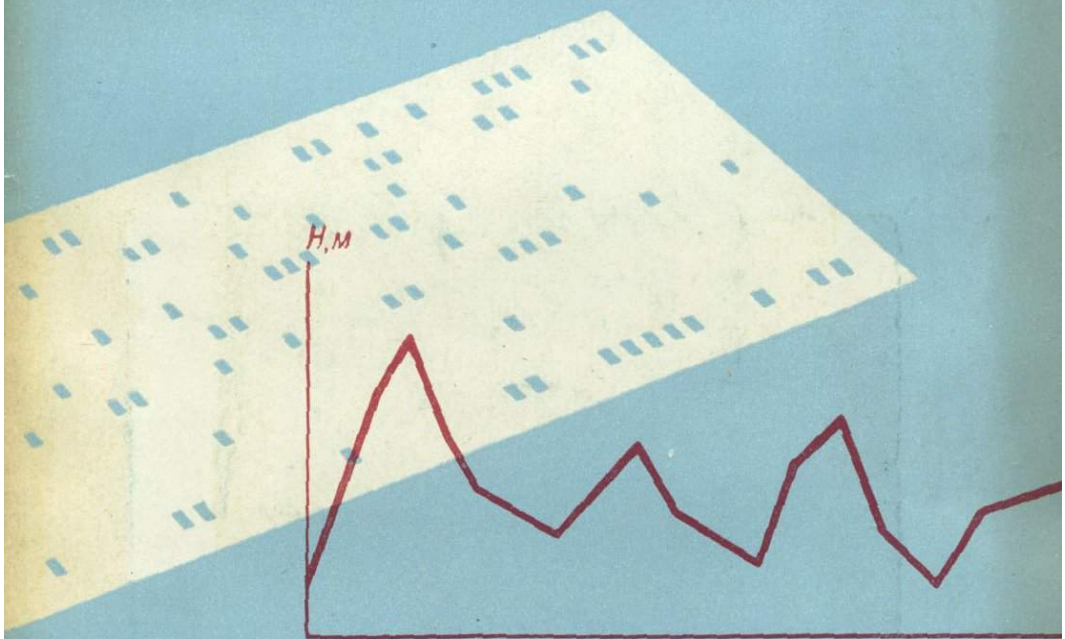


Э. А. ЗАЛЬЦБЕРГ

Статистические  
методы  
прогноза  
естественного  
режима уровня  
грунтовых вод



Недра

Э. А. ЗАЛЬЦБЕРГ

551.49:513

1818

Статистические  
методы  
прогноза  
естественного  
режима уровня  
грунтовых вод



Ленинград  
„НЕДРА“  
Ленинградское отделение  
1976



**Зальцберг Э. А.** Статистические методы прогноза естественного режима уровня грунтовых вод. Л., «Недра», 1976. 101 с.

В книге обобщены результаты работ по применению вероятностно-статистических методов для прогноза сезонного и многолетнего режима уровня грунтовых вод. Рассмотрены методы парной и множественной корреляции, гармонического анализа и авторегрессионные модели. Указаны трудности и ограничения, возникающие при применении этих методов к конкретной гидрорежимной информации. Приведены примеры расчетов прогнозных зависимостей и дана оценка устойчивости их коэффициентов по результатам наблюдений в представительных скважинах. Проведена проверка прогнозных уравнений на независимом фактическом материале. Рассмотрены вопросы, связанные с распространением прогнозов по территории.

Книга рассчитана на широкий круг гидрогеологов, гидрологов, мелиораторов, строителей и других специалистов, использующих в своей деятельности данные о режиме грунтовых вод. Табл. 17, ил. 30, прил. 1, список лит. 75 назв.

