

Методология рационализации региональных метеорологических сетей**Главная геофизическая обсерватория им. А.И.Воейкова , СПб**

О.М. Покровский

В последние годы возникла необходимость в обосновании развертывания новых и рационализации существующих систем мониторинга, обеспечивающих решение широкого круга задач охраны окружающей среды [8]. Для подразделений Росгидромета это касается, прежде всего, наземных, наблюдательных сетей (ННС). Проблемы с финансированием привели к заметному сокращению состава метеорологической и других сетей за последние два десятилетия. К сожалению, этот процесс подчас носил стихийный и неупорядоченный характер. В первую очередь сокращались станции не по соображениям их информативности, а в силу удобств снабжения и обеспечения кадрами. Это привело к закрытию ряда станций, имеющих достаточно высокий уровень информативности. В определенной степени спутниковая информация играет замещающую роль. Но, к сожалению, это касается не всех наблюдаемых параметров. Например, большая часть приземных величин, наблюдаемых на метеорологической сети, пока не замещается данными спутниковых наблюдений. Это не могло не сказаться на снижении оправдываемости прогнозов погоды, особенно регионального масштаба. Сейчас пришло время осмысления произошедших изменений в составе сети и постановки задачи ее восстановления в той ее части, где был ослаблен контроль погодообразующих процессов и их развитием на территории регионов. ГГО им. А.И. Воейкова традиционно занимается вопросами рационализации ННС как в части касающейся разработки методов, так и по их практическому применению. Методики ГГО в той или иной форме были связаны с методами численного анализа метеорологических полей [3, 6, 7, 11]. А в качестве критерия рационализации использовалась минимизация погрешности результатов анализа в узлы регулярной сетки [7]. Нельзя не отметить другие работы, посвященные данной проблеме [5, 12, 23]. При таком подходе имелась возможность сравнивать несколько вариантов состава

сети и определять наилучший среди них. Понятно, что такой подход позволит определить оптимальную сеть только в случае, если число станций не превышает 10-20. Если же в состав сети входит 80-100 станций, то обычный перебор всех возможных комбинаций станций, например, по 20-30 единиц станет практически нереализуемым. В этом случае поиск оптимума должен быть основан на более эффективных подходах. Численные методы оптимального планирования эксперимента предоставляют такую возможность [24]. Для их применения необходимо только создание информационной модели ННС. В предыдущих работах [13-21] нами рассмотрены отдельные элементы решения задачи рационализации ННС. К настоящему времени сформировался общий подход, определяющий алгоритм решения данной задачи на уровне УГМС. В данной работе изложены основы этого подхода и в качестве иллюстрации представлены некоторые результаты решения задачи для Камчатского УГМС.

1. Методология

Процедура рационализации региональной наземной, наблюдательной сети (ННС) включает в себя несколько обязательных этапов:

1. Объективное статистическое районирование и выделение информативно-однородных зон (ИОЗ)
2. Количественная оценка информативности ИОЗ и пунктов наблюдений (ПН)
3. Упорядочение ИОЗ и ПН по их информативности
4. Построение минимальной, наблюдательной сети
5. Анализ приращения информационного вклада каждой последующей станции и построение достаточной сети.

Первый этап предполагает разбиение области анализа (территории УГМС) на подобласти, ИОЗ [19], в пределах которых метеорологические поля важнейших величин в первом приближении ведут себя как однородные и изотропные, что позволяет применять стандартный статистический аппарат оценивания и анализа в каждой ИОЗ. Применение информационной

модели ННС [14-17], полученной для определенного набора метеорологических величин, позволяет осуществить количественную оценку информативности каждой ИОЗ и ПН (этап 2). Это открывает возможность на следующем этапе ранжировать ИОЗ и ПН по их информационной значимости. Одновременно может быть построена кривая информационного насыщения, которая позволяет определить номер ИОЗ или ПН в соответствующем списке. График этой кривой показывает вклад информации, поступающей от последующих ИОЗ или ПН. И если вклад дополнительной информации становится малым и сравнимым с влиянием ошибок наблюдений, которые также заложены в информационной модели ННС, то такие ИОЗ или ПН целесообразно исключить из дальнейшего анализа. После того как перечень информативных ИОЗ и пунктов наблюдений получен, можно перейти к построению вариантов оптимальной сети. С практической точки зрения важны два варианта сокращенной сети: минимальная и достаточная. Минимальная сеть предполагает минимальное представительство ИОЗ [20-21]: один ПН от каждой ИОЗ (этап 4). Понятно, что минимальная сеть не является оптимальной, но она, по крайней мере, обеспечивает равномерное представительство каждой из информативно значимых однородных зон и исключает потерю информации по какой-либо из ИОЗ. Этот вариант сети оправдан в случае ограниченного финансирования. Достаточная сеть [20, 21], по существу, является оптимальной (этап 5), поскольку в ней уже отсутствует какая-либо избыточность информации. Информативность ИОЗ определяется, прежде всего, средней изменчивостью рассматриваемых метеорологических величин на ее территории. Чем больше их изменчивость, тем больше пунктов наблюдений должно функционировать для слежения за погодообразующими процессами. Таким образом, достаточная сеть имеет неравномерное представительство по ИОЗ. Наиболее значимые зоны содержат несколько ПН, наименее значимые только один ПН.

2. Модели

Современный аппарат теории планирования эксперимента применим к регрессионной статистической модели системы наблюдений [24]. В работах [14, 16-18] предложен вариант построения регрессионной модели для системы наблюдений данного метеорологического поля с использованием базиса эмпирических ортогональных функций (ЭОФ) этого поля. В [14] для решения задачи отбора наиболее (наименее) информативных станций сети за основу взят метод оптимальной интерполяции, который имеет весьма широкое распространение в метеорологической практике. В данном разделе дается сравнение обоих подходов и оценка их эффективности в задачах оптимизации наблюдательных сетей.

2.1. Оптимизация сети на основе оптимальной интерполяции. Планировщик сети чаще всего сталкивается с задачами расширения или сокращения сети. Для решения таких задач необходимо иметь численный алгоритм, который мог бы ранжировать существующие станции по их информативности или определять наиболее благоприятное местоположение для новых станций. Поскольку для перевода данных наблюдений с сети на регулярную сетку обычно используется метод оптимальной интерполяции, можно использовать этот подход для восстановления значений метеорологического элемента непосредственно в точках наблюдений по данным остальных станций. Наименее информативной следует считать ту станцию сети, в которой указанное восстановление методом оптимальной интерполяции оказалось наиболее успешным (минимальная погрешность восстановления) [3, 14]. Эта станция и должна сокращаться в первую очередь. При расширении сети предпочтение должно быть отдано той точке, в которой погрешность восстановления поля (по данным имеющихся станций) максимальна. Формализуем постановку задачи. Пусть X - случайное поле исследуемого метеорологического элемента. Будем считать, что значение метеорологического элемента $x(w)$ в точке с координатами w представляет собой случайную величину с известным средним $\bar{x}(w) = E[x(w)]$. Будем также предполагать, что для любой совокупности точек w_1, \dots, w_k известна ковариационная матрица Σ_x с элементами

$$\{\sigma_x^2\}_{ij} = E[(x(w_i) - \bar{x}(w_i))(x(w_j) - \bar{x}(w_j))] \quad (i, j = 1, \dots, p).$$

Возьмем произвольную точку w_i и вычислим дисперсию ошибки оценки оптимальной интерполяции метеорологического элемента x в ней $\hat{x}(w_i/\Omega_{p/i})$ по данным остальных станций.

