этим возникают, однако, и важные вопросы методического характера: каким способом целесообразно определять параметры выбранной модели, какова выборочная погрешность определения этих параметров и т. д.

В [1, 2] этот круг вопросов рассматривается на примере задачи, традиционно решаемой с помощью использования вероятностных моделей, а именно, задачи определения расчетных скоростей ветра редкой повторяемости. Оценка экстремальных скоростей ветра по данным ограниченных выборок производится обычно либо на ос-

новации распределения скорости ветра во всем диапазоне ее изменения с последующей оценкой требуемых квантилей распределения, либо по данным о максимальных годовых значениях скорости ветра, описания этих данных с помощью известных распределений крайних членов выборки и определения вероятностей интересующих экстремальных значений. И в том и в другом случае для объективизации процедуры аппроксимации данных наблюдений и получения параметров распределения применяется, как правило, метод наименьших квадратов. Использование этого метода требует задания эмпирических вероятностей в ранжированной выборке, для определения которых предлагался ряд формул. В [1, 2] выполнена детальная оценка точности определения квантилей распределения при использовании различных формул, а также реализован метод максимального правдоподобия, при применении которого вообще отпадает необходимость задания эмпирических вероятностей.

Полученные результаты позволили выявить формулу эмпирических вероятностей, обеспечивающую минимальную систематическую и случайную погрешности квантилей, а также убедительно показали, что с помощью метода максимального правдоподобия можно добиться еще большей точности статистической экстраполяции.

На возможность использования метода максимального правдоподобия в задаче определения расчетных скоростей ветра указывалось еще в 1954 г. [9]. Однако этот метод не получил распространения, вероятно, в силу больших (по сравнению с методом наименьших квадратов) вычислительных трудностей. В настоящее время наличие мощной базы ЭВМ, а также полученные количественные результаты, подтверждающие предпочтительность применения метода максимального правдоподобия при статистической экстраполяции распределений Вейбулла и Фреше, позволяют рекомендовать использование этого метода в задаче определения расчетных скоростей ветра.

Заметим, что выполненные в [1, 2] оценки основаны на применении метода статистического моделирования. Следует обратить внимание на целесообразность использования этого метода при решении различных методических вопросов, отработки алгоритмов, получения оценок точности и т. д.

Ряд вопросов методики климатологической обработки данных приобретает в наши дни особую актуальность в связи с необходимостью осуществления мониторинга климата. Трудности решения этой задачи в значительной мере обусловлены характером метеорологических данных: нестабильностью сети станций, наличием пропусков в рядах наблюдений и др. Естественна поэтому наметившаяся в последнее время тенденция к глобальной регуляризации метеорологической информации. Одним из проявлений этой тенденции является создание банков данных, полученных путем пространственной интерполяции значений метеорологических элементов на станциях в узлы координатной сетки. При анализе

такой информации, однако, необходимо иметь в виду, что интерполяция приводит к искажению статистической структуры исходного поля, причем это искажение зависит от густоты сети станций, меняющейся как во времени, так и в пространстве. В [7] представлены некоторые оценки завышения корреляции, занижения дисперсии исходного поля и дисперсии градиентов, а также искажения распределения негауссовских величин при интерполяции. Показано, что в количественном отношении влияние этих эффектов может быть весьма существенным и приводит к значительным ложным пространственным неоднородностям, фиктивным ходам характеристик изменчивости, затрудняющим выявление реальных климатических закономерностей и осуществление мониторинга климата.

Затронутый вопрос является частью более крупной задачи по разработке методов корректного определения различных параметров, характеризующих глобальный климат, и оценки точности этих методов. Круг вопросов должен быть расширен и касаться не только интерполированных величин, но и характеристик, полученных путем осреднения, дифференцирования, а также сочетания различных изменяющих структуру исходного поля действий.

Первоочередной интерес для целей мониторинга климата представляет рассмотрение осредненных по заданному региону значений интерполированных величин и изучение характеристик их изменчивости. Некоторые оценки изменчивости средних широтных величин, полученных путем использования предварительно проинтерполированных данных, содержатся в [6].

В дальнейшем, однако, необходимо систематическое исследование изменения (под влиянием интерполяции, осреднения) не только вторых моментов какого-либо одного метеорологического элемента и его производных, но и рассмотрение смешанных моментов величин, относящихся к разным метеорологическим элементам.

В настоящее время широко обсуждается вопрос о подготовке Официального свода климатических параметров (Кадастра по климату СССР) и об организации Службы климата (подобно Службе погоды) и намечается разработка специального проекта. На Службу климата будут возложены функции, имеющие большое социальное и народнохозяйственное значение:

- сбор, обработка и хранение данных;
- информационное обслуживание народнохозяйственных и научных организаций данными о климате;
- оценка текущего состояния климатической системы и вероятностная оценка степени аномальности этого состояния;
- прогноз масштабов ожидаемых изменений климата.

Из примеров, приведенных в настоящей статье, следует, что для успешного функционирования Службы климата необходимы наряду с организационными мероприятиями дальнейшие серьезные научные исследования, направленные на разработку ее методических основ.

СПИСОК ЛИТЕРАТУРЫ

1. Каган Р. Л. О точности статистической экстраполяции величин, распределенных в соответствии с законом Фреше.—Труды ГГО, 1982, вып. 447, с. 3—16.
2. Каган Р. Л. О точности экстраполяции статистических характеристик вейбулловских последовательностей по выборочным данным.—Труды ГГО, 1981, вып. 437, с. 119—134.
3. Каган Р. Л., Сибир Е. Е. О выборе сроков наблюдений при обработке метеорологических данных.—Труды Всесоюзной конференции «Исследование взаимодействия мезо- и макропроцессов в атмосфере и применение статистических методов в метеорологии», 1981, Алма-Ата (тезисы) — 68 с.
4. Каган Р. Л., Сибир Е. Е. О корректной оценке средних суточных значений метеорологических элементов (см. настоящий сборник).
5. Каган Р. Л., Сибир Е. Е. О точности осреднения периодически коррелированных величин.—Труды ГГО, 1980, вып. 431, с. 90—105.
6. Каган Р. Л., Хлебникова Е. И. К вопросу о стационарности средних зональных значений метеорологических полей (см. настоящий сборник).
7. Каган Р. Л., Хлебникова Е. И. О влиянии густоты сети станций на характеристики изменчивости интерполированных значений.—Метеорология и гидрология, 1981, № 5, с. 39—47.
8. Федорченко Е. И. О влиянии связности метеорологических рядов на точность выборочных моментов.—Труды ГГО, 1974, вып. 336, с. 25—47.
9. Thom H. C. S. Frequency of maximum wind speeds.—Proc. Am. Soc. Civ. Eng., 1954, vol. 80, N 539. 11 p.

Р. Л. Каган

ОБ АППРОКСИМАЦИИ ХАРАКТЕРИСТИК ПРОСТРАНСТВЕННОЙ СТРУКТУРЫ ОСРЕДНЕННЫХ ВЕЛИЧИН

1. При оценке точности различных способов пространственного осреднения гидрометеорологических полей и выборе наиболее эффективных вариантов важно иметь представление о статистических характеристиках не только поля осредняемого параметра, но и искомым осредненным величинам. Зная дисперсию осредненных по площади значений поля, корреляцию осредненных величин с исходными неосредненными величинами, можно решать ряд методических задач, связанных с оптимизацией системы наблюдений.

В тех случаях, когда статистическая структура осредняемого поля нам известна, характеристики структуры осредненной по площади величины могут быть рассчитаны теоретически путем повторного интегрирования корреляционных функций точечных величин. Методика такого расчета подробно изложена в [2], где приводятся также примеры использования этих характеристик для решения конкретных задач.

Численное интегрирование, необходимое для этих расчетов, является довольно трудоемким и отнимает много времени на ЭВМ, особенно при необходимости рассмотрения многих гипотетических вариантов, что часто приходится делать при методических оцен-

ках. В этих условиях желательно использовать подходящие аппроксимационные формулы, которые достаточно точно воспроизводили бы реальное поведение искомых характеристик.

Применительно к однородному и изотропному полю величин с экспоненциально убывающей корреляционной функцией некоторые аппроксимационные формулы были предложены автором в [2]. Однако они относятся к случаю, когда расстояния между пунктами, корреляция между которыми (или между площадями, центрами которых они являются) не превышают масштаба области осреднения. На больших расстояниях корреляция осредненных величин быстро приближается к корреляции неосредненного поля. Разумеется, при определении средней по площади с использованием данных в пунктах, расположенных за пределами области осреднения, необходимо задание единой, притом возможно более гладкой, аппроксимации характеристик структуры.

Ниже приводятся аппроксимационные формулы, которые в какой-то мере решают эту задачу, и обсуждаются возможности их использования.

2. Пусть мы исследуем однородное и изотропное поле величины $f(\rho)$, корреляционная функция которой имеет вид

$$r(\rho) = \exp(-\rho/\rho_0). \quad (1)$$

Здесь ρ — радиус-вектор точки в принятой системе координат, ρ — расстояние между пунктами, ρ_0 — масштаб корреляции.

Очевидно, что при осреднении величины f по площади S дисперсия осредненных величин σ_s^2 связана с дисперсией σ^2 величины f соотношением

$$\mu_s(0) = \frac{\sigma^2 S}{\sigma^2} = \iint_{S(\rho)} \iint_{S(\rho')} r(\rho, \rho') dS dS', \quad (2)$$

а взаимная корреляция между средними по площади значениями f_s и значениями f_i в точке ρ_i описывается формулой

$$\omega_{s_i} = \frac{\overline{f_s f_i} - \bar{f}_s \bar{f}_i}{\sigma^2} = \iint_{S(\rho)} r(\rho, \rho_i) dS, \quad (3)$$

где черта сверху означает осреднение по статистическому ансамблю.

Для случая осреднения на плоскости по площади квадрата со стороной a функции $\mu_s(0)$ и ω_s затабулированы в [2, 4]. Там же приводятся и таблицы корреляционной функции $\mu_s(\rho)$, характеризующие связь между величинами f_s , полученными путем осреднения по квадратам площади S с центрами в различных пунктах. Выполненные оценки ω_s и μ_s показывают, что при небольшой вытянутости области осреднения эти функции сравнительно мало зависят от формы области и поле осредненных величин может с достаточным приближением считаться однородным и изотропным. Поэтому упомянутыми таблицами, полученными для областей

квадратной формы, можно пользоваться, полагая условно в качестве величины a сторону равновеликого квадрата, а в качестве начала координат — центр тяжести области осреднения.

В [2] показано, что в случае площадей квадратной формы для пунктов, находящихся в пределах квадрата или на расстоянии друг от друга, меньшем стороны квадрата, выполняются следующие соотношения

$$\mu_s(k) = \exp[-0,49\beta(1+k^2)] \text{ при } k \leq 1, \quad (4)$$

$$\omega_s(k) = \exp[-0,37\beta(1+2k^2)] \text{ при } k \leq \frac{\sqrt{2}}{2}, \quad (5)$$

где $\beta = a/\rho_0$ — масштаб осреднения (полагается, что $\beta < 2$), $k = \rho/a$ — безразмерное расстояние.

При больших k , как уже упоминалось, имеют место приближенные соотношения

$$\mu_s(k) \approx \omega_s(k) \approx \exp(-k\beta). \quad (6)$$

При оперативных расчетах желательно иметь аппроксимацию функций μ_s и ω_s аналитическими формулами, которые обеспечивали бы не только их непрерывность, но и достаточную гладкость на переходных участках. Для этого естественно потребовать непрерывности их первых производных. Легко проверить, что этим требованиям удовлетворяют аппроксимации

$$\mu_s(k) = \begin{cases} \exp\left(-\frac{\beta}{2}(1+k^2)\right) & 0 \leq k \leq 1, \\ \exp(-\beta k), & k > 1, \end{cases} \quad (7)$$

$$\omega_s(k) = \begin{cases} \exp\left(-\frac{\beta}{\sqrt{8}}(1+2k^2)\right) & 0 \leq k \leq \frac{\sqrt{2}}{2}, \\ \exp(-\beta k) & \frac{\sqrt{2}}{2} < k. \end{cases} \quad (8)$$

Соотношение между аппроксимационными формулами μ_s и ω_s и корреляционной функцией r иллюстрирует рис. 1, где в полулогарифмическом масштабе представлена их зависимость от расстояния.

Коэффициенты, входящие в формулы (7) и (8), несколько отличаются от коэффициентов в формулах (4) и (5). Однако они находятся в пределах точности аппроксимации, которая составляет около 5% аппроксимируемых величин. Такая точность воспроизведения характеристик структуры осредненных величин открывает возможность для статистического моделирования полей с гауссовской корреляционной функцией путем сглаживания предварительно моделированных полей с экспоненциальной корреляцией.

3. Менее очевидна возможность использования формул (7), (8) для оценки более детальных характеристик, при расчете которых приходится иметь дело с комбинациями этих функций.

Рассмотрим это на примере оценки ошибки распространения данных наблюдений в пункте ρ_i на всю площадь S .

Дисперсия ошибки распространения для рассматриваемого нами однородного и изотропного поля (при условии отсутствия

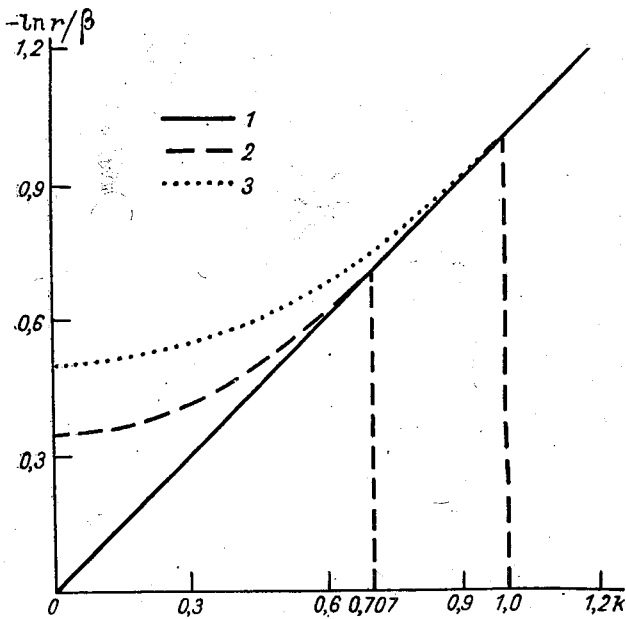


Рис. 1. Аппроксимация корреляционных функций μ_S и ω_S .
 1 — $r(k, \beta)$; 2 — $\omega_S(k)$ (формула (8)); 3 — $\mu_S(k)$ (формула (7)).

случайных ошибок наблюдений и исключения систематических ошибок) составляет

$$\sigma_\sigma^2(\rho_i) = \sigma^2 [1 + \mu_S(0) - 2\omega_S(\rho_i)]. \quad (9)$$

Или, вводя относительную среднюю квадратическую ошибку $\varepsilon = \sigma_\sigma / \sigma$, имеем

$$\varepsilon^2(\rho_i) = 1 + \mu_S(0) - 2\omega_S(\rho_i). \quad (10)$$

В частности, при измерениях в центре площади S

$$\varepsilon^2(0) = 1 + \mu_S(0) - 2\omega_S(0). \quad (11)$$

Подставляя в (11) формулы (7) и (8), получаем

$$\varepsilon^2(0) = 1 + \exp(-0,5\beta) - 2 \exp(-\beta/\sqrt{8}). \quad (12)$$

Разложение (12) в степенной ряд дает выражение

$$\varepsilon^2(0) = 0,207\beta - 0,030\beta^3 + \dots, \quad (13)$$

которое при небольших размерах площади ($\beta \leq 2$) с достаточной точностью может аппроксимироваться формулой

$$\varepsilon^2(0) \approx 0,21\beta. \quad (14)$$

Формула (14) дает заниженные на 10 % значения дисперсии ошибки сопоставления по сравнению с эмпирической формулой, предложенной ранее автором в [1] (в которой коэффициент при β составляет не 0,21, а 0,23). Соответственно, средние квадратические ошибки занижаются на 5 %. Учитывая большую неопределенность в задании исходных характеристик структуры, это занижение представляется не очень существенным.

При наблюдениях на периферии квадрата, а тем более за его пределами, распространение данных на весь квадрат характеризуется ошибками, дисперсия которых описывается более общей формулой (10).

В рамках нашей аппроксимации она имеет вид

$$\varepsilon^2_s(k) = \begin{cases} 1 + \exp\left(-\frac{\beta}{2}\right) - 2 \exp\left(-\frac{\beta}{\sqrt{8}}(1 + 2k^2)\right), & k \leq \frac{\sqrt{2}}{2}, \quad (15a) \\ 1 + \exp\left(-\frac{\beta}{2}\right) - 2 \exp(-kp), & k > \frac{\sqrt{2}}{2}. \quad (15b) \end{cases}$$

Сравнение результатов, полученных по этой формуле, с точными результатами показывает, что за исключением малых значений β (осреднение по площадям малого размера), где значения ε_s^2 при малых k могут оказаться заниженными до 15—20 %, аппроксимация (15) дает удовлетворительные результаты при любых k до значений $\beta \sim 2$.

Более того, в пределах $0 < \beta < 1$, $0 < k < 1$ удовлетворительные результаты дает формула, полученная путем разложения в ряд формулы (15_a):

$$\varepsilon^2_s(k) \approx \beta \frac{\sqrt{2}-1}{2} + \beta k^2 \left[\sqrt{2} - \frac{\beta}{2}(1+k^2) \right]. \quad (16)$$

Как и формула (14), она дает занижение ошибки примерно на 10 %. Однако во многих практических задачах это окупается простотой формулы и наглядностью, с которой она описывает уменьшение репрезентативности наблюдений по мере удаления их от центра области.

Практический интерес может представить также корреляция оценок, полученных для различных квадратов. Пусть мы интересуемся средними для одинаковых по размерам и одинаково ориентированных квадратов S и S_1 , расположенных на расстоянии $\rho = ka$ друг от друга. Среднюю по каждому из квадратов мы оцениваем по наблюдениям в центре. Точность такой оценки характеризуется формулой (10). Данные о корреляции соответствующих

ошибок могут оказаться полезными при решении ряда методических вопросов.

При принятых предположениях корреляция ошибок в направлении средней линии квадрата описывается формулой

$$r_0(\rho) = \frac{\varepsilon^2(k, \beta)}{\varepsilon^2(0, \beta)}, \quad (17)$$

где

$$\varepsilon^2(k, \beta) = r(k\beta) + \mu_s(k) - 2\omega_s(k). \quad (18)$$

Подставив в (18) формулы (7) и (8), получим:

$$\varepsilon^2(\beta, k) = \begin{cases} \exp(-\beta k) + \exp\left(-\frac{\beta}{2}(1+k^2)\right) - \\ - 2 \exp(-\beta/\sqrt{8}(1+2k^2)) & \text{при } 0 < k < \frac{\sqrt{2}}{2}, \\ \exp\left(-\frac{\beta}{2}(1+k^2)\right) - \exp(-\beta k) & \text{при } \frac{\sqrt{2}}{2} < k \leq 1, \\ 0 & \text{при } k > 1. \end{cases} \quad (19)$$

Формулы (19) довольно громоздки. Более простыми, наглядными и обеспечивающими непрерывность величины $r_0(\beta, k)$ вместе с ее первой производной во всем диапазоне изменения параметров β и k являются выражения, полученные путем разложения (18) в ряд:

$$\varepsilon^2(\beta, k) = \begin{cases} \beta \left[\frac{\sqrt{2}-1}{\sqrt{2}} - k + \left(\sqrt{2} - \frac{1}{2}\right) k^2 \right] & \text{при } 0 \leq k < \frac{\sqrt{2}}{2}, \\ \beta \left[k - \frac{1+k^2}{2} \right] & \text{при } \frac{\sqrt{2}}{2} \leq k < 1, \\ 0 & \text{при } k > 1. \end{cases} \quad (20)$$

Представление о том, какую точность оценки корреляции осредненных величин обеспечивает использование формул (19) и (20), дает рис. 2. Из рисунка видно, что эта точность вполне удовлетворительна при значительном перекрытии квадратов ($k < 0,3$ или $\rho < 0,3a$). При меньшем же их перекрытии ($0,3 < k < 0,8$) аппроксимации (20) и особенно (19) дают область довольно заметной отрицательной корреляции ошибок распространения, которая практически отсутствует при расчете по точным формулам (кривая 1).

Следует сразу оговориться, что сравнение с кривой 1 является несколько условным, поскольку она характеризует частный случай — корреляцию ошибок в одном направлении. Элементарные оценки показывают, что при рассмотрении корреляции в направлении их диагонали фактическая область отрицательной корреляции ошибок была бы заметно больше. Тем не менее существенное занижение корреляционной функции r_0 при расчете по аппроксимационным формулам (19) и (20) следует учитывать.

С практической точки зрения наиболее важно фактическое отсутствие корреляции между ошибками распространения для неперекрывающихся площадей, которое существенно облегчает оценки точности по большим площадям, на которых располагается много станций.

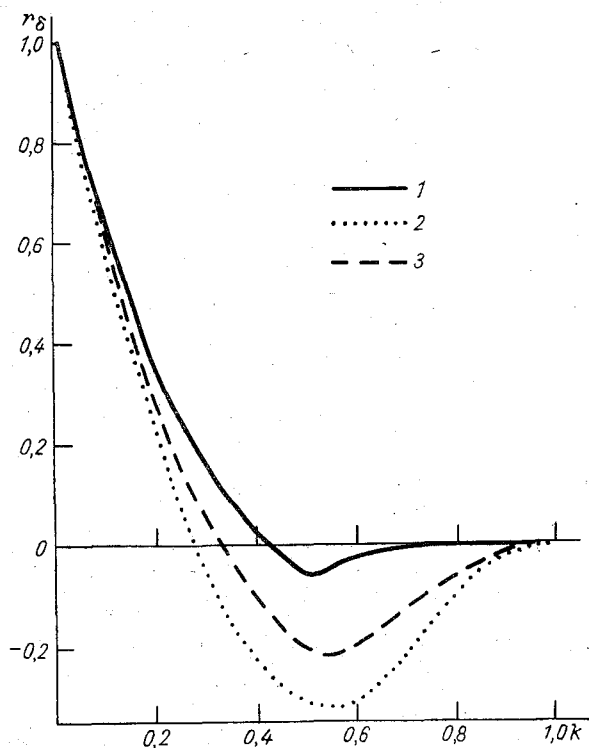


Рис. 2. Пространственная корреляция ошибок представления средних по квадратам данными в их центрах.

1 — фактическая корреляция в направлении средней линии квадрата; 2, 3 — корреляция с аппроксимацией (19) и (20) соответственно.

В заключение укажем, что предложенная аппроксимация отнюдь не исчерпывает всех возможностей приближенного описания статистической структуры осредненных величин. В частности, альтернативный вариант такой аппроксимации применительно к осреднению на поверхности Земли, рассматривается в [3]. Вопрос о предпочтительности того или иного варианта требует специального рассмотрения в каждом конкретном случае.

СПИСОК ЛИТЕРАТУРЫ

1. Каган Р. Л. О точности определения средней по площади по данным точечных измерений. — Труды ГГО, 1965, вып. 175, с. 117—131.
2. Каган Р. Л. Осреднение метеорологических полей. — Л.: Гидрометеоздат, 1979. 213 с.

3. Каган Р. Л., Лугина К. М. О глобальном анализе средних месячных полей общего содержания озона. — См. настоящий сборник.

4. Полищук А. И., Каган Р. Л. О статистической структуре осредненных значений метеорологических элементов. — Труды ГГО, 1966, вып. 191, с. 92—121.

Р. Л. Каган, К. М. Лугина

О ГЛОБАЛЬНОМ АНАЛИЗЕ СРЕДНИХ МЕСЯЧНЫХ ПОЛЕЙ ОБЩЕГО СОДЕРЖАНИЯ ОЗОНА В АТМОСФЕРЕ

1. Проблема глобального мониторинга малых примесей в атмосфере, в частности общего содержания озона (ОСО), является важной составной частью общего климатического мониторинга. Редкая сеть станций, на которых производятся измерения ОСО, и большие погрешности единичных измерений делают необходимым временное и пространственное осреднение этих характеристик, которое позволяет уменьшить выборочные погрешности измерений. Как показано в работе [6], в климатологических исследованиях наибольший интерес представляют средние месячные значения ОСО.

Выбор площадей пространственного осреднения определяется характеристиками естественной изменчивости осредняемого поля, количеством и расположением станций, данные измерений которых могут быть использованы.

В настоящее время разработана методика априорной оценки точности пространственного осреднения с использованием сведений о пространственной статистической структуре осредняемых величин. Она нашла широкое применение в исследованиях по проблеме мониторинга глобального термического режима атмосферы [4, 5]. С помощью этой методики получены некоторые оценки точности различных способов осреднения ОСО для отдельных широтных зон северного полушария [6].

Целью настоящей работы является сравнительная оценка информативности существующей в настоящее время сети наблюдений за ОСО в северном и южном полушариях.

2. При оценке точности осреднения использовалась пространственная корреляционная функция средних месячных значений ОСО $r(\rho)$ из работы [7].

Расчеты $r(\rho)$ средних месячных значений ОСО были выполнены по данным 65 озонометрических станций, расположенных на территории Советского Союза (19), Западной Европы (19), Южной Азии (11) и Северной Америки (16), за период 1961—1977 гг. для января, апреля, июля и октября с последующим осреднением. На 46 станциях измерения ОСО выполнялись с помощью спектрофотометров Добсона, на 19 — с помощью фильтрового озонометра М-83.

Корреляционная функция, рассчитанная в предположении однородности и изотропии, аппроксимирована экспоненциальной зависимостью вида

$$r(\rho) = e^{-\rho/1,6}, \quad (1)$$

где ρ — расстояние (тыс. км).

Мера ошибки наблюдений, представляющая собой отношение дисперсии случайной погрешности измерений к дисперсии средних месячных величин ОСО, принимается равной 0,1.

Разработанная методика осреднения позволяет использовать и более детальные сведения о характеристиках пространственной статистической структуры (неоднородность поля дисперсии осредняемой величины, анизотропию, сезонные различия и т. д.). Однако имеющиеся в настоящее время данные об ОСО не позволяют получить такие детальные характеристики с достаточной надежностью. Кроме того, как показано в выполненных ранее исследованиях, относящихся к другим метеорологическим элементам [2, 3], даже столь схематичные сведения о статистической структуре позволяют сделать вполне определенные выводы об информативности той или иной сети станций и о соотношении между точностью различных методов осреднения. Поэтому на данном этапе полученная в [7] пространственная корреляционная функция средних месячных значений ОСО использована для оценки точности пространственного осреднения как в северном так и в южном полушариях.

Точность осреднения, как и в [6], мы будем характеризовать двумя величинами: ε — отношение средней квадратической ошибки осреднения (E) к среднему квадратическому отклонению точечных значений осредняемой величины (σ) и ε_s — отношение средней квадратической ошибки осреднения (E) к среднему квадратическому отклонению осредненной по площади величины (σ_s).

Рассмотрены следующие методы осреднения:

оптимальное осреднение, при котором весовые множители для данных каждой станции определяются из условия минимума средней квадратической погрешности (E);

оптимальное осреднение с нормированием весов (условие равенства единице суммы весов осреднения);

метод полигонов (веса равны отношению площади, тяготеющей к данной станции, ко всей площади осреднения);

арифметическое осреднение.

3. Сети станций, на которых производится измерение ОСО в северном и южном полушариях существенно различаются. В северном полушарии, по данным [14], насчитывается 173 станции. Однако распределены эти станции по территории очень неравномерно. Большинство озонметрических станций расположено в Северной Америке, в Западной Европе, в Советском Союзе и в некоторых странах юго-восточной Азии (Япония, Индия). На акваториях океанов в северном полушарии озонметрические станции практически отсутствуют. Следует иметь в виду, что измерение

общего содержания озона на различных станциях производится различными приборами, что часто затрудняет сравнение результатов. Наиболее распространенным прибором для измерения ОСО на зарубежных станциях является спектрофотометр Добсона. В Советском Союзе используются озонметры М-83, данные наблюдений по которым в 60-х годах, насколько можно судить по литературе [1, 12], плохо согласовывались с добсоновскими. В последние годы в результате использования в М-83 новых фильтров, качество наблюдений существенно повысилось и полученные данные стали сопоставимы с данными наблюдений по спектрофотометру Добсона [10].

В южном полушарии регистрация ОСО производится на 29 станциях, 10 из которых расположены южнее 60° ю. ш. (рис. 1 б). Распределение остальных 19 станций по территории более равномерное, чем в северном полушарии. Расстояния между станциями в низких и умеренных широтах южного полушария, как правило, превышают 1000 км, а на акватории Тихого океана между 80 и 170° з. д. нет ни одной станции. На большинстве этих станций ОСО измеряется спектрофотометром Добсона, на советских антарктических станциях Мирный и Восток установлены озонметры М-83.

При оценке возможной точности мониторинга общего содержания озона в атмосфере важно учитывать также то обстоятельство, что измерения выполнялись не на всех существующих озонметрических станциях. Так, например, в 1975 г. работало 13 (из 29) станций в южном полушарии и 65 (из 173) — в северном. Поэтому априорные оценки точности осреднения вычислялись как с учетом всех возможных станций, так и по более редкой сети. Для южного полушария значения ϵ рассчитывались по данным 34 станций (5 из которых расположены в экваториальной зоне северного полушария и 29 в зоне южного) и для 15 станций (2 из них в северном полушарии), которые характеризуют сеть наблюдений за ОСО в 1975 г. Эти результаты представлены в табл. 1. Там же для сравнения приведены значения ϵ для различного количества станций в северном полушарии [6].

Наряду с оценками средних квадратических погрешностей для полной озонметрической сети, состоящей из 173 станций, получены значения ϵ для 68 станций, данные которых регулярно публиковались в [14] в течение 1961—1977 гг., и для 32 станций, использованных в работе Энджела и Коршовера [11] при исследовании трендов общего содержания озона в атмосфере. Неравномерность сети станций в табл. 1 характеризуется параметром α , представляющим собой коэффициент вариации площадей, тяготеющих к каждой станции, или, что то же самое, коэффициент вариации весов осреднения методом полигонов. Из табл. 1 видно, что как в северном, так и в южном полушариях при существующей в настоящее время сети станций погрешности осреднения средних месячных значений ОСО для всех рассмотренных методов осреднения, кроме оптимального, сравнимы, а в ряде случаев даже превышают

среднее квадратическое отклонение осредненной величины (σ_s). При оптимальном осреднении погрешности минимальны, однако и в этом случае они составляют 0,6—0,8.

Из анализа табл. 1 следует также, что для повышения точности осреднения необходимо в южном полушарии увеличить коли-

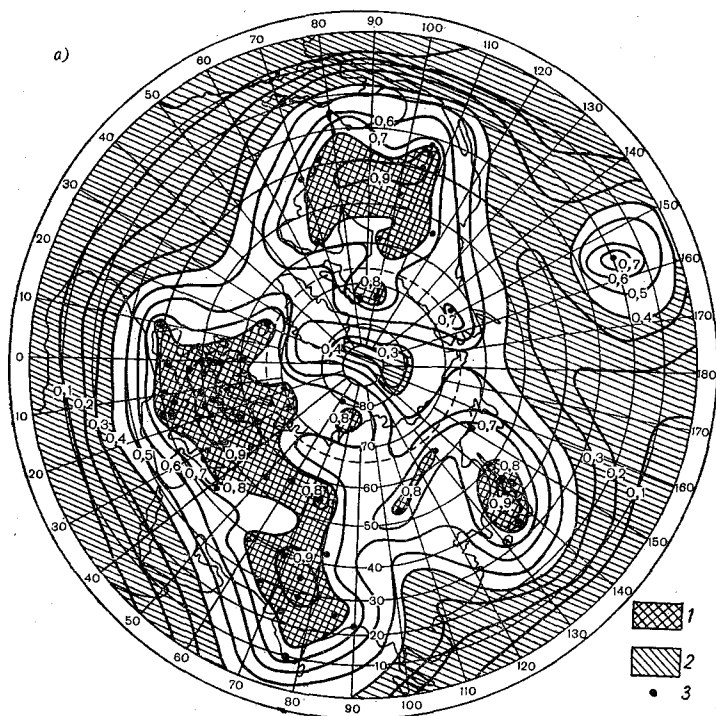


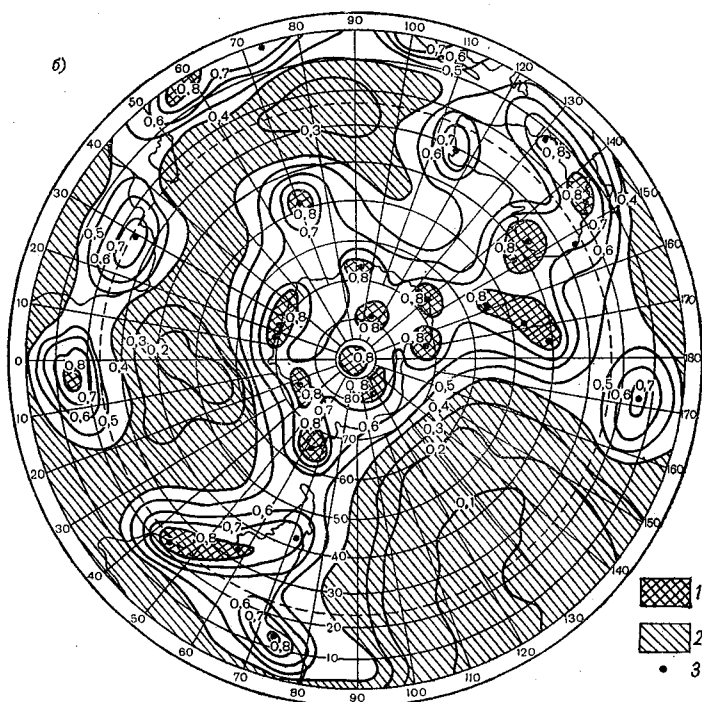
Рис. 1. Мера искажения стандарта σ для интерполированных ре северного полушария (а) и южного полушария (б).

1) $\sigma > 0,8$; 2) $\sigma < 0,4$; 3) озонметрические станции,

чество станций, а в северном полушарии более равномерно распределить имеющиеся станции. Значительный эффект может быть достигнут за счет некоторого сокращения сети на хорошо обжитых территориях и организации наблюдений в плохо освещенных данными озонметрических наблюдениях районах, в первую очередь на акваториях океанов.

В табл. 1 также представлены оценки погрешностей осреднения по отдельным широтным зонам для южного полушария. Сравнение их с аналогичными результатами для северного полушария из работы [6] показывает, что в тропических и умеренных широтах обоих полушарий средние значения ОСО могут быть получены по существующей сети станций с недостаточной для целей мониторинга точностью.

Иначе обстоит дело в полярных районах. Наличие пункта наблюдений на Южном полюсе (станция Амундсен-Скотт) и достаточно равномерное распределение остальных 9 станций на Антарктическом материке позволяет получить средние значения ОСО при использовании статистически оптимальных методов осредне-



средних месячных значений общего содержания озона в атмосфере-
данные которых учитывались при расчетах.

ния с погрешностью, не превышающей 50 % естественной изменчивости осредненной величины σ_S в зонах 60—80, 70—90 и 80—90° ю. ш. (см. табл. 1). В северном полушарии средние квадратические погрешности осреднения $\bar{\epsilon} \leq 0,5 \sigma_S$ могут быть получены лишь для зоны 70—90° с. ш. Напомним, что при применении метода оптимального осреднения необходимо иметь достаточно надежные статистические характеристики общего содержания озона в атмосфере, и в первую очередь сведения о нормах, дисперсиях и пространственных корреляционных функциях.