---

## Предисловие

Прогнозы естественного режима уровня грунтовых вод приобретают важное значение в различных областях народного хозяйства, в связи с чем вопросами прогнозирования занимается все большее число организаций и специалистов-гидрогеологов. В 1967 г. была опубликована работа А. А. Коноплянцева «Применение методов математической статистики для анализа и прогноза режима уровня подземных вод», в которой обобщены первые результаты применения некоторых статистических методов для прогнозов естественного режима уровня. За прошедшие с этого времени годы повсеместно накопились новые и весьма интересные фактические материалы наблюдений за режимом уровня грунтовых вод. Известные успехи были достигнуты в деле разработки новых и усовершенствования использованных ранее методов прогноза. В частности, в последние годы наряду с методами парной и множественной корреляции для прогнозов естественного режима уровня начали применяться автокорреляционные функции, авторегрессионные модели, методы гармонического анализа.

В настоящей работе сделана попытка обобщить как уже известные, так и сравнительно новые методы прогноза естественного режима уровня грунтовых вод. Приводятся основные расчетные зависимости, которые рекомендуется использовать для составления прогнозов. Значительное внимание уделено ранее практически не рассматривавшимся в гидрогеологической литературе вопросам о соотношении продолжительности гидрорежимных наблюдений и устойчивости найденных прогнозных зависимостей.

Автор не ставил своей главной задачей рекомендовать для тех или иных конкретных районов и областей наиболее рациональный метод прогноза, поскольку сделать это, вообще говоря,

довольно трудно. Условия формирования естественного режима грунтовых вод даже в пределах сравнительно небольших территорий настолько разнообразны и изменчивы, что указать для них какой-либо один, наиболее подходящий, метод прогноза, как правило, не представляется возможным. В связи с этим специалистам-гидрогеологам необходимо владеть комплексом методов и применять тот или иной из них в зависимости от конкретных природных условий и объема имеющейся гидрорежимной информации.

Предлагаемые в работе методики прогноза иллюстрируются конкретными расчетами, при этом наиболее широко использованы наблюдения за режимом уровня грунтовых вод на территории северо-западных областей РСФСР, которые проводит Северо-Западная гидрогеологическая партия (бывшая станция), а также Северо-Западное управление гидрометеорологической службы (ГМС) и Валдайская научно-исследовательская гидрологическая лаборатория им. В. А. Урываева. Помимо этого использованы данные наблюдений за режимом уровня на территории Калининградской, Московской, Воронежской, Горьковской областей РСФСР и Молдавской ССР. Всем сотрудникам гидрогеологических партий, оказавшим помощь и содействие в процессе работы над книгой, автор выражает глубокую благодарность.

## Краткая классификация прогнозов естественного режима уровня грунтовых вод, оценка их точности и эффективности

Классификация прогнозов естественного режима уровня грунтовых вод может проводиться по следующим признакам 1) по целевому назначению прогнозов, 2) по составу прогнозируемых уровней, 3) по заблаговременности прогнозов, 4) по методам прогнозов. Говоря о целевом назначении прогнозов режима уровня грунтовых вод, следует отметить, что уже теперь они отвечают на весьма разнообразные запросы народного хозяйства, а в дальнейшем сфера их применения будет все более расширяться. В настоящее время эти прогнозы составляются и используются для следующих целей:

- 1) решения ряда инженерных задач, связанных с проектированием и возведением промышленных и гражданских сооружений, дорожным строительством, проведением газопроводов и т. д.;
- 2) проектирования и проведения мелиоративных мероприятий на обширных площадях;
- 3) планирования севооборота сельскохозяйственных культур;
- 4) решения вопросов, связанных с водоснабжением, водоприпусками в горные выработки и др.