Здесь $\Omega_{p/i} = (w_1, \dots, w_{i-1}, w_{i+1}, \dots, w_p)$. Согласно классическим результатам (см. [1, 22, 24]),

имеем $\hat{x}(w_i/\Omega_{p/i}) = \bar{x}(w_i) + \Sigma_x(w_i, \Omega_{p/i}) \cdot \Sigma_x^{-1}(\Omega_{p/i}) \cdot [x(\Omega_{p/i}) - \bar{x}(\Omega_{p/i})]$. Здесь использованы

следующие обозначения: $x(\Omega_{p/i}) = [x(w_1), \dots, x(w_{i-1}), x(w_{i+1}), \dots, x(w_p)]^T$; $\Sigma_x(\Omega_{p/i})$ -

ковариационная матрица вектора $x(\Omega_{p/i})$; $\Sigma_x(w_i, \Omega_{p/i})$ - вектор-строка, составленная из

ковариаций случайной величины $x(w_i)$ и компонентов вектора $x(\Omega_{p/i})$. Здесь и ниже T -

знак транспонирования векторов и матриц. Дисперсия интересующей нас оценки $\hat{x}(w_i/\Omega_{p/i})$

имеет вид $\{\sigma_{\hat{x}}^2\}_{ii} = \{\sigma_x^2\}_{ii} - \Sigma_x(w_i, \Omega_{p/i}) \cdot \Sigma_x^{-1}(\Omega_{p/i}) \cdot \Sigma_x^T(w_i, \Omega_{p/i})$. Путем перенумерования

наблюдательных станций добьемся того, чтобы рассматриваемая точка наблюдений была

последней в наборе, т.е. именовалась w_p . Разделим симметричную ковариационную матрицу

на блоки. В этом случае правый нижний блок будет соответствовать выбранной точке

наблюдений. Теперь, используя известную формулу Фробениуса (см. [2]) для обращения

симметричной матрицы Σ_x , предварительно разбитой на блоки, можно найти соотношение

$$\{\sigma_x^{-2}\}_{pp} = 1/\{\sigma_{\hat{x}}^2\}_{pp}, \quad (1)$$

$(\{\sigma_x^{-2}\}_{ij})$ - элементы обратной матрицы Σ_x^{-1} : $\Sigma_x = \Sigma_x(\Omega_p); \Omega_p = (w_1, \dots, w_p)$, которое дает

простой метод для ранжирования наблюдательных станций по их информативности.

Действительно, согласно предыдущим соображениям, такое ранжирование должно

осуществляться по убыванию значений дисперсий погрешности оптимальной интерполяции

$\{\sigma_{\hat{x}}^2\}_{ii}$. Формула (1) позволяет упорядочить станции по возрастанию значений диагональных

элементов $\{\sigma_x^2\}_{ii}$ обратной матрицы Σ_x^{-1} после вычисления последней. Кроме того, указанный

подход позволяет построить алгоритм последовательного исключения наименее

информативных станций сети. Номер i -ой станции, подлежащей исключению из совокупности p станций, определяется по формуле $i = \arg \max \{\sigma_x^{-2}\}_{ll}$.

Рассмотрим теперь задачу расширения сети. Предположим, что станции w_1, \dots, w_{p-1} уже отобраны. Необходимо определить станцию w_p . С учетом вышеизложенного местоположение новой станции w_p необходимо определить из условия максимизации $\{\sigma_x^2\}_{pp}$. Из (1) следует, что задача отбора сводится к нахождению w_p , для которого достигается минимум $\{\sigma_x^{-2}\}_{pp}$.

Таким образом, решение обеих задач свелось к анализу диагональных элементов обратной матрицы Σ_x^{-1} . Рассмотрим более подробно структуру Σ_x . Результаты наблюдений $x(w)$ представляют собой сумму истинных значений метеорологического элемента и погрешностей измерений $x(w) = x_r(w) + \varepsilon(w)$. Можно считать, что $x_r(w)$ и $\varepsilon(w)$ статистически независимы. Объединим $x(w_i)$ в вектор $x(\Omega_p) = [x(w_1), \dots, x(w_p)]^T$. Тогда

$$\Sigma_x(\Omega_p) = \Sigma_{x_r}(\Omega_p) + \Sigma_\varepsilon(\Omega_p). \quad (2)$$

Поскольку наблюдения на сети проводятся однотипными приборами, обычно считают, что $\Sigma_\varepsilon(\Omega_p) = \sigma_\varepsilon^2 I$. Здесь I – единичная матрица. Последнее означает, что измерения равноточные и погрешности измерений статистически независимы. В дальнейшем для простоты изложения будем основываться на указанной модели $\Sigma_\varepsilon(\Omega_p)$. Тем не менее, все излагаемое ниже распространяется на случай матрицы $\Sigma_\varepsilon(\Omega_p)$ произвольной структуры. С помощью метода главных компонент [1, 22] ковариационная матрица $\Sigma_{x_r}(\Omega_p)$ может быть представлена в виде

$$\Sigma_{x_r} = F \cdot \Lambda \cdot F^T. \quad (3)$$

В (3) Λ - диагональная матрица $p \times p$, образованная собственными числами $\lambda_1, \dots, \lambda_p$ матрицы Σ_{x_r} , а матрица F образована векторами-столбцами f_1, \dots, f_p , являющимися соответствующими собственными векторами той же матрицы. В таком случае, используя (2), (3) и применяя удобные формулы для альтернативного представления обратных

ковариационных матриц для оценок метода наименьших квадратов [22] (см. подробности на с.24 в [13]), имеем

$$\Sigma_x^{-1} = (\Sigma_\varepsilon + F \cdot \Lambda \cdot F^T)^{-1} = \Sigma_\varepsilon^{-1} - \Sigma_\varepsilon^{-1} \cdot F \cdot (\Lambda^{-1} + F^T \cdot \Sigma_\varepsilon^{-1} \cdot F)^{-1} \cdot F^T \cdot \Sigma_\varepsilon^{-1}. \quad (4)$$

Как указывалось выше, $F = (f_1 | \dots | f_p)$. При этом вектор-столбец имеет вид $f_i = [f_i(w_1), \dots, f_i(w_p)]^T$. Обозначим через $f(w_i)$ - i -ю строку матрицы F . В этом случае

$$f(w_i) = [f_1(w_i), \dots, f_p(w_i)], \quad F = \begin{bmatrix} f(w_1) \\ \dots \\ f(w_i) \\ \dots \\ f(w_p) \end{bmatrix}.$$