4. В некоторых работах (см., например, [13]) данные об ОСО представлены в виде интерполированных в узлы регулярной сетки значений, которые затем используются для расчета характеристик изменчивости поля [9]. При этом следует иметь в виду, что сгла-

Таблица 1

Средние квадратические относительные погрешности осреднения (ε и $\bar{\varepsilon}$) средних месячных значений ОСО для отдельных широтных зон и полушарий в целом при различном количестве озонметрических станций

Широтная зона	$\frac{\sigma_S}{\sigma}$	Число станций	α	Характеристика	Осреднение			
					оптимальное	оптимальное с нормированием весов	по полигонам	арифметическое

Южное полушарие

По данным станций в зоне

20° с. ш.— 20° ю. ш.	0,23	12	0,7	ε $\bar{\varepsilon}$	0,18 0,79	0,29 1,24	0,32 1,40	0,30 1,30
20—60	0,27	12	0,7	ε $\bar{\varepsilon}$	0,20 0,74	0,30 1,13	0,32 1,22	0,33 1,26
60—80	0,45	8	0,4	ε $\bar{\varepsilon}$	0,23 0,50	0,28 0,62	0,29 0,64	0,29 0,65
0—90	0,22	34	0,8	ε $\bar{\varepsilon}$	0,14 0,64	0,18 0,82	0,19 0,86	0,20 0,91
		15	0,7	ε $\bar{\varepsilon}$	0,18 0,82	0,28 1,27	0,30 1,36	0,30 1,36

С привлечением станций из других зон

20—60	0,27	34	1,1	ε $\bar{\varepsilon}$	0,18 0,69	0,21 0,80	0,24 0,89	0,25 0,92
60—80	0,45	34	1,7	ε $\bar{\varepsilon}$	0,20 0,45	0,21 0,46	0,25 0,56	0,35 0,77
70—90	0,57	10	0,6	ε $\bar{\varepsilon}$	0,22 0,38	0,22 0,39	0,24 0,43	0,25 0,44
80—90	0,71	10	1,7	ε $\bar{\varepsilon}$	0,30 0,43	0,30 0,43	0,41 0,58	0,61 0,86

Северное полушарие

0—90	0,22	173	1,7	ε $\bar{\varepsilon}$	— —	— —	0,14 0,64	0,27 1,23
0—90	0,22	68	1,2	ε $\bar{\varepsilon}$	0,13 0,59	0,18 0,82	0,20 0,91	0,26 1,18
0—90	0,22	32	0,8	ε $\bar{\varepsilon}$	0,15 0,68	0,22 1,00	0,24 1,09	0,26 1,18
70—90	0,57	14	1,0	ε $\bar{\varepsilon}$	0,25 0,44	0,26 0,45	0,27 0,47	0,33 0,58
80—90	0,71	14	—	ε $\bar{\varepsilon}$	0,43 0,61	0,44 0,62	0,52 0,74	0,50 0,71

живание поля при применении процедуры интерполяции приводит к искажению статистических характеристик исследуемого поля и, в частности, к занижению средних квадратических отклонений. Мера занижения (κ), представляющая собой отношение стандарта интерполированных ($\tilde{\sigma}$) значений метеорологических элементов к стандарту значений, измеренных на станциях (σ), зависит от используемого метода интерполяции и от расположения станций, на которых производились измерения. Для метода оптимальной интерполяции, как показано в [8], значение κ связано с относительной средней квадратической погрешностью интерполяции ($\epsilon_{\text{опт}}$) простым соотношением

$$\kappa = \frac{\tilde{\sigma}}{\sigma} = \sqrt{1 - \epsilon_{\text{опт}}^2} \quad (2)$$

где $\epsilon_{\text{опт}}$ — мера ошибки интерполяции.

Располагая данными о точности оптимальной интерполяции при конкретном расположении сети станций, можно описать географическое распределение меры занижения стандарта интерполированных величин κ .

На рис. 1 приведены карты распределения значений κ для среднего месячного содержания озона в атмосфере, построенные на основе оценок точности оптимальной интерполяции по данным озонметрической сети станций в северном и южном полушариях. В северном полушарии (рис. 1 а) использовались данные 68 станций. В южном полушарии (рис. 1 б) расчеты выполнены для 34 станций, из которых 5 расположены в тропической зоне северного полушария.

Из рис. 1 а видно, что наряду с большими значениями κ (малым занижением стандарта) в Европе, Японии, Северной Америке имеются обширные зоны в полярных и экваториальных районах, на акваториях океанов, которые характеризуются значениями κ составляющими 50 % и менее.

В южном полушарии область относительно небольшого занижения ($\kappa \geq 0,7$) охватывает большую часть антарктического материка и прилегающих к нему океанических районов, что связано с относительно густой сетью озонметрических станций в южной полярной области. На остальной части южного полушария сеть станций настолько редкая, что области больших значений κ концентрируются вокруг отдельных станций.

Таким образом, из приведенных выше оценок точности осреднения и значений κ следует, что существующая в настоящее время сеть станций, измеряющих ОСО, позволяет получить достаточно надежную информацию о средних месячных значениях ОСО лишь в полярных районах южного полушария. В северном полушарии средние по площади значения ОСО со средней квадратической погрешностью, не превышающей 50 % естественной изменчивости, могут быть получены в широтной зоне 70—90° с. ш. только с помощью статистически-оптимальных методов осреднения. Заметим также, что эти оценки справедливы лишь в том случае, когда на

всех станциях, использованных в расчетах, выполняются регулярные измерения ОСО, позволяющие достаточно корректно определить средние месячные значения.

СПИСОК ЛИТЕРАТУРЫ

1. Гушин Г. П. Озонометрическая сеть СССР. — Метеорология и гидрология, 1979, № 3, с. 111—116.
2. Каган Р. Л. Осреднение метеорологических полей. — Л.: Гидрометеоздат, 1979. — 213 с.
3. Каган Р. Л., Лугина К. М. О точности зонального осреднения метеорологических полей. — Труды III Всесоюзного симпозиума по применению статистических методов в метеорологии. — М.: Гидрометеоздат, 1978, с. 206—211.
4. Каган Р. Л., Лугина К. М. О точности пространственного осреднения поля температуры воздуха. — Труды ГГИ, 1979, вып. 257, с. 50—66.
5. Каган Р. Л., Лугина К. М. О зональном осреднении метеорологических полей. — Труды ГГИ, 1981, вып. 271, с. 51—61.
6. Каган Р. Л., Лугина К. М. Некоторые оценки возможностей мониторинга общего содержания озона в атмосфере. — Труды ГГО, 1983, вып. 466, с. 81—92.
7. Каган Р. Л., Салль И. А., Щербакова Н. Н. О пространственной корреляции средних месячных значений общего содержания озона в атмосфере. — Труды ГГО, 1983, вып. 466, с. 18—26.
8. Каган Р. Л., Хлебникова Е. И. О влиянии густоты сети станций на характеристики изменчивости интерполированных значений. — Метеорология и гидрология, 1981, № 5, с. 39—47.
9. Кайгородцев А. Е., Покровский О. М. О естественных составляющих общего содержания озона в северном полушарии. — Метеорол. и гидрол., 1980, № 8, с. 52—57.
10. Общее содержание атмосферного озона и спектральная прозрачность атмосферы/Под ред. Г. П. Гушина. — Л.: Гидрометеоздат, 1978. — 157 с.
11. Angell J. K., Korshover J. Quasi-bienial variations in temperature, total ozone and tropopause height. — Journ. Atm. Sci., 1964, vol. 21, N 5, p. 479—492.
12. Bojkov M. D. Differences in Dobson spectrophotometer and filter ozonometer measurements of total ozone. — Journ. Appl. Meteorol., 1969, vol. 8, p. 362—368.
13. London J., Bojkov R. D., Oltmans S. Atlas of the global distribution of total ozone July 1957—June 1967. — NCAR Techn. Note, Boulder, Colorado, 1976, 276 p.
14. Ozone data for the world. — WMO, Toronto, Ontario, Canada.

Р. Л. Каган, Е. И. Хлебникова

К ВОПРОСУ О СТАЦИОНАРНОСТИ СРЕДНИХ ЗОНАЛЬНЫХ ЗНАЧЕНИЙ МЕТЕОРОЛОГИЧЕСКИХ ПОЛЕЙ

В силу большой пространственно-временной изменчивости метеорологических элементов для осуществления мониторинга климата приходится прибегать к осреднению метеорологических параметров как по времени, так и по пространству. В частности, предметом рассмотрения многих исследований, посвященных изучению изменений климата, являются средние зональные значения

метеорологических полей (см., например, [1, 3]). Для корректной оценки векового хода этих величин необходимо иметь представление об их изменчивости, которая определяется как естественной изменчивостью метеорологических элементов, так и способом получения средних зональных значений. Один из распространенных способов определения средних зональных значений состоит в использовании предварительно проинтерполированных в узлы регулярной сетки значений метеорологических элементов. Как показано в работе [2], различие густоты сети станций в разные периоды времени приводит к значительной неоднородности таких рядов. В настоящей статье обращается особое внимание на стационарность средних зональных значений, получаемых путем использования проинтерполированных данных, и приводятся количественные оценки, относящиеся к средней годовой температуре воздуха.

Аналогично [2] при оценках предполагается, что рассматриваемое метеорологическое поле f является однородным и изотропным с известной корреляционной функцией r , дисперсией σ^2 и мерой ошибок данных наблюдений η^2 .

При интерполяции вида

$$\tilde{f}_i = \sum_{k=1}^n p_k^i f_k, \quad (1)$$

где n — число станций, привлекаемых к интерполяции в i -й узел сетки, f_k — данные наблюдений на станциях, p_k^i — веса интерполяции, \tilde{f}_i — интерполированные значения, ковариация интерполированных значений \tilde{R} , их дисперсия $\tilde{\sigma}^2$, корреляция \tilde{r} , а также дисперсия средних широтных $\tilde{\sigma}^2(\bar{T}_\varphi)$, полученных путем использования проинтерполированных данных, описываются формулами

$$\tilde{R}_{ij} = \sigma^2 \left(\sum_{k=1}^n \sum_{l=1}^n p_k^i p_l^j r_{kl} + \eta^2 \sum_{k=1}^n p_k^i p_k^j \right), \quad (2)$$

$$\tilde{\sigma}^2 = x_i^2 \sigma^2 = \sigma^2 \left(\sum_{k=1}^n \sum_{l=1}^n p_k^i p_l^i r_{kl} + \eta^2 \sum_{k=1}^n (p_k^i)^2 \right), \quad (3)$$

$$\tilde{r}_{ij} = \frac{1}{x_i x_j \sigma^2} \tilde{R}_{ij}, \quad (4)$$

$$\tilde{\sigma}^2(\bar{T}_\varphi) = \iint_{(x,y) \in \Omega_\varphi} \tilde{R}(x,y) dx dy, \quad (5)$$

где $r_{kl} = r(\rho_{kl})$, ρ_{kl} — расстояние между станциями k и l , Ω_φ — круг широты φ .

Для выявления нестационарности рядов средних широтных значений средней годовой температуры воздуха, полученных путем использования банков данных в узлах регулярной сетки точек, нами были выполнены расчеты по формулам (2)–(5) при задании фактического расположения станций в разные периоды времени.

Приведем некоторые результаты, относящиеся к характеристикам изменчивости термического режима в полярных зонах северного и южного полушария, мониторинг которого в силу его особой чувствительности к возможным изменениям климата представляет собой важную практическую задачу.

Таблица 1

Статистическая структура интерполированных и истинных значений средней годовой температуры воздуха

Характеристика	Широта (разность широт)	Арктика				Антарктика			
		1900—1909	1936—1945	1966—1975	Истинные	1945	1955	1965	Истинные
$\frac{\sigma}{\sigma}$	75°	0,67	0,96	0,98	1,00	0,27	0,61	0,89	1,00
	80	0,64	0,92	0,97	1,00	0,19	0,61	0,90	1,00
	85	0,61	0,87	0,94	1,00	0,14	0,58	0,92	1,00
	90	0,60	0,85	0,92	1,00	0,12	0,56	0,98	1,00
$r_{\lambda}(\Delta\varphi)$	5	0,98	0,94	0,94	0,91	0,93	0,94	0,91	0,87
	10	0,94	0,84	0,82	0,76	0,78	0,81	0,71	0,63
	15	0,87	0,68	0,67	0,59	0,65	0,65	0,48	0,40
$\sigma_{\Gamma\varphi}$	80	0,32	0,58	0,62	0,69	0,16	0,38	0,66	0,86
$\sigma_{\Gamma\lambda}$	80	0,27	0,49	0,56	0,69	0,18	0,35	0,68	0,86
σ_{Γ}	80	0,42	0,76	0,84	0,98	0,24	0,51	0,95	1,22

При задании статистической структуры средней годовой температуры воздуха в полярных зонах мы воспользовались сведениями о корреляционных функциях, полученных в ГГИ К. М. Лугойной и Н. А. Сперанской. При расчетах для арктического бассейна задавалась корреляционная функция вида

$$r(\rho) = \exp(-0,34\rho^{1,617}). \quad (6)$$

Мера ошибок наблюдений η^2 полагалась при этом равной 0,021. Оценки для антарктической зоны были получены с использованием корреляционной функции

$$r(\rho) = \exp(-0,723\rho^{1,68}) \quad (7)$$

при $\eta^2=0,042$.

Расчеты производились для узлов широтно-долготной сетки в зонах севернее 70° с. ш. и южнее 70° ю. ш. с шагом по широте $\Delta\varphi=5^\circ$ и по долготе $\Delta\lambda=20^\circ$. В табл. 1 представлены некоторые данные, осредненные по всем 18 узлам каждой широтной зоны, которые получены при использовании способа оптимальной интерполяции. Заметим, что, как показано в [2], применение других способов интерполяции приводит к незначительным изменениям оценок характеристик изменчивости интерполированных величин.

Для арктического бассейна нами были выполнены оценки, относящиеся к трем периодам: 1900—1909, 1936—1945 и 1966—1975 гг. В первый период в районе севернее 70° с. ш. наблюдения за тем-

пературой воздуха вели лишь 3 станции, во второй период число их составляло уже более 30, в третий период их число еще увеличилось более чем в два раза.

Представленные в табл. 1 оценки для антарктической зоны относятся также к трем периодам: 1945, 1955, 1965 гг. В 1945 г. в районе южнее 70° ю. ш. действовало 7 станций, в 1955 г. — 23 станции, в 1965 г. — 33 станции.

В табл. 1 $\tilde{r}_\lambda(\Delta\varphi)$ представляет собой корреляционную функцию интерполированных значений в меридиональном направлении для расстояний, соответствующих $5, 10$ и 15° меридиана; $\tilde{\sigma}_{T_\varphi}$ и $\tilde{\sigma}_{T_\lambda}$ — средние квадратические значения градиента аномалий температуры в зональном и меридиональном направлении при расстоянии, соответствующем 10° меридиана; $\tilde{\sigma}_r$ — среднее квадратическое значение модуля этого градиента ($\Gamma = \sqrt{\Gamma_\varphi^2 + \Gamma_\lambda^2}$).

Из табл. 1 видно, что в антарктической зоне еще более резко, чем в Арктике, проявляется ложный временной тренд и стандартов, и градиентов интерполированных значений, что обусловлено как расположением станций, так и характером затухания корреляционной функции.

Обращает на себя внимание изменение дисперсии с широтой: если в Арктике для всех трех периодов занижение дисперсии увеличивается с ростом широты, то в Антарктике в последний из рассматриваемых периодов наблюдается обратная закономерность. Этот факт имеет простое объяснение: он связан с вводом в действие в этот период ст. Амундсен-Скотт, расположенной вблизи полюса.

Более детальное представление об искажении характеристик изменчивости средней годовой температуры воздуха в антарктической зоне за счет влияния густоты сети станций дает рис. 1, на котором для двух из рассматриваемых периодов даны карты-схемы меры занижения стандарта κ .

В табл. 2 представлены стандарты средних широтных значений средней годовой температуры воздуха, относящиеся к разным периодам наблюдений. Из этой таблицы хорошо видно, что как в Арктике, так и в Антарктике ряды средних широтных значений, полученные путем использования проинтерполированных данных, отличаются большой нестационарностью относительно среднего квадратического отклонения. Например, в антарктической зоне в начальный период наблюдений стандарт средних широтных величин составлял лишь 15—30 % от истинного его значения. Хотелось бы также отметить то обстоятельство, что после 1965 г. число станций, ведущих в Антарктике наблюдения за температурой, начало уменьшаться; это должно повлечь за собой усложнение формы тренда стандарта средних широтных значений, вызванного изменением числа станций.

Приведенные выше оценки получены с учетом упрощающих предположений об однородности и изотропности изучаемых полей, но при задании фактического расположения станций и сведений

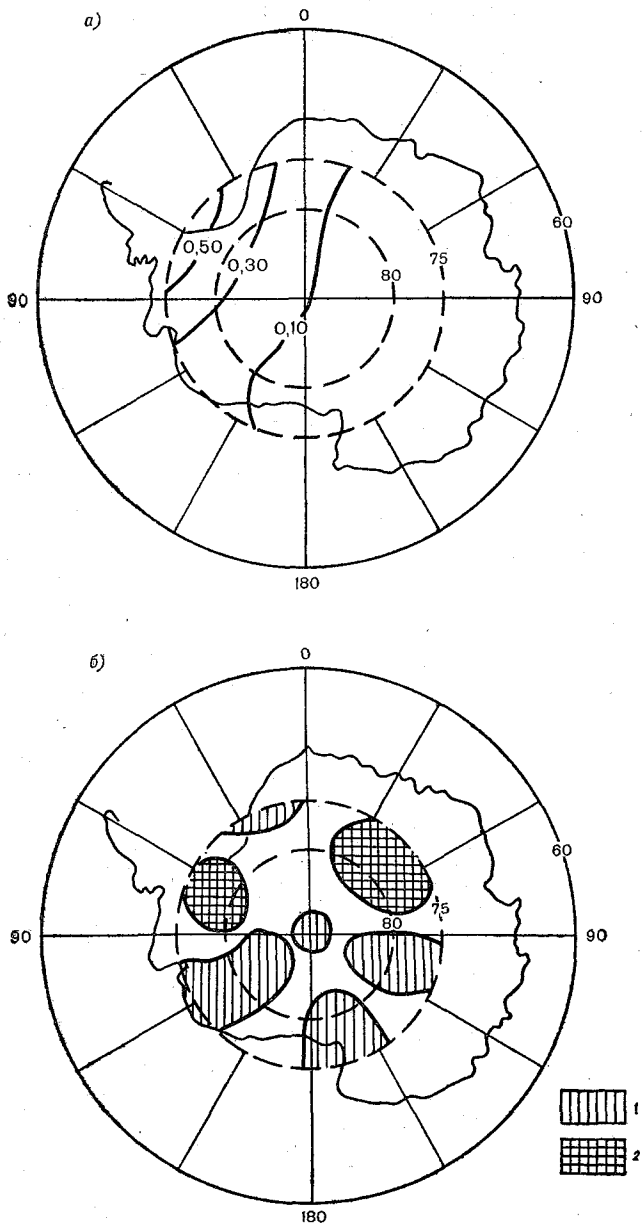


Рис. 1. Мера занижения стандарта интерполированных значений средней годовой температуры воздуха в антарктической зоне для 1945 г. (а) и 1965 г. (б).

1) $\kappa > 0,9$, 2) $\kappa < 0,7$.

о статистической структуре изучаемого поля в изучаемых районах. Эти оценки дают количественное представление о возможном тренде дисперсии средних широтных значений, вызванном методическими причинами. Они свидетельствуют о том, что степень нестационарности рядов средних зональных значений метеорологических полей достаточно велика и ее следует учитывать при их анализе.

Таблица 2

Нормированные стандарты ($\sigma(\bar{T}_\varphi)/\sigma$) средних зональных значений средней годовой температуры воздуха, рассчитанных по интерполированным данным

Период	φ°		
	75	80	85
Арктика			
1900—1909	0,53	0,57	0,59
1936—1945	0,69	0,76	0,80
1966—1975	0,70	0,80	0,89
Истинные значения	0,70	0,81	0,93
Антарктика			
1945	0,17	0,14	0,12
1955	0,43	0,51	0,55
1965	0,58	0,72	0,87
Истинные значения	0,58	0,72	0,89

СПИСОК ЛИТЕРАТУРЫ

1. Бudyко М. И. Климат в прошлом и будущем. — Л.: Гидрометеоздат, 1980. — 350 с.
2. Каган Р. Л., Хлебникова Е. И. О влиянии густоты сети станций на характеристики изменчивости интерполированных значений. — Метеорология и гидрология, 1981, № 5, с. 39—47.
3. Barnett T. P. Estimating variability of surface air temperature in the Northern Hemisphere. — Mon. Wea. Rev., 1978, vol. 106, N 9, p. 1353—1367.

Р. Л. Каган, Е. Е. Сибир

О КОРРЕКТНОЙ ОЦЕНКЕ СРЕДНИХ СУТОЧНЫХ ЗНАЧЕНИЙ МЕТЕОРОЛОГИЧЕСКИХ ПАРАМЕТРОВ

При решении многих задач климатологии, в том числе при исследовании изменений климата, следует иметь четкое представление о точности используемых статистических параметров метео-

рологических элементов. Возможности надежного получения этих характеристик ограничены сравнительно небольшим объемом исходной информации. Нередко делают попытки компенсировать небольшую длину ряда за счет использования более детальных данных путем перехода от наблюдений, производимых один раз в сутки (аэрологические наблюдения), к двухразовым и даже четырехразовым наблюдениям или от четырех сроков в сутки (наземные метеорологические наблюдения) к восьмисрочным и даже ежечасным наблюдениям. При этом возникает целый ряд аспектов которые следует учитывать, в частности наличие суточного и годового хода статистических параметров метеорологических элементов, неучет которого при обработке может свести на нет преимущества более детальных данных.

В первую очередь это относится к расчету средних суточных значений метеорологических параметров, которые очень широко используются на практике. Действующими в настоящее время инструкциями был предусмотрен переход от четырехсрочных наблюдений в климатологические сроки к восьмисрочным наблюдениям в синоптические сроки. При этом рекомендуется определение средних суточных значений путем арифметического осреднения данных наблюдений. Изменение сроков при наличии суточного хода приводит как к систематическим, так и к случайным ошибкам. Если систематические ошибки довольно детально оценивались при изменении системы наблюдений, то случайным ошибкам уделялось гораздо меньше внимания.

Нами были выполнены оценки таких ошибок на примере наземной температуры воздуха — элемента наиболее широко используемого в работах по мониторингу климата. Некоторые оценки возможных ошибок были получены авторами ранее путем аналитических расчетов на основании использования несколько идеализированной модели временной корреляции температуры воздуха [1, 2]. В настоящей работе оценки выполнены путем обработки фактических данных наблюдений на ст. Ленинград, ГМО, без использования каких-либо предположений о статистической структуре этих данных.

В расчетах использовались данные обработки самописцев за 22 года (1944—1965 гг.) для четырех центральных месяцев различных сезонов (января, апреля, июля и октября). В качестве «истинной» средней суточной температуры мы принимали среднюю за сутки температуру, рассчитанную методом трапеций по данным ежечасных наблюдений:

$$\bar{\theta}_{24} = \frac{1}{24} \left(\frac{\theta_0 + \theta_{24}}{2} + \theta_1 + \dots + \theta_{23} \right). \quad (1)$$

Значения средней суточной температуры вычислялись также по восьми-, четырех- и трехсрочным наблюдениям методом средней арифметической, причем время первого срока варьировалось, но таким образом, чтобы время наблюдений не выходило за пределы периода осреднения. Рассчитывались средние квадратические

ошибки суточного осреднения при использовании различного числа измерений и различном размещении их на периоде осреднения. В табл. 1 приведены средние за месяц значения средних квадратических ошибок суточного осреднения. Здесь t_1 — время первого измерения; ϵ_0 — средняя квадратическая ошибка осреднения при

Таблица 1

Средние за месяц значения средней квадратической ошибки (°C) суточного осреднения температуры воздуха

t_1	Число сроков n					
	3	4	8	3	4	8
	Январь			Апрель		
0	0,84	0,63	0,31	0,59	0,46	0,21
1	0,65	0,43	0,12	0,48	0,34	0,11
2	0,48	0,27	0,12	0,42	0,26	0,11
3	0,32	0,16		0,36	0,22	
4	0,22	0,27		0,30	0,26	
5	0,31	0,44		0,36	0,33	
6	0,46			0,41		
ϵ_0	0,22	0,16	0,12	0,30	0,22	0,11
ϵ_1	0,84	0,63	0,31	0,59	0,46	0,21
	июль			октябрь		
0	0,50	0,39	0,19	0,55	0,40	0,20
1	0,44	0,32	0,12	0,45	0,30	0,09
2	0,38	0,26	0,12	0,34	0,20	0,09
3	0,36	0,24		0,26	0,15	
4	0,35	0,27		0,22	0,20	
5	0,36	0,32		0,26	0,30	
6	0,39			0,34		
ϵ_0	0,35	0,24	0,12	0,22	0,15	0,09
ϵ_1	0,50	0,39	0,19	0,55	0,40	0,20

наиболее благоприятном размещении сроков на периоде осреднения, когда начало периода осреднения соответствует середине интервала между измерениями и ϵ_1 — средняя квадратическая ошибка осреднения при наиболее неблагоприятном размещении сроков, когда начало (либо конец) периода осреднения совпадает со сроками измерений. Из табл. 1 видно, что ошибки осреднения для всех месяцев при благоприятном размещении сроков значительно меньше, чем при неблагоприятном их размещении. Эти различия больше для января и октября из-за меньшего влияния нестационарности, что является следствием малого суточного хода дисперсии в эти месяцы (см. рис. 1).

При сравнении погрешностей осреднения для различных месяцев (при одинаковом выборе сроков наблюдений) следует учитывать, что температура воздуха в различные месяцы характеризуется разной временной связностью [3]. Это обеспечивает меньшую

относительно среднего значения дисперсии погрешность осреднения в январе и октябре, для которых характерна большая временная связность рядов температуры воздуха. Кроме того, дисперсия температуры воздуха в различные месяцы также очень различна (см. рис. 1); так, зимой она в несколько раз больше, чем летом,

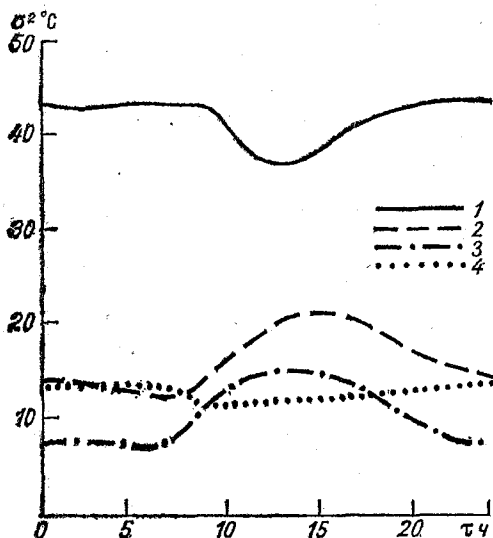


Рис. 1. Суточный ход средних месячных значений дисперсии температуры воздуха на ст. Ленинград, ГМО.
1 — январь, 2 — апрель, 3 — июль, 4 — октябрь.

и существенно больше, чем в переходные сезоны. В результате ошибка осреднения для января может по абсолютному значению превышать ошибку осреднения для других месяцев, что и имеет место при неблагоприятном размещении сроков наблюдений, когда средняя квадратическая погрешность осреднения для трехсрочных, четырехсрочных и даже восьмисрочных наблюдений для января примерно в 1,5 раза больше, чем для июля и других месяцев. При оптимальном же выборе сроков погрешность осреднения в январе существенно меньше, чем в июле.

Следует также учитывать суточный ход дисперсии. Как показано в [2], наибольшие различия в точности осреднения по оптимально и неблагоприятно расположенным срокам измерений наблюдаются в случае, когда начало периода осреднения совпадает с максимумом дисперсии исходной величины; в нашем случае это примерно соответствует январю и октябрю. Эти результаты хорошо подтверждают выводы, полученные ранее в работе [2] с помощью довольно грубой модели.

В качественном отношении при использовании применяемых в настоящее время восьмисрочных наблюдений различия в точности осреднения при использовании оптимально расположенных измерений и рекомендуемых в настоящее время неблагоприятно расположенных сроков (00, 03, 06, 09, 12, 15, 18, 21 ч) могут быть не столь велики — от 0,1 до 0,2 °C — и для многих задач ими, веро-

ятно, можно пренебречь. Но при меньшем числе измерений на периоде осреднения, например при трехсрочных измерениях, эти различия очень существенны (для января ошибка возрастает от 0,2 °С при благоприятном размещении сроков до 0,8 °С при неблагоприятном) и, вероятно, требуют учета.

СПИСОК ЛИТЕРАТУРЫ

1. Каган Р. Л., Сибир Е. Е. К вопросу о выборе интервала осреднения периодически коррелированных величин. — Труды ГГО, 1980, вып. 431, с. 74—89.

2. Каган Р. Л., Сибир Е. Е. О точности осреднения периодически коррелированных величин. — Труды ГГО, 1980, вып. 431, с. 90—105.

3. Федорченко Е. И. О временной структуре ежечасных наблюдений за температурой воздуха в Ленинграде. — Труды ГГО, 1975, вып. 348, с. 112—122.

М. И. Юдин

К ОПРЕДЕЛЕНИЮ КАЧЕСТВА МЕТЕОРОЛОГИЧЕСКОЙ ИНФОРМАЦИИ И МАКСИМАЛЬНО ВОЗМОЖНОГО ЭФФЕКТА ОТ ЕЕ ИСПОЛЬЗОВАНИЯ

Среди множества подходов к проблеме оценки качества метеорологической информации наиболее привлекательным представляется введение показателя качества, который был бы тесно связан с максимально возможным эффектом от использования информации в народном хозяйстве. Однако на этом пути возникают серьезные трудности, связанные прежде всего с тем, что прогнозы погоды и климатические справочники не рассчитаны на какого-либо определенного потребителя, а являются многоцелевыми. Требования к качеству информации разных групп потребителей различаются между собой. Действительно, каждая группа потребителей характеризуется некоторым диапазоном метеоусловий (одним или несколькими), в котором они должны выбирать между существенно различными решениями. Экономический или социальный эффект метеорологической информации для каждой группы в основном определяется ошибками в диапазоне максимальной зависимости этих потребителей от качества информации. Отсюда следует, что невозможно определить показатель качества, одинаково приемлемый для всех потребителей.

С другой стороны, если не установить приоритет определенных целевых функций, оценка качества прогнозов погоды, а также описаний текущих характеристик климата и тенденций его изменения становится крайне неопределенной задачей. Для иллюстрации степени неопределенности можно упомянуть работу [1], где были сопоставлены результаты оценок прогноза барического поля по 10 разным показателям качества. Было показано, что многие из этих оценок оказываются практически некоррелированными

между собой, так что прогноз, хороший по нескольким показателям, будет оценен как неудовлетворительный по другим показателям.

В значительной мере эти трудности могут быть преодолены, если ориентировать прогноз не на потребителя, управляющего отдельным хозяйством, а на потребителя высокого уровня. Здесь под потребителем высокого уровня понимается организация, управляющая деятельностью множества предприятий или хозяйств, различающихся матрицами потерь при осуществлении различных фактических метеорологических ситуаций при разных формулах прогноза. Следует подчеркнуть, что тесная связь между показателем качества информации и эффективностью ее использования может быть получена лишь при условии, что потребитель имеет достаточные представления о матрицах потерь и осуществляет оптимальную стратегию управления хозяйством с учетом прогнозов (или справочных данных). Разумеется, ошибочная стратегия потребителя может свести на нет пользу даже от высокоточных прогнозов или справочных данных. Поэтому в дальнейшем, говоря об эффективности прогнозов, будем рассматривать максимально возможный эффект, соответствующий оптимальной стратегии потребителя.

Начнем с рассмотрения случая альтернативного прогноза. Будем в основном опираться на построения и выводы известной работы А. М. Обухова [2]. Предполагается, что прогноз может быть дан в форме Π_1 (наличие некоторого явления) или Π_0 (отсутствие явления). Потребитель принимает решение S_1 в случае прогноза Π_1 и S_0 в случае прогноза Π_0 . Полагается, что эффекты истинного или ошибочного прогноза могут быть выражены в количественной форме (не обязательно денежной). Запишем величины эффекта в форме матрицы $\{u_{ji}\}$:

$$\begin{array}{cc} & \Phi_0 & \Phi_1 \\ \Pi_0 & u_{00} & u_{01} \\ \Pi_1 & u_{10} & u_{11} \end{array}$$

Здесь Φ_0 означает, что фактически явление отсутствовало, а Φ_1 — что оно имело место. Значения u_{ji} могут быть как положительными (доход, уменьшение заболеваемости и т. п.), так и отрицательными (убыток, рост заболеваемости и т. п.). Замечая, что $u_{10} < u_{00}$, $u_{01} < u_{11}$, введем обозначения

$$u_{00} = u_{00} - v_0 \quad u_{01} = u_{11} - v_1, \quad (1)$$

где v_0, v_1 — потери потребителя за счет ошибочного прогноза. Успешность прогноза характеризуется матрицей $\{n_{ij}\}$ числа случаев n_{ij} , когда наблюдалось Φ_i при прогнозе Π_j :

$$\left. \begin{array}{ccc} & \Pi_0 & \Pi_1 & n_{00} + n_{01} = n_0 \\ \Phi_0 & n_{00} & n_{01} & n_{10} + n_{11} = n_1 \\ \Phi_1 & n_{10} & n_{11} & n_0 + n_1 = n \end{array} \right\} \quad (2)$$

Тогда общий экономический эффект, который получает потребитель в среднем от одного прогноза, равен

$$U = \frac{1}{n} \sum_{i,j=0}^1 n_{ij} u_{ji}. \quad (3)$$

Заметим, что этот эффект зависит как от матрицы $\{n_{ij}\}$, так и от матрицы $\{u_{ji}\}$. Поэтому он весьма сложным образом зависит от качества прогнозов. Обухов [2] ввел понятие об идеальном потребителе, что дает возможность однозначно определить качество альтернативных прогнозов.

Пусть выборка из n прогнозов настолько велика, что отношение n_1/n можно приравнять к вероятности осуществления явления p

$$\frac{n_1}{n} = p, \quad \frac{n_0}{n} = 1 - p. \quad (4)$$

Обозначим далее долю ошибочных прогнозов при осуществившемся явлении α_1 и при его отсутствии α_0 . Тогда имеем

$$\begin{aligned} \frac{n_{00}}{n} &= (1 - \alpha_0)(1 - p); & \frac{n_{01}}{n} &= \alpha_0(1 - p); \\ \frac{n_{10}}{n} &= \alpha_1 p; & \frac{n_{11}}{n} &= (1 - \alpha_1)p. \end{aligned} \quad (5)$$

С учетом введенных обозначений перепишем (3) в виде

$$U = (1 - p) u_{00} + p u_{11} - \alpha_0 (1 - p) v_0 - \alpha_1 p v_1. \quad (6)$$

Как уже говорилось выше, количественный эффект не обязательно рассматривать в экономических терминах, но в дальнейшем это обстоятельство не будет оговариваться.

Что может делать потребитель, если он не желает пользоваться прогнозом? У него есть на выбор две «чистых» стратегии — принимать всегда решение S_1 или всегда S_0 . Кроме того, он может применять «смешанную» стратегию, чередуя решения в каком-то порядке независимо от прогноза. Определим значения эффекта $U^{(1)}$ и $U^{(0)}$ для «чистых» стратегий:

$$\begin{aligned} U^{(1)} &= (1 - p) u_{00} + p u_{11} - (1 - p) v_0, \\ U^{(0)} &= (1 - p) u_{00} + p u_{11} - p v_1. \end{aligned} \quad (7)$$

Сравнивая $U^{(1)}$ и $U^{(0)}$, видим, что при $(1 - p) v_0 < p v_1$ предпочтительна стратегия S_1 , а при $(1 - p) v_0 > p v_1$ — стратегия S_0 . Определим преимущество стратегии, основанной на использовании прогноза, над лучшей из «чистых» стратегий:

$$\begin{aligned} \delta U &= U - U^{(1)} = (1 - \alpha_0 - \alpha_1)(1 - p) v_0 - \\ &\quad - \alpha_1 (U^{(1)} - U^{(0)}) \text{ при } U^{(1)} > U^{(0)}, \\ \delta U &= U - U^{(0)} = (1 - \alpha_0 - \alpha_1) p v_1 - \\ &\quad - \alpha_0 (U^{(0)} - U^{(1)}) \text{ при } U^{(0)} > U^{(1)}. \end{aligned} \quad (8)$$

Будем рассматривать взвешенное среднее величин v_0 и v_1

$$\bar{v} = (1 - p)v_0 + pv_1$$

как масштаб экономической зависимости потребителя от погоды. При постоянстве этого масштаба наибольший экономический эффект от прогноза достигается при условии

$$(1 - p)v_0 = pv_1 = \bar{v}/2.$$

Максимальный эффект равен $(1 - \alpha_0 - \alpha_1)(v/2)$, причем первый множитель по определению зависит только от качества прогнозов.

Эти выводы раскрывают смысл двух введенных Обуховым понятий — идеального потребителя, для которого не существует преимущественной «чистой» стратегии $U^{(1)} = U^{(0)}$, и критерия точности прогнозов

$$Q = 1 - \alpha_0 - \alpha_1. \quad (9)$$

С первого взгляда может показаться, что эффективность альтернативных прогнозов для реального потребителя связана с критерием точности лишь косвенно. Однако фактически эта связь довольно тесная. В самом деле, в случаях когда модуль разности $|U^{(1)} - U^{(0)}|$ велик, потребитель почти ничего или совсем ничего не может извлечь из прогноза. Главная часть экономического эффекта от прогноза приходится на потребителей, которые близки к идеальному в том смысле, что

$$|U^{(1)} - U^{(0)}| \ll \bar{v}.$$

Из формул (8) следует, что при

$$\frac{pv_1}{\bar{v}} \leq \frac{\alpha_0}{1 + \alpha_0 - \alpha_1}, \quad \frac{pv_1}{\bar{v}} \geq \frac{1 - \alpha_0}{1 - \alpha_0 + \alpha_1}$$

стратегия, основанная на использовании прогноза, не дает преимущества над лучшей из «чистых» стратегий.