В зависимости от назначения и целей прогнозов можно предсказывать положение различных уровней. Наибольший практический интерес представляет прогнозирование следующих уровней:

- 1) минимальных предвесенних и максимальных весенних (для провинции с устойчивым зимним промерзанием зоны аэрации);
  - 2) максимальных зимних или зимне-весенних (для провинции со спорадическим промерзанием или отсутствием промерзания зоны аэрации);
  - 3) летних или летне-осенних минимальных;
  - 4) среднегодовых уровней как текущего, так и ряда следующих лет;
  - 5) уровней на определенную дату или определенный период.
- Заблаговременность прогноза зависит от характера колебаний уровня, от примененной методики прогноза, а также от

пожеланий потребителей. Все перечисленные причины в той или иной степени связаны между собой. Так, например, характер колебаний уровня, а также заданная потребителем заблаговременность прогноза в значительной мере определяют выбор той или иной методики. По степени заблаговременности прогнозов режима уровня можно различать прогнозы краткосрочные (с заблаговременностью до 1 месяца), долгосрочные (с заблаговременностью от 1 месяца до 1 года) и сверхдолгосрочные (с заблаговременностью свыше 1 года). Независимо от примененной методики достоверность и точность прогнозов с увеличением заблаговременности уменьшаются.

Прогнозы режима уровня составляются с использованием ряда методов, среди которых условно можно выделить следующие: 1) аналитические, 2) балансовые, 3) гидрогеологических аналогий, 4) моделирования, 5) вероятно-статистические.

Аналитические методы применяются в тех случаях, когда известны гидрогеологические параметры водоносного горизонта (коэффициенты фильтрации, недостатка насыщения или водоотдачи, мощность) и установлена функциональная зависимость между изменениями уровней и одного или нескольких аргументов, в числе которых, как правило, фигурирует и время  $t$ . С аналитическими методами самым тесным образом связаны методы прогноза грунтовых вод с использованием моделирования.

Среди исследователей, внесших значительный вклад в развитие аналитических методов, следует назвать Майе, Бусинеска, Г. Н. Каменского, П. Я. Полубаринову-Кочину, С. Ф. Аверьянова, А. И. Силина-Бекчурина, Н. Н. Биндемана, Ф. М. Бочевера, В. М. Шестакова и других.

Балансовые методы основаны на решении уравнения водного баланса грунтового потока или почвенно-грунтовой призмы относительно величины изменения уровня  $\Delta h$  за определенный промежуток времени. Элементы баланса определяются либо экспериментальным путем с использованием специальных приборов и установок, либо рассчитываются по формулам гидродинамики. И в том и в другом случае устанавливается функциональная или корреляционная зависимость между величиной изменения уровня  $\Delta h$  и теми или иными элементами баланса. Фундаментальные разработки этих методов можно найти в трудах Г. Н. Каменского, А. В. Лебедева, П. А. Киселева, Д. М. Каца и других.

Методы гидрогеологических аналогий заключаются в том, что закономерности режима уровня, выявленные для определенной территории за многолетний период наблюдений (продолжительность циклов, характер изменения уровней внутри циклов и т. д.), экстраполируются во времени и в пространстве. Достоверность таких прогнозов возрастает, если изменения режима уровня удастся связать с изменениями какого-либо регионального или глобального режимобразующего фактора, например, атмосфер-

ной циркуляции, солнечной активности и др. Используя метод аналогии, можно прогноз режима, составленный каким-либо методом для территории, обеспеченной достаточным количеством наблюдательных точек, перенести на территорию со сходными условиями формирования режима уровня, где этих точек либо недостаточно, либо нет вовсе. Методы гидрогеологических аналогий успешно использовали для прогнозов режима грунтовых вод Н. А. Кенесарин, А. А. Коноплянцев, В. А. Коробейников и ряд других исследователей режима подземных вод.

При использовании вероятностно-статистических методов с помощью корреляционного анализа устанавливаются не функциональные, а корреляционные зависимости между изменениями уровня и одним или несколькими факторами. С вероятностно-статистическими методами тесно связаны методы гармонического анализа, с помощью которых выявляются количественные закономерности в колебаниях уровней за период режимных наблюдений, которые экстраполируются за пределы этого периода.

В данной работе рассматриваются некоторые из вероятностно-статистических методов прогноза естественного режима уровня грунтовых вод, а именно: методы парной и множественной корреляции, автокорреляция, простейшие авторегрессионные модели и гармонический анализ рядов уровней.