С учетом введенных обозначений можно выписать выражение для i -го диагонального элемента матрицы Σ_x^{-1} . Используя (4), имеем

$$\{\sigma_x^{-2}\}_{ii} = \{\sigma_x^2\}_{ii}^{-2} \left[\{\sigma_x\}_{ii} - f(w_i) \cdot (\Lambda^{-1} + F^T \cdot \Sigma_\varepsilon^{-1} \cdot F)^{-1} \cdot f^T(w_i) \right]. \quad (5)$$

Из (5) следует, что процедура отбора наименее информативной точки w_i сводится к минимизации

$$S(w_i) = f(w_i) (\Lambda^{-1} + F^T \cdot \Sigma_\varepsilon^{-1} \cdot F)^{-1} \cdot f^T(w_i). \quad (6)$$

При решении задачи расширения сети необходимо искать точку w , доставляющую максимум $S(w_i)$. Таким образом, задача анализа диагональных элементов обратной матрицы Σ_x^{-1} может быть заменена проблемой нахождения экстремумов функционала вида (6). В дальнейшем будет показано, что выражение (6) характерно для представления "функции отклика" в теории планирования эксперимента для регрессионных моделей [24].

2.2. Оптимизация сети на основе построения регрессионной модели наблюдательной системы. Оптимальная интерполяция представляет собой метод объективного анализа метеорологических полей на сетке. В последние годы определенный интерес проявляется к методу спектрального анализа, предполагающему представление полей в виде разложения по определенной ортонормированной системе функций [9]. Остановимся на существовании метода.

Будем считать, что вектор $x(\Omega_p) = [x(w_1), \dots, x(w_p)]^T$ сформирован из результатов наблюдений, проведенных на p станциях w_1, \dots, w_p . Предположим, что в качестве базиса пространственного представления поля метеорологического элемента X применяется совокупность ортонормированных функций $\{g_i(w)\}_{i=1}^{\infty}$. Очевидно, что при численной реализации метода можно использовать только конечное число базисных функций $\{g_i(w)\}_{i=1}^N$. Введем в рассмотрение прямоугольную матрицу $G = \{g_{ji}\} = \{g_i(w_j)\}_{i=1}^N \{j=1}^p$ размерности $p \times N$. В таком случае задача спектрального анализа заключается в нахождении вектора спектральных коэффициентов $c = (c_1, \dots, c_N)^T$ из данных наблюдений, входящих в вектор $x(\Omega_p) = [x(w_1), \dots, x(w_p)]^T$. Неизвестные коэффициенты находят из системы уравнений:

$$x(\Omega_p) = G \cdot c + \varepsilon(\Omega_p). \quad (7)$$

В равенство (7) входит случайный вектор $\varepsilon(\Omega_p) = [\varepsilon(w_1), \dots, \varepsilon(w_p)]^T$, представляющий сумму погрешностей наблюдений и ошибки представления метеорологического поля конечной суммой базисных функций. При использовании достаточного числа членов разложения соответствующие ошибки становятся меньше погрешностей наблюдений. Будем считать, что $E[\varepsilon(\Omega_p)] = 0$ и $\Sigma_{\varepsilon}(\Omega_p)$ - ковариационная матрица $\varepsilon(\Omega_p)$. Уравнение (7) представляет собой регрессионную модель наблюдательной сети.

Методы теории планирования эксперимента предназначены для применения к функционалам от информационной матрицы Фишера (ИМФ). В случае модели (7) ИМФ имеет вид $M = G^T \cdot \Sigma_{\varepsilon}^{-1} \cdot G$. Матрица $M = M(\Omega_p)$ непосредственно связана с ковариационной матрицей Σ_{ε} для оценки вектора коэффициентов $\hat{c} = M^{-1} \cdot \Sigma_{\varepsilon}^{-1} \cdot G \cdot x(\Omega_p)$. Оценка \hat{c} получена из (7) методом наименьших квадратов [1]. Если c - набор неизвестных детерминированных параметров, то $\Sigma_{\hat{c}}^{-1} = M$. Однако нас в большей мере интересует случай, когда c - случайный вектор, для которого считаются известными вектор средних значений \bar{c} и ковариационная

матрица Σ_c . Дело в том, что, располагая информацией об $\bar{x}(\Omega_p)$ и $\Sigma_x(\Omega_p)$, можно перейти к соответствующим характеристикам для спектральных коэффициентов. В этом случае $\Sigma_c = (\Sigma_c^{-1} + M)^{-1}$. Матрица $\Sigma_c = \Sigma_c(\Omega_p)$, поскольку $M = M(\Omega_p)$. Тогда задача оптимального планирования наблюдательной сети заключается в нахождении Ω_p^* , реализующего минимум функционала, который в случае A -критерия оптимизации имеет вид: $\Phi = \Phi(\Sigma_c) = tr(\Sigma_c)$ (tr -след или сумма диагональных элементов матрицы). Часто рассматривается [24] также D -критерий оптимизации, являющийся одним из наиболее распространенных критериев. В этом случае $\Phi = \det(\Sigma_c)$.

Для построения алгоритмов оптимизации введем в рассмотрение "функцию отклика" $\varphi(w, \Omega_p)$ в произвольной точке с координатами w при фиксированном наборе станций сети Ω . В случае D -критерия "функция отклика" имеет вид [24]:

$$\varphi(w, \Omega_p) = g^T(w) (\Sigma_c^{-1} + G^T \cdot \Sigma_c^{-1} \cdot G)^{-1} \cdot g(w), \quad (8)$$

где $g(w) = [g_1(w), \dots, g_N(w)]^T$.

По аналогии с предыдущим разделом рассмотрим две задачи: отбора наименее информативных станций и определения местоположения новых станций. В первом случае численная процедура заключается в нахождении элемента $w_i \in \Omega_p$, реализующего минимум $\varphi(w, \Omega_p)$. Решение второй задачи основано на определении w , реализующего максимум $\varphi(w, \Omega_p)$.

2.3. Сравнительный анализ методов оптимизации. Численные процедуры оптимизации сходны по структуре для обоих подходов к задаче планирования наблюдательной сети. Они содержат механизм прямой и обратной пошаговой оптимизации. В случае оптимальной интерполяции ищутся экстремумы функционала $S(w)$ (см.(6)), при использовании регрессивной модели $\varphi(w, \Omega_p)$ - (см.(8)). Сравним оптимизируемые функционалы. Структура

обоих функционалов сходна. Различие заключается в использовании вектор-функций f и g в формулах (6) и (8). Другое отличие состоит в использовании матриц F и G , Λ и Σ_c .