Доля потребителей, для которых прогноз может быть полезным, также является характеристикой его качества. В первом приближении можно принять, что эта доля пропорциональна ширине интервала

$$\left(\frac{\alpha_0}{1 + \alpha_0 - \alpha_1}; \frac{1 - \alpha_0}{1 - \alpha_0 + \alpha_1} \right),$$

которая равна

$$\Delta = \frac{1 - \alpha_0 - \alpha_1}{1 - (\alpha_1 - \alpha_0)^2}. \quad (10)$$

При обычных значениях α_1 , α_0 знаменатель близок к единице и

$$\Delta \approx Q. \quad (11)$$

На рис. 1 в качестве примера изображена зависимость эффекта, который дает применение прогноза, от величины pv_1/\bar{v} для случая $\alpha_0 = 0,2$ и $\alpha_1 = 0,3$. Суммарный эффект для всех потребителей, которым прогноз приносит пользу, примерно пропорционален пло-

щади треугольника, выделенного на рисунке жирными линиями, т. е. величине $Q^2/2$.

Обобщение понятия идеального потребителя. Прежде чем ввести конкретное определение идеального потребителя для неальтернативного прогноза, целесообразно напомнить основную

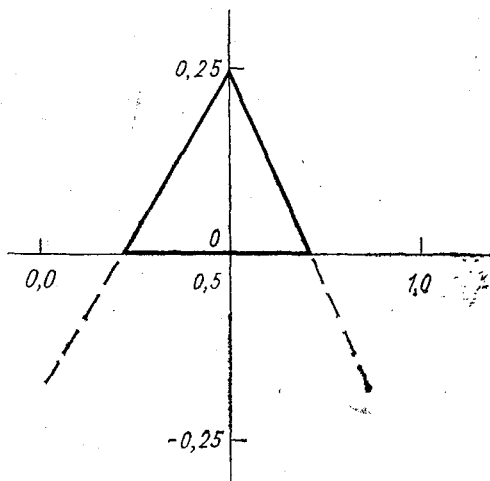


Рис. 1. Зависимость $\delta U/\bar{\sigma}$ (ось ординат) от $P\sigma_1/\bar{\sigma}$ (ось абсцисс).

идею работы [3]. В ней отмечалось, что при невысокой оправданности прогнозов решения о их применении не должны приниматься на уровне отдельных предприятий или хозяйств. В этом случае неизбежны ошибки, которые могут не только свести на нет возможный экономический эффект от применения прогнозов, но даже нанести серьезный ущерб хозяйству. Однако при управлении деятельностью целой отрасли народного хозяйства или крупного экономического района учет умеренно успешных прогнозов может дать существенный экономический эффект. Такой потребитель высокого уровня, дающий указания множеству отдельных потребителей и имеющий возможности осуществления крупных мероприятий, перераспределения ресурсов и т. п., может принимать оптимальные решения с учетом различных обстоятельств. Одним из наиболее важных аспектов принятия решения является ориентирование на тех подведомственных управляющему органу потребителей, которые не имеют преимущественной чистой стратегии. В этом смысле потребитель высшего уровня приближается к идеальному потребителю.

Решение потребителя будет зависеть от значения некоторого комплексного метеорологического параметра M . Для простоты будем рассматривать его как скаляр. В постановке работы [3] принимается, что точность прогноза можно охарактеризовать коэффициентом корреляции R между предсказанным и фактическим зна-

чением параметра, определяемым по достаточно большой экзаменационной (а не обучающей) выборке. Тогда при упрощающих предположениях, достаточно близких к действительности, расчет полного эффекта от прогноза при оптимальной стратегии его применения можно свести к вычислению двойного интеграла (формула (6) в работе [3]). Зависимость эффекта от показателя точности R оказывается квадратичной, как и в случае альтернативного прогноза.

Рассмотрим теперь вопросы оценки эффективности прогноза градаций комплексного метеорологического параметра (в частности, он может характеризовать распределение аномалий температуры и осадков по территории), когда решение принимает потребитель высокого уровня. Определим I градаций фактических значений параметра $\Phi_1, \dots, \Phi_i, \dots, \Phi_I$. Прогноз дается в тех же градациях $\Pi_1, \dots, \Pi_j, \dots, \Pi_I$ и при прогнозе Π_j принимается решение S_j . По отношению к потребителю высокого уровня можно считать, что все решения S_j различаются между собой. Пусть u_{ji} — экономический эффект, соответствующий решению S_j при осуществлении градации Φ_i , и n_{ij} — число случаев, когда это происходило. Тогда средний экономический эффект описывается формулой

$$U = \frac{1}{n} \sum_{ij=1}^I n_{ij} u_{ji}, \quad (12)$$

частным случаем которой является (3). Введем теперь в рассмотрение стратегии, соответствующие принятию решения S_k независимо от прогноза. Средний экономический эффект $U^{(k)}$ при стратегии S_k равен

$$U^{(k)} = \sum_{i=1}^I p_i u_{ki}, \quad (13)$$

где p_i — вероятность градации Φ_i .

В ряде случаев мы имеем основание приравнять вероятности p_i эмпирическим частотам, т. е. принять

$$p_i = \sum_{j=1}^I \frac{n_{ij}}{n}. \quad (14)$$

Однако более точным является определение значений p_i по обучающей выборке, по которой отработывалась формула прогноза.

Условия, характеризующие идеального потребителя и состоящие в отсутствии оптимальной «чистой» стратегии, можно записать в форме

$$U^{(k)} = \sum_{i=1}^I p_i u_{ki} = C. \quad (15)$$

Заметим еще, что доходы и убытки потребителя могут отсчитываться от переменного исходного уровня. В частности, можно выбрать исходный уровень из условия $C=0$. Но тогда условия (15) совпадают с соотношениями, введенными в работе [4] для определе-

ния меры успешности прогноза. Их смысл состоит в том, что в случаях когда распределение вероятностей градаций при прогнозах любых градаций одно и то же, мера успешности равна нулю. Как видим, такой подход применим лишь в случае идеального потребителя. Дальнейшие рассуждения и выводы работы [4] основываются на том, что мера успешности прогноза связывается с потерей информации по сравнению с точным прогнозом. Эти рассуждения представляются справедливыми при рассмотрении свойств идеального потребителя высокого уровня. Поскольку в его задачу входит принятие сложных решений, которые относятся к множеству отдельных предприятий и хозяйств, а также часто к разным видам хозяйственной деятельности, то трудно указать более общий показатель, характеризующий его потери по сравнению с точным прогнозом, чем потери информации. В связи с этим в простейшем случае можно записать дополнительные соотношения, характеризующие свойства идеального потребителя, в форме (см. [4]).

$$u_{ii} - u_{ki} = V \left[P(i, k) \lg P(i, k) - \sum_{l=\min(i, k)}^{\max(i, k)} p_l \lg p_l \right], \quad (16)$$

где

$$P(i, k) = \sum_{l=\min(i, k)}^{\max(i, k)} p_l. \quad (17)$$

Коэффициент пропорциональности V имеет смысл масштаба экономической зависимости потребителя от прогноза. Если он постоянен, то предложенная в [4] мера успешности метода прогноза Y прямо пропорциональна эффективности использования прогноза идеальными потребителями низшего уровня. Рассуждая далее так же, как и в случае альтернативного прогноза, заметим, что чем выше мера успешности прогноза, тем больше потребителей, несколько отличающихся от идеального, могут получить реальную пользу от прогноза. Доля таких потребителей в первом приближении пропорциональна Y , а полная эффективность прогноза пропорциональна Y^2 .

Вернемся теперь к вопросу о справедливости условия $V = \text{const}$. Представляется, что такой подход справедлив для оценки экономической эффективности большинства видов прогнозов многоцелевого характера, предназначенных для использования на высоком уровне, а также эффективности климатологической информации. Однако в ряде случаев решение потребителя очень тесно связано с переходом прогноза через определенную границу. В этих случаях более приемлемым может оказаться прием частичного упорядочения градаций, описанный в конце статьи [4], или прием варьирования значений коэффициента V для разных значений i .

В заключение заметим, что соотношения (16) всегда означают, что $u_{ii} > u_{ki}$, если $k \neq i$. Однако при сравнении u_{ii} с u_{ij} ($j \neq i$) аналогичные неравенства выполняются не всегда. В отдельных случа-

ях u_{ij} , соответствующее прогнозу более крайнего состояния, чем осуществившееся, превышает значение u_{ii} . Это обстоятельство впервые отметил Н. А. Багров применительно к матрице показателей информации Y_{ij} и счел его парадоксальным. По отношению к матрице u_{ij} возможность увеличения эффекта при прогнозе крайних градаций представляется вполне очевидной, так как крайняя формулировка прогноза заставляет принимать нестандартное решение большее число потребителей. Для матрицы показателей информации объяснение парадокса не столь очевидно. Превышение отдельных $Y_{ij} (j \neq i)$ над Y_{ii} объясняется тем, что показатели информативности являются относительными (информативность случайного прогноза принимается равной нулю) и необходимость этого условия может быть поставлена под сомнение. Однако если признать, что условие тесной связи между показателями качества прогнозов и эффективности их использования является важным, то высказанное выше соображение следует считать весьма существенным подтверждением правомерности этого условия.

СПИСОК ЛИТЕРАТУРЫ

1. Житорчук А. Т., Житорчук Ю. В., Кондратович К. В. О возможности использования человеко-машинных процедур при оценке качества прогнозов. Экспресс-информация. ВНИИГМИ—МЦД.—Метеорология. 1977, вып. 1 (48), с. 14—18.
2. Обухов А. М. К вопросу об оценке успешности альтернативных прогнозов.—Изв. АН СССР, сер. геофиз., 1955, № 4, с. 72—81.
3. Юдин М. И. Долгосрочный прогноз погоды, как средство управления хозяйственной деятельностью.—В кн.: Применение статистических методов в метеорологии.—Л.: Гидрометеиздат, 1977, с. 5—10.
4. Юдин М. И. Информационный подход к задаче опознавания градаций.—Докл. АН СССР, 1977, т. 237, № 5, с. 1031—1034.

В. М. Мирвис, М. И. Юдин

ПРИМЕНЕНИЕ ИНФОРМАЦИОННОГО ПОДХОДА К ИССЛЕДОВАНИЮ ИЗМЕНЕНИЙ КРУПНОМАСШТАБНЫХ ХАРАКТЕРИСТИК КЛИМАТА

В настоящее время в отделе динамической метеорологии ГГО имеется архив рядов (начиная с 1891 г.) данных крупномасштабных процессов (коэффициенты разложения метеорологических полей по е. о. ф., индексы циркуляции, ледовитость и др.), которые используются для исследования статистических связей в целях долгосрочного прогноза погоды [3, 4]. Анализ нестационарности этих рядов представляется важным прежде всего с точки зрения использования их в прогнозе, а также в качестве получения дополнительных сведений о характере флуктуаций климата.

Действительно, рассматриваемые нами переменные описывают режим обширных регионов и характеризуются значительной дис-

персией. Изучение климатических изменений этих переменных может привести к новым статистически обоснованным выводам, которые трудно обнаружить при рассмотрении данных по отдельным пунктам. С другой стороны ряды коэффициентов разложения метеорологических полей по е. о. ф. содержат значительно больше климатической информации, чем, например, средние значения температуры по полушарию.

Таблица 1

Характеристики, исследованные на наличие временного тренда

Функция	Форма представления	Период
Давление воздуха по секторам (Атлантико-Европейскому (I), Американскому (II) и Азиатскому (III))	Коэффициенты (a_1, a_2, a_3) разложения среднемесячных значений по е. о. ф.	1891—1978
Аномалии температуры воздуха по северному полушарию	Коэффициенты (a_1, a_2, a_3) разложения среднемесячных значений по е. о. ф.	1891—1978
Повторяемость типов циркуляции Г. Я. Вангенгейма (W, E, C).	Число дней с типами (W, E, C) за месяц	1891—1978
Повторяемость типов циркуляции Г. Я. Вангенгейма — А. А. Гирса ($З, M_1, M_2$)	Число дней с типами ($З, M_1, M_2$) за месяц	1900—1978
Ледовитость Баренцева моря	Средние месячные значения	1934—1978*

* Период с 1934 по 1978 был разбит на 9 равновероятных градаций.

Как и в работе [6], в настоящей работе для анализа нестационарности ряда используется метод информационной оценки опознавания градаций, предложенный в [5]. Этот метод не требует никаких допущений о характере распределения переменных и позволяет наряду с интегральными оценками информативности для гипотезы линейного тренда получить дифференциальные оценки значимости изменения переменной от одного временного интервала к другому.

В данной работе приводятся более полные, по сравнению с [6] выводы, полученные при использовании информационного подхода к анализу временных рядов крупномасштабных характеристик климата (табл. 1).

Напомним коротко основные положения используемого метода. Пусть переменная x в каждом конкретном случае относится к одной из градаций x_i с вероятностью p_i

$$x_1 < x_2 < x_3 < \dots < x_M.$$

Если в ряду x имеется временной тренд, то знание временного интервала, к которому относятся данные, должно нести информа-

цию о наиболее вероятной градации исследуемой переменной в этом временном интервале.

Мера информации ошибочного определения градации $Y(i, k)$, в отличие от классической теории информации, дифференцирована в зависимости от величины ошибки. $Y(i, k)$ убывает по мере уменьшения энтропии системы H , включающей объединение градаций от ошибочной x_i до истинной x_k , по сравнению с энтропией первоначальной системы градаций H_0 :

$$Y(i, k) = Y(i, i) - \underbrace{\left[p(i, k) \lg p(i, k) - \sum_{l=\min(i, k)}^{\max(i, k)} p_l \lg p_l \right]}_{\Delta H = H_0 - H}, \quad (1)$$

$$p(i, k) = \sum_{l=\min(i, k)}^{\max(i, k)} p_l.$$

Информационная мера случайного определения градаций принимается равной нулю

$$\sum_{k=1}^M p_k Y(i, k) = 0, \quad i = 1, 2, \dots, M. \quad (2)$$

Из (1) и (2) определяется

$$Y(i, i) = \sum_{j=1}^M p_j \left[p(i, j) \lg p(i, j) - \sum_{l=\min(i, j)}^{\max(i, j)} p_l \lg p_l \right]. \quad (3)$$

Если мы имеем матрицу повторяемостей $n(l, z)$ градаций переменной x_z в различных временных интервалах t_l , то на основании матриц $Y(i, k)$ и $n(l, z)$ можно рассчитать матрицу $I(m, j)$ средних мер информативности предсказания x_j градации в m -ом временном интервале:

$$I(m, j) = \frac{\sum_{\alpha=1}^M n(m, \alpha) Y(j, \alpha)}{\sum_{\alpha=1}^M n(m, \alpha)}. \quad (4)$$

Конфигурация элементов матрицы $I(m, j)$, существенно отличающихся от случайных, определяет характер зависимости x от t . Значимость отдельных элементов матрицы $I(m, j)$ определяется путем сопоставления их с доверительными пределами для $I(m, j)$ в предположении об отсутствии зависимости между x и t . Распределение $I(m, j)$ при достаточно больших N можно считать нормальным с математическим ожиданием, равным нулю, и дисперсией

$$\sigma_j^2 = \sum_{\alpha=1}^M p_\alpha [Y(j, \alpha)]^2. \quad (5)$$

В случае, когда элементы матрицы $I(m, j)$ существенно отличающиеся от случайных, располагаются вдоль одной из главных диагоналей, мы имеем прямую или обратную линейную связь (ли-

нейный тренд). Среднее из диагональных членов матрицы $I(m, j)$ будет оценкой информативности гипотезы линейного возрастания I_+ или линейного убывания I_- величины x во времени:

$$I_+ = \frac{\sum_{\alpha=1}^M I(\alpha, \alpha) N_\alpha}{\sum_{\alpha=1}^M N_\alpha}, \quad I_- = \frac{\sum_{\alpha=1}^M I(M+1-\alpha, \alpha) N_\alpha}{\sum_{\alpha=1}^M N_\alpha}, \quad (6)$$

$$N_\alpha = \sum_{\beta=1}^M n(\alpha, \beta).$$

Помимо оценки значимости связи x и t в целом представляется интересным оценить значимость изменения x в соседних временных интервалах. В частности, интересно проверить гипотезу о возможности скачкообразных изменений климатических характеристик вследствие перехода от одного устойчивого состояния к другому. С этой целью рассмотрим разность элементов матрицы $I(m, j)$, относящихся к соседним временным интервалам:

$$\Delta I_j = I(m, j) - I(m+1, j). \quad (7)$$

При условии независимости $I(m, j)$ и $I(m+1, j)$ величина $\sigma \Delta I_j$ для случайного диагноза может быть определена

$$\sigma^2 \Delta I_j = \sigma^2_j / N_m + \sigma^2_j / N_{m+1}. \quad (8)$$

Теперь, установив некоторый доверительный уровень, можно определить, превосходит ли ΔI_j соответствующие доверительные пределы. Перечень характеристик, исследованных нами на наличие трендов, приведен в табл. 1.

На первом этапе проверялась гипотеза линейных трендов. С этой целью по всем рядам вычислялись показатели I_+ и I_- (ряды 1891—1978 гг. были разбиты на 11, а ряды 1900—1978 гг. — на 10 равновероятных градаций). На рис. 1 приведены графики изменения I_+ и I_- по месяцам. Как видно из графиков, в рядах многих характеристик ($a_{1д}$ и $a_{3д}$ среднемесячного давления по секторам; чисел дней с формами циркуляции W , E , M_1 , M_2 ; $a_{1г}$ аномалии температуры воздуха северного полушария; ледовитость Баренцова моря) имеются четко выраженные временные тренды, статистическая значимость которых даже для отдельных месяцев превышает 5 %-ный уровень значимости. Оказалось, что для каждой из перечисленных характеристик можно выделить периоды длительностью несколько месяцев, полгода или даже год, в течение которых знак тренда сохраняется. В этих случаях рассчитывались осредненные по периодам характеристики и производилась оценка тренда. Осреднение проводилось для значений x , выраженных в непрерывной шкале, а затем значения осредненных характери-

стик вновь разбивались на градации. Перед осреднением каждый ряд нормировался по формуле

$$x' = (x - \bar{x}) / \sigma_x,$$

где \bar{x} — среднее арифметическое, σ_x — среднее квадратическое отклонение x . Значимость тренда для некоторых из осредненных ха-

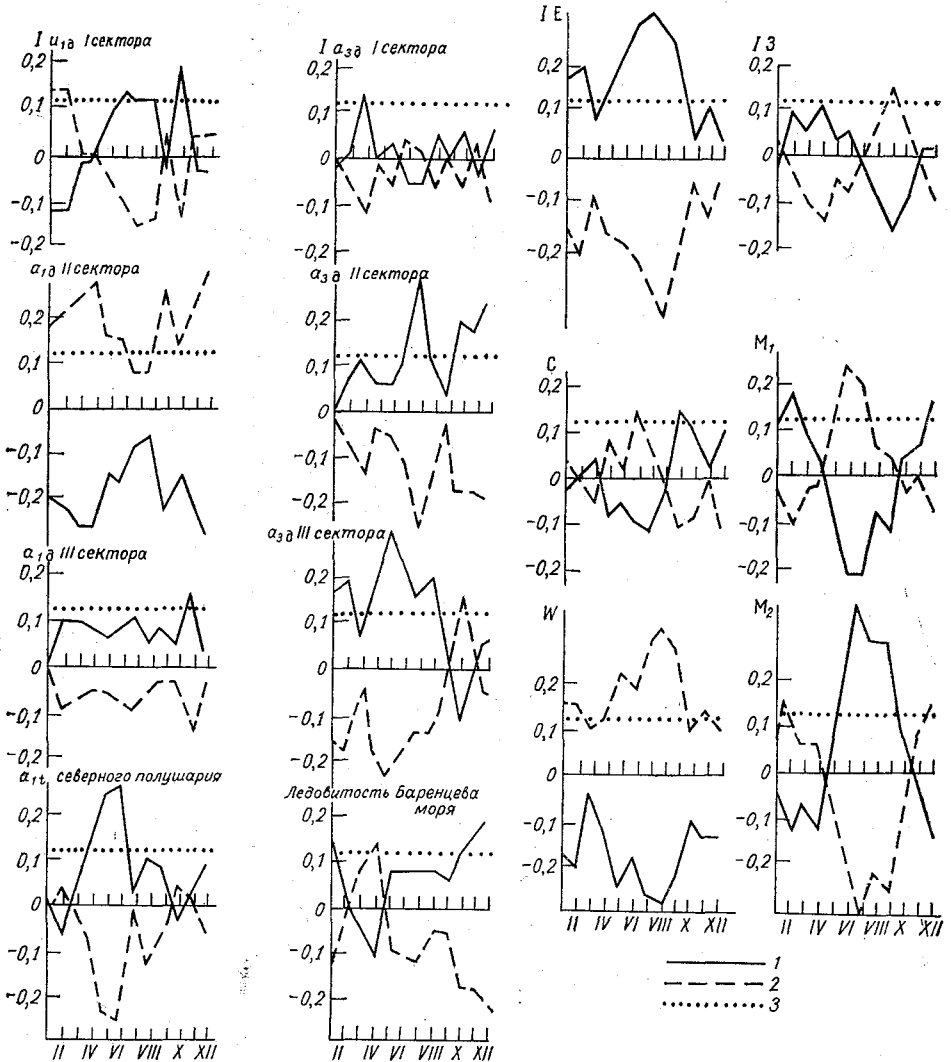


Рис. 1. Изменение показателей I_+ (1) и I_- (2) по месяцам для некоторых крупномасштабных характеристик атмосферных процессов. 5% доверительный уровень (3).

рактических характеристик получилась исключительно высокой. Так, в рядах $a_{1д}$ давления II сектора, осредненных за теплое и холодное полу-

годия, $a_{3д}$ II сектора за холодный период, $a_{3д}$ III сектора за теплый период, чисел дней с формами циркуляции W и E за год, чисел дней с формами M_1 и M_2 за теплое полугодие значимость трендов превосходит 4σ .

Интерпретация отмеченных трендов $a_{1д}$ и $a_{3д}$ давления возможна на основе карт распределения соответствующих естественных

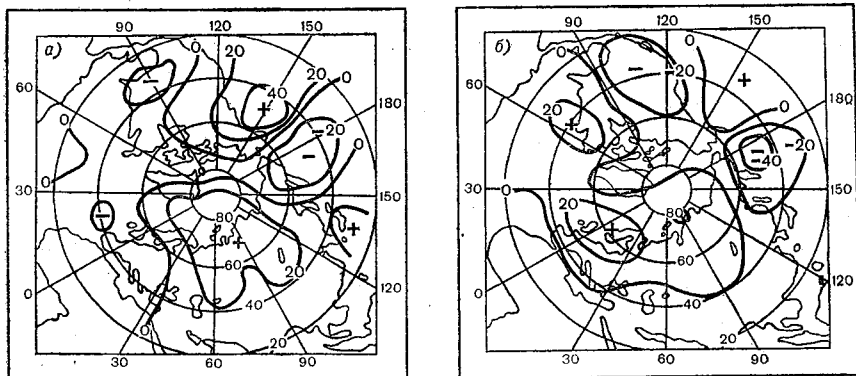


Рис. 2. Поля третьей естественной составляющей среднего месячного давления для теплого (а) и холодного (б) периодов.

функций и с учетом вклада их в суммарную дисперсию. Так, например, тенденция к уменьшению $a_{1д}$ в холодное полугодие, наиболее четко выраженная в Американском секторе, но проявляющаяся и в других секторах, свидетельствует об ослаблении зональной составляющей циркуляции (карты распределения первой е. о. ф. опубликованы в [6])*. Изменения, свойственные $a_{3д}$, отражают, по-видимому, изменения интенсивности муссонной циркуляции и смещения отдельных центров действия атмосферы (рис. 2). Для того чтобы судить о росте или уменьшении давления в конкретной точке полушария, необходимо рассматривать все коэффициенты разложения в совокупности. Поэтому здесь можно говорить лишь об относительном изменении давления в соответствии с вкладом $a_{3д}$ в суммарную дисперсию. Так, положительный тренд $a_{3д}$ Американского и Азиатского секторов в холодный период свидетельствует об относительном увеличении давления в северо-восточной части Тихого и в северо-западной части Атлантического океана по сравнению с относительным падением давления в северо-западной части Тихого океана, а также в Северо-Американском и Гренландском секторах Арктики. Возрастание $a_{3д}$ в теплый период также указывает на относительный рост давления в северо-восточной части Тихого океана, по сравнению с падением давления в северо-западной его части. Область относительного увеличе-

* Разложение полей давления по е. о. ф. выполнено Н. А. Болдыревой и Н. А. Гулиной.

ния давления захватывает также северо-западную половину Северо-Американского континента.

Положительный тренд в ходе числа дней с формой E и отрицательный тренд для W находятся в полном соответствии с выводами А. А. Гирса о смене циркуляционных эпох [1]. Для ряда числа дней с формой M_1 характерно уменьшение в теплый период и увеличение в холодный; для чисел дней с формой M_2 картина противоположная. Тот факт, что в рамках одной и той же циркуляционной эпохи в различные месяцы могут преобладать различные формы циркуляции, нашел свое отражение в работах А. А. Гирса как деление на внутригодовые стадии [2]. Выводы, касающиеся тренда в ходе a_{17} аномалии температуры северного полушария, мало отличаются от опубликованных нами в [6], поэтому здесь на них не останавливаемся.

На втором этапе по матрицам $I(m, j)$, рассчитанным для осредненных характеристик, была проведена оценка значимости изменения показателей информативности $I(m, 1)$ и $I(m, M)$ соответствующих крайним градациям x при переходе от одного временного интервала к другому. Показатели информативности для крайних градаций должны быть особенно чувствительны к изменениям в повторяемости крупных аномалий рассматриваемой характеристики. Ход $I(m, 1)$ и $I(m, M)$, как правило, прямо противоположен. Выводы, полученные при анализе изменений $I(m, 1)$ и $I(m, M)$ оказались практически одинаковы, поэтому ниже рассмотрим только $I(m, M)$.

На рис. 3 приведены примеры графиков, показывающих изменение $I(m, M)$ в зависимости от m , столбчатой диаграммой показаны значения ΔI_M . Из графиков видно, что в ряде случаев величина ΔI_M превосходит $\sigma_{\Delta I_M}$ более чем в два раза. В этом случае с большой степенью достоверности можно говорить о резком изменении климатической характеристики на рубеже двух рассматриваемых временных интервалов. Поскольку границы временных интервалов не варьировались нами и они могут не совпадать по времени с моментами наибольшего изменения в значениях x , имеет смысл рассматривать и относительно малые изменения $I(m, M)$.

В табл. 2 приведена сводка наличия изменений $I(m, M)$, для которых $\Delta I_M > 1,5 \sigma_{\Delta I_M}$. Изменение в сторону возрастания $I(m, M)$ показано знаком «+», в сторону убывания — знаком «-», звездочкой отмечены случаи, для которых $\Delta I_M > 2\sigma_{\Delta I_M}$. Из табл. 2 видно, что в определенные периоды повторяемость значительных изменений $I(m, M)$ явно превышает ожидаемую при случайном характере этих изменений. Так, для 19 рассмотренных характеристик, при условии их независимости, можно ожидать, что $\Delta I_M > 1,5 \sigma_{\Delta I_M}$ встретится в 2—3 случаях. На рубеже 20-х и 30-х годов отмечено 10 случаев $\Delta I_M > 1,5 \sigma_{\Delta I_M}$. Такая повторяемость даже при учете связности характеристик вряд ли является случайной.

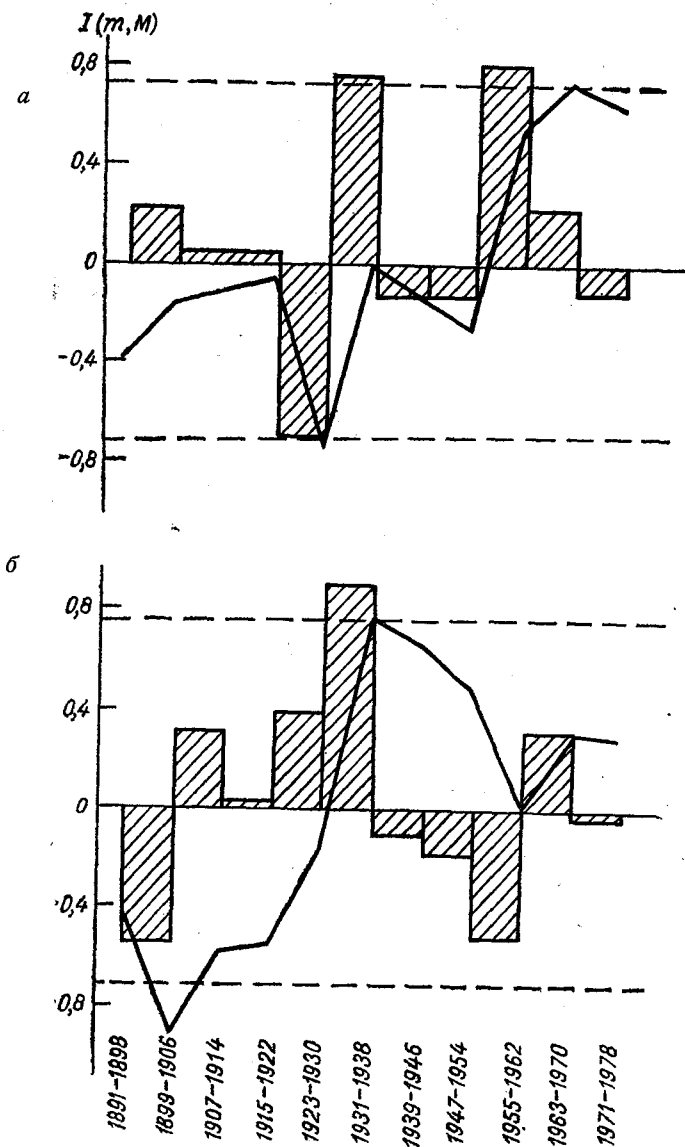


Рис. 3. Изменение показателя $I(m, M)$ по восьмилетним временным интервалам.

a — для Δ давления Атлантико-Европейского сектора за теплый период, *b* — для Δ аномалии температуры северного полушария в теплый период; пунктиром показан 5% доверительный уровень для ΔI_M .

Следует отметить также повышенную повторяемость резких изменений в соседних с 30-ми годами временных интервалах. По-видимому, некоторые характеристики начали меняться ранее, а некоторые — несколько позднее. Для большинства характеристик, по нашим данным, резкие изменения отмечены именно на рубеже 20-х и 30-х годов. Большая повторяемость значительных изменений $I(m, M)$ приходится на начало 50-х годов. Если судить по левой части табл. 2, полученной по данным начиная с 1891 г., то следует отметить также некоторый перелом в конце предыдущего и начале нынешнего столетия.

Последние годы (1971—1978 гг.) не выделяются большой повторяемостью резких изменений, однако в отдельных характеристиках наблюдаются изменения $I(m, M)$, превосходящие по величине $2\sigma_{\Delta I/M}$.

СПИСОК ЛИТЕРАТУРЫ

1. Гирс А. А. Многолетние колебания атмосферной циркуляции и долгосрочные гидрометеорологические прогнозы. — Л.: Гидрометеониздат, 1971. — 280 с.
2. Гирс А. А. Макроциркуляционный метод долгосрочных метеорологических прогнозов. — Л.: Гидрометеониздат, 1974. — 477 с.
3. Мещерская А. В., Леднева К. В., Блажевич В. Г. Характеристика дополнительной гидрометеорологической информации, используемой в физико-статистическом прогнозе. — Труды ГГО, 1975, вып. 353, с. 115—122.
4. Юдин М. И., Мещерская А. В., Блажевич В. Г. Характеристика гидрометеорологической информации, использованной при долгосрочном прогнозе температуры и осадков для районов недостаточного увлажнения. — Труды ГГО, 1970, вып. 236, с. 45—63.
5. Юдин М. И. Информативный подход к задаче опознавания градаций. — Докл. АН СССР, 1977, т. 237, № 5, с. 1031—1034.
6. Юдин М. И., Блажевич В. Г., Голод М. П., Мирвис В. М. Применение информационного подхода к исследованию долгопериодных атмосферных процессов и изменений климата. — Труды III Всесоюзного симпозиума по применению статистических методов в метеорологии. 1978, с. 223—233.

А. В. Мещерская, В. Г. Блажевич, М. П. Голод, И. Г. Белякина:

СТАТИСТИЧЕСКАЯ ОЦЕНКА ИНФОРМАТИВНОСТИ ПРЕДИКТОРОВ В ЗАВИСИМОСТИ ОТ ЗАБЛАГОВРЕМЕННОСТИ

В настоящее время разработан и внедрен в оперативную практику ряд методов, позволяющих составлять сверхдолгосрочные метеорологические прогнозы с заблаговременностью от нескольких месяцев до одного-двух лет [1, 2, 3, 4, 5, 8, 16, 21]. Сверхдолгосрочные прогнозы базируются на самых разных принципах: изучении типов макроциркуляционных процессов и их преобразований [5], статистической экстраполяции временных рядов [21], асинхронных статистических связей между предикторами и предиктантами [1—4, 12], учете гелиогеофизических факторов [16] и др. Хотя ме-

тоды составления сверхдолгосрочных прогнозов различны, уровень их успешности по единодушному мнению специалистов достаточно низок. Однако оправдываемость сверхдолгосрочных прогнозов не ниже оправдываемости долгосрочных прогнозов погоды (при общем низком уровне успешности).

Этот вывод стимулирует разработки, направленные на дальнейшее увеличение заблаговременности прогнозов до нескольких лет и даже десятилетий (климатические прогнозы). Однако число таких разработок не очень велико. Обнадеживающие результаты получены О. А. Дроздовым, А. С. Григорьевой и Л. Г. Полозовой на основе учета цикличности атмосферных процессов [6, 7, 8, 17]. Это направление оказалось перспективным лишь для прогноза метеорологических параметров, сильно осредненных во времени (5—10 лет). Попытки давать тем же путем прогнозы на отдельные месяцы или части сезона пока не увенчались успехом [6].

Наиболее последовательно идею сверхдолгосрочных месячных прогнозов на несколько лет вперед развивает Е. В. Воробьева, используя в качестве предикторов интенсивность атмосферной циркуляции [3] и некоторые энергетические характеристики атмосферы [4]. Многолетние архивы метеорологических данных, используемых в физико-статистическом методе долгосрочных прогнозов погоды [15, 24], позволяют оценить обеспеченность сверхдолгосрочных метеорологических прогнозов на большом статистическом материале.

Из архивов исходных данных [15, 24] было отобрано 190 длиннорядных предикторов, данные по которым имелись с 1891 г., так как длиннорядные предикторы более информативны [14; 18]. Большая часть длиннорядных предикторов характеризует крупномасштабные особенности полей наземного среднего месячного давления в трех секторах северного полушария (108 предикторов) и средней месячной температуры (36 предикторов). Давление и температура представлены в виде трех первых коэффициентов разложения каждого поля по естественным ортогональным функциям (е. о. ф). Кроме того, в архив длиннорядных предикторов включена повторяемость за месяц числа дней с формами циркуляции E и W по Вангенгейму (24 предиктора), планетарный индекс K_p (15 предикторов) за каждый месяц и осредненный по трехмесячным интервалам (апрель—июнь, июль—сентябрь и октябрь—декабрь), а также средний месячный общий индекс зональности (I_3) Атлантико-Европейского сектора с октября по апрель (7 предикторов). Общий индекс зональности, предложенный А. Л. Кацем [10], рассчитан в Гидрометцентре СССР за период с 1938 г. и по настоящее время. За более ранние годы значения I_3 восстановлены Манделем [11].

Предиктантами являются первые два коэффициента разложения по е. о. ф. средней месячной температуры и месячных сумм осадков за апрель—июль в основных сельскохозяйственных районах ЕТС, Северного Казахстана и Западной Сибири.

На базе перечисленных материалов проведены статистические эксперименты по оценке информативности предикторов для сверхдолгосрочного прогноза температуры и осадков с заблаговременностью от одного до семи лет.

Первый этап статистического эксперимента состоял в расчете корреляционных связей между предикторами и предиктантами с целью составления независимых прогнозов на 10 лет (1969—1978 гг.). Для каждого года рассчитывались корреляционные связи, при этом длина исходных рядов несколько менялась в зависимости от заблаговременности предикторов и года, на который составлялся прогноз. Условно будем считать прогнозы по данным за март прошлого — февраль текущего года прогнозами с годовой заблаговременностью, прогнозы по данным за март позапрошлого — февраль прошлого года прогнозами с двухлетней заблаговременностью и т. д. Ряды предикторов двух-, трех- и семилетней заблаговременности получены путем временного сдвига соответственно на один, два и шесть лет рядов предикторов годовой заблаговременности.

При расчете корреляционных связей на 1969 г. с годовой заблаговременностью использовались ряды с 1891 по 1968 г. (78 лет), при расчете на 1970 г. — с 1891 по 1969 г. (79 лет) и т. д. При увеличении заблаговременности на один год длина исходных рядов укорачивается на единицу. В общей сложности длина рядов, по которым рассчитываются корреляционные связи, колеблется от 72 до 88 лет.