При использовании той или иной прогнозной зависимости критерием точности прогноза является его стандартная ошибка:

$$S_y = \sqrt{\frac{\sum_{i=1}^n (y_i - \tilde{y}_i)^2}{n}}, \quad (1)$$

где  $y_i$  — уровень по данным фактических замеров;  $\tilde{y}_i$  — уровень, вычисленный по прогнозной зависимости;  $n$  — число членов ряда.

Вычисленная по формуле (1) стандартная ошибка отличается от ее истинного значения для генеральной совокупности в тем большей степени, чем меньше объем выборки  $n$ . Для оценки стандартной ошибки прогноза для выборки ограниченного объема можно вычислить величину приведенной ошибки по формуле

$$S_y = \sqrt{\frac{\sum_{i=1}^n (y_i - \tilde{y}_i)^2}{n - m}}, \quad (2)$$

где  $m$  — число, зависящее от вида уравнения регрессии и равное количеству определяемых в нем коэффициентов.

Оценка эффективности прогнозной зависимости должна характеризовать соотношение между ошибками прогноза, полученными при использовании данной зависимости, и отклонениями прогнозируемого уровня от его среднееголетнего значения

либо от какой-либо иной величины. Такой оценкой является величина корреляционного отношения:

$$\eta = \sqrt{1 - \frac{S^2}{\sigma^2}}, \quad (3)$$

где  $S$  — стандартная ошибка, характеризующая отклонение фактических уровней от их значений, полученных по прогнозной зависимости;  $\sigma$  — стандартное отклонение прогнозируемого уровня.

С помощью корреляционного отношения можно оценивать связи и, следовательно, прогнозные зависимости как линейные, так и нелинейные.

Помимо корреляционного отношения для оценки эффективности прогнозов можно использовать индекс детерминации

$$d = 1 - \frac{S^2}{\sigma^2}, \quad (4)$$

также отношения

$$S^2/\sigma^2 \text{ и } S/\sigma. \quad (5)$$

Чем в большей степени изменения прогнозируемого уровня связаны с изменениями одного или нескольких факторов, использованных в прогнозной

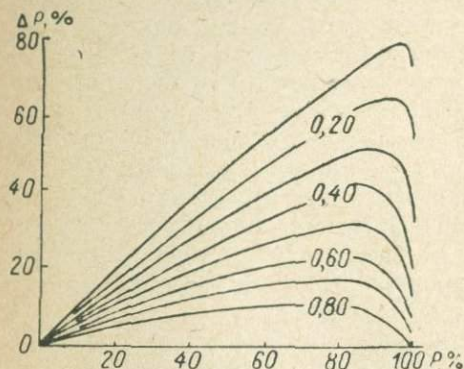


Рис. 1. Изменение превышения обеспеченности ошибок прогноза над обеспеченностью равных им отклонений от нормы ( $\Delta P, \%$ ) в зависимости от обеспеченности ошибок прогноза ( $P, \%$ ) и отношения  $S/\sigma$  (цифры у кривых).

ной зависимости, тем больше величина  $\eta$  и  $d$  и тем меньше отношения  $S^2/\sigma^2$  и  $S/\sigma$ .

Для оценки эффективности той или иной корреляционной зависимости необходимо определить превышение обеспеченности ошибок прогноза над обеспеченностью равных им отклонений от нормы в зависимости от величин  $\eta$ ,  $d$  или  $S/\sigma$ . Результаты подобных расчетов для случая нормального распределения прогнозируемого уровня и ошибок прогноза представлены на рис. 1, заимствованном из работы Е. Г. Попова [55].