До сих пор выбор базисных функций $\{g_i(w)\}$ не был конкретизирован. Для того чтобы провести содержательное сравнение двух рассмотренных ранее подходов, необходимо предполагать, что в обоих случаях мы располагаем одинаковыми исходными данными. Поскольку в методе оптимальной интерполяции использовалась ковариационная матрица, при построении регрессионной модели в качестве базисных функций целесообразно применять ЭОФ [12, 13, 16]. Для конечного множества точек наблюдений w_1, \dots, w_p ЭОФ совпадают с собственными векторами ковариационной матрицы $\Sigma_x(\Omega_p)$. Поэтому можно считать, что $g_j(w_i) = f_j(w_i)$. Тогда $G = F$. Кроме того, $\Sigma_c = \Lambda$, где Λ - матрица, определенная в (3). Теперь формулы (6) и (8) полностью идентичны. Это означает, что оба подхода полностью равнозначны. Следует, однако, остановиться на различиях, связанных с численной реализацией двух методов.

При использовании метода оптимальной интерполяции для ранжирования станций сети по информативности необходимо произвести вычисление диагональных элементов матрицы $\Sigma_x^{-1}(\Omega_p)$. Обращение ковариационной матрицы является обычно неустойчивой процедурой, которая приводит к получению значений, содержащих большие погрешности. Для устойчивого обращения ковариационной матрицы могут быть применены некоторые регуляризующие процедуры. Однако их применение, как правило, сопряжено с искажением структуры корреляционных связей, описываемых матрицей Σ_x .

При использовании второго подхода центр тяжести вычислительных трудностей переносится на расчет ЭОФ, т.е. собственных векторов ковариационной матрицы Σ_x . Эта процедура также в общем случае является неустойчивой, особенно для собственных векторов, отвечающих близким к нулю собственным значениям. Тем не менее, именно возможность разделения ЭОФ на устойчивую и неустойчивую части (с вычислительной точки зрения)

позволяет осуществить регуляризацию регрессионной модели. Этот подход позволяет также регуляризовать процедуру численной оптимизации, что гарантирует получение устойчивых результатов решения оптимизационной задачи.

2.4. Модификация для схемы четырехмерного усвоения. В последние годы разрабатываются методы четырехмерного усвоения данных наблюдений [9]. При этом вычислительная процедура, как правило, разбивается на 2 этапа: анализ и прогноз. Понятно, что результаты краткосрочного прогноза позволяют получать начальные поля для анализа, которые в среднем меньше отклоняются от истинных значений, чем средние климатические поля, используемые в традиционном объективном анализе. Рассмотрим модификацию ковариационной матрицы Σ_x для набора точек метеорологического поля $\Sigma_x(\Omega_p)$ или для совокупности пространственных гармоник Σ_c , которые могут быть получены в рамках схемы усвоения с помощью фильтра Калмана [13]. Для упрощения обозначений здесь опущены индексы, связанные с пространственными координатами или спектральными компонентами, рассмотренными выше. Дискретный фильтр Калмана включает системы уравнений прогноза и наблюдения по временной координате j :

$$\begin{aligned} x(j+1) &= A \cdot x(j) + R \cdot \xi(j) \\ y(j) &= B \cdot x(j) + \varepsilon(j) \end{aligned} \tag{9}$$

В (9) использованы следующие традиционные обозначения: x – вектор, описывающий состояние атмосферы, y – совокупность наблюдений, A – линеаризованный оператор прогностической модели, B – для системы наблюдений, $\xi(j), \varepsilon(j)$ – случайные векторы, описывающие шумы прогностической модели и системы наблюдений, соответственно. Система (9) решается в условиях, когда заданы не только начальные условия x_0 для вектора x , но и ковариационные матрицы: $\Sigma_{x_0}, \Sigma_\varepsilon, \Sigma_\xi$. Имеется в виду, что $\Sigma_{x_0} = \Sigma_x$. Ковариационная матрица оценок поля $\Sigma_x(j+1/j)$, полученная после одного цикла усвоения анализ-прогноз, имеет следующий вид:

$$\begin{aligned}\Sigma_x(j+1/j) &= A \cdot \Sigma_x(j/j) \cdot A^T + R \cdot \Sigma_\xi(j) \cdot R^T; \\ \Sigma_x(j/j) &= \Sigma_x(j/j-1) - \Sigma_x(j/j-1) \cdot B^T \cdot [B \cdot \Sigma_x(j/j-1) \cdot B^T + \Sigma_\varepsilon]^{-1} \cdot B \cdot \Sigma_x(j/j-1)\end{aligned}\quad (10)$$

Второе уравнение (10) показывает, насколько уменьшилась ковариационная матрица при усвоении информации на j -ом шаге ассимиляции. Первое уравнение (10), наоборот, демонстрирует ее увеличение на этапе прогноза при переходе от момента времени j к $j+1$. Если схема усвоения является эффективной, то ковариационная матрица $\Sigma_x(j+1/j)$ описывает меньшую неопределенность поля, чем матрица Σ_x , характеризующая его естественные колебания. Поэтому можно было бы думать использовать второе уравнение из (10) для получения функционала для оптимизации системы наблюдений. Однако, очевидно, что матрица вида $\Sigma_x(j+1/j)$ согласно (10) зависит от ряда величин, которые либо сами носят изменчивый характер (например, оператор A зависит от $x(j/j)$), либо зависят от неопределенных величин (например, ковариационные матрицы $\xi(j), \varepsilon(j)$). Более того, схема усвоения зависит от прогностической модели, которая, как предполагается, находится в состоянии совершенствования. Но конфигурация системы наблюдений, по определению, должна быть более консервативной, чтобы отвечать требованиям ряда пользователей информации, включая, например, несколько моделей прогноза. Поэтому большее практическое значение имеют модели и методы, развитые в предыдущих разделах.

3. Применение

В качестве иллюстрации для применения методов рационализации сети рассмотрим метеорологическую сеть Камчатского УГМС.

3.1. Районирование. Согласно физико-географическому районированию [3] Камчатский регион состоит из 5 климатических провинций: Западно-Камчатской, Срединной, Центрально-Камчатской, провинции Вулканического нагорья и провинции Командорские острова. Анализ результатов объективного статистического районирования Камчатского УГМС на основе 4-х элементного (температура, влажность, осадки, скорость ветра) анализа (см.[16, 17]) показывает

(рис. 1), что картина распределения зон имеет значительное сходство с физико-географическим районированием этой территории. Однако, число ИОЗ значительно превышает число