Оценка информативности предикторов была выполнена двумя уже известными способами [13, 20, 22]. Первый состоял в подсчете числа корреляционных связей (q) в процентах от возможного, превышающих по абсолютному значению заданный уровень $k \left(\left| \frac{z}{\sigma_z} \right| \geq k \right)$, где z — преобразование Фишера, σ_z — средняя квадратическая ошибка z . В данном эксперименте k было принято равным 2,15. Полученное для каждого варианта расчетов значение q сравнивалось со значением q_c при нулевой гипотезе, предполагающей наличие лишь случайных связей между предикторами и предиктантами:

$$q_c = 1 - 2 \Phi \left(\frac{z}{\sigma_z} \right), \quad (1)$$

где Φ — нормированная функция Лапласа. Для $|z/\sigma_z| \geq 2,15$ $q_c = 3,2$ %. Оценка по q была впервые применена в физико-статистическом методе в 1972 г. [22].

В 1981 г. Ю. В. Житорчук [9] предложил еще один критерий информативности

$$\varphi = \frac{1}{n} \sum_{i=1}^n \left(\frac{z_i}{\sigma_{z_i}} \right)^2. \quad (2)$$

В оценке по φ учитываются все корреляционные связи, независи-

мо от их тесноты. В отличие от q статистика ϕ является несмещенной. Однако оценка по ϕ имеет и один существенный недостаток. Указывая лишь на наличие прогностической информации в данной совокупности предикторов, она не дает возможности выявить преимущества каждого конкретного предиктора.

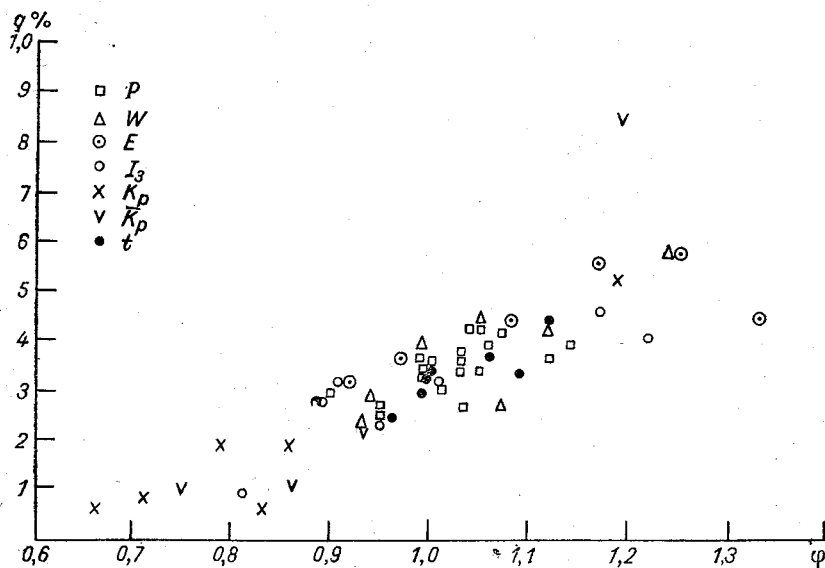


Рис. 1. Взаимосвязь двух оценок информативности q и ϕ , рассчитанных по единому исходному материалу (1891—1977 гг.) P — наземное давление, W и E — число дней с формой W и E ; I_3 — индекс зональности; t — средняя месячная температура, K_p и K_p — индексы геомагнитной возмущенности, средние соответственно за месяц и три месяца.

Оценки по q и ϕ достаточно близки. Об этом свидетельствует рис. 1, где приведена зависимость между q и ϕ для разных групп предикторов и разной заблаговременности. Вычисления ϕ и q сделаны на одном и том же исходном материале за период с 1891 по 1977 г. Нетрудно видеть, что большому значению q (%) соответствует более высокое значение ϕ , и наоборот.

Для оценки информативности по q в общей сложности было рассчитано 425 600 корреляционных связей (190 предикторов, 32 предиктанта, 10 лет, заблаговременность от 1 до 7 лет). Следует иметь в виду, что имеет место дублирование числа корреляционных связей за счет их расчета для соседних лет. Согласно [23], из 100 значимых корреляционных связей, полученных для прогноза осадков на n -й год по 40-летним рядам данных, 86 % связей остаются значимыми при прогнозе на $n+1$ -й год и 80 % — при прогнозе на $n+2$ -й год. Для длинных рядов данных доля сохранившихся значимых связей будет еще выше. Но даже если допустить, что общее число корреляционных связей, рассчитанных за

10 лет, эквивалентно числу независимых корреляционных связей за один год, объем проведенных вычислений будет на два порядка превосходить объем ранее выполненных расчетов [3, 4].

В зависимости от заблаговременности число значимых связей N распределилось по годам неравномерно (табл. 1). Наибольшие значения соответствуют ($N=2352$) годовой, шестилетней ($N=2129$) и пятилетней ($N=2110$) заблаговременности.

Из 425 600 корреляционных связей 13 978 оказались значимыми (q по всей совокупности составило 3,3 %), что лишь на 0,1 % выше q_c . Для заблаговременностей в один год, пять и шесть лет значения q были выше случайного и равнялись соответственно 3,9; 3,5 и 3,5 %. Для остальных интервалов заблаговременности $q < 3,2$ (табл. 1, рис. 2).

Помимо годовых значений q представляют интерес значения q для каждого месяца (рис. 2). Отметим, что средние месячные значения q характеризуют информативность предикторов конкретного месяца по отношению к придиктантам за все весенне-летние месяцы вместе (апрель—июль). В этом состоит одно из отличий от работы [22], где информативность предикторов за отдельные месяцы рассчитывалась по отношению к придиктантам за каждый месяц весенне-летнего периода. Рис. 2 показывает, что колебания q от месяца к месяцу велики и, по-видимому, носят случайный характер. На этом фоне очень высокой информативностью выделяется январь текущего года, для которого $q=7,3$, т. е. в два раза больше q_c . Более детальный анализ распределения q по отдельным группам предикторов показал, что высокая информативность января текущего года обусловлена большей долей значимых связей для индекса зональности I_3 ($q=14,6$ %), числа дней с формой W (10,3 %) и давления ($q=6,5$ %).

Наряду с январем большие значения q пришлось на декабрь и август прошлого года. Эти три месяца (особенно январь) определили повышенное число значимых связей в процентах от возможного для предикторов годовой заблаговременности в целом. Таким образом, подтверждается сделанный ранее вывод [22] о большей информативности (по q) ближайшей предыстории. Аналогичный результат получен в работе [9] при использовании критерия ϕ .

Расчет доли значимых связей в зависимости от заблаговременности был выполнен также дифференцированно по группам предикторов (табл. 2, рис. 3). Распределение q по годам получилось довольно изменчивым, как и распределение числа значимых связей по заблаговременности в работах [3, 4]. Максимальные значения q для температуры, числа дней с западной формой циркуляции (W) и зонального индекса (I_3) пришлось на годовую заблаговременность, для числа дней с восточной формой циркуляции (E) — на двухлетнюю.

Отличным от других групп предикторов оказалось распределение по годам доли значимых связей для индекса K_p . Наибольшее значение q соответствует пяти и шестилетней заблаговременности. Причем при пятилетней заблаговременности $q=7,3$, что является

Оценка прогнозов коэффициентов разложения по е.о.ф. полей осадков и температуры в зависимости от заблаговременности

Оценка прогнозов	Заблаговременность прогнозов											
	1 год			2 года			3 года			4 года		
	п	р	Р	п	р	Р	п	р	Р	п	р	Р
очень хорошие хорошие удовлетворительные неудовлетворительные плохие N q	331	14,1	14,1	268	14,5	14,5	249	13,8	13,8	264	13,9	13,9
	583	24,8	38,9	462	25,0	39,4	428	23,7	37,5	454	23,8	37,7
	745	31,7	70,5	618	33,4	72,8	574	31,8	69,3	628	33,0	70,7
	658	28,0	98,5	479	25,9	98,7	529	29,3	98,6	529	27,8	98,5
	35	1,5	100	24	1,3	100	26	1,4	100	29	1,5	100
	2552		1851			1806			1904			
	3,9		3,0			3,0			3,1			
Оценка прогнозов	Заблаговременность прогнозов											
	5 лет			6 лет			7 лет			Вся совокупность		
	п	р	Р	п	р	Р	п	р	Р	п	р	Р
	309	14,6	14,6	334	15,7	15,7	268	14,7	14,7	2023	14,5	14,5
	544	25,8	40,4	551	25,9	41,6	444	24,3	39,0	3466	24,8	39,3
697	33,0	73,5	665	31,2	72,8	608	33,3	72,3	4535	32,4	71,7	
533	25,3	98,7	548	25,7	98,5	482	26,4	98,7	3758	26,9	98,6	
27	1,3	100	31	1,5	100	24	1,3	100	196	1,4	100	
	2110		2129			1826			13978			
	3,5		3,5			3,0			3,3			

Примечание. п — число прогнозов по градациям; N — общее число прогнозов; q — число значимых связей в процентах от возможного; р — повторяемость прогнозов в процентах; Р — накопленная повторяемость прогнозов в процентах.

абсолютным максимумом по всем вариантам расчета. При заблаговременности от одного года до четырех лет и при семилетней заблаговременности q значительно меньше q_c .

Для группы наземного давления амплитуда изменения q по годам невелика ($q=3,4$ при четырехлетней заблаговременности; $q=3,0$ при двух- и семилетней заблаговременности). Полагая, что это связано со слишком большой численностью этой группы предикторов, расчеты q были выполнены отдельно для наземного давления по каждому из трех секторов северного полушария: Атлантико-Европейскому (I), Американскому (II) и Азиатскому (III). При этом амплитуда колебаний q по годам возросла не очень сильно. Наземное давление Атлантико-Европейского сектора в целом более информативно (по всем интервалам заблаговременности $q=3,6$), чем Американского ($q=3,2$) и особенно Азиатского ($q=3,1$) секторов.

Рассмотрим теперь информативность предикторов по оценке успешности независимых прогнозов коэффициентов разложения предиктантов, основанных на количественных параметрах аналогичности (A). Качество каждого прогноза определяется по отношению средней квадратической ошибки прогноза к квадрату ошибки случайного прогноза по аналогу [13, 20]. При

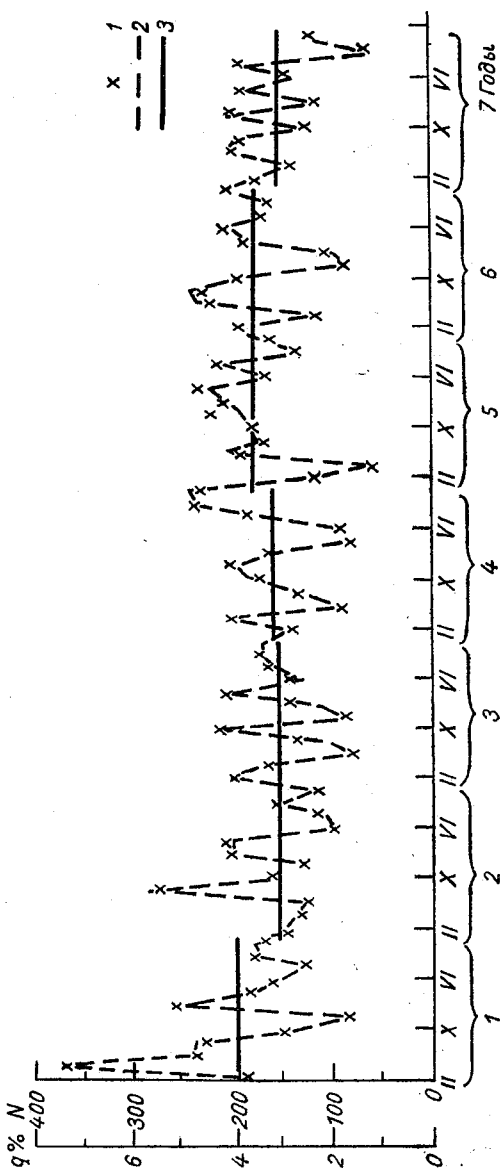


Рис. 2. Изменение доли (%) значимых связей q (1) и их числа N (2) в зависимости от заблаговременности по месяцам и годам, а также средний уровень q по годам (3).

Таблица 2
Повторяемость оценок прогнозов коэффициентов разложения по е.о.ф. полей осадков и температуры, составленных по разным группам предикторов в зависимости от заблаговременности

Оценка прогнозов	Число дней с формами циркуляции по Г. Я. Вангенгейму		Средняя месячная температура		Среднемесячное давление				Индекс зональной циркуляции		Индекс геомагнитной возмущенности					
	E		T		P _I		P _{II}		P _{III}		J ₃		K _p			
	п	р	п	р	п	р	п	р	п	р	п	р	п	р		
	W		Г		P _I		P _{II}		P _{III}		J ₃		K _p			
Заблаговременность 1 год																
очень хорошие	27	18,5	31	14,2	65	12,2	61	13,3	52	14,1	66	16,3	21	15,2	8	9,5
хорошие	30	39,0	64	43,6	137	37,8	105	36,2	81	36,1	102	41,5	35	40,6	29	44,0
удовлетворительные	42	67,8	56	69,3	177	71,0	156	70,2	124	69,8	132	74,1	36	66,7	22	70,2
неудовлетворительные	47	100	64	98,6	149	98,9	133	99,1	105	98,4	95	97,5	40	95,6	25	100
плохие	0	100	3	100	6	100	4	100	6	100	10	100	6	100	0	100
N	146		218		534		459		368		405		138		84	
q	3,8		5,7		4,6		4,0		3,2		3,5		6,8		1,8	
Заблаговременность 2 года																
очень хорошие	35	19,0	3	4,2	62	14,9	56	16,6	58	13,4	39	13,9	8	11,8	7	11,3
хорошие	40	40,8	17	27,8	107	40,7	92	43,9	95	35,4	73	39,9	21	42,6	17	38,7
удовлетворительные	58	72,3	30	69,4	136	73,5	103	74,5	151	70,4	85	70,1	26	80,9	29	85,5
неудовлетворительные	48	98,4	22	100	105	98,8	79	97,9	125	99,3	78	97,9	13	100	9	100
плохие	3	100	0	100	5	100	7	100	3	100	6	100	0	100	0	100
N	184		72		415		337		432		281		68		62	
q	4,8		1,9		3,6		2,9		3,8		2,4		3,0		1,3	
Заблаговременность 3 года																
очень хорошие	25	15,1	8	8,2	43	11,7	64	14,7	50	14,8	49	15,8	8	9,5	2	20,0
хорошие	36	36,7	34	43,3	75	32,2	104	38,7	81	38,8	74	39,5	23	36,9	1	30,0
удовлетворительные	55	69,9	28	72,2	139	70,2	135	69,8	97	67,5	88	67,8	30	72,6	2	50,0
неудовлетворительные	47	98,2	27	100	105	98,9	123	98,2	105	98,5	97	99,0	20	96,4	5	100
плохие	3	100	0	100	4	100	8	100	5	100	3	100	3	100	0	100

N	166	97	366	434	338	311	84	10
q	4,3	2,5	3,2	3,8	2,9	2,7	3,8	0,2
Заблаговременность 4 года								
очень хорошие	15	12,7	13,2	37	15,2	68	4	6,1
хорошие	23	32,2	35,7	59	41,9	99	15	28,8
удовлетворительные	36	62,7	70,5	81	72,2	138	22	62,1
неудовлетворительные	41	97,4	100	101	98,2	124	23	97,0
плохие	3	100	100	5	100	100	2	100
N	118	129	283	475	435	370	66	28
q	3,1	3,4	2,4	4,1	3,8	3,2	2,9	0,6

Заблаговременность 5 лет								
очень хорошие	9	10,0	9,7	51	15,1	53	2	33,3
хорошие	27	40,0	36,9	90	41,8	110	0	33,3
удовлетворительные	38	82,2	74,8	103	72,4	131	1	50,0
неудовлетворительные	16	100	100	89	98,8	107	2	83,3
плохие	0	100	100	4	100	4	1	100
N	90	103	337	405	392	427	6	350
q	2,3	2,7	2,9	3,5	3,4	3,7	0,3	7,3

Заблаговременность 6 лет								
очень хорошие	18	14,6	16,3	56	13,3	58	3	7,3
хорошие	31	39,8	39,4	114	40,4	119	10	31,7
удовлетворительные	41	73,2	73,7	124	69,8	141	12	61,0
неудовлетворительные	30	97,6	100	118	97,9	106	12	90,2
плохие	3	100	100	9	100	4	4	100
N	123	160	421	428	339	351	41	266
q	3,2	4,2	3,6	3,7	2,9	3,0	1,8	5,5

Заблаговременность 7 лет								
очень хорошие	20	12,6	15,6	49	13,2	51	10	15,9
хорошие	39	37,1	38,5	92	38,1	100	10	31,7
удовлетворительные	42	63,5	75,6	116	69,5	121	22	66,7
неудовлетворительные	52	96,2	98,5	110	99,2	91	20	98,4
плохие	6	100	100	3	100	7	1	100
N	159	135	370	360	315	380	63	44
q	4,1	3,5	3,2	3,1	2,7	3,3	2,8	0,9

этом весь диапазон изменения ошибок прогноза делился на пять градаций (очень хорошие, хорошие, удовлетворительные, неудовлетворительные и плохие) и подсчитывалась повторяемость прогнозов по каждой градации.

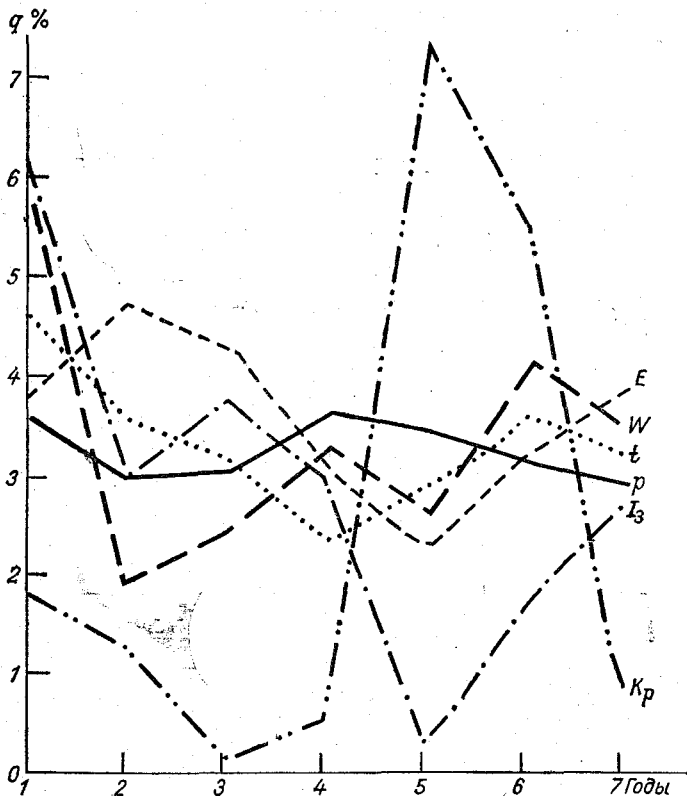


Рис. 3. Изменение доли значимых связей (q) для различных групп предикторов в зависимости от заблаговременности (K_p — индекс геомагнитной возмущенности за все интервалы осреднения).

Усл. обозначения см. рис. 2.

Под прогнозом здесь понимается прогноз коэффициентов разложения предиктанта по одночленному уравнению регрессии, которое составляется для каждой значимой корреляционной связи.

В последние годы М. И. Юдиным разработана еще одна мера успешности Y прогнозов коэффициентов разложения предиктантов [18], основанная на соображениях теории информации. Ее отличительная черта состоит в том, что повышенное значение придается оценке прогнозов крайних значений предиктантов. Однако в регрессионных схемах прогностические указания на аномальные явления встречаются редко, поэтому оценка по Y оказалась достаточно близкой к оценке по A , что показано в работе [14].

В табл. 1 и 2 данной статьи приведены результаты оценок по A независимых прогнозов за 10 лет. Анализ оценок прогнозов удобно проводить по накопленным повторностям ($P_{\text{опр}}$) прогнозов первых трех градаций (очень хорошие, хорошие и удовлетворительные), которые в дальнейшем будем называть оправдавшимися прогнозами.

Из табл. 1 и рис. 4 следует, во-первых, что монотонное уменьшение успешности прогнозов с ростом их заблаговременности отсутствует, во-вторых, что наибольшая повторяемость оправдавшихся прогнозов приходится на пяти-, шестилетнюю и двухлетнюю заблаговременность. По методике, представленной в [13, 18], оценена статистическая значимость превышения $P'_{\text{опр}}$ для этих интервалов заблаговременности над уровнем $P_{\text{опр}}$ для всей совокупности. Разность $\Delta P_{\text{опр}} = P'_{\text{опр}} - P_{\text{опр}}$ сравнивалась со средним квадратическим отклонением (σ) этой разности [27]:

$$\sigma = \sqrt{\frac{P'_{\text{опр}}(1 - P'_{\text{опр}})}{n} \frac{n_2}{n_1}}, \quad (3)$$

где n — общее число случаев во всей совокупности, а n_1 — число случаев в частной совокупности (например, при конкретной заблаговременности), причем $n = n_1 + n_2$.

Гипотеза о неслучайности различий повторяемостей хорошо подтверждается лишь для пятилетней заблаговременности, для которой $(P'_{\text{опр}} - P_{\text{опр}}) > 2,3 \sigma$, при 2 и 6-летней заблаговременности разность $P'_{\text{опр}} - P$ близка к σ . По абсолютному значению разности повторяемостей оправдавшихся прогнозов для отдельных интервалов заблаговременности и по всей совокупности не превышают нескольких процентов. Для всей совокупности $P_{\text{опр}} = 71,7\%$, для самых неудачных прогнозов с трехлетней заблаговременностью $P'_{\text{опр}} = 69,3\%$ ($\Delta P_{\text{опр}} = -2,4\%$), для пятилетней заблаговременности $P'_{\text{опр}} = 73,5\%$ ($\Delta P_{\text{опр}} = 1,8\%$).

Несколько повышенное число оправдавшихся прогнозов при двух-, пяти- и шестилетней заблаговременности, возможно, связано с цикличностью атмосферных процессов. Квазидвухлетняя цикличность является наиболее распространенной в гидрометеорологических рядах. Циклы в 5—6 лет, близкие к циклам солнечной природы (кратные 11 годам), также обнаруживались неоднократно. Их анализ, в частности, дан в работах [7, 19].

Рассмотрим теперь повторяемость оправдавшихся прогнозов в зависимости от заблаговременности для разных групп предикторов (табл. 2, рис. 5). Отметим, что на рис. 5 оценка прогнозов коэффициентов разложения предиктантов сделана по совокупности данных для трех секторов наземного давления, а в табл. 2 — для каждого сектора отдельно. Анализ рис. 5 и табл. 2 показывает, что по числу оправдавшихся прогнозов некоторые группы предикторов при определенных интервалах заблаговременности резко выделяются. Так, при двухлетней заблаговременности из 68 прогнозов коэффициентов разложения предиктантов, составлен-

ных по значениям индекса зональной циркуляции, 80,9 % прогнозов оказались оправдавшимися. При использовании в качестве предиктора индекса K_p (62 прогноза) повторяемость оправдавшихся прогнозов составила 85,5 %. При пятилетней заблаговременности высокая оправдываемость прогнозов соответствует двум груп-

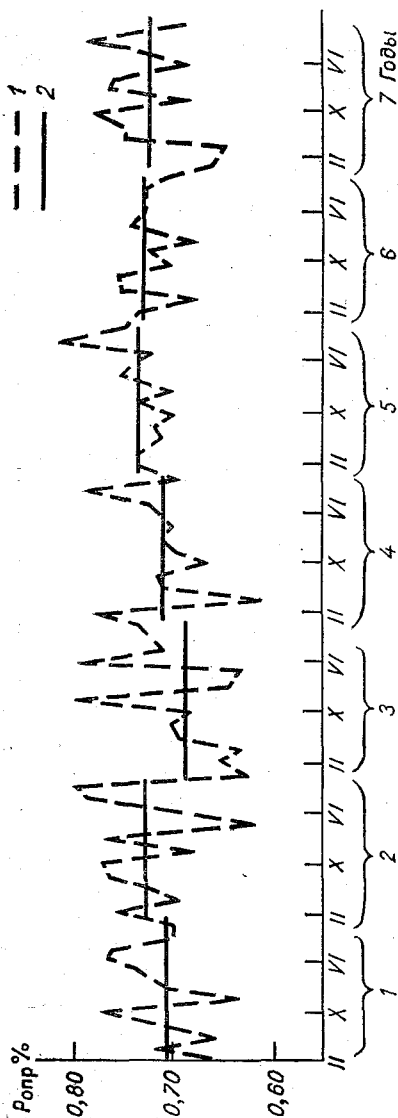


Рис. 4. Повторяемость (%) оправдавшихся прогнозов ($P_{\text{опр}}$) в зависимости от заблаговременности по месяцам (1) и средний уровень $P_{\text{опр}}$ по годам (2).

пам предикторов: индексу K_p ($P'_{\text{опр}} = 80,9 \%$) и числу дней с восточной (E) формой циркуляции ($P'_{\text{опр}} = 82,2 \%$).

Статистическая значимость $P'_{\text{опр}}$ по этим (и некоторым другим) группам предикторов оценивалась путем нахождения разностей $P'_{\text{опр}} - P_{\text{опр}}$ (где $P_{\text{опр}} = 71,7 \%$ — повторяемость оправдавшихся прогнозов по всей совокупности из 13978 прогнозов) и сравнения со средним квадратическим отклонением σ . Значение последнего, как следует из (3), существенно зависит от n_1 и n .

Будем считать разности $P'_{\text{опр}} - P_{\text{опр}} > 2\sigma$ статистически значимыми. Укажем значения $P'_{\text{опр}}$, соответствующие этому условию и приведенные в табл. 2. К таким группам предикторов относится при двухлетней заблаговременности индекс K_p ($\Delta P_{\text{опр}} = 2,4\sigma$), при пятилетней заблаговременности индекс K_p ($\Delta P_{\text{опр}} = 3,9\sigma$) и число дней с формой E ($\Delta P_{\text{опр}} = 2,2\sigma$), при шестилетней заблаговременности индекс K_p ($\Delta P_{\text{опр}} = 2,6\sigma$) и при семилетней заблаговременности давление II сектора ($\Delta P_{\text{опр}} = 2,9\sigma$). Обращает на себя внимание высокая информативность индекса геомагнитной возмущенности K_p при некоторых интервалах заблаговременности. Можно предположить, что в какой-то мере это вызвано сильной временной

корреляцией рядов индексов K_p за соседние месяцы. Поскольку связность предикторов в данном варианте расчетов не учитыва-

лась, число успешных и неуспешных прогнозов может искусственно завышаться или занижаться. Однако более глубокий анализ результатов показал, что при заблаговременности 5—6 лет устойчивую успешность прогнозов дали не средние месячные индексы K_p , а осредненные за неперекрывающиеся трехмесячные интервалы,

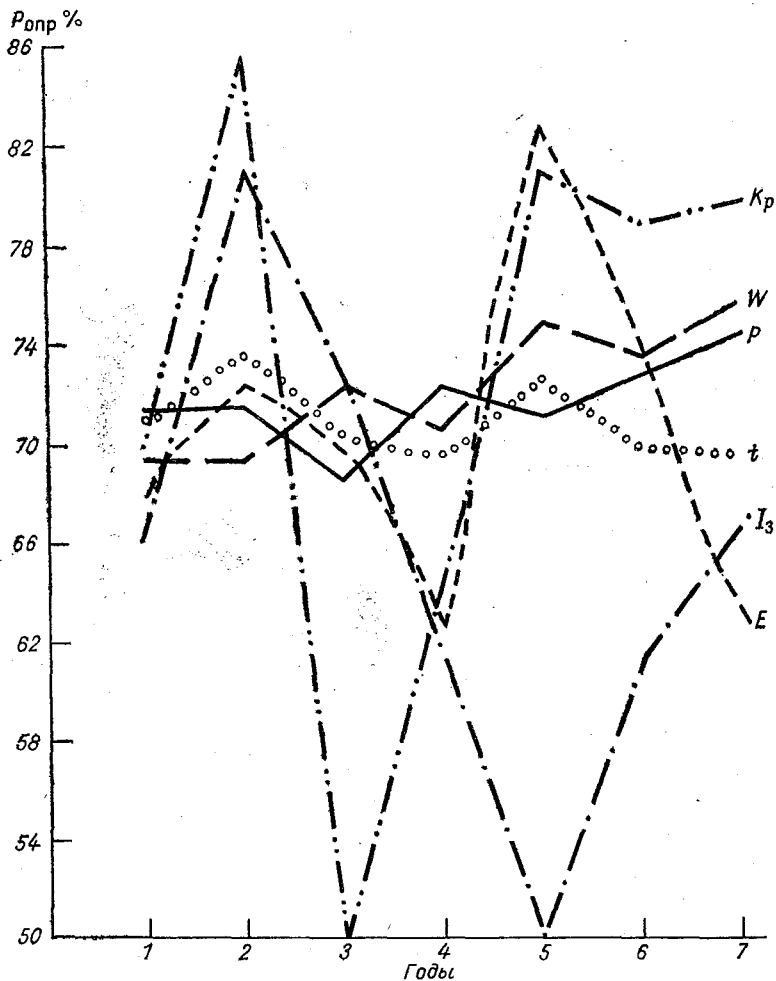


Рис. 5. Повторяемость оправдавшихся прогнозов ($P_{опр}$) в зависимости от заблаговременности для разных групп предикторов.
Усл. обозначения см. рис. 1.

Последнее обстоятельство, возможно, отражает установленное М. И. Юдиным и В. М. Мирвис [26] соотношение между интервалами осреднения предикторов и их заблаговременностью.

Индекс геомагнитной возмущенности с успехом использовался Т. В. Покровской [16] для прогноза засух с заблаговременностью

в 1—2 года. Расчеты, представленные в данной статье, показывают, что прогностические возможности этого индекса не исчерпаны.

Необходимо сопоставить информативность предикторов в зависимости от заблаговременности по критерию q и по успешности независимых 10-летних прогнозов, оцененных по принципам аналогичности (A).

Напомним, что по q выделились большей информативностью предикторы 5-, 6-летней и, особенно, годовой заблаговременности. Оценки независимых прогнозов по A подтвердили большую информативность предикторов пяти-, шестилетней заблаговременности. Информативность предикторов годовой заблаговременности (по A) оказалась даже ниже информативности всей совокупности предикторов. Таким образом, большая доля значимых корреляционных связей при той или иной заблаговременности еще не гарантирует большую успешность независимых прогнозов, составленных по предикторам той же заблаговременности (во всяком случае, для прогнозов на период с 1969 по 1978 г.).

Этот вывод подтверждается при разбиении заблаговременности по месяцам. При сравнении рис. 2 и 4 нетрудно заметить отсутствие четкой зависимости между q и $P_{\text{опр}}$. В частности, максимальное значение q , полученное по предикторам за январь текущего года ($q=7,3$) на материалах независимых прогнозов, не обеспечило превышение $P'_{\text{опр}}=71,9\%$ над средним уровнем $P_{\text{опр}}=71,7\%$ всей совокупности.

Сравнение рис. 3 и 5, где представлены оценки по q и $P'_{\text{опр}}$ разных групп предикторов в зависимости от заблаговременности, также свидетельствует об отсутствии однозначной согласованности этих двух оценок. Так, по числу дней с формой W выделилась большей долей значимых связей годовая заблаговременность (рис. 3), а в повторяемости оправдавшихся прогнозов этот максимум исчез (рис. 5). По числу дней с формой E максимальное значение q пришлось на двухлетнюю заблаговременность и минимальное на пятилетнюю (см. рис. 3). В повторяемости $P_{\text{опр}}$ для этой группы предикторов вторичный максимум при двухлетней заблаговременности сохранился, но основной максимум $P_{\text{опр}}$ пришелся на пятилетнюю заблаговременность, при которой значение q было минимальным (рис. 5) и т. д.

В связи с рассмотренными результатами расчетов встает вопрос о том, на какую оценку (q или $P_{\text{опр}}$, как частный случай A) следует ориентироваться в практической работе при составлении оперативных долгосрочных прогнозов погоды? Очевидно, целесообразнее руководствоваться, как это обычно принято, оценками независимых прогнозов по A или Y . Но при этом встает проблема устойчивости результатов экспериментов, т. е. вопрос о том, в какой мере выводы, полученные на материалах 10-летних независимых прогнозов за 1969—1978 гг., будут справедливы в последующие годы?

К сожалению, положительный ответ здесь не гарантирован, на что указывает статистические эксперименты, проведенные ранее

[14, 18]. В работе [18] представлены результаты статистических испытаний информативности предикторов на независимых прогнозах за два периода: 1965—1969 и 1970—1976 гг. Оказалось, что лишь некоторые группы предикторов (по длинным рядам E , W и высота снежного покрова) оказались устойчиво информативными (по Y) и в первом и во втором периоде.

Подводя итоги, можно отметить следующее.

1. Рассчитанная по длинным рядам данных (70—90 лет) и на большом статистическом материале ($425,6 \cdot 10^3$ корреляционных связей) доля q значимых корреляционных связей, превышающих заданный уровень, имеет основной максимум для предикторов годовой заблаговременности (ближняя предыстория) и вторичный максимум для предикторов пяти-, шестилетней заблаговременности.

2. Большая (меньшая) доля значимых связей (q) не всегда обеспечивает большую (меньшую) успешность прогнозов коэффициентов разложения предиктантов на независимом материале.

3. По материалам независимых прогнозов за 10 лет (1969—1978 гг.) не выявлено уменьшения успешности прогнозов с ростом заблаговременности. Установлена несколько более высокая успешность прогнозов коэффициентов разложения полей осадков и температуры по совокупности всех предикторов с заблаговременностью в 2 года и 5—6 лет. Однако статистически значимой является лишь более высокая успешность прогнозов с заблаговременностью 5 лет.

4. На материалах независимых прогнозов установлено, что при заблаговременности 2 года и 5—6 лет значимо информативен индекс K_p . Доля оправдавшихся прогнозов ($P_{\text{опр}}$) по K_p при этих интервалах заблаговременности составила соответственно 85,5; 80,9 и 78,9 %. Значимо информативными оказались также число дней с формой E при пятилетней заблаговременности ($P_{\text{опр}}=82,2$ %) и давление Американского сектора при семилетней заблаговременности ($P_{\text{опр}}=79$ %). Эти выводы могут быть использованы при составлении как долгосрочных, так и сверхдолгосрочных метеорологических прогнозов.

5. Необходимо продолжение статистических экспериментов для решения вопроса об устойчивости полученных выводов на материалах независимых прогнозов за другие периоды.

СПИСОК ЛИТЕРАТУРЫ

1. Борисенков Е. П., Борисова Л. Е. Физико-статистическая схема предвычисления средних месячных аномалий температуры воздуха на полушарии. — Труды ААНИИ, 1967, т. 275, с. 310—318.

2. Борисенков Е. П., Борисова Л. Е. Уточненная физико-статистическая схема сверхдолгосрочного прогноза средних месячных аномалий температуры воздуха для северного полушария. — Труды ГГО, 1981, вып. 443, с. 3—11.

3. Воробьева Е. В. Некоторые пути увеличения заблаговременности месячных и сезонных прогнозов погоды. — Труды ГГО, 1975, вып. 354, с. 23—29.