Из приведенных на рис. 1 графиков следует, что при использовании для прогноза той или иной зависимости существенное преимущество по сравнению с теми случаями, когда в качестве прогноза берется норма прогнозируемого уровня, наблюдается при  $S/\sigma \leq 0,8$ . По аналогии с гидрологическими прогнозами условно можно принять, что методика прогноза и, следовательно, эффективность и точность прогнозной зависимости являются хорошими, если отношение  $S/\sigma \leq 0,5$ , удовлетворительными, когда  $0,51 \leq S/\sigma \leq 0,8$ , и плохими или неудовлетворительными, когда  $S/\sigma > 0,8$ . Все эти соотношения справедливы при числе наблюдений  $n \geq 25$ . В случае, если число членов ряда  $15 \leq$

$\leq n \leq 25$ , приведенные выше значения  $S/\sigma$  следует уменьшить на 0,05, при  $n < 15$  — на 0,1.

Условным критерием оправдываемости каждого прогноза или их совокупности является допустимая ошибка. Прогноз считается оправдавшимся, если его ошибка, т. е. разность между предсказываемым и фактическим значением, меньше или равна допустимой.

Величина допустимой ошибки может быть различной. При нормальном или близком к нему распределении прогнозируемого уровня и ошибок прогноза за допустимую принимается ошибка с обеспеченностью 50%, или вероятная ошибка, которая вычисляется по формуле

$$\Delta_{\text{доп}} = 0.674\sigma, \quad (6)$$

где  $\Delta_{\text{доп}}$  — допустимая, или вероятная, ошибка;  $\sigma$  — стандартное отклонение уровня.

С помощью допустимой ошибки, вычисляемой по формуле (6), можно сравнивать оправдываемость различных прогнозов.

Таким образом, если эффективность и точность методики прогноза были оценены как хорошие или удовлетворительные, то при массовой проверке на независимом фактическом материале следует ожидать, что ошибки прогноза по меньшей мере в 50% случаев будут меньше допустимых.

## 2

### Применение парной и множественной корреляции для составления прогнозов естественного режима уровня грунтовых вод

#### 2.1 — расчет линейного уравнения парной регрессии

При применении парной линейной корреляции отыскиваются линейные зависимости вида

$$\bar{y} = a + bx, \quad (7)$$

где  $\bar{y}$  — зависимая переменная величина — предиктант;  $x$  — независимая переменная величина — предиктор.

Применительно к задачам прогноза режима уровня грунтовых вод  $y$  является тот или иной прогнозируемый уровень,  $x$  — режимобразующий фактор или уровень, предшествующий прогнозируемому. Задача нахождения коэффициентов  $a$  и  $b$  по наблюдаемым значениям  $y$  и  $x$  решается методом наименьших квадратов. Согласно основному положению этого метода уравнение (7) будет наилучшим образом аппроксимировать исходные данные в том случае, если сумма квадратов отклонений наблюдавшихся (эмпирических) значений  $y$  от его значений, вычисленных по уравнению (7), будет минимальной, т. е.

$$u = \sum_{i=1}^n (y_i - \bar{y}_i)^2 = \sum_{i=1}^n (y_i - a - bx_i)^2 = \min, \quad (8)$$

где  $y_i$  — наблюдавшиеся значения переменной  $y$ ;  $\bar{y}_i$  — значения переменной, вычисленные по уравнению регрессии;  $n$  — число наблюдений.

Функция  $u$  принимает минимальное значение в том случае, если частные производные  $\frac{\partial u}{\partial a}$  и  $\frac{\partial u}{\partial b}$  обращаются в 0, т. е.

$$\begin{aligned} -\frac{1}{2} \frac{\partial u}{\partial a} &= \sum_{i=1}^n (y_i - a - bx_i); \\ -\frac{1}{2} \frac{\partial u}{\partial b} &= \sum_{i=1}^n (y_i - a - bx_i) x_i. \end{aligned} \quad (9)$$

Приравняв левые части уравнений 0 и проведя необходимые преобразования, получим систему нормальных уравнений вида

$$\sum_{i=1}^n y_i - an - b \sum_{i=1}^n x_i = 0;$$

$$\sum_{i=1}^n y_i x_i - a \sum_{i=1}^n x_i - b \sum_{i=1}^n x_i^2 = 0. \quad (10)$$