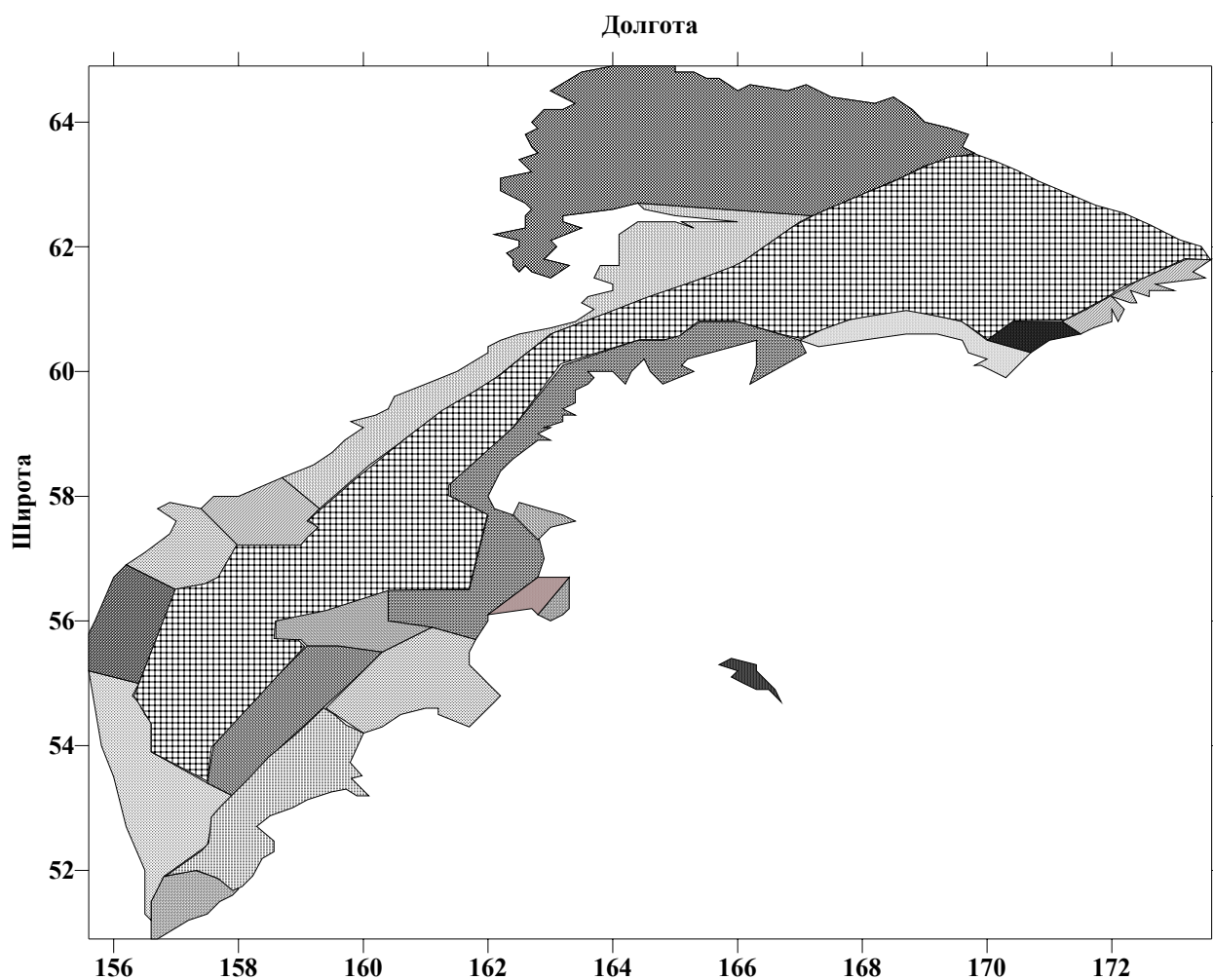


Рис. Метеорологическое районирование Камчатского УГМС

Рис.1. Метеорологическое районирование Камчатского УГМС

климатических зон и обеспечивает более детальное районирование территории Камчатки. Расположение станций по территории климатических провинций весьма неоднородно, что говорит о некоторой несбалансированности всей сети. Так, Западно-Камчатская провинция имеет на своей территории 8 станций, Центрально-Камчатская – 7 ПН, провинция «вулканического нагорья» -13 станций, Командорские острова - 1 ПН, Срединная провинция не представлена вообще. Тем не менее, статистическое районирование позволяет выделить внутри каждой провинции, за исключением Срединной, где метеорологические данные отсутствуют, несколько более мелких зон. Значительная часть этих зон представлена одной станцией. В Западно-Камчатской провинции выделяются 6 зон, из которых только две (№ 5, 12) имеют более

одной (по две) станции в зоне. Центрально-Камчатская провинция (долина реки Камчатки) включает 3 ИОЗ: № 2 (4 станции в верхнем течении реки), № 6 (2 станции в среднем течении и притоке - р. Быстрой), а также ИОЗ №7, растянутой вдоль восточного побережья северной части полуострова вплоть до мыса Корф. Зона № 21, начинаясь от устья реки Камчатки, узкой полосой тянется на юг вдоль побережья средней части полуострова до Кроноцкой сопки. Здесь полоса переходит в зону № 1, состоящую из четырёх станций, которая на самом юге примыкает к мысу Лопатка, который сам представляет собой отдельную зону, состоящую из одной станции. Полоса вдоль восточного побережья является низменной частью климатической провинции «вулканическое плато», возвышенная часть которого также практически не имеет станций, только на юге полуострова имеются две зоны - № 16 (ст. Начики) и № 20 (Паужетские ключи). В северной части этой провинции станция Озерный мыс также образует отдельную зону - № 15. Провинция Командорских островов, имея одну станцию на о. Беринга (Никольское), входит в общую зону № 3 вместе со станцией на мысе Африка на восточной окраине полуострова, вытянутой в сторону Командорских островов.

Объективное метеорологическое районирование позволяет осуществить более подробное (по сравнению с физико-географическим) разбиение территории УГМС на ИОЗ. Выполнение этапа районирования обеспечивает получение некоторых наметок для определения структуры оптимальной, наблюдательной сети.

3.2. Оценка информативности ИОЗ и ПН. В качестве основы для расчета показателей информативности ИОЗ и ПН наблюдений нами использовалась схема 4-элементного объективного анализа полей температуры, влажности, осадков, скорости ветра, на территории УГМС (см.[16, 17]). Изменение относительной средней дисперсии точности оценок полученных в узлах регулярной сетки ($0,25^\circ \times 0,25^\circ$) при исключении данной станции рассматривались нами в качестве индикатора информативности этой станции. Оказалось, что самой значимой и наиболее информативной станцией является ПН, мыс Лопатка, расположенный на южной оконечности полуострова. Ее вклад в изменение средней дисперсии оценок составляет 1,35%.

Следующим за ним является мыс Шипунский - 0,32%. Он находится на юго-востоке Камчатки. Значимость остальных ПН оказывается ниже уровня 0,2%. Наименее информативные ПН имеют уровень около 0,03%. Наиболее информативные ИОЗ и станции расположены вдоль восточного побережья полуострова, на южной оконечности и на северо-восточной его части. Это объясняется тем, что здесь именно здесь обнаруживается наибольшая изменчивость наблюдаемых метеорологических элементов. Станции, расположенные на западном побережье имеют меньшую информативность. Информативность ПН и ИОЗ соответствуют друг другу с точки зрения их упорядоченности по убыванию величин.