4. Воробьева Е. В., Приемова Л. П. Информативность некоторых характеристик атмосферы для сверхдолгосрочного прогнозирования. — Труды ГГО, 1979, вып. 428, с. 102—110.
5. Гирс А. А. Макроциркуляционный метод долгосрочных метеорологических прогнозов. — Л.: Гидрометеиздат, 1974, — 486 с.
6. Дроздов О. А. О возможности климатических прогнозов на основании учета цикличности, обусловленной космическими и земными факторами. — Труды ГГО, 1971, вып. 274, с. 3—25.
7. Дроздов О. А., Григорьева А. С. Многолетние циклические колебания атмосферных осадков на территории СССР. — Л.: Гидрометеиздат, 1971. — 157 с.
8. Дроздов О. А., Полозова Л. Г., Рубинштейн Е. С. Прогноз многолетнего хода температуры воздуха и осадков по территории СССР на 2—3 десятилетия. — В кн.: Климатология и сверхдолгосрочный прогноз. — Л.: Гидрометеиздат, 1977, с. 10—19.
9. Житорчук Ю. В. О предсказуемости крупномасштабных атмосферных процессов с большой заблаговременностью. — Труды ГГО, 1981, вып. 446, с. 82—93.
10. Кац А. Л. Сезонные изменения общей циркуляции атмосферы и долгосрочные прогнозы. — Л.: 1960 — 270 с.
11. Мандель Д. Г. Интенсивность зональной циркуляции в Атлантико-Евразийском секторе и связь ее с колебаниями температуры воздуха в умеренных широтах Евразии. Автореф. дисс. на соискание учен. степен канд. геогр. наук, Ленинград, 1967. — 14 с.
12. Мещерская А. В. Опыт сверхдолгосрочного прогноза температуры и осадков для юга ЕТС и Северного Казахстана физико-статистическим методом. — В кн.: Климатология и сверхдолгосрочный прогноз. — Л., 1977, с. 26—33.
13. Мещерская А. В., Блажевич В. Г. К оценке некоторых методических положений физико-статистического метода. — Труды ГГО, 1975, вып. 329, с. 54—65.
14. Мещерская А. В., Голод М. П. Статистический эксперимент по проверке целесообразности учета трех форм циркуляции Г. Я. Вангенгейма в физико-статистических долгосрочных прогнозах осадков и температуры. — Труды ГГО, 1981, вып. 446, с. 24—34.
15. Мещерская А. В., Леднева К. В., Блажевич В. Г. Характеристика дополнительной гидрометеорологической информации, используемой в физико-статистическом прогнозе. — Труды ГГО, 1975, вып. 353, с. 115—122.
16. Покровская Т. В. Синоптико-климатологические и гелиогеофизические долгосрочные прогнозы погоды. — Л.: Гидрометеиздат, 1969. — 269 с.
17. Полозова Л. Г. Предполагаемый знак отклонения от нормы средней месячной температуры воздуха в период 1976—1980 гг. на западной половине территории СССР. — Труды ГГО, 1975, вып. 354, с. 137—143.
18. Предварительные выводы об информативности прогностических соотношений Юдин М. И., Блажевич В. Г., Голод М. П., Чувашина И. Е. — Труды ГГО, 1981, вып. 446, с. 12—23.
19. Сазонов Б. И. К прогнозированию экстремальных погодных явлений. — Труды ГГО, 1981, вып. 443, с. 60—63.
20. Юдин М. И., Блажевич В. Г. Оценка значимости предикторов на основе комплексного статистического испытания. — Труды ГГО, 1975, вып. 329, с. 41—53.
21. Юдин М. И., Голод М. П. Применение теории статистических решений к сверхдолгосрочному прогнозу осадков. — Труды ГГО, 1975, вып. 329, с. 74—80.
22. Юдин М. И., Мещерская А. В. Результаты применения физико-статистического метода прогноза осадков и температуры с большой заблаговременностью. — Труды V Всесоюзного метеорологического съезда, том. II, 1972, с. 83—96.
23. Юдин М. И., Мещерская А. В. Некоторые оценки естественных составляющих как предикторов и предиктантов. — Труды ГГО, 1972, вып. 273, с. 3—15.

24. Юдин М. И., Мещерская А. В., Блажевич В. Г. Характеристики гидрометеорологической информации, использованной при долгосрочном физико-статистическом прогнозе температуры и осадков для районов недостаточного увлажнения. — Труды ГГО, 1969, вып. 236, с. 45—63.

25. Юдин М. И., Мещерская А. В. Комплексный физико-статистический метод прогноза большой заблаговременности. — Метеорология и гидрология, 1977, № 1, с. 3—12.

26. Юдин М. И., Мирвис В. М. Применение информационного подхода к задаче долгосрочного прогноза. — Труды ГГО, 1981, вып. 446, с. 3—11.

27. Юл Д. Ж. Эдни, Кендэл М. Дж. Теория статистик. — М.: Госстатиздат, ЦСУ СССР, 1960, с. 778.

А. Т. Житорчук, Ю. В. Житорчук

К ВОПРОСУ О СМЕЩЕННОСТИ СПЕКТРА СОБСТВЕННЫХ ЧИСЕЛ ВЫБОРОЧНЫХ КОРРЕЛЯЦИОННЫХ МАТРИЦ

Спектр собственных чисел выборочной корреляционной матрицы является смещенной оценкой соответствующего истинного спектра. Первые собственные числа завышены, а последние занижены по сравнению с их истинными значениями.

Следовательно,

$$C_i^{(n)} = M \{ \lambda_i^{(n)} \} - \lambda_i^{(\infty)} \neq 0, \quad (1)$$

где $M \{ \lambda_i^{(n)} \}$ — математическое ожидание i -го собственного числа, рассчитанного по выборке объема n ; $C_i^{(n)}$ — смещение i -го собственного числа, рассчитанного по выборке объемом n .

Несмотря на то что с качественной точки зрения картина смещенности спектров собственных чисел выборочных корреляционных матриц довольно ясна, аналитическая оценка этого эффекта является сложной задачей. Поэтому для ее решения в настоящей статье будет использован метод статистических испытаний.

С помощью датчика случайных чисел генерировался марковский процесс первого порядка

$$x_{t+1} = \beta x_t + \xi_{t+1}. \quad (2)$$

Для расчета автокорреляционной матрицы порядка m использовалась выборка объемом $n+m$, где n — количество членов, по которым рассчитывался каждый из входящих в матрицу коэффициентов корреляции. Далее, с помощью стандартной программы, имеющейся в составе математического обеспечения ЭВМ, рассчитывались m собственных чисел автокорреляционной матрицы, и вся процедура повторялась l раз. В результате были получены осредненные по l реализациям оценки математического ожидания собственных чисел.

Для сравнения был рассчитан истинный спектр собственных чисел. Для этого использовалось известное выражение для автокорреляционной функции марковского процесса первого порядка

$$R(\tau) = \beta^\tau. \quad (3)$$

Сравнение осредненных спектров собственных чисел, полученных при различных длинах выборки (рис. 1), выявляет существенные систематические ошибки. Так, например, при $\beta=0,2$ даже при длине выборки $n=100$ первое собственное число оказывается завышенным по сравнению с его истинным значением в 2 раза, а

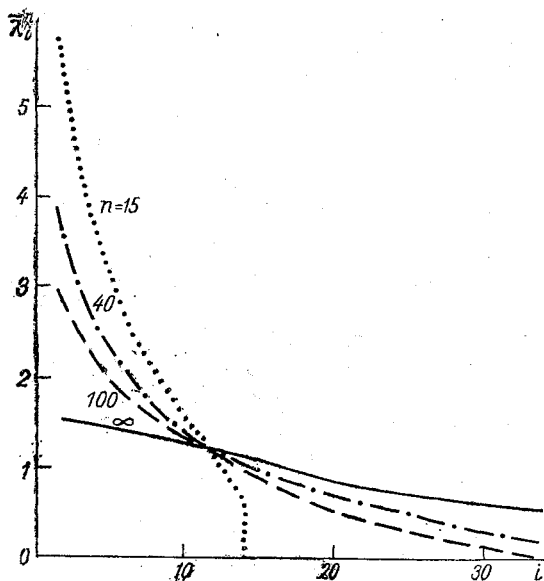


Рис. 1. Осредненные по 100 реализациям спектры собственных чисел выборочных автокорреляционных матриц при различных длинах n выборки.

последнее — заниженным в 2,5 раза. Таким образом, в этом случае число Тодда (λ_1/λ_m), характеризующее степень обусловленности матрицы, завышается в 5 раз.

При исключении смещенности спектра собственных чисел улучшается обусловленность корреляционной матрицы, а следовательно, увеличивается точность регрессионного анализа. Для того чтобы реализовать такой подход необходимо построить модель, описывающую зависимость смещения λ_i от объема выборки.

На большом материале при различных значениях порядка матрицы, коэффициента β , а также номера собственного числа было установлено, что зависимость $\ln(C_i^{(n)})$ от $\ln n$ носит линейный характер (рис. 2) и, следовательно, может быть удовлетворительно описана степенным уравнением

$$C_i^{(n)} = \alpha_i n^{-\gamma_i}, \quad \gamma_i > 0, \quad (4)$$

где α_i и γ_i — неизвестные пока параметры.

Поскольку сумма собственных чисел корреляционной матрицы независимо от n равна ее следу, то

$$\sum_{i=1}^m C_i^{(n)} = \sum_{i=1}^m \alpha_i n^{-\gamma_i} = 0. \quad (5)$$

Равенство (5) выполняется лишь при условии, что $\gamma_1 = \gamma_2 = \dots =$

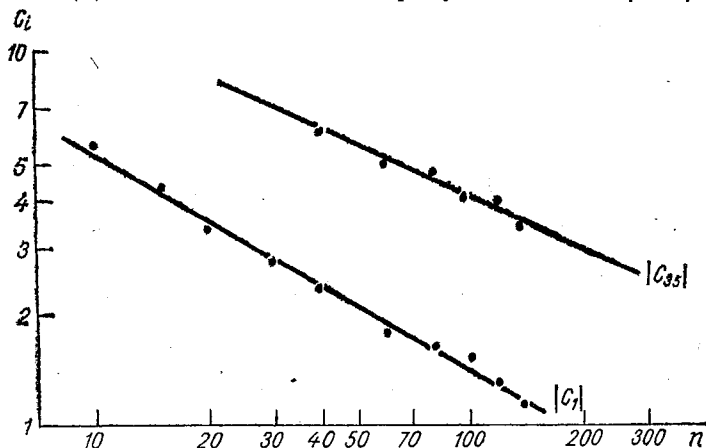


Рис. 2. Зависимость абсолютной величины смещенности первого (C_1) и последнего (C_{35}) собственных чисел от объема выборки.

$= \gamma_m = \gamma$. С учетом этого несмещенная оценка спектра собственных чисел имеет вид

$$\hat{\lambda}_i^{(\infty)} = \lambda_i^{(n)} - \alpha_i n^{-\gamma}. \quad (6)$$

Рассмотрим теперь методику оценки параметров модели (6). Пусть имеется k спектров собственных чисел корреляционных матриц, рассчитанных по данным выборок объемом $n_1 < n_2 < \dots < n_k$. Введя обозначение $z_j = n_j^{-\gamma}$, запишем (6) в виде линейного уравнения

$$\lambda_i^{(n_j)} = \lambda_i^{(\infty)} + \alpha_i n_j^{-\gamma} + \varepsilon = \lambda_i^{(\infty)} + \alpha_i z_j + \varepsilon, \quad (7)$$

где ε — случайная ошибка $\lambda_i^{(n_j)}$.

Далее, задавая то или иное значение γ при помощи методов наименьших квадратов, рассчитаем соответствующие ему значения параметров:

$$\alpha_i = \frac{\sum_{j=1}^k (\lambda_i^{(n_j)} z_j) - \sum_{j=1}^k \lambda_i^{(n_j)} \sum_{j=1}^k z_j}{\sum_{j=1}^k z_j^2 - \left(\sum_{j=1}^k z_j \right)^2}, \quad (8)$$

$$\hat{\lambda}_i^{(\infty)} = \frac{1}{k} \sum_{j=1}^k (\lambda_i^{(n_j)} - \alpha_i n_j^{-\gamma}), \quad (9)$$

а также оценку точности модели, соответствующую выбранному значению γ :

$$f(\gamma) = \sum_{i=1}^m \sum_{j=1}^k \frac{(\lambda_i^{(n,j)} - \hat{\lambda}_i^{(\infty)} - \alpha_i n_j^{-\gamma})^2}{(\lambda_i^{(n,j)})^2}. \quad (10)$$

Методом направленного поиска выбирается такое значение γ , при котором наблюдается $\min \{f(\gamma)\}$, а соответствующие этому значению γ оценки $\hat{\lambda}_i^{(\infty)}$ принимаются за несмещенную оценку спектра собственных чисел корреляционной матрицы.

Заметим, что уравнение (6) описывает лишь основные черты зависимости смещенности собственных чисел от объема выборки. В то же время некоторые экспериментальные факторы не могут быть описаны в рамках этой модели. Прежде всего это относится к тому, что для собственных чисел, значение которых близко к единице, зависимость λ_i от n не является монотонно возрастающей или убывающей функцией. Кроме того, в зависимости от номера собственного числа наблюдается некоторое изменение величины γ .

Для того чтобы описать эти эффекты, необходимо ввести в (6) по крайней мере еще один член

$$\lambda_i^{(n)} = \lambda_i^{(\infty)} + \alpha_{1i} n^{-\gamma_1} + \alpha_{2i} n^{-\gamma_2}. \quad (11)$$

Однако оценка параметров, входящих в (11), весьма сложна, а их надежность невелика. Поэтому для коррекции уравнения (6) вблизи $\lambda_i^{(n)} \approx 1$ был введен следующий фильтр:

$$\alpha_i^* = \alpha_i [1 - e^{-(3\lambda_i^{(n)} - 3)^2}] \text{sign}(\alpha) \text{sign}(\lambda_i^{(n)} - 1). \quad (12)$$

Эксперименты по несмещенной оценке спектра собственных чисел показали, что значение $\gamma_{\text{опт}}$ лежит в пределах от 0,3 до 1,0.

На рис. 3 а представлены результаты расчетов исходного, несмещенного и истинного спектров собственных чисел взаимокорреляционной матрицы белого шума при объеме выборки $n=87$. Сравнение этих кривых показывает, что процедура коррекции спектра $\lambda_i^{(n)}$ позволяет существенно уменьшить смещение. В то же время наблюдается недостаточная коррекция последних собственных чисел. В отличие от исходного в скорректированном спектре появляются локальные максимумы, обусловленные статистической погрешностью параметров, входящих в (6).

Кроме анализа модельных рядов, представляет интерес оценка эффектов смещенности собственных чисел реальных атмосферных процессов. На рис. 3б представлены спектры собственных чисел корреляционной матрицы месячных сумм апрельских осадков для 35 районов ЕТС при длине выборки $n=87$.

В этом случае смещенность первых собственных чисел незначительна и составляет примерно 10%. В то же время высокие номера $\lambda_i^{(87)}$ занижены по крайней мере в 3—4 раза.

Рассмотрим теперь использование процедуры несмещенной оценки в задаче регрессионного анализа. Для этого воспользуем-

ся спектральным разложением матрицы по главным компонентам

$$\bar{\Sigma}_{xx} = \sum_{i=1}^m \lambda_i^{(n)} v_i v_i', \quad (13)$$

где $\bar{\Sigma}_{xx}$ — корреляционная матрица предикторов порядка $m \times m$
 v_i — i -ый собственный вектор матрицы $\bar{\Sigma}_{xx}$.

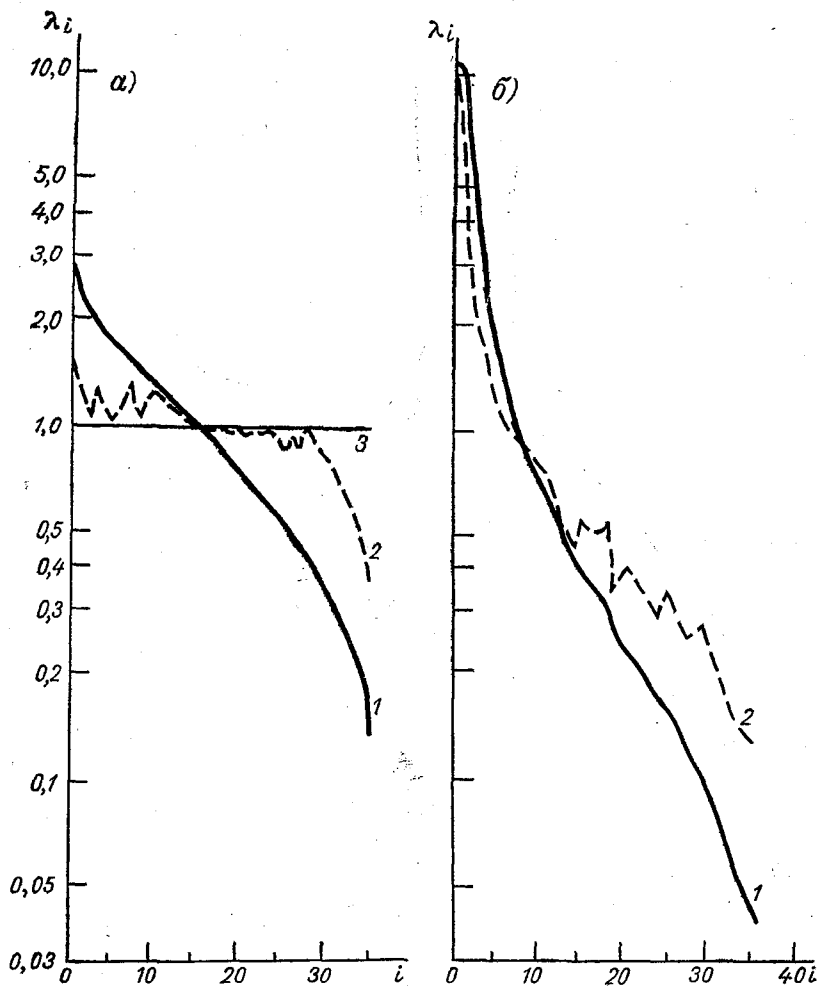


Рис. 3. Спектры собственных чисел взаимокорреляционных матриц случайных процессов типа белого шума (а) и месячных сумм осадков на ЕТС (б).

1 — выборочные спектры, 2 — истинные спектры, 3 — несмещенная оценка λ_i .

При исключении смещенности собственных чисел корректированная корреляционная матрица может быть записана в следующем виде:

$$\begin{aligned} \hat{\Sigma}_{xx} &= \sum_{i=1}^m (\lambda_i^{(n)} - C_i^{(n)}) v_i v_i' = \sum_{i=1}^m \lambda_i^{(n)} v_i v_i' + \\ &+ \sum_{i=1}^m C_i^{(n)} v_i v_i' = \bar{\Sigma}_{xx} - C, \end{aligned} \quad (14)$$

где C — корректирующая матрица порядка $m \times m$.

Поскольку $C_i^{(n)}$ представляют из себя собственные числа матрицы C , то, следовательно,

$$\text{tr}(C) = 0, \quad (15)$$

а сама матрица является знаконеопределенной.

Коэффициенты регрессии в этом случае определяются как

$$a(C) = (\bar{\Sigma}_{xx} - C)^{-1} Q_{xy}, \quad (15)$$

где Q_{xy} — вектор коэффициентов корреляции между предикторами и предиктантом.

Нетрудно видеть связь оценки (16) с известными методами гребневой оценки [1] и методов регуляризации Тихонова [2].

Действительно, гребневая оценка коэффициентов регрессии записывается в общем виде как

$$a(K) = (\bar{\Sigma}_{xx} + K)^{-1} Q_{xy}, \quad (17)$$

где K — некоторая неотрицательно определенная матрица порядка $m \times m$.

Суть оценки (17) ясна: добавление к корреляционной матрице неотрицательно определенной матрицы K делает ее лучше определенной, а оценки более устойчивыми.

В то же время полученные таким образом коэффициенты регрессии являются смещенными и существенно зависят от произвола, связанного с выбором вида и параметров матрицы K . Эти недостатки исключаются при использовании предлагаемой методики, поскольку улучшение степени обусловленности корреляционной матрицы достигается за счет использования в качестве естественного стабилизатора знаконеопределенной матрицы, след которой равен нулю. Именно за счет этого свойства устойчивость оценок достигается без внесения в их значения какой-либо дополнительной смещенности.

Можно предположить, что благодаря своим фильтрующим свойствам естественные стабилизаторы найдут применение в спектральном и дискриминантном анализе, а также в методе канонических корреляций.

СПИСОК ЛИТЕРАТУРЫ

1. Демиденко Е. З. Линейная и нелинейная регрессии. — М.: Финансы и статистика, 1981. — 302 с.
2. Василенко Г. И. Теория восстановления сигналов. — М.: Советское радио, 1979. — 269 с.

СТАТИСТИЧЕСКИЕ СВЯЗИ МЕЖДУ ИНДЕКСОМ ИНТЕНСИВНОСТИ ЗОНАЛЬНОЙ ЦИРКУЛЯЦИИ И НЕКОТОРЫМИ МЕТЕОРОЛОГИЧЕСКИМИ ПАРАМЕТРАМИ

В настоящей работе рассматривается возможность установления статистических связей между градуированными значениями различных метеорологических параметров и индексом зональной циркуляции (I_3) по Е. Н. Блиновой [1]. Ранее при разработке метода прогноза аномалий средней декадной температуры ряда значений некоторых параметров с 1949 по 1977 г. (отдельно для весны и осени) были разбиты на четыре типа: I тип — значительно больше нормы, II — больше нормы, III — меньше нормы, IV — значительно меньше нормы. В результате для индексов I_3 за весенние месяцы (апрель) получено:

Тип	I	II	III	IV
I_3	$\geq 4,3$	1,0...4,2	0,9...-4,0	$\leq -4,1$

Для индексов I_3 за осенние месяцы (октябрь) получено:

Тип	I	II	III	IV
I_3	$\geq 3,0$	0,1...2,9	0...-3,2	$\leq -3,3$

Такой же принцип разбивки по четырем типам был применен и для других параметров.

Разбивка по четырем типам была произведена для следующих метеоэлементов: температуры воздуха, общей облачности, границы снежного покрова, а также для меридионального и гидродинамического индексов циркуляции атмосферы. В дальнейшем будут приведены данные только для I и IV типов.

Температура воздуха и общая облачность были представлены рядами данных в 42 узлах сетки. Разбиение на типы сделано отдельно для каждого узла. На рис. 1 представлено распределение по рассматриваемой территории граничных значений I, IV типов по температуре для весны (а, б) и осени (в, г). Аналогичные карты были построены для общей облачности.

Из анализа этих карт можно сделать вывод, что типы I и IV для температуры (а также и для облачности) существенно различаются. Следует иметь в виду, что перед значениями температуры на картах при типе I должен стоять знак больше, а при типе IV следует ставить знак меньше. Положение границы снежного покрова было взято от 30 до 60° в. д. (через 5° широты).

В табл. 1 представлены граничные значения аномалий положения границы снежного покрова для I и IV типов в апреле и октябре, рассчитанные по исходным рядам за период с 1947 по 1977 г.

Из табл. 1 видно, что первый тип соответствует значительным положительным аномалиям в положении границы снежного покрова, четвертый тип — отрицательным.

В качестве меридионального индекса (I_M) был использован индекс, введенный М. И. Юдиным и А. А. Рождественским, который характеризует тепловые различия между океанами и материками [3]. Под гидродинамическим индексом (I_i) следует понимать индекс, пропорциональный циркуляции скорости по замкнутому контуру [4].

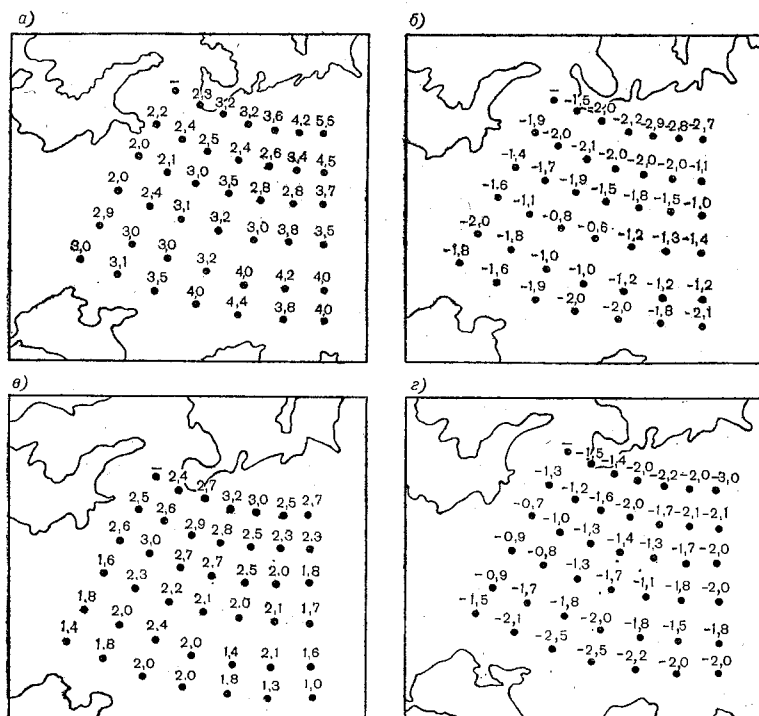


Рис. 1. Распределение граничных значений аномалий средней декадной температуры для I типа (а, в) и IV типа (б, г) за апрель (а, б) и октябрь (в, г).

а, в — положительные аномалии, б, г — отрицательные аномалии.

В табл. 2 приведены граничные значения аномалии индексов I_M и I_i .

После того как было выяснено, какие аномалии соответствуют I и IV типам рассмотренных выше величин, нужно было установить, имеется ли какая-либо связь между ними и типами I_3 .

Из табл. 3 видно, что при различных типах I_3 (I и IV) наблюдаются определенные типы рассмотренных нами метеоэлементов. Например, I типу I_3 весной соответствуют положительные аномалии облачности (I, II типы), отрицательные аномалии температуры, границы снежного покрова, меридионального и гидродинамического индексов (III, IV типы). При повышенной интенсивности

Таблица 1

Граничные значения аномалии положения границы снежного покрова
в градусах широты при I и IV типах

Тип	λ°						
	30	35	40	45	50	55	60
Апрель							
I IV	>3 <-2	>3 <-4	>4 <-3	>3 <-3	>4 <-2	>3 <-3	>3 <-3
Октябрь							
I IV	>5 <-4	>4 <-4	>5 <-5	>5 <-4	>4 <-4	>5 <-5	>5 <-4

Таблица 2

Граничные значения аномалии индексов I_M и I_i для
весны и осени при I и IV типах

Индекс	Тип	Месяц	
		Апрель	Октябрь
I_M	I	$>1,5$	$>2,8$
	IV	$<-1,6$	$<-2,7$
I_i	I	$>31,3$	$>32,4$
	IV	$<-31,2$	$<-32,8$

Таблица 3

Связь между типами I_3 и некоторыми параметрами

Тип	Параметр				
	облачность	температура	граница снежного покрова	I_i	I_M
Апрель					
I IV	I, II III, IV	III, IV III, IV	III, IV I, II	III, IV I, II	III, IV I
Октябрь					
I IV	I, II II, III	I, II III, IV	III, IV I, II	III, IV I, II	III, IV I, II

зональной циркуляции следует ожидать повышенной облачности, низких температур, ослабления меридионального обмена и циркуляции по замкнутому контуру, а также смещения к северу границы снежного покрова. Таким образом, имея значения ожидаемого индекса интенсивности зональной циркуляции, можно судить о преобладающем типе аномалий температуры, облачности и границы снежного покрова в следующей декаде (см. табл. 3).

Значения ожидаемого индекса интенсивности зональной циркуляции атмосферы можно вычислить по прогностическим картам (на 72 ч), составляемыми в Национальном метеорологическом центре США под руководством Шумана.

СПИСОК ЛИТЕРАТУРЫ

1. Блинова Е. Н. Гидродинамическая теория волн давления, температурных волн и центров действия атмосферы. — ДАН СССР, 1943, т. 39, № 3, с. 284—287.
2. Есакова Н. П., Афанасьева В. Б. О способах характеристики аномалий облачности, снежного покрова и радиационных потоков. — Труды ГГО, 1962, вып. 143, с. 104—112.
3. Рождественский А. А. Исследование синоптических процессов, связанных с осуществлением карских воздействий на основе изучения процессов циркуляции. — Труды ГГО, 1955, вып. 55, с. 36—50.
4. Богданова Н. П. Изучение барико-циркуляционного режима европейской территории СССР для осеннего периода. — Труды ГГО, 1960, вып. 114, с. 114—120.

Ю. М. Либман

О НЕКОТОРЫХ ОСОБЕННОСТЯХ ОПТИМАЛЬНОЙ ИНТЕРПОЛЯЦИИ ПРИМЕНИТЕЛЬНО К ЗАДАЧАМ УСВОЕНИЯ МЕТЕОРОЛОГИЧЕСКИХ ДАННЫХ

Метод оптимальной интерполяции уже давно применяется в СССР в оперативной практике объективного анализа метеорологических полей [1]. В последние годы этот метод получил развитие во многих прогностических центрах мира [7—10]. Он широко используется, в частности, для усвоения данных Глобального метеорологического эксперимента.

В связи с этим целесообразно обратить внимание на некоторые особенности метода оптимальной интерполяции, недостаточный учет которых может приводить к существенным трудностям при применении этого метода для усвоения данных метеорологических наблюдений. Мы будем иллюстрировать рассматриваемые особенности простыми примерами, относящимися к двумерной одноэлементной оптимальной интерполяции. Разумеется, те же особенности могут проявляться и в более общих случаях многоэлементной и трех- или четырехмерной интерполяции.

Ради удобства будем считать, что интерполяции подвергаются отклонения f' метеорологического параметра f от его среднего значения \bar{f} , деленные на средние квадратические значения σ этих отклонений, т. е. что интерполяция производится по формуле

$$\hat{f}'_0 = \sigma_0 \sum_{i=1}^n p_i \bar{f}'_i / \sigma_i, \quad (1)$$

где \hat{f}'_0 — интерполированное значение f' в интересующей нас точке с номером 0, \bar{f}'_i — наблюдаемые значения f' , а p_i — весовые множители, определяемые из условия минимума величины

$$\varepsilon^2 = \overline{(\hat{f}'_0 - f'_0)^2} / \sigma_0^2. \quad (2)$$

Осреднение в формуле (2) производится по некоторому ансамблю, который может быть различным. Например, если это ансамбль всех возможных в данный месяц значений, то \bar{f} представляет собой среднее в климатическом смысле (норму), а если имеется в виду ансамбль возможных значений при данном численном прогнозе на рассматриваемый момент времени, то \bar{f} является прогностическим полем. Последующие формулы верны безотносительно к спецификации ансамбля осреднения. От нее зависит лишь конкретный вид корреляционной функции

$$\mu_{ij} = \overline{f'_i f'_j} / (\sigma_i \sigma_j). \quad (3)$$

Будем считать, что функция μ зависит лишь от расстояния r_{ij} между точками (гипотеза однородности и изотропии). Что касается случайных ошибок наблюдений

$$\delta_i = \bar{f}_i - f_i, \quad (4)$$

то их статистическими характеристиками являются аналогичная ε^2 величина

$$\eta_i^2 = \overline{\delta_i^2} / \sigma_i^2, \quad (5)$$

именуемая мерой ошибок наблюдений, и коэффициенты корреляции μ_{ij} между ошибками в различных точках.

Если надлежащим образом отбирать «влияющие» пункты наблюдений для оптимальной интерполяции в каждую точку сетки (т. е. брать пункты, равномерно окружающие точку со всех сторон), то обычно достаточно весьма небольшого их числа n : даже теоретические оценки (расчет ε) показывают, что увеличение n сверх 5—6 не приводит к заметному повышению точности. Однако организовать достаточно совершенный поиск влияющих пунктов непросто. Далее, при малых n в районах с редкой сетью наблюдений существует опасность получения фиктивно больших градиентов вследствие изменения набора «влияющих» пунктов при переходе от одной точки сетки к другой. Поэтому существует тенденция к использованию значительно больших значений n . Однако при этом возникает другая опасность, связанная с плохой обусловленностью систем уравнений для определения весов p_i .

Она проявляется в виде так называемого эффекта экранирования. Этот эффект может иметь место при наличии двух и более станций, расположенных по одну сторону от узла, в который производится интерполяция. В таком случае ближайшие к узлу станции как бы заслоняют собой дальние и их весовые множители

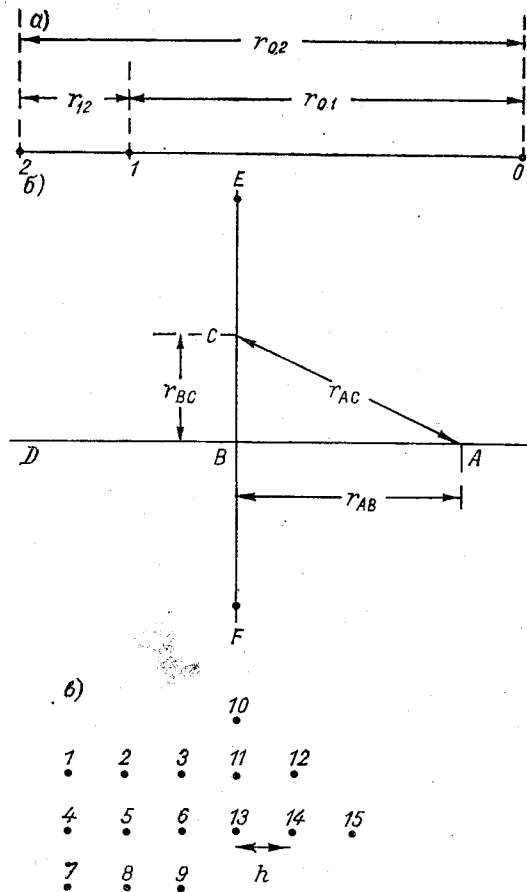


Рис. 1. Схемы расположения пунктов.

оказываются существенно выше. Более того, множители дальних станций могут быть отрицательными и притом значительными по величине.

Такое распределение весов не приведет к неприятностям, если, как это обычно бывает, значения f_1 и f_2 на близких друг к другу станциях не очень отличаются. Однако в тех редких случаях, когда различия между f_1 и f_2 заметны, экранирование может быть причиной неправильных результатов анализа. Именно эта редкость проявления составляет опасную особенность эффекта экра-

нирования. Наиболее отчетливо экранирование сказывается применительно к гауссовской корреляционной функции

$$\mu(r) = \exp(-r^2/a^2), \quad (6)$$

где a — радиус корреляции. Такая функция нередко используется в работах зарубежных авторов [6, 7, 10].

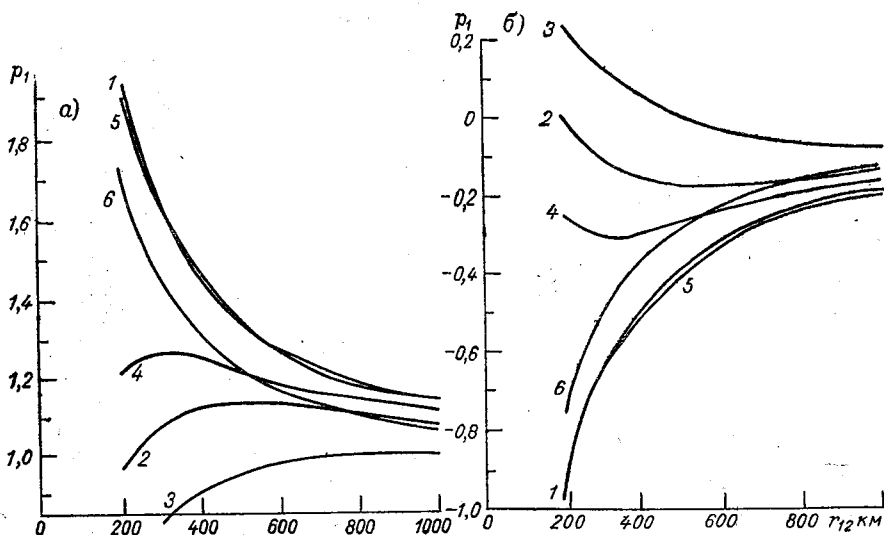


Рис. 2. Зависимости весов p_1 (а) и p_2 (б) от расстояния r_{12} .

Проиллюстрируем эффект экранирования на простейшем примере одномерной интерполяции в узел 0 из пунктов 1 и 2 (рис. 1 а). Уравнения для определения весов, вытекающие из требования минимума величины ε^2 , имеют в этом случае вид

$$\begin{aligned} (1 + \eta^2_1) p_1 + [\mu(r_{12}) + \eta_1 \eta_2 \nu(r_{12})] p_2 &= \mu(r_{01}), \\ [\mu(r_{12}) + \eta_1 \eta_2 \nu(r_{12})] p_1 + (1 + \eta^2_2) p_2 &= \mu(r_{02}). \end{aligned} \quad (7)$$

На рис. 2 представлены зависимости p_1 и p_2 от r_{12} , рассчитанные по формулам (6) и (7) при $a=2000$ км и $r=400$ км. Кривые 1 на рис. 2 а, б соответствуют случаю $\eta=0$. Мы видим, что p_1 значительно превосходит единицу, а p_2 существенно отрицателен, причем этот эффект тем сильнее, чем меньше r_{12} . Кривые 2 и 3 соответствуют случаям $\eta^2=0,02$ и $\eta^2=0,05$ при $\nu(r_{12})=0$. Они показывают, что учет случайных ошибок наблюдений в предположении некоррелированности последних позволяет значительно снизить эффект экранирования. В частности, с уменьшением r_{12} веса не растут безгранично по абсолютному значению, а приближаются к одному и тому же значению $\mu(r_{01})/(2+\eta^2)$.

Однако учет коррелированности ошибок, характерной, как известно, для данных косвенного зондирования атмосферы, снова

приводит к росту влияния экранирования. Это видно из кривых 4 и особенно 5, рассчитанных соответственно при $\nu(r_{12})=0,5$ и $\nu(r_{12})=1$ ($\eta^2=0,02$).

Другой способ воздействия на эффект экранирования связан с выбором более реалистической аппроксимации корреляционной функции $\mu(r)$, например, с переходом от формулы (6) к предложенной М. И. Юдиным [4] формуле

$$\mu(r) = (1 + r/b) \exp(-r/b) \quad (8)$$

или к одной из формул, рекомендованных Тьебо [11]. Влияние этого выбора на проявление экранирования может быть проиллюстрировано кривыми 6 на рис. 2, построенными для случая $\eta=0$ при использовании аппроксимации (8) с $b=900$ км.