3.3. Варианты минимальной сети. Основная идея построения минимальной сети заключается в определении минимального числа ПН, обеспечивающих минимальное представительство однородных зон, которые покрывают территорию УГМС. Это означает, что каждая ИОЗ должна быть представлена, по крайней мере, одной станцией. В основу проведенных расчётов положены результаты районирования, представленные на **рис. 1**. Отбор станций в минимальную сеть осуществляется численными методами оптимизации по заданному критерию. Нами рассматривались 2 количественных критерия, основанных на применении формул (5) и (6), описывающих функцию отклика модели наблюдательной сети. Первый - связан с минимумом дисперсии погрешностей аппроксимации метеорологических полей в точках, соответствующих совокупности ПН. Второй - характеризует среднюю дисперсию погрешностей оценок метеорологических полей на регулярной сетке ($0,25^\circ \times 0,25^\circ$). Для каждого критерия было рассмотрено два случая. В первом случае все отбираемые станции являются равноправными, во втором случае приоритет в отборе получают станции, являющиеся реперными, их дублёрами, климатическими и т. п. (см. знаки + на **рис. 2**). Общее число приоритетных станций равно 10. Распределение показателей информативности (критерий 1) для станций принадлежащих разным ИОЗ представлено на **рис.2**. Отсюда следует, что вклад ПН в результаты объективного анализа оказывается различным. В каждой ИОЗ можно выделить наиболее информативные станции. Обращает на себя внимание тот факт, что 6

приоритетных станций имеет высокие значения информативности, тогда как остальные 4 ПН можно отнести к числу малоинформативных ПН. Причем, 3 приоритетные станции являются

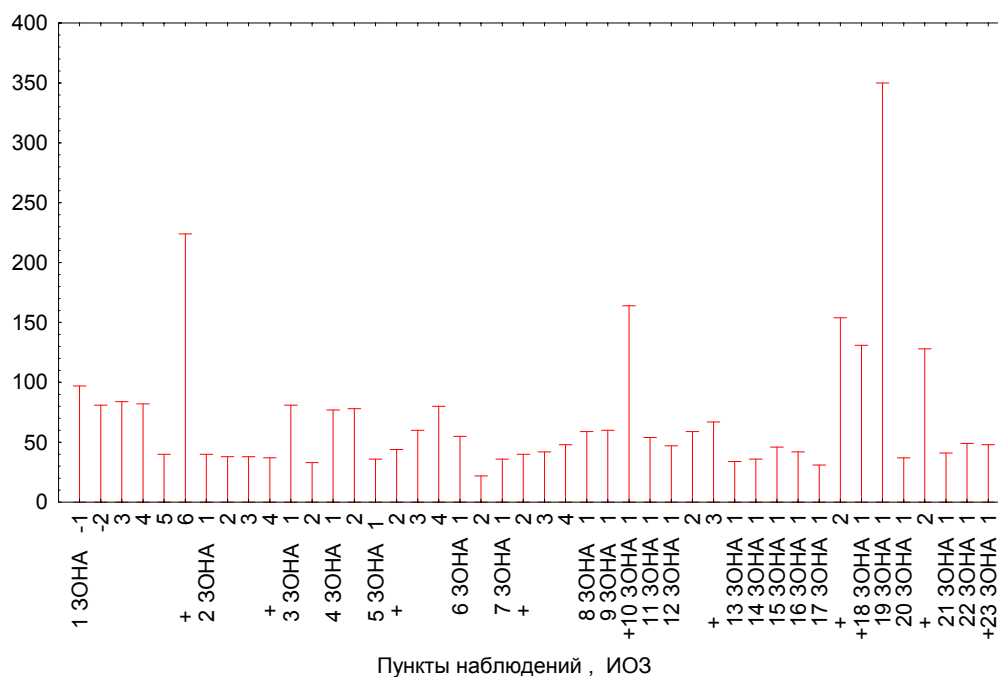


Рис. 2. Информативность (критерий 1) пунктов наблюдений и их распределение по ИОЗ. Знак + означает, что данная станция является реперной, климатической или обеспечивает международный обмен.

малоинформативными внутри ИОЗ, где имеются более информативные неприоритетные ПН. Таким образом, минимальная сеть, построенная с учетом приоритетных станций, отличается от варианта, в котором приоритетность не учитывалась на 3 ПН. На **рис.3** представлен вариант минимальной сети, включающий приоритетные станции. Отбор произведен по критерию 1. Общее число ПН, входящих в достаточную сеть составляет 20.

3.4. Достаточная сеть. В случае, когда проблемы финансирования ННС не являются критическими для существования сети, появляются возможности оптимизации сети за счет ее перераспределения и расширения. В этом случае возникает вопрос о том, каким образом можно построить сеть, которая бы действительно была оптимальной, т.е. обеспечивала получение полной информации для целей анализа и прогноза полей основных метеорологических элементов. Понятно, что наблюдения, проводимые на соседних станциях, могут содержать и часто содержат значительный объем дублирующей (избыточной) информации. Косвенно об

этом свидетельствует тот факт, что информативность отдельных станций, характеризующая небольшую долю независимой информации, получаемой в данном пункте наблюдений,

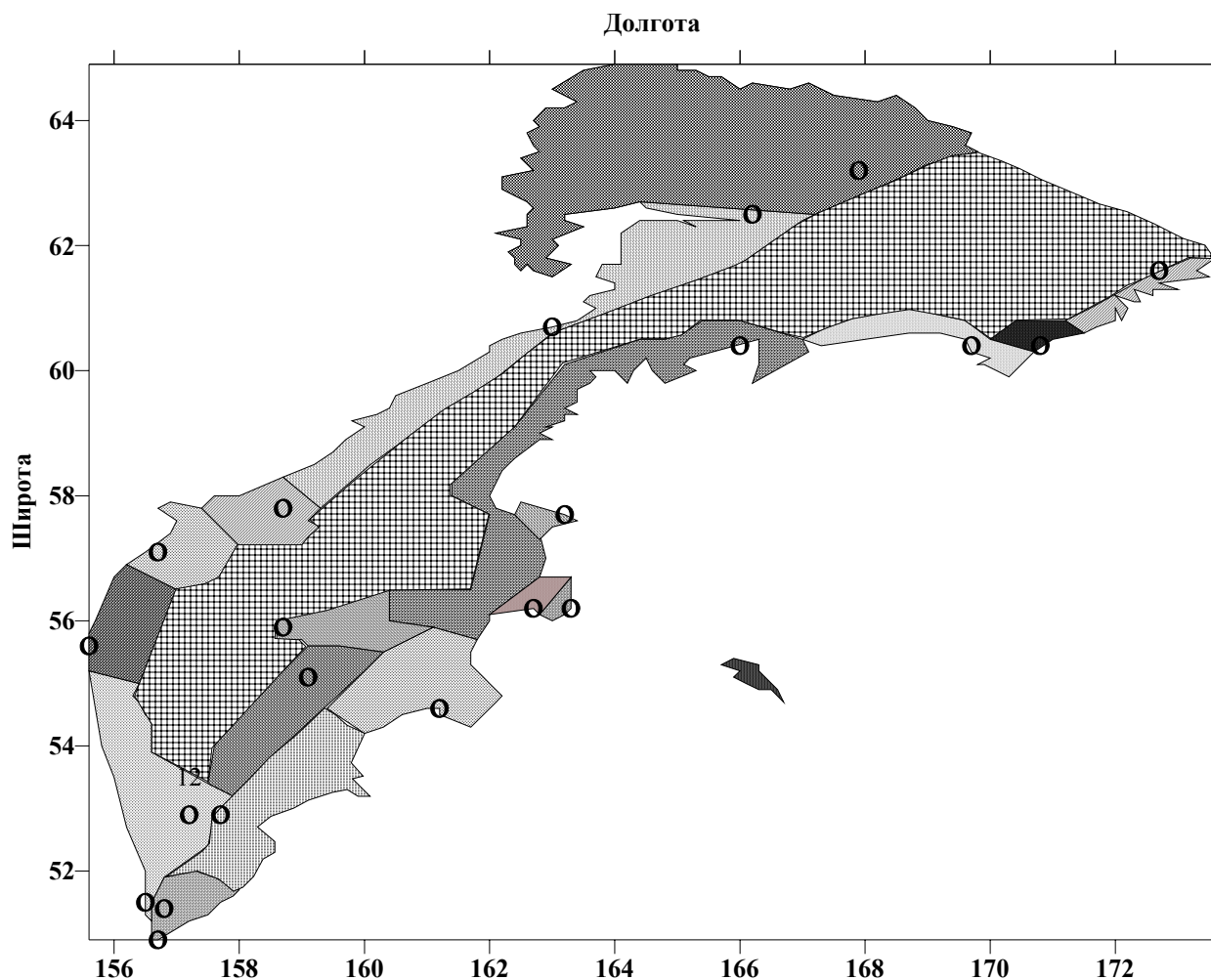


Рис. Метеорологическое районирование и минимальная сеть Камчатского УГМС

Рис.3. Метеорологическое районирование и минимальная сеть Камчатского УГМС

составляется более, чем 0,2%. С другой стороны, учитывая общее количество станций сети, равное всего 42, предположение о наличии информационной избыточности требует проверки. Чтобы получить более объективные оценки необходимо специальное исследование региональной сети, как информационной системы. С этой целью нами был использован подход, основанный на оценках информационной системы с помощью математического аппарата, использующего меру Шеннона для оценки систем с многоканальными информационными потоками [15]. В этом случае вводится понятие обобщённого отношения "сигнал/шум" для

информационной системы. При небольшом числе пунктов наблюдения такое отношение существенно превышает единицу. Включение в систему нового пункта наблюдений приводит,

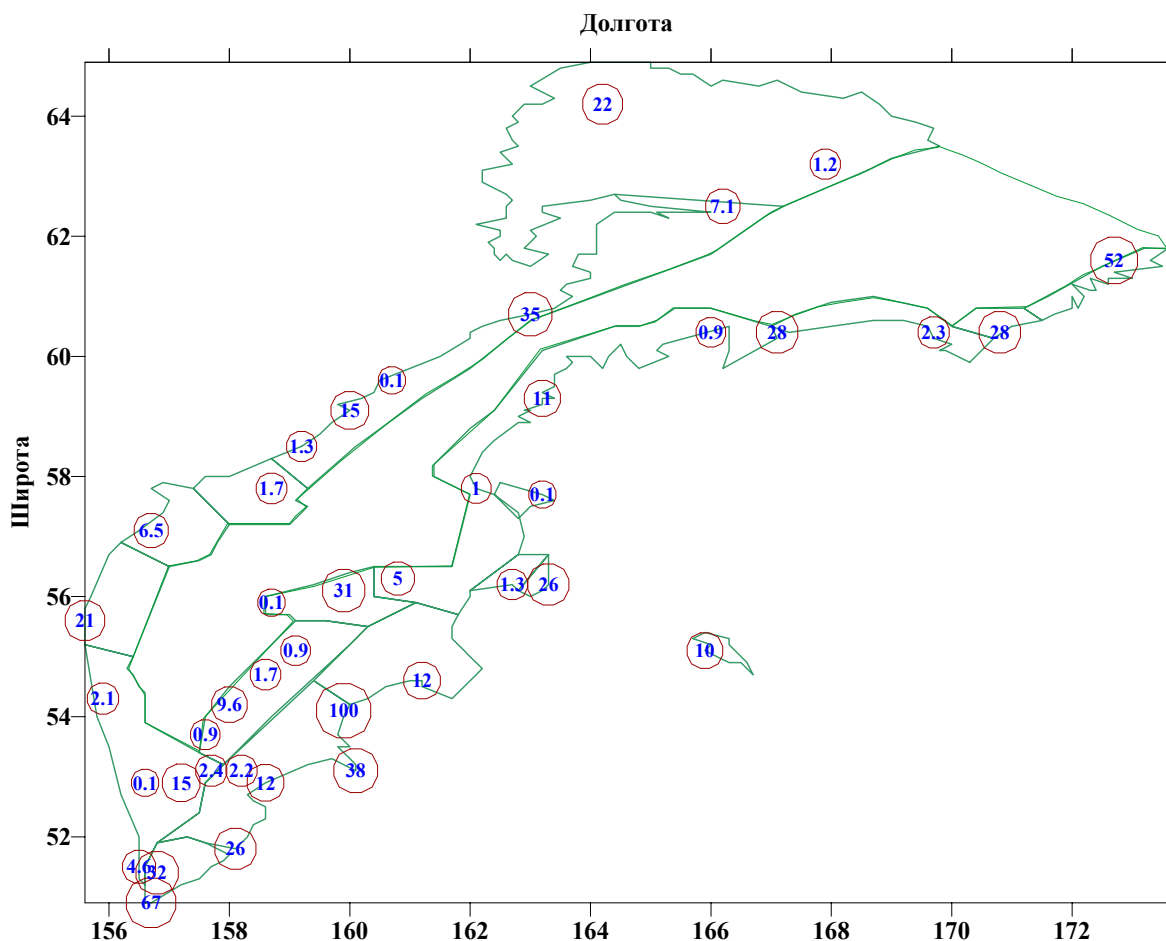


Рис. . Информативность (отношение сигнал/шум) ПН Камчатского УГМС (указана в кружках)

Рис.4. Информативность пунктов наблюдений Камчатского УГМС (указана в кружках)

как правило, к уменьшению отношения "сигнал/шум" из-за значительного вклада дублирующей информации. После отбора определённого числа пунктов наблюдений приращение "полезной" информации может сравниться с приращением "шума", что свидетельствует о том, что дальнейшее расширение наблюдательной сети ведёт лишь к её информационному перенасыщению. Имеется удобное пороговое значение для отношения "сигнал/шум", равное единице и являющееся индикатором достаточности выбранного множества пунктов наблюдений для региональной сети. Результаты расчетов отношения "сигнал/шум" для ПН Камчатского УГМС представлены на **рис. 4**. Семь ПН имеют значение этого критерия меньше единицы. Общее число ПН равно 41. Следовательно, достаточная сеть включает 34 ПН. Это

более, чем в полтора раза превышает состав минимальной сети. Следовательно, после решения задач определения состава минимальной сети, обоснование достаточной (оптимальной) сети становится практически значимой проблемой рационализации региональных ННС.