Наиболее радикальный способ уменьшения опасности проявления эффекта экранирования состоит в использовании надлежащего алгоритма поиска влияющих пунктов наблюдений. Этот алгоритм должен обеспечивать не только близость влияющих пунктов к точке сетки, но и ее сравнительно равномерное окружение со всех сторон. Можно, например, разделить окрестность точки сетки на четыре квадранта и требовать, чтобы из каждого квадранта бралось не более и, по возможности, не менее одного влияющего пункта. Этот подход, применимый для одноэлементной двумерной интерполяции, нетрудно обобщить на трехмерную интерполяцию. Аналогичный подход целесообразно разработать и для многоэлементной интерполяции. Как показывает опыт, дополнительные затраты машинного времени, необходимые при усложненном поиске влияющих данных, будут более чем компенсированы экономией времени благодаря уменьшению числа влияющих данных [2, 3].

Еще более наглядно эффект экранирования при использовании корреляционной функции (6) проявляется в двумерном случае.

Рассмотрим задачу определения значения параметра в пункте A на рис. 1 б по его точным значениям в пунктах B и C (прямые AD и EF перпендикулярны) в случае корреляционной функции (6). Веса p_B и p_C определяются из аналогичных (7) уравнений

$$\begin{aligned} p_B + \mu(r_{BC}) p_C &= \mu(r_{AB}), \\ \mu(r_{BC}) p_B + p_C &= \mu(r_{AC}). \end{aligned} \quad (9)$$

Однако в силу (6)

$$\mu(r_{AC}) = \mu(r_{AB}) \mu(r_{BC}). \quad (10)$$

Подставляя (10) в (9), сразу находим решение

$$p_B = \mu(r_{AB}), \quad p_C = 0, \quad (11)$$

так что точка C оказывается невлияющей. Легко видеть, что то же верно при наличии влияющей точки B для любого количества точек, расположенных на прямой EF , и даже для континуума

всех точек этой прямой. Более того, если представить себе, что значения f известны во всех точках полуплоскости левее прямой EF , то влияющими будут лишь точки, расположенные на луче BD , так что получается экранирование в чистом виде.

Таблица 1

Веса для интерполяции в пункты 10—15 при использовании корреляционной функции (6) ($h=0,2a$; $\eta=0$)

Пункт	P_1	P_2	P_3	P_4	P_5	P_6	P_7	P_8	P_9
10	2,37	-6,83	7,11	-2,27	6,56	-6,83	0,79	-2,27	2,37
11	0,89	-2,56	2,67	0	0	0	0	0	0
12	2,18	-5,61	4,38	0	0	0	0	0	0
13	0	0	0	0,89	-2,56	2,67	0	0	0
14	0	0	0	2,18	-5,61	4,38	0	0	0
15	0	0	0	3,31	-7,98	5,54	0	0	0

Для иллюстрации сказанного в табл. 1 приведены веса для показанной на рис. 1 в схемы расположения пунктов 1—9, где значения параметра известны, и пунктов 10—15, в которых они разыскиваются.

Данные табл. 1 наглядно иллюстрируют как отмеченные выше свойства функции (6), так и проявления эффекта экранирования. Возможные последствия экранирования можно видеть на рис. 3а, где представлены результаты расчета с весами из табл. 1 значений элемента в точках 10—15 и симметричных им по значениям в точках 1—9 (в произвольных единицах). Разумеется, большие положительные значения элемента в ряде точек совершенно не оправданы в свете информации о точках 1—9.

Как уже упоминалось, учет случайных некоррелированных между собой в различных точках ошибок наблюдений позволяет ослабить эффект экранирования. Это можно видеть из табл. 2, где веса гораздо меньше по абсолютному значению. Интересно, что примерно того же результата можно добиться, если, не учитывая случайных ошибок, просто заменить при расчетах аппроксимацию корреляционной функции (6) на аппроксимацию (8) (табл. 3). Рис. 3б, в показывают, что применение весов из табл. 2 и 3 дает гораздо более правдоподобные результаты, чем использование весов из табл. 1.

На основании сказанного можно утверждать, что применение аппроксимации вида (6) при объективном анализе нежелательно. Взамен можно, в частности, рекомендовать аппроксимацию (8).

Известно, что корреляционные матрицы, используемые при расчетах методом оптимальной интерполяции, должны удовлетворять требованию положительной определенности. Однако эта положительная определенность отнюдь не обеспечивается сама собой. Для эмпирически рассчитанной матрицы она может нарушаться в случае использования различного объема данных для

разных ее компонент (что практически неизбежно ввиду наличия пропусков данных наблюдений). Положительная определенность может также не иметь места, если в качестве $\mu(r)$ взяты значения произвольным образом описанной корреляционной функции. В этом случае могут быть получены даже абсурдные результаты, например отрицательная дисперсия ошибок интерполяции ε^2 .

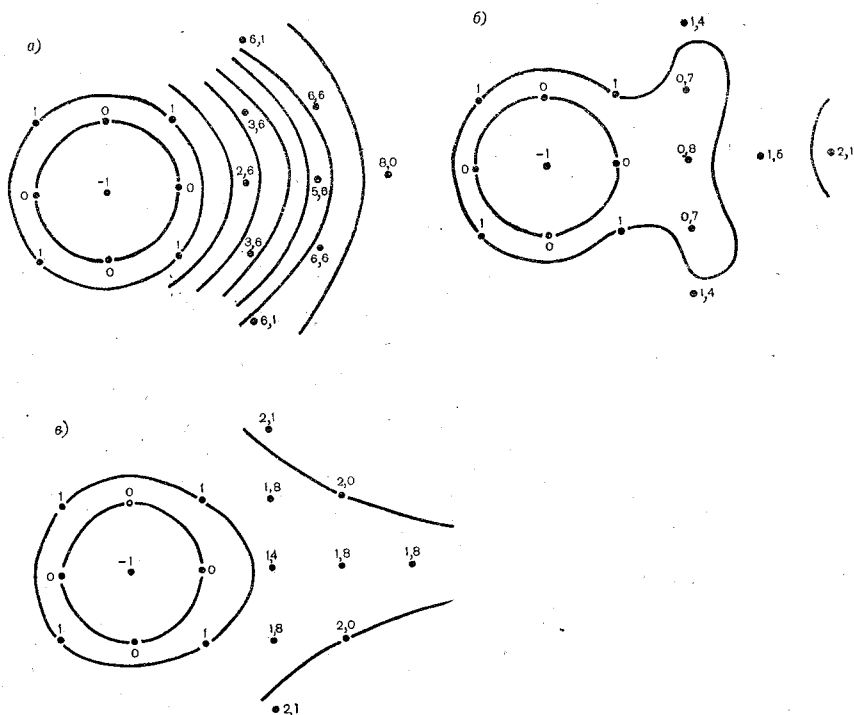


Рис. 3. Поля, рассчитанные при использовании весов из табл. 1—3 (соответственно а—в).

Удобное средство обеспечения положительной определенности корреляционных функций базируется на аппроксимации эмпирических корреляционных функций формулами, обеспечивающими положительный спектр [5]. Для двумерного случая спектр определяется формулой

$$S(k) = \int_0^{\infty} \mu(r) I_0(kr) r dr, \quad (12)$$

где I_0 — бесселева функция первого рода. Применительно к аппроксимациям (6) и (8) эта формула дает выражения

$$S = \frac{a^2}{2} \exp\left(-\frac{a^2 k^2}{4}\right) \quad (13)$$

и

$$S = 3b^2 (1 + b^2 k^2)^{-5/2}, \quad (14)$$

положительные при всех значениях k . Это означает, что обе рассмотренные аппроксимации имеют положительный спектр.

Линейная комбинация положительно определенных функций с положительными коэффициентами является положительно опре-

Таблица 2

Веса для интерполяции в пункты 10—15 при корреляционной функции (6)
($h=0,2a$; $\eta^2=0,02$)

Пункт	p_1	p_2	p_3	p_4	p_5	p_6	p_7	p_8	p_9
10	-0,13	0,22	1,23	-0,22	-0,30	0,22	0,16	-0,22	-0,13
11	-0,14	0,02	0,88	-0,09	-0,14	0,47	0,02	-0,22	0,11
12	-0,18	-0,27	1,08	-0,05	-0,39	0,64	0,12	-0,42	0,23
13	-0,09	-0,14	0,47	-0,04	-0,08	0,57	-0,09	-0,14	0,47
14	-0,05	-0,39	0,64	-0,02	-0,34	0,74	-0,05	-0,39	0,64
15	0,02	-0,58	0,73	0,05	-0,56	0,82	0,02	-0,58	0,73

Таблица 3

Веса для интерполяции в пункты 10—15 при корреляционной функции (8)
($h=0,5a$; $\eta=0$)

Пункт	p_1	p_2	p_3	p_4	p_5	p_6	p_7	p_8	p_9
10	0,14	-0,44	1,60	0,00	-0,21	-0,44	0,01	0,00	0,14
11	0,11	-0,56	1,28	0,08	-0,25	0,20	0,02	-0,11	0,13
12	0,12	-0,64	1,16	0,12	-0,33	0,19	0,04	-0,24	0,31
13	0,06	-0,32	0,44	0,12	-0,39	0,83	0,06	-0,32	0,44
14	0,08	-0,46	0,65	0,15	-0,39	0,46	0,08	-0,46	0,65
15	0,08	-0,45	0,64	0,13	-0,33	0,23	0,08	-0,45	0,64

деленной. Это следует хотя бы из аналогичного свойства, касающегося положительности спектров. Поэтому дополнительный учет некоррелированных случайных ошибок не нарушает положительной определенности матрицы, а обусловленность системы уравнений для определения весов при этом возрастает. В случае коррелированности ошибок положительная определенность также не нарушается, если только корреляционная функция ошибок $v(r)$ тоже имеет положительный спектр в нужном числе измерений. Однако обусловленность системы в этом случае падает, и может возникнуть фиктивное нарушение положительной определенности под влиянием ошибок округления.

В случае же, когда часть используемых данных обладает некоррелированными ошибками, а часть — коррелированными, вопрос об условиях положительной определенности вообще неясен. Поэтому если среди наблюдений, используемых для интерполяции в рассматриваемую точку сетки, имеются как данные косвенного

зондирования, ошибки которых коррелированы между собой, так и данные радиозондирования, то существует реальная опасность нарушения положительной определенности. Во избежание ее можно проводить интерполяцию отдельно по данным прямого и по данным косвенного зондирования с последующим осреднением результатов.

Заметим, что появление абсурдных отрицательных значений ϵ^2 , обусловленное нарушением положительной определенности матриц, лишь изредка приводит к плохим результатам самой интерполяции. В этом отношении эффект нарушения положительной определенности аналогичен эффекту экранирования.

СПИСОК ЛИТЕРАТУРЫ

1. Багров А. Н. Оперативная схема объективного анализа метеорологической информации для Северного полушария. — Труды Гидрометцентра СССР, 1978, вып. 196, с. 3—11.

2. Либерман Ю. М. Базовая программа объективного анализа метеорологических полей и ее применение для анализа ветра. — Труды ГГО, 1980, вып. 412, с. 13—19.

3. Либерман Ю. М. О поиске влияющих станций при объективном анализе метеорологических полей. — Труды Гидрометцентра СССР, 1977, вып. 170, с. 81—86.

4. Юдин М. И. Некоторые закономерности структуры поля геопотенциала. — Труды ГГО, 1961, вып. 121, с. 3—18.

5. Яглом А. М. Введение в теорию стационарных случайных функций. — УМН, вып. 5(51), 1952, с. 3—168.

6. Bengtsson L. Four-dimensional assimilation of meteorological observations. — GARP Publication Ser., 1975, N 15. 76 p.

7. Bergman K. H. Multivariate analysis of temperatures and winds using optimum interpolation. — Mon. Wea. Rev., 1979, vol. 107, N 11, p. 1424—1444.

8. Lorence A., Rutherford J., Larsen G. The ECMWF analysis and data-assimilation scheme: analysis of mass and winds fields. — ECMWF Tech. Rep., N 6, 1977, 47 p.

9. Rutherford J. D. An operational 3-dimensional multivariate statistical analysis scheme. — GARP Programme on Numerical Experimentation. Rep. N 11, 1976, p. 98—121.

10. Schlatter T. W. Some experiments with a multivariate statistical objective analysis scheme. — Mon. Wea. Rev., 1975, vol. 103, p. 246—257.

11. Thiebaut H. J. Experiments in correlation representations for objective analysis. — Mon. Wea. Rev., 1975, vol. 103, N 7, p. 617—627.

Ю. М. Либерман, В. П. Тараканова

ОБ ИСПОЛЬЗОВАНИИ ПРОГНОСТИЧЕСКОЙ ИНФОРМАЦИИ ПРИ АНАЛИЗЕ МЕТЕОРОЛОГИЧЕСКИХ ПОЛЕЙ ЮЖНОГО ПОЛУШАРИЯ

Методы оптимальной интерполяции позволяют априори оценить точность анализа метеорологических полей, если известна статистическая структура анализируемого элемента и координа-

ты наблюдательных станций. Характеристикой точности служит средний квадрат абсолютной ошибки интерполяции

$$E^2 = \varepsilon\sigma, \quad (1)$$

где ε^2 — мера ошибки интерполяции, σ^2 — дисперсия элемента.

Таблица 1

Доля площади полушария (%) по градациям ε и E

ε	Полушарие		E	Полушарие	
	северное	южное		северное	южное
0,0—0,2	36,7	4,8	0—2	37,0	30,1
0,2—0,4	28,9	29,4	2—4	24,1	25,2
0,4—0,6	19,4	28,7	4—6	10,7	14,0
0,6—0,8	8,8	31,5	6—8	5,3	9,0
0,8—1,0	6,2	5,7	8	22,9	21,7
Среднее	0,343	0,510	Среднее	4,2	5,0

Таблица 2

Уточнение анализа при использовании прогностической информации

Средняя квадратическая ошибка прогноза, гп. дам	ε	ε^*	$\frac{\varepsilon - \varepsilon^*}{\varepsilon} \%$	E	E^*	$\frac{E - E^*}{E} \%$
4	0,510	0,357	30,0	4,97	2,73	45,0
6	0,510	0,409	19,6	4,97	3,39	31,7
8	0,510	0,443	13,7	4,97	3,87	22,0

В свою очередь ε^2 легко определить по формуле

$$\varepsilon^2 = 1 - \sum_{i=1}^n p_i \mu_{0i}. \quad (2)$$

Здесь p_i — интерполяционный вес, выражающий влияние станции i на точку анализа 0; μ — автокорреляционная функция.

Сотрудник Национального центра атмосферных исследований США Р. Дженни любезно предоставил ГГО сведения о географическом распределении изменчивости геопотенциала σ в южном полушарии. Пользуясь этими сведениями, мы рассчитали поля ε и E применительно к геопотенциалу поверхности 500 гПа. Результаты расчета показывают, что наибольшие значения E имеют место в умеренных широтах океанов, где изменчивость велика, а сеть станций практически отсутствует. Эти максимумы на больших площадях составляют 12—16 гп. дам, что примерно втрое больше значений E на тех же широтах океанов северного полушария [2]. Такое различие обусловлено в первую очередь ред-

кой аэрологической сетью (в южном полушарии имеется лишь 131 аэрологическая станция, в северном — около 700). Большой контраст точности анализа в двух полушариях подтверждается данными табл. 1.

Проблема повышения точности анализа метеорологических полей в южном полушарии становится первоочередной в связи с глобальным численным прогнозом. Привлечение прогностической информации для анализа может служить одним из путей решения этой трудной проблемы. При этом необходимо учитывать статистические свойства привлекаемой информации, иначе анализ может не улучшиться, а даже стать хуже [3]. Теоретические основы привлечения дополнительной информации с учетом статистических свойств ее ошибок изложены в работе [1].

Как показано в работе [1], при выполнении некоторых не слишком ограничительных предположений мера ошибки интерполяции с учетом прогноза равна

$$\epsilon^{*2} = \frac{\epsilon^2 \eta_0^2}{\epsilon^2 + \eta_0^2 - \epsilon^2 \eta_0^2}, \quad (3)$$

где η_0^2 — мера ошибки прогноза. Результаты расчетов, выполненных на основе соотношения (3) (осреднение производилось по южному полушарию), приведены в табл. 2. Эти результаты, как и географическое распределение разностей $\epsilon - \epsilon^*$ и $E - E^*$, подтверждают существенную ценность прогностических данных для целей диагностического анализа даже при сравнительно малой точности прогноза.

Дальнейшему уточнению анализа, по-видимому, будет способствовать привлечение спутниковой информации.

СПИСОК ЛИТЕРАТУРЫ

1. Каган Р. Л., Шахмейстер В. А. Об использовании прогностической информации при объективном анализе барического поля. — Труды ГГО, 1977, вып. 397, с. 107—115.
2. Либерман Ю. М. О точности анализа поля геопотенциала над северным и южным полушариями. — Труды ГГО, 1968, вып. 228, с. 41—48.
3. Либерман Ю. М. Эксперимент по использованию прогностической информации в объективном анализе. — Труды ГГО, 1977, вып. 397, с. 99—106.

Л. П. Клягина, Ю. М. Либерман

ГОРИЗОНТАЛЬНО-ВРЕМЕННОЙ КОНТРОЛЬ ПРИЗЕМНОГО ДАВЛЕНИЯ

В настоящей статье излагаются результаты экспериментов по комплексному контролю данных о приземном давлении, выполненных в 1980—1981 гг. Эти эксперименты были направлены на решение следующих вопросов:

определение параметров временного и горизонтального контроля;

комплексация блоков временного и горизонтального контроля в рамках единой схемы;

оценка возможностей комплексного горизонтально-временного контроля на реальных данных и разработка основных принципов будущего алгоритма принятия решения (АПР).

Как временной, так и горизонтальный контроль строятся на сравнении так называемой невязки контроля δ с ее предельным допустимым значением Δ : нарушение условия $|\delta| \leq \Delta$ сигнализирует о вероятном грубом искажении данных. Фактическая невязка находится как разность

$$\delta = \hat{p} - p \quad (1)$$

интерполированного значения давления \hat{p} и проверяемого наблюдения p . Предельное допустимое значение

$$\Delta = K\tilde{E} = K\sigma\tilde{\epsilon} \quad (2)$$

представляет собой теоретическую ошибку сопоставления интерполированного и наблюдаемого значений \tilde{E} , увеличенную в K раз. Ошибка Δ зависит от климатической изменчивости давления σ и от меры ошибки сопоставления

$$\tilde{\epsilon}^2 = \epsilon^2 + \eta^2, \quad (3)$$

которая в свою очередь есть сумма мер ошибок интерполяции ϵ^2 и наблюдения η^2 . Как известно, на сети метеорологических станций давление измеряется по барометру с относительно высокой степенью точности, тогда как на многочисленных судах наблюдения по anerоиду характеризуются заметно большей случайной погрешностью. Имея это в виду и учитывая, вместе с тем, микроклиматическую изменчивость давления, примем для средней квадратической ошибки наблюдения η значение 1,0 гПа. Тогда

$$\eta^2 = \frac{\gamma^2}{\sigma^2} = \frac{1,0}{\sigma^2}. \quad (4)$$

Для того чтобы реализовать контроль давления, необходимо задать входящие в (2) параметры K и $\tilde{\epsilon}$. Рассмотрим сначала вопрос об определении $\tilde{\epsilon}$. При горизонтальном контроле характеристика $\tilde{\epsilon}$ существенно локальна: она зависит от конкретного взаимного расположения проверяемой станции и всех влияющих станций, наблюдения которых интерполируются на проверяемую. Поэтому переход от одной проверяемой станции к другой каждый раз требует расчета индивидуального значения $\tilde{\epsilon}$. При временном контроле интервал между сроками является стандартным. Как правило, он равен 3 ч, а там, где цикл приема и обработки повторяется лишь дважды в сутки, увеличивается до 12 ч. Таким образом, оказывается возможным априорное задание $\tilde{\epsilon}$.

Для любой интерполяции, линейной относительно наблюдений, имеет место общая формула

$$\tilde{\varepsilon}^2 = 1 - 2 \sum_{i=1}^n b_i \mu_{01} + \sum_{i=1}^n \sum_{j=1}^n b_i b_j \mu_{ij} + \eta^2 \sum_{i=1}^n b_i^2, \quad (5)$$

где b — интерполяционный вес, μ — автокорреляционная функция, n — число интерполируемых данных. В двухточечном случае, когда $n=2$, она принимает упрощенный вид

$$\tilde{\varepsilon}^2 = 1 - 2(\mu_{31} b_1 + \mu_{02} b_2) + 2\mu_{12} b_1 b_2 + (1 + \eta^2)(b_1^2 + b_2^2). \quad (6)$$

Здесь индексы 0, 1, 2 означают соответственно проверяемый и два влияющих срока. При интерполяции на середину временного интервала длиной $2t$ с его концов имеем ещё более простое соотношение

$$\tilde{\varepsilon}^2 = \frac{3}{2} - 2\mu(t) + \frac{1}{2}\mu(2t) + \frac{1}{2}\eta^2. \quad (7)$$

Наряду с двухточечной интерполяцией для нас представляет интерес случай односточечной экстраполяции давления на последующий срок; в этом случае формула (6) принимает вид

$$\tilde{\varepsilon}^2 = 1 - 2b_1 \mu_{01} + b_1^2 (1 + \eta^2), \quad (8)$$

Рассмотрим следующие четыре варианта двухточечной интерполяции, различающиеся размещением точек на временной оси:

- 1) $t_1 = t_0 - 3$, $t_2 = t_0 + 3$;
- 2) $t_1 = t_0 - 3$, $t_2 = t_0 + 6$ или $t_1 = t_0 - 6$, $t_2 = t_0 + 3$;
- 3) $t_1 = t_0 - 6$, $t_2 = t_0 + 6$;
- 4) $t_1 = t_0 - 12$, $t_2 = t_0 + 12$.

Учитывая (2), (3) и (4), представим теоретическую ошибку сопоставления в виде

$$\tilde{E} = \sqrt{A\sigma^2 + B} \quad (9)$$

и рассчитаем константы A и B для указанных вариантов применительно к аппроксимации временной корреляционной функции давления в виде

$$\mu(t) = \left(1 + \frac{t}{30}\right) e^{-t/30}, \quad (10)$$

рекомендованной в работе [2]. Результаты расчета приведены в табл. 1.

При режимной обработке данных обычно встречается вариант 1. Оценим для этого варианта изменение \tilde{E} по северному полушарию. Поскольку σ колеблется на полушарии зимой в пределах от 4 до 17 гПа, а летом в значительно меньших пределах, то характеристика \tilde{E} оказывается географически весьма устойчивой: она меняется от 1,23 гПа до 1,29 гПа. Иными словами, при интервале между наблюдениями $\tau=3$ ч в любом пункте полушария можно положить $\Delta \approx 1,25$ К. В отличие от этой ситуации при переходе к интервалу $\tau=12$ ч необходимо учитывать географическое положение

проверяемой станции. В этом случае A много больше и потому значение \bar{E} на полушарии может меняться в довольно широких пределах: от 1,4 гПа до 3,1 гПа. Что касается экстраполяции, то при $\tau = 12$ ч ее теоретическая ошибка \bar{E} , а значит, и предельная невязка Δ приблизительно вдвое больше соответствующих интерполя-

Таблица 1

Константы A и B для разных вариантов двухточечной интерполяции

Вариант	Интервал ($t_2 - t_1$)	A	B гПа ²
1	6	$660 \cdot 10^{-6}$	1,50
2	9	$158 \cdot 10^{-5}$	1,56
3	12	$433 \cdot 10^{-5}$	1,50
4	24	$275 \cdot 10^{-4}$	1,50

ционных значений. По этой причине временной контроль в оперативных условиях, когда возможна лишь экстраполяция, гораздо менее эффективен. Однако при режимной обработке данных его возможности весьма велики.

Значительный потенциал временного контроля сочетается с предельной простотой его реализации. Вполне допустимое пренебрежение суточным ходом давления позволяет оперировать непосредственно с наблюдениями, а не с их аномалиями. Использование весов, линейно зависящих от времени, практически не снижает точность контроля по сравнению с оптимальными весами. Наконец, все требуемые параметры могут быть рассчитаны заранее.

Перейдем к вопросу об определении эмпирического коэффициента K . Этот коэффициент используется в дальнейшем как практический критерий для выделения грубых ошибок из совокупности грубых и случайных ошибок рассматриваемой выборки данных. Случайные ошибки, свойственные любому конкретному наблюдению, распределены по нормальному закону. Под грубыми ошибками мы понимаем ошибки, обусловленные просчетами наблюдателей или искажениями, возникшими при передаче данных и их автоматической обработке. Ошибки подобного рода свойственны лишь малому числу элементов выборки. Можно построить основанную на тех или иных гипотезах статистическую модель грубых ошибок и оценить ее параметры.

Для оценки параметра K мы воспользовались способом, который назвали «усечением по эксцессу». Здесь имеется в виду многократное последовательное усечение совместного распределения грубых и случайных ошибок. Поясним смысл такого усечения. Естественно предположить, что случайные ошибки, небольшие по абсолютному значению, заключены в некотором интервале $(-a, a)$, тогда как значительные грубые ошибки находятся вне этого интервала. В реальной задаче некоторая доля случайных ошибок по абсолютному значению превышает уровень a и потому при усечении

попадает в класс грубых ошибок. Вместе с тем, некоторые грубые ошибки, значение которых меньше a , попадают в класс случайных. Уровень a следует задавать таким, чтобы указанные неизбежные погрешности были бы минимальными. В работе [3] описан опыт субъективной оценки этого уровня на основе кривой распределения. Более корректна количественная оценка, примененная в настоящем эксперименте.

Пусть известна дисперсия совместного распределения грубых и случайных ошибок. Будем выполнять усечение исходного полного распределения при разных a , вычисляя каждый раз эксцесс Ex усеченного распределения. Тогда большим значениям a будет соответствовать положительный эксцесс, а малым — отрицательный. Найдем приближенно то значение $a=K$, при котором $Ex=0$. Наконец, сопоставим дисперсию усеченного распределения с дисперсией случайных ошибок; в случае их близости можно утверждать, что в результате усечения удалось удовлетворительно отделить грубые ошибки от случайных.

Способ усечения по эксцессу применим для определения параметра K как горизонтального, так и временного контроля. Проиллюстрируем его результатами расчета, выполненного применительно к горизонтальному контролю.

Исходная выборка содержит данные о давлении на 505 станциях Европы, Западной Сибири и Средней Азии за восемь сроков, разделенных трехчасовыми интервалами. Первый срок серии — 15 ч 3 января 1977 г. Эти данные, извлеченные из оперативных синоптических телеграмм Ленинградского Гидрометцентра, подвергнуты горизонтальному контролю по описанной ниже программе. При этом получена совокупность относительных ошибок контроля $\beta=\delta/E$, образующая статистический ряд из 4040 элементов. В интервале от $-8,0$ до $8,0$ находится 3977 элементов, т. е. 98,5 %. Этот интервал разделен на 65 градаций шириной 0,250.

Преобразование исходного распределения выполняется последовательным отсечением градаций с левого и правого концов интервала. На каждой стадии определяются выборочные значения средней ошибки $\bar{\beta}$, дисперсии D , коэффициента асимметрии S и эксцесса Ex . При смене знака S совершается переход в отсечении с левого конца интервала на правый или с правого на левый. Появление первого отрицательного значения Ex служит признаком прекращения циклического процесса.

В табл. 2 содержатся значения $\bar{\beta}$, D , S и Ex , отвечающие начальному и конечному состояниям и некоторым промежуточным стадиям.

На рис. 1 представлены три кривых, первая из которых — эмпирическое начальное распределение относительных ошибок, а вторая и третья — теоретические кривые распределения, входными параметрами для построения которых служат данные из табл. 2 соответственно начальной и конечной стадий усечения. Анализ рис. 1 и табл. 2 показывает, что нечетные моменты распределения

близки к нулю, а среднее квадратическое отклонение относительных ошибок составляет 1,22. Аппроксимируя эмпирическое распределение нормальным законом, получаем теоретическую кривую 2. Однако она не совсем точно описывает фактическое распределение. Применяя метод усечения по эксцессу, отсекаем тем самым около 14 % грубых ошибок. Внутри оставшихся градаций число случаев

Таблица 2

Статистические параметры на различных стадиях усечения

Стадия усечения	$\beta \cdot 10^2$	D	$S \cdot 10^2$	Ex	Λ
начальная (1)	-0,7	1,49	82,4	17,13	3977
5	-2,5	1,13	33,5	13,73	3954
10	-3,8	1,04	-18,1	12,54	3947
15	-3,6	0,96	-14,2	11,29	3940
20	-2,7	0,82	28,5	9,98	3926
25	-3,2	0,72	-13,1	7,47	3910
30	-2,6	0,66	7,1	6,11	3899
35	-2,5	0,59	6,7	4,74	3882
40	-2,6	0,50	0,2	2,95	3854
45	-2,6	0,45	-3,6	2,06	3829
50	-0,9	0,35	13,7	1,09	3751
55	-1,3	0,25	4,6	0,18	3591
конечная (57)	-1,2	0,20	5,6	-0,17	3456

составляет 3456. В конечном распределении все нечетные моменты также сохраняются около нуля, а среднее квадратическое отклонение уменьшается до 0,45. Усечение прерывается, так как именно в этом случае эксцесс переходит через 0. Используя параметры β и \sqrt{D} из табл. 2 для конечной стадии усечения, строим теоретическую кривую распределения. Из рис. 1 видно, что эта кривая достаточно хорошо описывает эмпирическую и практически все отклонения от среднего (99,7 %) лежат в интервале $(-3\sqrt{D}, 3\sqrt{D})$, т. е. искомый коэффициент $K=3 \cdot 0,45=1,35$.

Горизонтальный контроль информации в этом случае считает подозрительной всякую ситуацию, в которой эмпирическая ошибка больше теоретической на 40 % и более. Однако при разработке комплексного контроля в рамках нашей задачи важно было учесть взаимодействие горизонтального контроля с временным. Кроме того с практической точки зрения на первой стадии исследований целесообразно несколько ограничить объем информации, подлежащей тщательному, всестороннему анализу. По этим причинам описываемые ниже эксперименты по горизонтальному контролю выполнены при удвоении значения K , т. е. при $K_r=2,8$, а $K_b=4,0$.

Комплексный горизонтально-временной контроль (ГВК) предполагает взаимодействие горизонтального контроля (ГК) и временного контроля (ВК) в рамках единого алгоритма. Его про-

граммная реализация оказывается весьма громоздкой: здесь сочетаются большой объём вычислений и множество логических операций. В основу ГВК положена базовая программа объективного анализа, составленная для ЭВМ БЭСМ-6 на языке АЛГОЛ [4]. Исходными данными служат координаты синоптических станций и

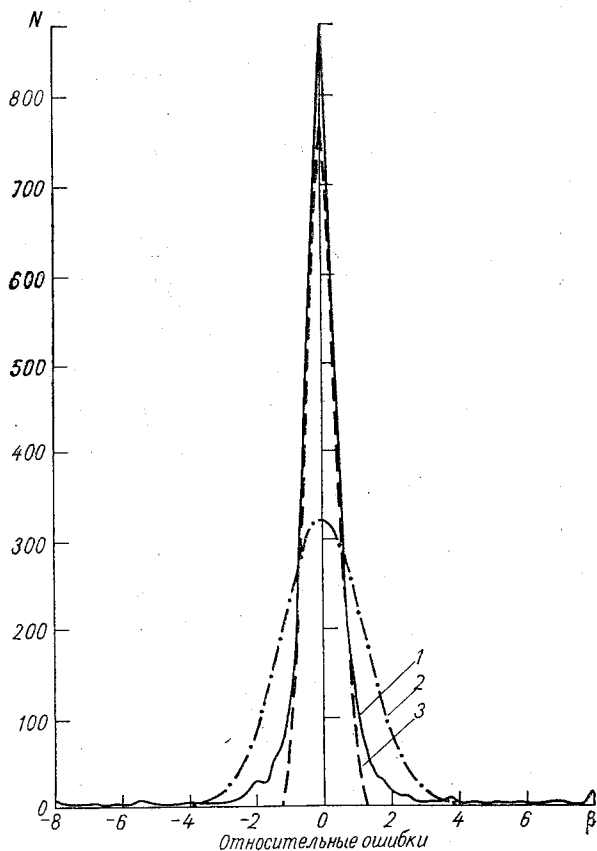


Рис. 1. Кривые распределения относительных ошибок.
 1 — эмпирическая, 2 — теоретическая по всему ряду,
 3 — теоретическая по усеченному ряду.

проверяемые значения давления p , зафиксированные на этих станциях в 5 последовательных сроках наблюдений с интервалом 3 ч. Дополнительно вводится климатическая информация по северному полушарию, именно нормы \bar{p} и среднее квадратическое отклонение давления σ в узлах широтно-долготной сетки с шагом 10° . Отсутствие данных на отдельных станциях отмечается признаком 999. Предполагается, что пространственная корреляция давления не отличается от используемой в работе [4] корреляции геопотенциала; временная корреляция аппроксимируется функцией (10).

По сравнению с [4] в блок ГК внесены следующие изменения. Интерполяция на проверяемую станцию производится независимо от густоты сети: даже в случае, когда одна из влияющих станций расположена очень близко от проверяемой. Напомним, что в [4] при расстоянии менее 50 км считалось, что значение давления в узле сетки равно значению на станции. Это предположение, естественное при объективном анализе, неприемлемо при контроле. Максимальное число станций в каждом квадрате машинной карты увеличено с 20 до 60, а во всей области поиска, содержащей $5 \times 5 = 25$ квадратов, со 100 до 300. Такое изменение обусловлено тем, что синоптическая сеть насчитывает на порядок больше станций по сравнению с аэрологической. Поэтому поиск 6 ближайших станций, влияющих на проверяемую, требует осмотра не 15—20 станций, как было в [4], а 50—80 и более. Уменьшить это число можно при условии, что сторона квадрата машинной карты уменьшается с 750 до 150 или 200 км. Но в таком случае резко увеличится потребность в памяти.

Сгущение синоптической сети по сравнению с аэрологической обуславливает ещё одну особенность: значительно чаще встречается ситуация экранирования. Здесь подразумевается соотношение интерполяционных весов влияющих станций. При определенном взаимном расположении некоторые станции имеют отрицательные веса. Алгоритм предусматривает исключение из числа влияющих любой станции с весом менее $-0,2$. В работе [4] при анализе аэрологических полей такое исключение было сравнительно редким. Напротив, в блоке ГК оно весьма часто, особенно на периферии области.

Существенно, что алгоритм ГВК построен по принципу итераций. Все станции, данные которых на некоторой итерации признаны ошибочными, исключаются из рассмотрения и в следующей итерации не участвуют. Такой подход предотвращает искажение верной информации в блоках ГК и ВК вследствие влияния ошибочных станций. Последней является та итерация, на которой не обнаружено ни одной ошибочной станции. Как показал опыт, для окончания ГВК обычно достаточно выполнить 2 итерации. Если число проверяемых станций около 500, то для этого требуется около 10 мин машинного времени.

Отработка алгоритма требует тщательного анализа разнообразных ситуаций, встречающихся при ГВК. С этой целью на печать выдается обширная информация по каждой станции, которая бракуется одним из блоков ГК и ВК или обоими блоками. Приведем перечень такой информации.

1. Общие сведения о станции:

- 1.1 порядковый номер.

- 1.2 географические и декартовы координаты.

- 1.3 значения \bar{p} , σ , η^2 .

2. Результаты оптимальной интерполяции и горизонтального контроля:

2.1. интерполированные значения аномалии давления p' и самого давления $P = \bar{p} + p'$

2.2. сумма интерполяционных весов;

2.3. номера влияющих станций, их координаты, веса, фактические аномалии и расстояния от проверяемой станции;

2.4. теоретические ошибки сопоставления \bar{E} и ε ;

2.5 фактическая невязка δ и её предельное допустимое значение Δ .

3. Результаты временного контроля:

3.1 фактические значения давления на проверяемой станции в 5 последовательных сроках;

3.2 интерполированные значения давления для трёх центральных из этих пяти сроков, полученные по двум ближайшим срокам;

3.3. фактическая невязка δ и допустимая невязка Δ ;

3.4 относительная невязка β .

Помимо перечисленной цифровой информации печатается диагноз в словесной форме.

Описанная программа использована для экспериментов по горизонтально-временному контролю данных, относящихся к срокам с 3 по 8. Крайние сроки 1, 2, 9, 10 не включены в выборку потому, что для них временной контроль невозможен или сильно ограничен. Таким образом, общее число проверенных наблюдений $505 \times 6 = 3030$. Результаты экспериментов иллюстрируются примерами, приведенными в табл. 3. Нижний индекс у величин, входящих в табл. 3, означает срок, а верхний индекс указывает, к какому именно контролю — горизонтальному или временному — относится данная величина. Рассмотрим эти примеры.

1. Алгоритм ГVK в его нынешнем виде обнаружил 17 явных ошибок (примеры 1—17). Достоверность диагноза во всех этих случаях подтверждается двумя признаками:

— близостью фактических невязок $\delta_n^Г$ и $\delta_n^В$;

— противоположностью знаков $\delta_n^В$ по отношению к $\delta_{n-1}^В$ и $\delta_{n+1}^В$.

Естественно ожидать, что изолированная ошибка в срок n даст приблизительно половинные невязки противоположного знака в сроки $n-1$, $n+1$. Действительно, во многих случаях обнаруживается такое соотношение невязок ВК (примеры 6, 8, 12, 13, 17). Однако иногда, особенно при очень значительных ошибках, оно оказывается замаскированным. Заслуживает внимания то обстоятельство, что зачастую удается уверенно обнаружить даже сравнительно небольшие ошибки, лишь немного превосходящие допустимые невязки (примеры 2, 4, 5, 7, 9, 16).

2. Вторую группу образуют случаи, когда результаты ГК и ВК противоречат и потому уверенный диагноз поставить не удастся.

2.1. Среди них выделяются случаи подозрения на ошибку по результатам ГК, причём фактическая ошибка $\delta^Г$ весьма близка к допуску $\Delta^Г$ или, иными словами, относительная фактическая ошибка $\beta^Г$ близка к принятому у нас значению $K^Г = 2,8$ (примеры 18—21). Как подсказывает большой опыт комплексного контроля аэро-

Примеры горизонтально-временного контроля давления

№ п/п	Станция	Срок п	p	\hat{p}_n	\hat{p}_n	$\Delta \Gamma$	$\beta \Gamma$	ΔB	βB	$\delta \Gamma_n$	δB_{n-1}	δB_m	δB_{n+1}
1	33	5	971,1	988,4	992,8	16,2	3,0	5,1	-17,1	17,3	-6,0	21,7	-13,7
2	51	3	1020,1	1026,8	1028,3	4,2	4,4	5,8	6,5	6,7	-4,8	8,3	-4,1
3	73	6	1057,0	1049,0	1050,1	4,6	-4,9	5,0	-5,5	-8,0	3,1	2,8	2,8
4	87	3	1055,8	1050,6	1049,6	4,2	-3,4	5,1	-4,9	-5,2	5,9	-6,2	2,9
5	88	3	1049,4	1055,6	1056,8	-4,4	3,8	5,0	5,9	6,1	-3,7	7,4	-3,8
6	92	6	1000,9	1044,2	1044,1	4,3	28,4	5,0	34,3	43,3	-21,5	43,2	-21,5
7	108	8	1040,4	1046,4	1046,9	4,3	3,9	5,1	5,1	6,0	3,8	6,5	-2,7
8	124	8	1000,9	1048,8	1052,1	4,7	28,8	5,0	41,2	47,9	-25,8	51,2	-25,8
9	147	5	1050,9	1045,1	1045,6	4,3	-3,8	5,0	-4,3	-5,8	3,5	-5,3	2,2
10	173	8	1047,7	1040,2	1041,1	4,9	-4,3	4,9	-5,3	-7,5	3,2	-6,6	2,9
11	186	3	1048,1	1040,1	1039,1	5,1	-4,4	5,0	-7,2	-8,0	5,8	-8,9	3,4
12	204	8	1000,9	1029,1	1027,6	4,4	18,2	5,0	21,6	28,2	-13,1	26,7	-13,8
13	215	3	997,6	982,6	983,3	4,8	-8,7	5,1	-11,1	-15,0	10,9	-14,3	6,1
14	344	6	1034,4	1018,5	1024,5	4,4	-10,2	5,1	-7,7	15,9	5,1	9,9	4,8
15	376	4	1037,2	1029,2	1031,2	4,2	-5,3	5,2	-4,7	-8,0	3,2	-6,0	3,0
16	377	3	1038,2	1033,2	1033,0	4,4	-3,2	5,1	-4,0	-5,0	2,7	-5,2	2,9
17	444	3	1030,4	1044,4	1044,7	4,3	9,0	5,1	11,3	14,0	-7,3	-14,3	-7,6
18	58	3	1031,5	1027,0	1031,5	4,4	-2,9	5,1	0,6	-4,5	0,1	0,8	-0,4
19	69	3	1049,2	1043,6	1050,0	5,4	-2,9	5,0	0,6	-5,6	-0,1	-0,3	-0,1
20	197	3	1035,5	1040,4	1035,4	4,7	3,6	5,0	0,2	4,7	2,4	-0,3	0,1
21	5	3	991,9	1006,1	994,0	10,9	3,6	5,1	1,7	11,8	1,3	2,2	-0,4
22	59	3	1021,5	1030,9	1021,7	4,4	6,0	5,1	0,2	9,4	-0,3	0,2	-0,9
23	83	6	1040,0	1031,3	1041,0	4,6	-5,8	5,1	0,1	-9,6	0,4	0,1	0,1
24	30	3	974,6	987,4	987,4	5,1	5,1	5,1	0,1	0,1	-24,2	12,9	0,0
25	215	3	986,7	992,8	992,8	5,1	5,1	5,1	-3,9	-5,3	-14,2	6,1	0,6
26	142	3	1056,1	1050,8	1051,2	4,4	-3,4	5,0	3,5	4,6	2,8	-4,9	2,5
27	95	5	1051,6	1056,2	1055,9	4,2	3,0	5,0	3,5	4,6	-2,4	4,4	-1,9

логической информации [5], целесообразно увеличивать в таких случаях K_r , скажем, на 20 % (процедура реабилитации).

2.2. Отрицательная реакция ГК при значительном превышении допуска (примеры 22, 23) может быть обусловлена сложной синоптической ситуацией (глубокий циклон, прохождение фронта). Такие случаи наиболее трудны. Здесь для постановки правильного диагноза с достаточной достоверностью очень желательно привлечение дополнительной информации о ветре и облачности [1]. С другой стороны, причиной отрицательной реакции ГК может оказаться грубая ошибка не на проверяемой станции, а на одной из влияющих станций. Искажающий вклад такой ошибки в проверяемое значение легко учесть: он равен произведению ошибки на вес влияющей станции.

2.3. Отрицательная реакция ВК срока n при положительной реакции ГК, как правило, является следствием ошибки в один из соседних сроков $n-1$ или $n+1$ (примеры 24, 25). В такой ситуации невязка срока n близка по абсолютной величине к половине невязки соседнего действительно ошибочного срока и противоположна ей по знаку. Диагноз подобных ошибок с одновременной реабилитацией проверяемого срока не представляет трудностей.

3. В третью группу входят реальные ошибки, которые вызывают отрицательную реакцию лишь у одного из двух комплексированных методов, тогда как второй метод «не замечает» этих ошибок. Как правило, более чувствительным оказывается блок ВК (примеры 26, 27). Понятно, что в подобных ситуациях данные необходимо признать ошибочными (процедура «обратной реабилитации»).

4. Алгоритм ГВК легко обнаруживает систематические ошибки на отдельных станциях, поскольку они не сказываются на результатах ВК, тогда как ГК в каждом сроке даёт отрицательную реакцию. В нашей выборке такие ошибки обнаружены на двух станциях, телеграммы которых содержали значения давления без приведения его к уровню моря.

5. Эксперименты показали, что интересующие нас грубые ошибки можно условно классифицировать на малые, средние и большие. К малым отнесём те ошибки, которые не вызывают выраженной реакции ни горизонтального, ни временного контроля. Понятно, что их нельзя обнаружить в рамках ГВК. Напротив, большие ошибки, инициирующие реакцию обоих компонентов ГВК, уверенно обнаруживаются. Наибольшие трудности связаны с диапазоном средних ошибок, на которые реагирует либо только ГК, либо только ВК. Именно в этой области совершенствование алгоритма ГВК обещает заметное повышение эффективности контроля.

СПИСОК ЛИТЕРАТУРЫ

1. Багров А. Н. Оперативная схема объективного анализа аэрологической информации для северного полушария. — Труды ГМЦ, 1978, вып. 196, с. 3—11.
2. Лугина К. М., Каган Р. Л. К вопросу о пространственно-временном анализе барического поля. — Труды ГГО, 1974, вып. 336, с. 75—94.

3. Клягина Л. П. Определение параметров комплексного контроля поля приземного давления. — Труды ГГО, 1980, вып. 431, с. 44—50.

4. Либерман Ю. М. Базовая программа объективного анализа метеорологических полей и ее применение для анализа ветра. — Труды ГГО, 1980, вып. 412, с. 13—19.

5. Либерман Ю. М., Тараканова В. П. Опыт статико-временного контроля аэрологической информации. — Труды ГГО, 1980, вып. 431, с. 31—43.

Л. П. Клягина, Ю. М. Либерман

О ВЕРТИКАЛЬНОМ СТАТИСТИЧЕСКОМ КОНТРОЛЕ ТЕМПЕРАТУРЫ В СВОБОДНОЙ АТМОСФЕРЕ

Вертикальный статистический контроль метеорологических данных сводится к интерполяции дискретных измерений на заданный изобарический уровень с последующим сравнением проинтерполированного и измеренного значений. Такой контроль весьма прост. Он особенно полезен в рамках комплексного подхода, при котором диагноз грубых ошибок и соответствующее исправление данных производится по результатам нескольких независимых методов контроля. Приведем основные соотношения вертикального контроля.

Пусть t — неизвестное истинное значение отклонения метеорологического элемента от климатической нормы, а γ — ошибка наблюдения. Пусть далее известны результаты наблюдения, например данные радиозондирования на n изобарических поверхностях

$$\tilde{t}_k = t_k + \gamma_k, \quad k = 1, 2, \dots, n. \quad (1)$$

Проинтерполировав аномалии \tilde{t} с уровней i, j на промежуточный уровень k , получим

$$\hat{t}_k = a_{ki}\tilde{t}_i + b_{kj}\tilde{t}_j, \quad (2)$$

где a_{ki}, b_{kj} — веса, отражающие влияние уровней i, j на уровень k ($i < k, j > k$). Разность

$$\delta_k = \hat{t}_k - \tilde{t}_k \quad (3)$$

представляет собой ошибку сопоставления проинтерполированного и измеренного значений. Сравнение фактической ошибки δ_k с некоторым предельно допустимым значением Δ_k позволяет судить о достоверности наблюдений: при нарушении условия

$$|\delta_k| \leq \Delta_k = A\bar{E}_k \quad (4)$$

возникает сомнение в их правильности. Здесь \bar{E}_k — средняя квадратическая ошибка сопоставления:

$$\bar{E}_k = \sqrt{\bar{\delta}^2} = \sqrt{(a_{ki}\tilde{t}_i + b_{kj}\tilde{t}_j - \tilde{t}_k)^2}, \quad (5)$$

A — эмпирический коэффициент.

Параметры вертикального контроля a_{ki} , b_{kj} , \bar{E}_k могут быть оценены априори, если известны ковариации наблюдений $\tilde{m}_{ij} = \tilde{t}_i \tilde{t}_j$. Действительно, в общем случае линейной интерполяции относительно наблюдаемых значений соотношение (5) принимает вид

$$\begin{aligned} \bar{E}_k^2 = & a^2_{ki} \tilde{m}_{ii} + b^2_{kj} \tilde{m}_{jj} + 2a_{ki} b_{kj} \tilde{m}_{ij} - \\ & - 2a_{ki} \tilde{m}_{ki} - 2b_{kj} \tilde{m}_{kj} + \tilde{m}_{kk}. \end{aligned} \quad (6)$$

В случае же оптимальной интерполяции вместо (6) получается более простая формула

$$\bar{E}_k^2 = \tilde{m}_{kk} - a_{ki} \tilde{m}_{ki} - b_{kj} \tilde{m}_{kj}. \quad (7)$$

Легко показать, что для входящих в (7) оптимальных весов справедливы выражения

$$\begin{aligned} a_{ki} &= (\tilde{m}_{ki} \tilde{m}_{jj} - \tilde{m}_{kj} \tilde{m}_{ij}) / C_k, \\ b_{kj} &= (\tilde{m}_{kj} \tilde{m}_{ii} - \tilde{m}_{ki} \tilde{m}_{ij}) / C_k, \end{aligned} \quad (8)$$

где

$$C_k = \tilde{m}_{ii} \tilde{m}_{jj} - \tilde{m}^2_{ij}. \quad (9)$$

На практике часто приходится иметь дело не с интерполяцией, а с экстраполяцией наблюдений на уровень k с вышележащего уровня j или с нижележащего уровня i . В общем случае экстраполяция сверху вниз

$$\hat{t}_k = b_{kj} \hat{t}_j \quad (j > k) \quad (10)$$

характеризуется ошибкой

$$\bar{E}_k^2 = b^2_{kj} \tilde{m}_{jj} + \tilde{m}_{kk} - 2b_{kj} \tilde{m}_{kj}, \quad (11)$$

а экстраполяция снизу вверх

$$\hat{t}_k = a_{ki} \hat{t}_i \quad (i < k) \quad (12)$$

— ошибкой

$$\bar{E}_k^2 = a^2_{ki} \tilde{m}_{ii} + \tilde{m}_{kk} - 2a_{ki} \tilde{m}_{ki}. \quad (13)$$

В случае оптимальной экстраполяции наряду с (11) и (13) справедливы соответственно соотношения

$$\bar{E}_k^2 = \tilde{m}_{kk} - b_{kj} \tilde{m}_{kj}, \quad (14)$$

$$\bar{E}_k^2 = \tilde{m}_{kk} - a_{ki} \tilde{m}_{ki}, \quad (15)$$

а для весов имеют место формулы

$$b_{kj} = \tilde{m}_{kj} / \tilde{m}_{jj}, \quad (16)$$

$$a_{ki} = \tilde{m}_{ki} / \tilde{m}_{ii}. \quad (17)$$

Заметим, что ковариации измеренных значений \tilde{m}_{ij} отличаются от истинных ковариаций m_{ij} . Первые завышены по сравнению со вторыми на ковариации ошибок наблюдений $\eta_{ij} = \gamma_i \gamma_j$:

$$\tilde{m}_{ij} = m_{ij} + \eta_{ij}. \quad (18)$$

Применительно к данным радиозондирования в одном пункте корреляция ошибок измерений на разных уровнях должна иметь место. Однако, поскольку в настоящее время мы не располагаем соответствующими данными, будем приближенно считать, что

$$\bar{m}_{ij} = \begin{cases} m_{ij} & \text{при } i \neq j, \\ m_{ii} + \bar{\gamma}^2 & \text{при } i = j. \end{cases} \quad (19)$$

Очевидно, это допущение о некоррелированности ошибок на разных уровнях может привести к некоторому завышению оценок невязок сопоставления \bar{E}_k .

В работе [1] приведены параметры вертикального контроля данных о геопотенциале и температуре, полученные на основе описанной выше методики. Они рассчитаны для стандартных изобарических поверхностей 1000, 850, 700, 500, 300, 200 и 100 гПа с использованием сведений о вертикальных корреляциях, опубликованных О. Б. Мерцаловой в 1967 и 1972 гг. [3, 4]. В 1978 г. появилась работа Р. И. Хвостовой [5], содержащая значительно более полные и детализированные сведения о корреляциях температуры над территорией СССР. Хвостова приводит корреляционные матрицы для января и июля в двух условно однородных районах, охватывающих почти всю территорию СССР. При этом в январе каждый район дополнительно делится на два подрайона. Таким образом, общее число приведенных в работе [5] корреляционных матриц равно шести. Для каждой из них мы рассчитали веса a , b и допустимые невязки Δ вертикального контроля. Расчеты выполнены при постоянном по высоте значении средней квадратической ошибки наблюдений $\sqrt{\bar{\gamma}^2} = 1,0$ и при $A = 2,5$. Использовались сглаженные сезонные профили изменчивости температуры (табл. 1). В отличие от [1] рассмотрены не только варианты контроля отдельных уровней по двум ближайшим, но и такие варианты, в которых влияющие уровни отделены от контролируемого промежуточной изобарической поверхностью.

Практический интерес представляет вопрос об устойчивости параметров контроля по отношению к вариациям корреляционной матрицы. Точность оптимальной интерполяции (экстраполяции) с весами (8), (16) и (17) характеризуется средним квадратом ошибки (7), (14) и (15). Понятно, что любое изменение весов по сравнению с оптимальными ведет к увеличению ошибки. Новое, увеличенное по сравнению с оптимальным, значение \bar{E}^2 можно оценить по формулам (6), (11) и (13). Подставляя в эти формулы веса, отвечающие различным корреляционным матрицам из имеющихся в работе [5], мы определили потерю точности по сравнению с весами, оптимальными для каждой матрицы. Оказалось, что рост \bar{E} , как правило, не превосходит 3—5 %. Лишь в редких случаях для уровней 400 гПа и выше при использовании отдаленных влияющих поверхностей потеря точности достигает 15—20 %. Это обстоятельство позволяет рекомендовать для практического использования обобщенные параметры контроля, приведенные в табл. 2, в

которой приняты следующие обозначения: k — контролируемый уровень; i, j — соответственно нижележащий и вышележащий влияющие уровни; a_{ki}, b_{kj} — интерполяционные веса уровней i, j ; Δ — предельная допустимая невязка.

Таблица 1

Осредненные профили изменчивости температуры ($^{\circ}\text{C}$)

Район	Широта	Поверхность, гПа								
		850	700	500	400	300	250	200	150	100
Январь										
1	60—90°	5,9	5,3	5,1	4,8	4,5	5,1	5,7	5,4	5,1
2	40—60	5,8	5,2	5,0	5,7	4,4	5,0	5,6	5,3	5,0
Июль										
1	60—90	4,8	4,3	4,2	3,9	3,7	4,1	4,7	4,5	4,2
2	40—60	3,8	3,4	3,3	3,1	2,9	3,3	3,7	3,5	3,3

Рассмотрим далее реакцию вертикального контроля на характерные случаи грубых ошибок.

1. Предположим, что полученное при радиозондировании значение температуры на промежуточном уровне \tilde{t}_k в процессе обработки и передачи по каналам связи искажено грубой ошибкой ζ_k , т. е. вместо \tilde{t}_k известно значение

$$\hat{t}_k = \tilde{t}_k + \zeta_k. \quad (20)$$

Тогда фактическая ошибка сопоставления оказывается увеличенной не только на уровне k , но также на смежных с ним уровнях $k-1$ и $k+1$:

$$\begin{aligned} \delta_k &= \hat{t}_k - t_k = \delta_k - \zeta_k, \\ \delta_{k-1} &= \hat{t}_{k-1} - \tilde{t}_{k-1} = \delta_{k-1} + b_{k-1, k} \zeta_k, \\ \delta_{k+1} &= \hat{t}_{k+1} - t_{k+1} = \delta_{k+1} + a_{k+1, k} \zeta_k. \end{aligned} \quad (21)$$

Перепишем соотношения (21) в виде

$$\begin{aligned} \delta_k + \zeta_k &= \delta_k, \\ \delta_{k-1} - b_{k-1, k} \zeta_k &= \delta_{k-1}, \\ \delta_{k+1} - a_{k+1, k} \zeta_k &= \delta_{k+1}. \end{aligned} \quad (22)$$

Здесь правые части представляют собой неизвестные невязки сопоставления, которые имели бы место при отсутствии искажения ζ_k . Наиболее вероятны нулевые значения невязок. Однако естественно вместо нулевых значений указать некоторые интервалы, в которых должны находиться абсолютные значения правых и левых частей (22). В качестве таких интервалов можно взять, например,

средние квадратические или предельно допустимые невязки сопоставления.

Теперь сформулируем следующие правила диагноза изолированной ошибки на промежуточном уровне $k=2, 3, \dots, n-1$.

Таблица 2

Обобщенные параметры вертикального контроля температуры

Уровни			a_{ki}	b_{kj}	Δ_k °C	
k	i	j			январь	июль
850	—	700	—	0,90	7,8	7,5
850	—	500	—	0,80	10,3	8,1
700	850	500	0,41	0,52	5,5	5,0
700	850	400	0,50	0,42	6,0	5,5
700	—	500	—	0,85	7,2	5,8
500	850	400	0,22	0,73	5,3	4,5
500	850	300	0,50	0,36	8,5	6,1
500	700	400	0,37	0,60	4,8	4,3
500	700	300	0,70	0,24	6,5	5,3
400	500	300	0,74	0,28	5,4	4,4
400	500	250	0,86	0,08	6,0	4,6
400	700	300	0,56	0,43	7,8	5,3
400	700	250	0,72	0,02	8,4	6,0
300	400	250	0,61	0,50	6,8	6,0
300	400	200	0,60	0,16	8,9	6,5
300	500	250	0,54	0,57	8,0	6,0
300	500	200	0,49	0,14	10,1	7,0
250	300	200	0,51	0,56	7,0	7,7
250	300	150	0,46	0,44	8,5	8,7
250	400	200	0,17	0,56	9,3	8,8
250	400	150	0,01	0,41	10,2	9,6
200	250	150	0,32	0,72	6,6	6,5
200	250	100	0,46	0,53	8,3	8,4
200	300	150	-0,08	0,82	7,4	6,9
200	300	100	-0,11	0,62	10,1	8,9
150	200	100	0,46	0,54	4,9	5,0
150	250	100	0,20	0,78	6,1	6,4
150	200	—	0,76	—	8,0	6,7
100	150	—	0,76	—	6,5	6,2
100	200	—	0,56	—	9,0	8,2

1.1. На трех смежных уровнях обнаружено превышение допустимой ошибки:

$$|\delta_k| > \Delta_k, \quad |\delta_{k-1}| > \Delta_{k-1}, \quad |\delta_{k+1}| > \Delta_{k+1}. \quad (23)$$

1.2. Знаки невязок δ_{k-1} и δ_{k+1} одинаковы и противоположны знаку δ_k . Это следует из (21), если принять $\delta_k = \delta_{k-1} = \delta_{k+1} = 0$ и учесть, что веса $a_{k+1, k}$ и $b_{k-1, k}$ положительны.

1.3. Между фактическими и допустимыми невязками имеют место вытекающие из (22) соотношения

$$\begin{aligned} |\delta_{k-1} - b_{k-1, k} \zeta_k| &\leq \alpha \Delta_{k-1}, \\ |\delta_{k+1} - a_{k+1, k} \zeta_k| &\leq \alpha \Delta_{k+1}, \end{aligned} \quad (24)$$

где α — эмпирический коэффициент, меньший единицы.

2. Если грубая ошибка ζ_1 возникла на крайнем нижнем уровне $k=1$, то фактические ошибки сопоставления изменяются на двух смежных уровнях:

$$\begin{aligned} \delta_1 &= \hat{t}_1 - t_1 = \delta_1 - \zeta_1, \\ \delta_2 &= \delta_2 + a_{21} \zeta_1. \end{aligned} \quad (25)$$

В этом случае диагноз производится по следующему правилу.

2.1. На двух нижних уровнях обнаружено превышение допустимой ошибки:

$$|\delta_1| > \Delta_1, \quad |\delta_2| > \Delta_2. \quad (26)$$

2.2. Знаки δ_1 и δ_2 противоположны.

2.3. Невязки подчиняются соотношениям

$$\begin{aligned} \delta_1 &\approx -\zeta_1, \\ |\delta_2 - a_{21} \zeta_1| &\leq \beta \Delta_2. \end{aligned} \quad (27)$$

3. Аналогичное рассуждение для случая грубой ошибки ζ_n на крайнем верхнем уровне $k=n$ приводит к следующему правилу диагноза.

3.1. Допустимые невязки превышены на двух верхних уровнях:

$$|\delta_n| > \Delta_n, \quad |\delta_{n-1}| > \Delta_{n-1}. \quad (28)$$

3.2. Знаки δ_n и δ_{n-1} противоположны.

3.3. Невязки подчиняются соотношениям

$$\begin{aligned} \delta_n &\approx -\zeta_n, \\ |\delta_{n-1} - b_{n-1, n} \zeta_n| &\leq \beta \Delta_{n-1}. \end{aligned} \quad (29)$$

4. Более сложной является ситуация, когда одновременно имеются две грубые ошибки на смежных уровнях $k, k+1$. Тогда происходит изменение фактических невязок даже не на трех, как в случае единственной ошибки, а на четырех уровнях. Можно показать, что при этом

$$\begin{aligned} \delta_k &= \delta_k - \zeta_k + b_{k, k+1} \zeta_{k+1}, \\ \delta_{k-1} &= \delta_{k-1} + b_{k-1, k} \zeta_k, \\ \delta_{k+1} &= \delta_{k+1} - \zeta_{k+1} + a_{k+1, k} \zeta_k, \\ \delta_{k+2} &= \delta_{k+2} + a_{k+2, k+1} \zeta_{k+1}. \end{aligned} \quad (30)$$

В таком случае целесообразно ограничиться при диагнозе проверкой условий

$$\begin{aligned} |\delta_k| &> \Delta_k, \quad |\delta_{k+1}| > \Delta_{k+1}, \\ |\delta_{k-1}| &> \Delta_{k-1}, \quad |\delta_{k+2}| > \Delta_{k+2}. \end{aligned} \quad (31)$$

Исправление обнаруженных грубых ошибок сводится к замене искаженных значений результатами интерполяции или экстраполяции по правильным данным.

Расчитанные параметры и сформулированные выше диагностические правила вертикального контроля нуждаются в проверке

Таблица 3

Примеры вертикального контроля

№ пп.	Уровень гПа	\hat{t}_k	t'_k	\hat{t}_k	ξ_k	δ_k	Δ_k	δ_{k-1}	Δ_{k-1}	δ_{k+1}	Δ_{k+1}
1	250	35,6	-35,6	-35,7	71,2	-71,3	6,8	19,9	5,4	36,4	7,0
2	150	56,0	-56,0	-61,4	112,0	-117,4	6,6	61,2	7,0	56,0	4,9
3	400	30,8	-30,8	-29,2	61,6	-60,0	4,8	29,7	5,5	44,6	5,4
4	150	60,8	-60,8	-56,7	121,6	-117,5	6,6	68,0	7,0	55,6	4,9
5	150	54,2	-54,2	-52,1	108,4	-106,3	6,6	60,9	7,0	48,8	4,9
6	300	21,0	-21,0	-21,5	42,0	-42,5	5,4	23,9	4,8	27,3	6,8
7	400	9,0	-9,0	-11,1	18,0	-20,1	4,8	10,4	5,5	14,4	5,4
8	300	32,0	-35,9	-38,7	67,9	-70,7	5,4	41,2	4,8	42,7	6,8
9	200	45,2	-45,2	-44,0	90,4	-89,2	7,0	47,4	6,8	21,3	6,6
10	200	52,6	-52,6	-51,3	105,2	-103,9	7,0	54,9	6,8	31,5	6,6
11	150	43,2	-43,2	-44,9	86,4	-88,1	6,6	47,9	7,0	39,5	4,9
12	400	-21,3	-30,7	-30,2	9,4	-8,9	4,8	3,3	5,5	7,0	5,4
13	400	-20,3	-30,9	-31,2	10,6	-10,9	4,8	5,1	5,5	10,5	5,4
14	250	-46,5	-36,2	-38,2	-10,3	8,3	6,8	-3,8	5,4	-1,4	7,0
15	250	-56,5	-42,2	-47,8	-14,3	8,7	6,8	-3,3	5,4	-3,8	7,0
16	200	-57,9	-40,6	-35,1	-17,3	22,8	7,0	-12,3	6,8	-17,6	6,6
17	300	-49,3	-38,5	-39,0	-10,8	10,4	5,4	-4,8	3,8	-5,5	6,8
18	400	-9,1	-18,6	-19,7	9,5	-10,6	4,8	4,6	5,5	8,2	5,4
19	150	-64,1	-54,1	-56,8	-10,0	7,3	6,6	-2,2	7,0	-6,1	4,9
20	250	-16,9	-26,9	-26,1	10,0	-9,2	6,8	2,4	5,4	4,5	7,0
21	400	-17,5	-27,5	-26,9	10,0	-9,4	4,8	5,5	5,5	7,2	5,4
22	300	-26,1	-35,2	-36,4	9,1	-10,3	5,4	6,0	4,8	3,8	6,8
23	200	-13,1	-51,4	-54,6	38,3	-41,6	7,0	18,9	6,8	21,6	6,6
24	150	-20,5	-56,8	-55,4	36,3	-34,8	6,6	24,5	7,0	15,8	4,9
25	200	-25,3	-53,8	-55,6	28,5	-30,3	7,0	13,0	6,8	17,0	6,6
26	400	-55,7	-14,0	-16,5	-41,7	39,2	4,8	-17,7	5,5	-31,2	5,4
27	300	-40,7	-13,1	-15,5	-27,6	25,2	5,4	-14,8	4,8	-15,5	6,8
28	400	-50,5	-3,9	-7,1	-46,6	43,4	4,8	-21,7	5,5	-32,6	5,4
29	200	-26,1	-58,0	-49,5	31,9	-23,4	7,0	13,8	6,8	5,8	6,6
30	150	-17,1	-60,0	-55,0	42,9	-37,9	6,6	21,7	7,0	21,2	4,9
31	200	26,0	-64,7	-55,1	90,7	-81,1	7,0	-41,1	6,8	-28,1	6,6

на фактическом материале. С этой целью мы воспользовались оперативными аэрологическими телеграммами, поступившими в Гидрометцентр СССР в сентябре 1976 г. Содержащиеся в них сведения о геопотенциале и температуре на изобарических поверхностях были предварительно подвергнуты комплексному статико-временному контролю (СВК) [2]. Проверенные и исправленные таким образом значения температуры t' мы условно приняли за истинные. По отношению к t' определялись искажения $\xi = \hat{t} - t'$ исходных значений температуры \hat{t} . С другой стороны, с \hat{t} и t' сравнива-

лись проинтерполированные по вертикали значения t и вычислялись невязки независимого вертикального контроля (ВК).

Всего было рассмотрено 67 случаев изолированных грубых ошибок, из которых 34 случая относились к промежуточным уровням и 33 — к крайним уровням 850 и 100 гПа. Диагноз ВК и СВК совпал в 64 случаях. В табл. 3 приведены примеры вертикального контроля промежуточных уровней. В первых 11 примерах искажен знак, в других примерах — значения температуры.

Разумеется, для суждения об эффективности вертикального контроля необходимы более широкие испытания и накопление опыта в оперативных условиях. Однако уже сейчас ясна его полезность как одного из методов в рамках комплексного подхода. По-видимому, вертикальный контроль целесообразно применять, в первую очередь, к ветру и влажности, поскольку температура и геопотенциал могут контролироваться значительно более мощным методом статического контроля.

СПИСОК ЛИТЕРАТУРЫ

1. Либерман Ю. М. Параметры статистического контроля вертикальных профилей геопотенциала и температуры. — Труды ЗСРНИГМИ, 1978, вып. 29, с. 62—67.
2. Либерман Ю. М., Тараканова В. П. Опытная оценка качества оперативной аэрологической информации. — Метеорология и гидрология, 1979, № 10, с. 68—71.
3. Мерцалова О. Б. Стандартизация вертикальных корреляционных связей температуры над северным полушарием. — Труды НИИАК, 1967, вып. 46, с. 129—148.
4. Мерцалова О. Б. Стандартизация вертикальных корреляционных связей давления. — Труды НИИАК, 1972, вып. 79, с. 37—53.
5. Хвостова Р. Н. Вертикальная статистическая структура поля температуры в свободной атмосфере над территорией СССР. — Труды ВНИИГМИ—МЦД, 1978, вып. 47, с. 14—34.

Л. Л. Брагинская

ОБ ЭКОНОМИЧЕСКИ ОПТИМАЛЬНОМ ПОДХОДЕ ПРИ ПЛАНИРОВАНИИ НОРМ ОРОШЕНИЯ ХЛОПЧАТНИКА

В условиях Средней Азии из-за острого дефицита влаги продуктивность хлопчатника зависит в первую очередь от обеспеченности растений влагой в течение вегетационного периода. При этом создание оптимальных условий водоснабжения посевов, обеспечивающих нормальный рост и развитие растений в течение всего периода вегетации, тесно связано с решением вопроса о сроках и нормах орошения.

Существующие в настоящее время оросительные системы не могут полностью удовлетворить хозяйственные потребности в воде. Поэтому в маловодные годы орошают в первую очередь особо ценные и нуждающиеся в воде культуры. В последние годы ведут-

ся большие работы, связанные с развитием оросительных систем и расширением площади орошаемого земледелия.

При планировании оросительных систем крайне важно знать, на какую мощность они должны быть рассчитаны, какие потребности в воде должны удовлетворять. В этой связи весьма важное значение приобретает разработка экономически оптимальных решений при выборе норм орошения. В работах [1, 3] такая задача рассмотрена применительно к планированию норм орошения яровой пшеницы в Кулундинской степи (юг Западной Сибири).

В [4] предложена модель, описывающая связь между хозяйственным урожаем хлопка-сырца и расходами влаги в течение вегетационного периода. Эта связь, как и для других культур, имеет параболическую форму. Хозяйственный урожай хлопка-сырца U растет по мере увеличения расходов влаги до 600—650 мм; при дальнейшем увеличении количества влаги урожай начинает уменьшаться, что обуславливается, по-видимому, избыточным увлажнением почвы, при котором происходит чрезмерное развитие вегетативной массы в ущерб плодообразованию.

В пределах расхода влаги E от $E_1=330$ мм до $E_2=750$ мм указанная связь хозяйственного урожая хлопка-сырца U (ц/га) с E (мм) за вегетационный период имеет вид

$$U = -0,00038 E^2 + 0,47 E - 111,5. \quad (1)$$

Будем исходить из того, что при значении $E \approx 600$ мм урожай достигает максимального значения U_m . Естественно, эти значения U_m существенно различны для разных хозяйств; примем его в соответствии с данными работы [4] для среднеспелых сортов хлопчатника равным 34 ц/га. Таким образом, зависимость урожая хлопка-сырца U от суммарного водопотребления выражается функцией вида

$$U = \begin{cases} 0 & \text{при } E < E_1, \\ -0,00038 E^2 + 0,47 E - 111,5 & \text{при } E_1 < E < E_2, \\ U_m & \text{при } E > E_2. \end{cases} \quad (2)$$

Примем, что доход от реализации урожая пропорционален величине урожая, а затраты на орошение пропорциональны величине оросительной нормы. Отсюда следует, что при любой естественной влагообеспеченности вегетационного периода чистый доход от реализации урожая максимален, если оросительную норму N планировать исходя из перспективы получения максимального урожая:

$$N = E_2 - E_0, \quad (3)$$

где E_0 — водопотребление, обеспечиваемое естественными условиями увлажнения.

Величина E_0 заранее не известна, так как для ее расчета необходимо знать запасы продуктивной влаги в начале и в конце вегетационного периода, а также сумму выпавших за этот период осадков. Поэтому вопрос о планировании оросительной нормы приходится решать в условиях неопределенности, т. е. ориенти-

руясь на некоторое предполагаемое значение $E = \bar{E}$ и вычисляя норму орошения

$$N = 600 - E. \quad (4)$$

В этом случае возможны два рода ошибок. Если фактическое водопотребление, обеспечиваемое естественными условиями увлажнения, оказалось меньше величины \bar{E} , то вследствие недостатка влаги урожай снизится на величину, пропорциональную разности $(\bar{E} - E)$. В случае когда фактическое водопотребление, обеспечиваемое естественными условиями увлажнения, оказывается выше предполагаемой величины \bar{E} , имеет место непроизводительный расход воды, равный разности $(E - \bar{E})$. При этом, как показано в [3], функция потерь из-за несоответствия фактических и естественных условий влагообеспеченности является разностной, линейной и несимметричной и описывается формулой

$$L(E, \bar{E}) = \begin{cases} A_1 (\bar{E} - E) & \text{при } E \leq \bar{E}, \\ A_2 (E - \bar{E}) & \text{при } E > \bar{E}, \end{cases} \quad (5)$$

где A_1 и A_2 — постоянные коэффициенты, характеризующие экономическую значимость отклонений разных знаков.

Априори ясно, что значение коэффициента A_1 во много раз превышает значение A_2 . По косвенным оценкам A_1 в десять раз больше A_2 , что хорошо согласуется с полученным в [3] соотношением коэффициентов в функции потерь из-за несоответствия фактических и предполагаемых условий влагообеспеченности яровой пшеницы в условиях Кулундинской степи.

В качестве критерия эффективности различных стратегий потребителей метеорологической информации наиболее часто используются показатели средних потерь и среднего выигрыша [2, 3]. Однако не всегда этот принцип является наиболее приемлемым. Во многих случаях целесообразно при выработке оптимальных хозяйственных решений потребителями метеорологической информации использовать подход, основанный на минимизации повторяемости крупных потерь. Этот подход был ранее использован в работе [1].

При таком подходе оптимальное значение метеозлемента, на которое рекомендуется ориентироваться потребителю, находится из условия, что вероятность превышения некоторой заданной величины потерь минимальна. В случае когда можно считать, что ежегодные колебания метеорологического параметра подчиняются нормальному закону распределения, а функция потерь описывается формулой (5), дифференциальная функция распределения потерь имеет вид

$$\gamma(\lambda, \xi) = \frac{1}{\sqrt{\pi}} \left[\lambda e^{-(\lambda \xi)^2} + \frac{1}{\xi} e^{-\left(\frac{\lambda}{\xi} + \xi\right)} \right], \quad (6)$$

где $\lambda^2 = A_2/A_1$ — отношение потерь при отрицательной и положительной ошибках; $\xi = (E - \bar{E})/\sqrt{2}\sigma$ — безразмерное значение E ; $\lambda = L/\sigma\sqrt{2A_1A_2}$ — безразмерное значение потерь L ; $\gamma = \sigma\sqrt{2A_1A_2}g$ — дифференциальная функция распределения безразмерных потерь λ .