Подводя итог вышеизложенному, отметим, что результаты объективного статистического районирования вполне соответствуют физико-географическому и климатическому делению региона. Не смотря на ограничения, связанные с отсутствием сети в провинции Срединного хребта и её нехваткой в ряде других мест, можно отметить более детальный характер выделения зон по сравнению с традиционным. Существующие климатические провинции, как правило, подразделяются на более мелкие территории, отличающиеся своеобразием статистических параметров наблюдений. Сокращение имеющейся сети на наш взгляд можно проводить с учетом "достаточной" сети. В случае необходимого сокращения, оно должно быть проведено таким образом, чтобы полностью сохранить минимальную сеть.

Заключение

В данной работе подведены определенные итоги исследований ГГО им. А.И Воейкова по проблематике рационализации региональных метеорологических сетей, проведенных за последние годы. В результате сформировалась методология практического решения этой задачи, которая представляет собой алгоритм поэтапного рассмотрения проблемы: статистическое районирование территории УГМС и выделение информативно-однородных зон (ИОЗ), оценка информативности ИОЗ и пунктов наблюдений (ПН), построение минимальной и достаточной сетей. Накопленный опыт показывает, что данная схема может быть перенесена на решение задачи рационализации состава и конфигурации других наблюдательных сетей и систем. В последние годы возрастает роль данных спутникового дистанционного зондирования атмосферы. Изложенный подход требует незначительной модификации в части построения более сложной информационной модели наблюдательной системы [15] для применения,

например, при решении задачи рационализации для аэрологической сети. Другое направление возможного приложения этого подхода это рассмотрение сетей, предназначенных для изучения крупномасштабных атмосферных процессов, например, для климатического мониторинга [14] или для слежения за вариациями поля общего содержания озона [12].

Литература:

1. Андерсон Т. Введение в многомерный статистический анализ. Физматгиз, М., 1963, 547 с
2. Гантмахер Ф.Р. Теория матриц. М, Наука, 1967, 412 с.
3. Гандин Л.С., Каган Р.Л. Статистические методы в метеорологии. Л., Гидрометеиздат, 1976, 274 с.
4. Гвоздецкий Н.А. (ред.) Физико-географическое районирование территорий СССР. Изд-во МГУ, 1960. 320с.
5. Дмитриев А.А. О репрезентативности наблюдений над температурой воздуха при неоднородной подстилающей поверхности. - Изв.АН СССР, сер.геофизич.- 1961, N 2, с.328-333.
6. Каган Р.Л. К оценке репрезентативности осадкомерных данных.- Труды ГГО, 1966, вып.191., с.22-34.
7. Каган Р.Л. Осреднение метеорологических полей.-Л., Гидрометеиздат, 1979, 213 с.
8. Кондратьев К.Я. (ред.) Экодинамика и экологический мониторинг Санкт-Петербургского региона в контексте глобальных измерений. 1996 , Наука, С-Пб., 442 с.
9. Машкович С.А., Вейль И.Г. Численные эксперименты по четырехмерному анализу на основе спектральной прогностической модели. Труды ГМЦ, 1972, вып.100, с.3-16.
10. Методические указания управлениям Гидрометслужбы N 14. Оценка характерности станций.- Л. Гидрометеиздат, 1953, 28 с.
11. Методические указания управлениям Гидрометслужбы по оценке репрезентативности реперных станций и их дублеров. - Л., ГГО, 1978, 22 с. (Ротапринт ГГО)

12. Могилевер О.М. Получение репрезентативных значений метеозлементов в приземном слое атмосферы как задача оптимальной фильтрации. В книге “Применение статистических методов в метеорологии”- Л., Гидрометеиздат, 1971,с.170-179.
13. Покровский О.М. Оптимизация метеорологического зондирования атмосферы со спутников. Л., Гидрометеиздат, 1984, 267 с.
14. Покровский О.М. Анализ эффективности методов оптимизации наземных наблюдательных сетей. - Труды ГГО, 1989, вып.528,с.82-89.
15. Покровский О.М., Малыгина А.К. О рационализации схемы сбора данных дистанционных измерений озона в северном полушарии. - Исследование Земли из Космоса, 1985, N 1, с.10-16.
16. Покровский О.М., Денисов С.Г. Информативность океанической сети для целей объективного анализа полей геопотенциала в северном полушарии.- Метеорология и гидрология 1985, N 10, с.37-43.
17. Покровский О.М., Кароль С.И. Об оптимальном отборе станций для климатического мониторинга приземной температуры. - Метеорология и гидрология 1988, N 9,
18. Покровский О.М. Современные методы оптимизации систем мониторинга окружающей среды.- Исследование Земли из Космоса,1995,N 2, с.106-120.
19. Покровский О.М., Пятигорский А.Г. Применение эмпирических ортогональных функций для районирования метеорологических полей. - Метеорология и гидрология, 1995, N 6, с.12-22.
20. Покровский О.М. О методе выбора репрезентативных пунктов метеорологических наблюдений.- Метеорология и гидрология, 1999, N 2, с.55-67.
21. Покровский О.М. О рационализации региональных наблюдательных сетей.- Метеорология и гидрология, 2000, N 8, с.5-21.
22. Рао С.Р. Линейные статистические методы и их приложения., М., Наука, 1968, 547 с.
23. Тихомиров Е.И. Репрезентативность метеорологических наблюдений и возможная степень их точности.- Метеорологический вестник, 1935, N 5-6, с.2-8.
24. Успенский А.Б., Федоров В.В. Вычислительные аспекты метода наименьших квадратов в анализе и планировании регрессионных экспериментов., 1975, М., изд. МГУ, 196 с.

Методология рационализации региональных метеорологических сетей

О.М. Покровский

Аннотация

Рассмотрены вопросы рационализации региональных метеорологических сетей. Изложена общая методология, включающая следующие основные этапы практического решения задачи: статистическое районирование территории УГМС и выделение информативно-однородных зон (ИОЗ), оценка информативности ИОЗ и пунктов наблюдений (ПН), построение минимальной и достаточной сетей. Представлены варианты информационных моделей наблюдательных сетей и проведено их сравнение. Получены выражения для функций отклика, используемой при численном решении задачи оптимизации. В качестве иллюстрации рассмотрена оптимизация сети Камчатского УГМС с изложением результатов основных этапов решения задачи. Представлены карты районирования территории УГМС. Даны схемы минимальной и достаточной, наблюдательной сети. Учтены приоритетные станции.