На рис. 1 приведена безразмерная дифференциальная функция распределения безразмерных потерь $\gamma(\lambda)$ при различных стратегиях потребителя для $\kappa=1$ и $\kappa=0,3$. Сравнивая между собой кривые 1 и 2, соответствующие симметричной функции потерь ($\kappa=1$), можно видеть, что в этом случае оптимальной является ориента-

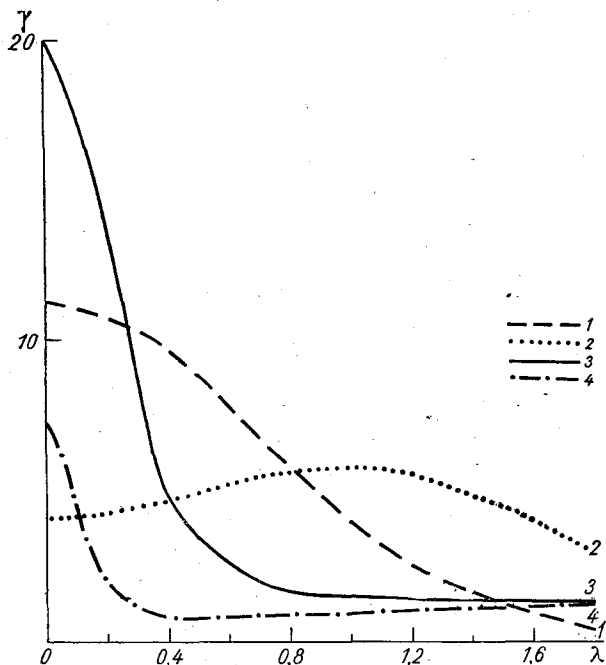


Рис. 1. Безразмерная дифференциальная функция распределения безразмерных потерь $\gamma(\lambda)$.
 1) $\kappa=1, S=0$; 2) $\kappa=1, S=1$; 3) $\kappa=0,3, S=0$; 4) $\kappa=0,3, S=1$.

ция на норму. При ориентации на норму увеличивается повторяемость малых потерь и уменьшается повторяемость крупных потерь.

В случае несимметричной функции потерь $\kappa=0,3$ (кривые 3 и 4) при ориентации на норму увеличивается повторяемость не только малых, но и очень крупных потерь. В этом случае целесообразно не ориентироваться на норму.

Рассмотрим подробно различные стратегии потребителя на примере ст. Бозсу. В качестве исходных данных примем $\bar{E}=110$ мм, $\sigma=62$ мм, $A_1=1$ (р.га)/мм, $A_2=0,1$ (р.га)/мм.

1. Потребитель ориентируется всегда на одно и то же значение оросительной нормы, считая, что водопотребление, обеспечиваемое естественными условиями увлажнения, будет равно среднему многолетнему водопотреблению

$$\bar{E} = \bar{E},$$

а оросительная норма, рассчитанная по формуле $N=600-\bar{E}$, соответственно составит $N=490$ мм.

Согласно [2], при такой стратегии средние потери R

$$R_n = (A_1 + A_2) \frac{\sigma}{\sqrt{2\pi}} \quad (7)$$

оказываются равными $R_n=27,2$ р./га.

2. В ряде работ показано, что значительно более рациональное планирование может быть осуществлено на основе оптимального использования агроклиматологической информации о распределении E . В этом случае планирование оросительной нормы ведут, ориентируясь на некоторое значение E_0 , при котором средние потери минимальны.

При нормальном распределении $f(E)$ климатологически оптимальное значение E_0 находится из условия

$$\operatorname{erf}(t) = \frac{A_2 - A_1}{A_2 + A_1}, \quad (8)$$

где $t = \frac{E_0 - \bar{E}}{\sqrt{2}\sigma}$ — выраженное в долях среднего квадратического отклонения смещение величины E_0 относительно \bar{E} , а $\operatorname{erf}(t)$ — функция Крампа

$$\operatorname{erf}(t) = \frac{2}{\sqrt{\pi}} \int_0^t e^{-U^2} dU. \quad (9)$$

Что касается знака отклонения оптимального значения от нормы, то он будет положительным в случае $A_2 > A_1$ и отрицательным при $A_2 < A_1$. Иначе говоря, если потери при отрицательных ошибках больше, чем при положительных, то значения E_0 , минимизирующие повторяемость средних потерь, оказываются больше нормы. Если же потери при положительных ошибках больше, чем при отрицательных, то значение E_0 меньше нормы \bar{E} .

Оказывается, $E_0=27$ мм, что существенно меньше нормы. Оросительную норму (мм) в этом случае следует планировать следующим образом:

$$N = 600 - 27 = 573.$$

При такой стратегии средние потери, рассчитываемые по формуле

$$R_{\text{кл}} = R_n \exp(-t^2) \quad (10)$$

равны 11,1 р./га.

При ориентации на климатологически оптимальное значение мы сокращаем средние потери больше чем в два раза.

3. Потребителю требуется минимизировать повторяемость крупных потерь.

На рис. 2 приведена интегральная функция распределения потерь потребителя для разных значений потерь. Как видно из рис. 2, при любом L существует единственное значение E , при котором

$G(L, \bar{E})$ максимально. С увеличением L величина E сдвигается в сторону меньших значений, т. е. чем больше потери мы минимизируем, тем на меньшее значение естественной увлажненности мы должны ориентироваться. При этом вероятность того, что потери не превысят это заданное значение с увеличением L растет.

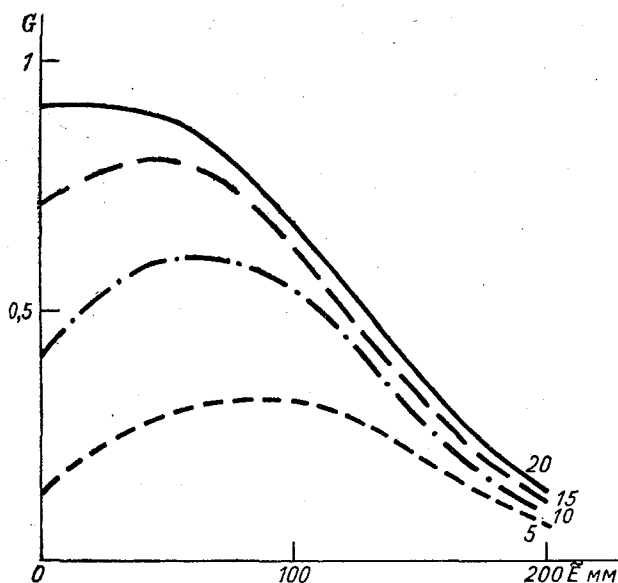


Рис. 2. Интегральная функция распределения потерь потребителя для разных значений потерь.

На рис. 3 приведены оптимальные значения \bar{E} (мм) и соответствующие им значения оросительной нормы N при разных значениях минимизируемой величины крупных потерь L .

Существует единственное значение $L=L_k$, такое, что минимизация повторяемости потерь, больших L_k , эквивалентна минимизации средних потерь. Значение L_k определяется из соотношения

$$\operatorname{erf}\left(\frac{A_2 - A_1}{2\sqrt{2} A_1 A_2 \sigma} L_k\right) = \frac{A_2 - A_1}{A_1 + A_2}. \quad (11)$$

Оказалось, что $L_k = 18,4$ р./га. Это означает, что при минимизации повторяемости потерь, больших 18,4 р./га, средние потери равны минимальной величине, составляющей 11,1 р./га.

Нужно ли задавать большее значение L_k , чтобы оградить потребителя от повторяемости очень крупных потерь? Представляется целесообразным ограничиться минимизацией средних потерь. Это объясняется тем обстоятельством, что в данном конкретном случае из-за существенной несимметричности функции потерь и значительной изменчивости естественной влагообеспеченности кли-

матологически оптимальное значение E_0 оказалось значительно ниже нормы, а ориентация на бoльшие значения крупных потерь приведет к сдвигу \tilde{E} в сторону еще меньших значений, что не имеет смысла. Следует учесть также и то, что вблизи насыщающих значений E зависимость урожая от водопотребления ослабевает.

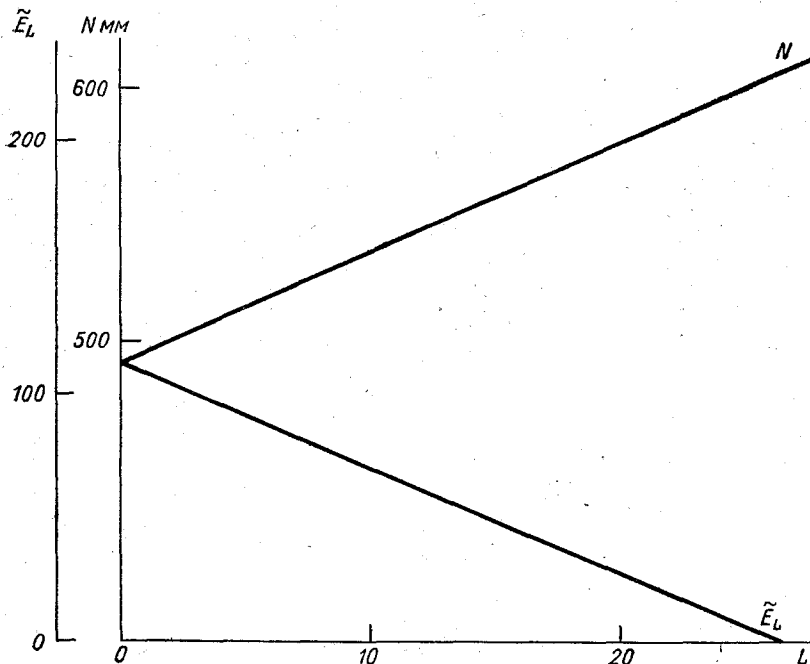


Рис. 3. Зависимость оптимального значения E_L и нормы N от величины крупных потерь L .

Такие результаты имеют место для данной культуры и конкретных агрометеорологических условий. Если бы исследовалась другая культура и другие агрометеорологические условия, результаты были бы иные.

СПИСОК ЛИТЕРАТУРЫ

1. Брагинская Л. Л. О применении принципа минимизации крупных потерь при назначении оросительных норм. — Труды ГГО, 1980, вып. 412, с. 119—127.
2. Жуковский Е. Е., Чудновский А. Ф. Методы оптимального использования метеорологической информации при принятии решений. — Л.: Гидрометеониздат, 1978. 52 с.
3. Жуковский Е. Е., Чудновский А. Ф. Оптимальные решения при назначении оросительных норм. — Труды УкрНИГМИ, 1973, вып. 128, с. 50—57.
4. Муминов Ф. А., Абдуллаев А. К. Агрометеорологическая оценка влагообеспеченности посевов хлопчатника. — Труды СарНИГМИ, 1974, вып. 18 (99). 84 с.

ОБ АВТОМАТИЧЕСКОЙ КЛАССИФИКАЦИИ ВЕРТИКАЛЬНЫХ ПРОФИЛЕЙ ВЕТРА НА ОСНОВЕ ГЛАВНЫХ КОМПОНЕНТ

Возможность оперировать с небольшим ансамблем признаков существенно облегчает решение таких задач, как выявление характерных особенностей элементов рассматриваемого архива случайных функций (полей) и разделение его на классы (типы), подбор аналогов и других. Для моделирования процесса принятия решений о схожести и различии объектов классификации весьма перспективным представляется сочетание алгоритмов самообучения [2, 6] и метода эмпирических ортогональных функций (ЭОФ) [1, 4]. Последний позволяет получить статистически оптимальное описание случайных полей метеорологических величин, т. е. сконцентрировать основную долю суммарной дисперсии исходной информации в сравнительно небольшом числе первых членов разложения и, таким образом, проводить классификацию по небольшому числу информативных признаков. Оптимальное представление полей особенно необходимо при изучении векторных величин.

В настоящей работе предпринята попытка осуществить объективную (автоматическую) типизацию вертикальных профилей вектора скорости ветра, описываемых коэффициентами разложения по ЭОФ, или главными компонентами.

Исходной информацией для выполнения исследования послужили результаты зондирования атмосферы проведенного в течение января в 3 ч за десятилетний период (1969—1978 гг.) на ст. Фрунзе, находящейся на высоте 760 м над уровнем моря. Измерения скорости и направления ветра производились на высоте флюгера и на семи изобарических поверхностях: 850, 700, 500, 400, 300, 200, 100 гПа. Вся совокупность данных разбивалась на три части так, чтобы не нарушалась положительная неопределенность ковариационной матрицы из-за высокой корреляции профилей за соседние даты. Первая выборка включала 96, вторая — 92, третья — 97 профилей зональной (u) и меридиональной (v) составляющих вектора скорости. По данным каждой выборки были построены корреляционные матрицы, описывающие межуровневую связь для названных составляющих, и вычислены их собственные элементы.

Один из принципиальных вопросов, связанных с применением метода ЭОФ, состоит в определении числа членов ряда, которые целесообразно учитывать. В связи с этим рассмотрим оценку точности сходимости выполненных разложений. В табл. 1 приведены оценки доли общей дисперсии (%), приходящейся на отдельные члены рядов разложения u и v по ЭОФ (см. параметр α_j , где j — порядковый номер слагаемого в ряду), и их нарастающая сумма β_j . Будем полагать, что при $\alpha_j < 5\%$ слагаемые описывают флуктуации метеорологических шумов, и ограничимся рассмотрением

первых пяти членов каждого ряда, которые в сумме несут 92—94 % общей дисперсии исходных данных о зональной составляющей ветра и 89—91 % — о меридиональной.

Учитывая, что каждая выборка обладает определенными статистическими особенностями, для u и v были построены унифициро-

Таблица 1

Оценка точности сходимости разложений зональной и меридиональной составляющих ветра в ряд по ЭОФ

j	Номер выборки					
	I		II		III	
	α	β	α	β	α	β
Зональная составляющая						
1	48,82	48,82	48,02	48,02	48,79	48,79
2	18,68	67,50	14,18	62,20	14,39	63,18
3	11,68	79,18	12,49	74,69	13,49	76,67
4	9,62	88,80	10,34	85,03	9,36	86,06
5	5,61	94,41	7,48	92,51	7,44	93,47
6	3,09	97,50	3,40	95,91	4,18	97,65
7	1,94	99,44	2,51	98,42	1,50	99,15
8	0,56	100,00	1,58	100,00	0,86	100,00
Меридиональная составляющая						
1	46,90	46,90	39,66	39,66	44,49	44,49
2	14,51	61,41	16,29	55,95	16,21	60,70
3	12,45	73,86	11,86	67,81	12,70	73,40
4	11,04	84,90	10,19	78,00	10,38	83,78
5	6,48	91,38	9,95	87,95	6,76	90,54
6	5,35	96,73	6,19	94,14	4,80	95,34
7	1,72	98,45	3,08	97,22	3,35	98,69
8	1,55	100,00	2,78	100,00	1,31	100,00

ванные матрицы, элементами которых являются средние из трех выборок коэффициенты межуровневой корреляции. Собственные векторы таких матриц послужили базовыми функциями для расчета коэффициентов разложения $T_{ig}^{u,v}$ (g — номер даты в общей выборке), содержащих информацию о размахе естественных колебаний и являющихся случайными функциями времени.

Систематизация эмпирических данных об изучаемых профилях ветра осуществлялась с помощью одной из адаптивных модификаций алгоритмов самообучения [5], позволяющей оптимальным образом учесть пространственно-временные масштабы естественных колебаний.

В самой общей форме задача типизации формулировалась, как обычно [6], следующим образом: требуется разбить совокупность профилей на классы так, чтобы профили, объединенные в один класс по некоторым признакам, отличались от профилей любого

другого класса, но внутри данного класса это отличие было бы несущественным. «Индивидуальность» каждого профиля в случае разложения в ряд по ЭОФ характеризуется коэффициентами $T_{jg}^{u,v}$ (главными компонентами), изменение которых во времени более закономерно, чем изменение данных наблюдений, не прошедших математическую обработку [4]. Следовательно, априори можно утверждать, что использование главных компонент в качестве признаков, по которым производится типизация, будет давать преимущество в получении более дифференцированных классов, чем при использовании «сырой» информации.

Таким образом, чтобы разбить архив N профилей вектора скорости ветра на заранее не заданное число типов, каждое состояние поля движения кодируется набором m признаков:

$$X_t = (x_{t1}, x_{t2}, \dots, x_{tm}) \in A, \quad (1)$$

где x_{tk} — k -я компонента (признак) вектора X_t , представляющая собой коэффициент $T_{jg}^{u,v}$ разложения исследуемого профиля по ЭОФ; A — множество N векторов X_t (выборка).

Для признаков, имеющих непрерывный интервал изменения, в качестве метрики m -мерного пространства, на основе которой изучается схожесть элементов архива, целесообразно использовать евклидово расстояние

$$\rho_{st} = \rho(X_s, X_t) = \left[\sum_{i=1}^m (x_{si} - x_{ti})^2 \right]^{1/2}.$$

В этом случае геометрическим образом типа является сфера в m -мерном пространстве коэффициентов разложения по ЭОФ. Следовательно, классификация заключается в построении покрытия элементов архива непересекающимися эталонами на основе оптимально подобранных m -мерных сфер.

Задание указанной меры близости профилей в пространстве признаков позволяет свести процесс их объективной группировки к поиску максимума функционала качества типизации $W(\Pi_r, A)$ на множестве покрытий $\{\Pi_r\}$ векторов архива A :

$$W^*(\Pi_r, A) = \sup_r W(\Pi_r, A), \quad (2)$$

где $W^*(\Pi_r, A)$ — оптимальное значение функционала качества типизации; r — параметр, характеризующий покрытие Π_r с учетом выбранной меры близости.

Синтез алгоритма автоматической классификации можно разбить на три последовательно решаемые задачи [5]. Первая задача включает вычисление значения параметра алгоритма r , в соответствии с которым элементы архива X_t относятся к различным типам, и задание способа его дальнейшего изменения при поиске экстремума критерия качества $W(\Pi_r, A)$. В предлагаемом алгоритме расчет начального значения параметра r_0 осуществляется автоматически на основе анализа максимальных и минимальных расстояний между векторами выборки, что позволяет добиться

адаптации алгоритма к рассматриваемому архиву. Последующие изменения r при поиске экстремума $W(\Pi_r, A)$ соответствуют процедуре стохастической аппроксимации.

Вторая задача синтеза алгоритма автоматической классификации заключается в построении покрытия Π_r . С этой целью нами использован эффективный и достаточно простой в вычислительном отношении метод ранжирования по параметру r треугольной матрицы расстояний между элементами архива.

Третья задача сводится к определению пригодности построенного разделения множества A на типы с точки зрения критерия $W(\Pi_r, A)$. При выборе критерия учитываются требования к качеству типизации [2, 3, 6].

Для оценки качества классификации профилей ветра использовался критерий I , введенный А. А. Дорофеевом [2]:

$$I = I_1 - I_2, \quad (3)$$

где I_1 — нормированное среднее расстояние между различными классами, которое рассчитывается для классов с числом векторов X_i не менее пяти; I_2 — усредненное по классам расстояние между векторами внутри классов. Максимальному значению критерия I соответствует оптимальное разделение выборки на классы.

Типизация профилей ветра проводилась нами на основе адаптивного алгоритма [5] последовательно в пять этапов: на первом этапе классификация производилась по первым коэффициентам разложения $T_{1g}^{u,v}$, на втором этапе — по первым и вторым коэффициентам совместно и т. д. На последнем этапе пространство признаков заполнялось пятью коэффициентами для составляющих u и v соответственно. В связи с ограниченными возможностями оперативной памяти ЭВМ М-222, на которой производились все расчеты, в типизируемую выборку было включено 150 случаев за 1969—1974 гг. Данные за 1975—1978 гг. (135 профилей) составили экзаменационную выборку.

На первом этапе максимальное значение критерия I (оптимальная классификация) было получено при разделении двумерного пространства $T_1^u \times T_1^v$ на четыре класса, когда практически вся выборка (132 профиля, т. е. 88 %) вошла в основной первый класс. Поскольку применение объективной типизации позволяет разбить совокупность векторов на наиболее компактные и удаленные друг от друга классы, то обширность первого класса объясняется, по-видимому, тем, что всей выборке при ее описании с помощью $T_{1g}^{u,v}$ присуща некоторая общая особенность. Действительно, первые коэффициенты разложения зональной и меридиональной составляющих вектора скорости ветра, содержат информацию о наиболее крупномасштабных процессах в поле движения — естественных колебаниях, описывающих климатические особенности пункта.

Внесение коэффициентов T_2^u и T_2^v в пространство признаков привело к разделению архива на 11 классов. В основной, третий, класс вошло 116 случаев, в остальные — от 13 до 1. Для графиче-

ского анализа результатов классификации на различных этапах в двумерном пространстве $T_1^u \times T_1^v$ было построено распределение элементов выборки по классам. С помощью рис. 1 можно проследить за трансформацией областей типов от этапа к этапу. Сравнение областей классов, полученных на первом и втором этапах, по-

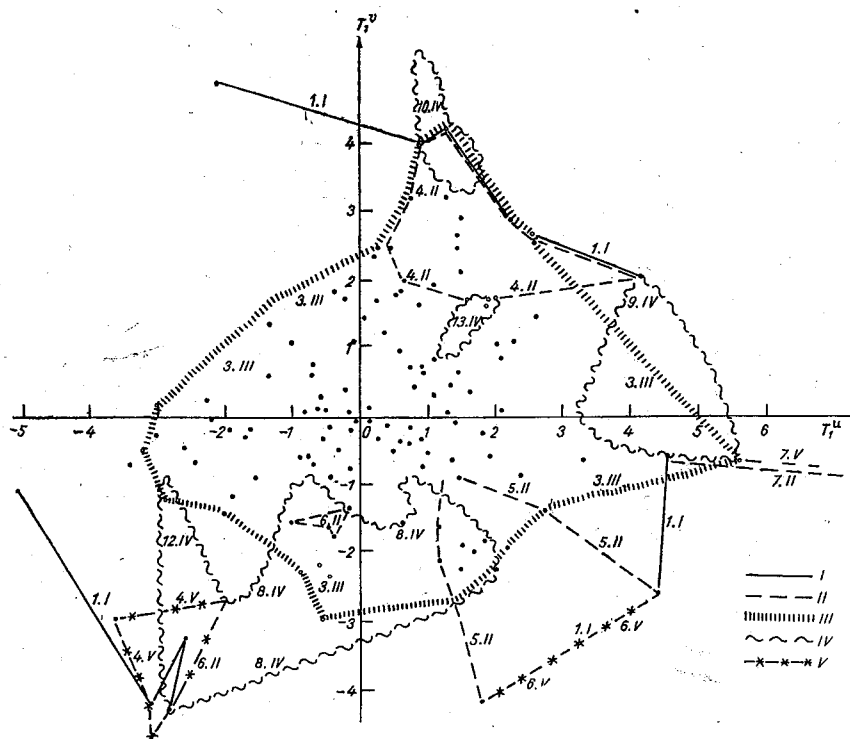


Рис. 1. Распределение векторов исходной выборки по классам в пространстве $T_1^u \times T_1^v$ на этапах I—V автоматической классификации по главным компонентам.

I) T_1^u, T_1^v ; II) $T_1^u - T_2^u, T_1^v - T_2^v$; III) $T_1^u \div T_3^u, T_1^v \div T_3^v$; IV) $T_1^u \div T_4^u, T_1^v \div T_4^v$; V) $T_1^u \div T_5^u, T_1^v \div T_5^v$.

Арабские цифры — классы, содержащие наибольшее число ситуаций, полученные на различных этапах (римские цифры) типизации.

зволяет сделать вывод о том, что первому классу этапа I соответствует третий класс этапа II, третьему — седьмой, четвертому — шестой; отмечается некоторое расширение малых классов за счет основного. Одновременно из основного класса выделяются четвертый и пятый типы.

При типизации по первым трем коэффициентам $T_{jg}^{u,v}$ выделилось восемь классов. Наиболее наполненным вновь оказался третий класс (130 профилей). Из рис. 1 видно, что в этом случае распределение профилей по классам очень схоже с типизацией на эта-

пе I , хотя имеет место некоторое смещение границ в сторону второго класса, полученного при систематизации профилей по первым коэффициентам разложения.

Учет коэффициентов $T_{4g}^{u, v}$ привел к разделению типизируемой выборки на 27 классов. Однако классы 1—6 и 8—27 содержат

Таблица 2

Оценки качества типизации по критерию А. А. Дорофеева

Параметры	Этапы типизации				
	<i>I</i>	<i>II</i>	<i>III</i>	<i>IV</i>	<i>V</i>
I_1	4,8	5,6	5,7	5,6	0,0
I_2	2,8	3,2	4,0	3,5	4,5
I	2,0	2,4	1,7	2,1	-4,5

лишь по одному профилю, большинство из которых являются, по видимому, своеобразными «выбросами», соответствующими нетипичным состояниям поля движения, поэтому рассматривать их нецелесообразно. На этапе *IV* произошла существенная трансформация областей типов в пространстве $T_1^u \times T_1^v$. В основной, седьмой, класс вошли 83 случая (55 %), в восьмой — 18, в остальные — от 6 до 1.

При использовании пяти коэффициентов T_{jg}^u, T_{jg}^v ($j=1, \dots, 5$) наблюдается резкое ухудшение типизации, выражающееся в размывании границ классов. Этот эффект объясняется, по нашему мнению, тем, что пятое естественное колебание описывает менее 10 % суммарной дисперсии и действие его при решении задачи типизации аналогично действию метеорологических шумов, т. е. приводит к искажению результатов.

В табл. 2 приведены оценки критерия качества I и его составляющих для всех этапов типизации. Поскольку максимальному значению I соответствует построение покрытия пространства R^m , в котором среднее расстояние между классами наибольшее, а элементы внутри классов сгруппированы компактно, то из табл. 2 следует, что при использовании евклидовой метрики на втором этапе получена наилучшая типизация. В принципе, этот вывод можно было ожидать, так как первые два члена разложения составляющих ветра несут 56—68 % суммарной дисперсии и содержат наиболее существенную информацию об изучаемых профилях.

Рассмотрим характерные особенности эталонных профилей, полученных на этапе *II* (рис. 2).

Отличительной чертой типа *I* является постепенный рост скорости ветра с высотой (с небольшим ослаблением в слое 300—200 гПа) и устойчивое западное направление. Для типа *II* характерно резкое усиление скорости ветра в слое 700—300 гПа и значительное ослабление ее в вышележащих слоях при устойчивом

юго-западном переносе практически во всей толще атмосферы. Для типа III свойственно неравномерное увеличение скорости западного переноса с высотой; для типа IV — слабое усиление скорости в слое 700—300 гПа и резкое — выше уровня 300 гПа, а также значительные изменения направления переноса от северо-восточно-

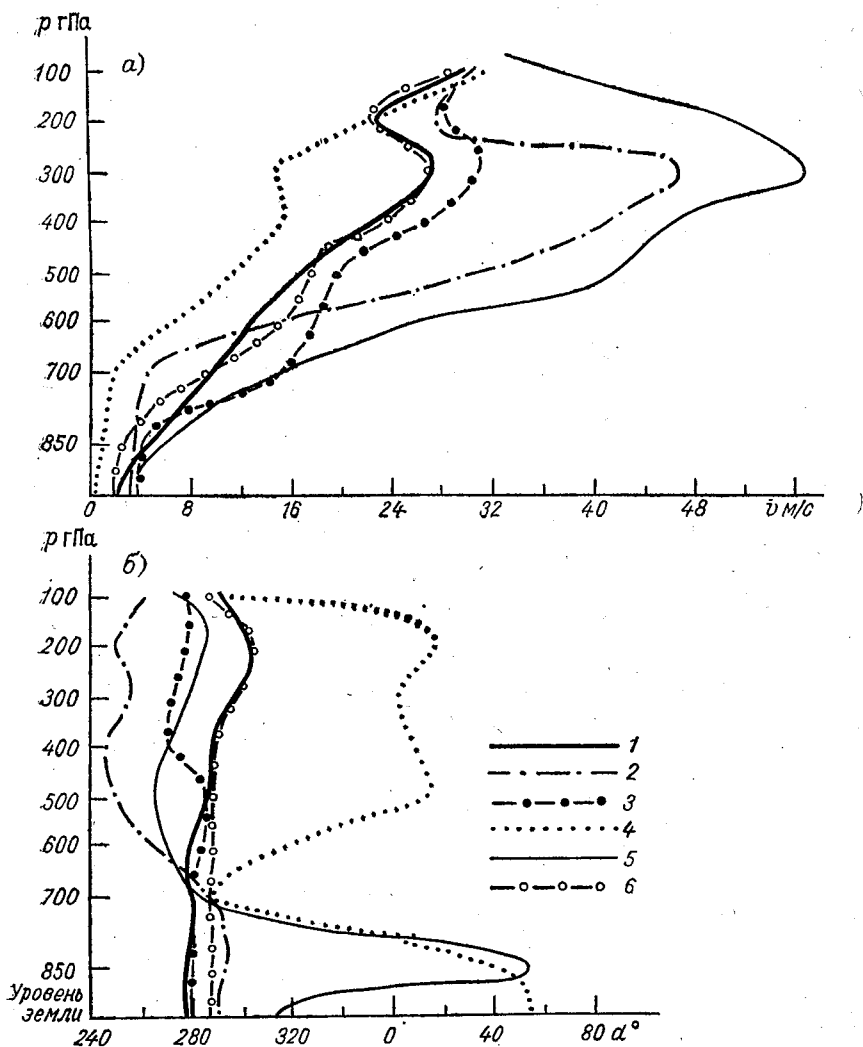


Рис. 2. Типовые вертикальные профили ветра, полученные при классификации по двум главным компонентам.

1, 2, ..., 6 — типовые вертикальные профили скорости ветра (а) и его направления (б) соответственно для первого—шестого классов.

го до западного в различных слоях атмосферы. Характерной чертой типа V является очень сильный рост скорости ветра начиная

от подстилающей поверхности вплоть до уровня 200 гПа и устойчивое западное направление в слое 700—100 гПа. Эталонный профиль типа VI очень близок к профилю типа I, имеют место лишь небольшие различия скорости и направления ветра от земли до поверхности 500 гПа.

Сравнение эталонных профилей, найденных на различных этапах классификации, показывает, что выделяются главным образом четыре класса, основной из которых включает в себя от 55 % (этап IV) до 88 % (этап I) типизируемой выборки. При переходе от этапа I к этапу II отмечается появление двух новых типов, характерные особенности которых описаны выше, т. е. происходит уточнение классификации за счет привлечения к решению задачи информации, описываемой вторым естественным колебанием — более мелкомасштабным, чем климатическое.

На этапе III мы получили результат, аналогичный классификации на этапе I. Указанный факт можно объяснить тем, что второе и третье естественные колебания находятся практически в противофазе и в значительной мере компенсируют друг друга в качестве весовых множителей.

Систематизация архива на этапе IV классификации проводилась с учетом еще более мелкомасштабных особенностей профилей ветра ($\alpha_4 = 10 \dots 12$ % суммарной дисперсии), поэтому число классов оказалось столь велико. Этап IV также характеризуется наличием основного седьмого класса, эталонный профиль которого хорошо согласуется с эталонными профилями основных классов на этапах I—III. Эталон девятого типа аналогичен эталону третьего при классификации по одной—трем главным компонентам. Кроме того, на этапе IV происходит организация новых классов (8—16), представляющих собой модификации основного класса, полученного на предыдущих этапах. Из них можно выделить группу с равномерным или неравномерным ростом скорости ветра с высотой (классы 8, 12, 14) и группу с резким усилением ветра до уровня 300 гПа и ослаблением его в более высоких слоях (классы 10, 13). Отдельно стоит одиннадцатый класс, отличающийся неустойчивыми колебаниями скорости в нижней и средней тропосфере и значительными усилениями ее в верхней тропосфере и стратосфере. Классы 15 и 19 тяготеют к девятому, поэтому их целесообразно объединить. Некоторые из остальных классов, содержащих по одному профилю, также можно отнести к той или иной уже описанной группе либо классифицировать как «выброс».

Итак, становится очевидным, что внесение в пространство признаков главных компонент с $j > 2$ приводит лишь к отысканию подтипов, к детализации оптимального варианта классификации.

Задавая в качестве эталонных профили, полученные при оптимальной классификации на основе первых двух главных компонент, мы подвергли систематизации экзаменационную выборку. Вне классификации осталось лишь 4 профиля из 135, т. е. не более 3 %, что также свидетельствует о представительности найденных типовых эталсонов.

Таким образом, сочетание методов ЭОФ и распознавания образов позволяет выделить классы вертикальных профилей ветра, отличающиеся друг от друга характером распределения скорости ветра с высотой, направлением и местоположением струйного течения. Оптимальным образом архив разбивается на классы при использовании в задаче типизации первых двух главных компонент $T_{jg}^{u,v}$ ($j=1, 2$) вертикальных профилей ветра, описывающих большую часть суммарной дисперсии.

Последовательное применение двух названных методов имеет перед оперированием непосредственно данными измерений следующие преимущества: производится фильтрация метеорологических шумов; существенно уменьшается число параметров, участвующих в типизации, или, иными словами, имеет место уменьшение размерности пространства признаков и, следовательно, сокращение объема вычислений при последующей классификации; типизации получают по признакам (естественным колебаниям) различного пространственного масштаба, что позволяет выбрать оптимальную из них с учетом необходимой детализации.

Разумеется, что адаптивный алгоритм типизации не изменится при классификации профилей на больших площадях, т. е. по некоторому множеству пунктов наблюдений. Главная трудность заключается в подготовке соответствующих исходных данных, а также в определении собственных значений корреляционной матрицы большого порядка.

СПИСОК ЛИТЕРАТУРЫ

1. Багров Н. А. Аналитическое представление последовательности метеорологических полей посредством естественных ортогональных составляющих. — Труды ЦИП, 1959, вып. 74, с. 3—15.
2. Дорофеев А. А. Алгоритмы автоматической классификации. — Автоматика и телемеханика, 1971, № 12, с. 73—113.
3. Еникеева В. Д. К вопросу о критериях качества автоматической типизации метеорологических ситуаций. — В кн.: Численное моделирование циркуляции в стратосфере, Новосибирск, 1976, с. 52—57.
4. Естественные составляющие метеорологических полей/Под ред. М. И. Юдина. — Л., Гидрометеоздат, 1970. — 200 с.
5. Солонин С. В., Еникеева В. Д. Адаптивные алгоритмы автоматической типизации гидрометеорологических явлений и процессов. — В кн.: Численное моделирование циркуляции в стратосфере, Новосибирск, 1976, с. 43—51.
6. Сонечкин Д. М. Математическая теория классификации и ее применение в метеорологии. — Метеорология и гидрология, 1969, № 12, с. 24—34.

СОДЕРЖАНИЕ

Л. В. Руховец, Е. И. Хлебникова. О научной деятельности Р. Л. Кагана	3
Список научных трудов Р. Л. Кагана	13
Г. В. Груза, Р. Л. Каган. Некоторые актуальные вопросы методики климатологической обработки данных	19
Р. Л. Каган. Об аппроксимации характеристик пространственной структуры осредненных величин	24
Р. Л. Каган, К. М. Лугина. О глобальном анализе средних месячных полей общего содержания озона в атмосфере	31
Р. Л. Каган, Е. И. Хлебникова. К вопросу о стационарности средних зональных значений метеорологических полей	38
Р. Л. Каган, Е. Е. Сибир. О корректной оценке средних суточных значений метеорологических параметров	43
М. И. Юдин. К определению качества метеорологической информации и максимально возможного эффекта от ее использования	47
В. М. Мирвис, М. И. Юдин. Применение информационного подхода к исследованию изменений крупномасштабных характеристик климата	54
А. В. Мещерская, В. Г. Блажевич, М. П. Голод, И. Г. Белянкина. Статистическая оценка информативности предикторов в зависимости от заблаговременности	63
А. Т. Житорчук, Ю. В. Житорчук. К вопросу о смещенности спектра собственных чисел выборочных корреляционных матриц	79
В. Б. Афанасьева, Н. П. Есакова, Р. В. Климентова. Статистические связи между индексом интенсивности зональной циркуляции и некоторыми метеорологическими параметрами	85
Ю. М. Либерман. О некоторых особенностях оптимальной интерполяции применительно к задачам усвоения метеорологических данных	88
Ю. М. Либерман, В. П. Тараканова. Об использовании прогностической информации при анализе метеорологических полей южного полушария	96
Л. П. Клягина, Ю. М. Либерман. Горизонтально-временной контроль приземного давления	98
Л. П. Клягина, Ю. М. Либерман. О вертикальном статистическом контроле температуры в свободной атмосфере	109
Л. Л. Брагинская. Об экономически оптимальном подходе при планировании норм орошения хлопчатника	116
Р. П. Релинская, В. Д. Еникеева. Об автоматической классификации вертикальных профилей ветра на основе главных компонент	123