Решая эту систему относительно неизвестных коэффициентов  $a$  и  $b$ , получаем:

$$b = \frac{n \sum_{i=1}^n x_i y_i - \sum_{i=1}^n x_i \sum_{i=1}^n y_i}{n \sum_{i=1}^n x_i^2 - \left( \sum_{i=1}^n x_i \right)^2}; \quad (11)$$

$$a = \frac{\sum_{i=1}^n y_i \sum_{i=1}^n x_i^2 - \sum_{i=1}^n x_i \sum_{i=1}^n x_i y_i}{n \sum_{i=1}^n x_i^2 - \left( \sum_{i=1}^n x_i \right)^2}. \quad (12)$$

Если значение коэффициента  $b$  уже найдено, коэффициент  $a$  определяется из выражения

$$a = \frac{\sum_{i=1}^n y_i - b \sum_{i=1}^n x_i}{n}. \quad (13)$$

Теснота линейной зависимости величин  $y$  и  $x$  устанавливается с помощью вычисления выборочного коэффициента корреляции

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}}, \quad (14)$$

где  $\bar{x}$ ,  $\bar{y}$  — средние значения переменных  $x$ ,  $y$ .

В случае нелинейной зависимости теснота связи определяется величиной корреляционного отношения, которое вычисляется по формуле (3).

Стандартная ошибка прогноза  $S_y$ , характеризующая разброс наблюдаемых значений  $y$  относительно найденной линии регрессии, определяется по формуле

$$S_y = \sqrt{\frac{\sum_{i=1}^n (y_i - a - bx_i)^2}{n-2}}. \quad (15)$$

Величину  $S_y$  можно вычислить и по другой формуле, а именно:

$$S_y = \sqrt{\frac{1}{n-2} \left\{ \sum_{i=1}^n (y_i - \bar{y})^2 - \frac{\left[ \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) \right]^2}{\sum_{i=1}^n (x_i - \bar{x})^2} \right\}}. \quad (16)$$

При большом числе наблюдений целесообразно пользоваться приближенной зависимостью

$$S_y = \sigma_y \sqrt{1 - r^2}, \quad (17)$$

где  $\sigma_y = \sqrt{\frac{\sum_{i=1}^n (y_i - \bar{y})^2}{n}}$  — стандартное отклонение  $y$ .

Величина  $S_y$  характеризует среднее отклонение от линии регрессии для всей выборки. Для каждого конкретного значения аргумента  $x_k$  стандартная ошибка прогноза определяется по формуле

$$S_{y_k} = \frac{S_y}{\sqrt{n-2}} \sqrt{1 + \frac{n(x_k - \bar{x})^2}{\sum_{i=1}^n (x_i - \bar{x})^2}}, \quad (18)$$

где  $S_{y_k}$  — стандартная ошибка прогноза величины  $y$  при  $x = x_k$ .

Как видно из формулы, ошибки прогнозов увеличиваются по мере отклонения независимой переменной  $x$  от ее среднего значения.

Для каждого значения  $x = x_k$  можно определить доверительные интервалы прогнозного значения  $y_k$  по формуле

$$G_{в.н} = y_k \pm \frac{t_{q, n-2} S_y}{\sqrt{n-2}} \sqrt{1 + \frac{n(x_k - \bar{x})^2}{\sum_{i=1}^n (x_i - \bar{x})^2}}, \quad (19)$$

где  $G_{в.н}$  — верхний и нижний доверительные интервалы;  $t_{q, n-2}$  — нормированное отклонение распределения Стьюдента, значения которого в зависимости от уровня значимости  $q$  и числа степеней свободы  $n$  за tabулированы.\*

Если известны  $\bar{x}$ ,  $\bar{y}$ ,  $\sigma_x$ ,  $\sigma_y$  и  $r$ , уравнение регрессии (7) можно представить в виде

$$\bar{y} - \bar{y} = b(x - \bar{x}). \quad (20)$$

В этом случае

$$b = r \frac{\sigma_y}{\sigma_x}. \quad (21)$$

\* Соответствующие таблицы можно найти в работах [35, 64].

Таблица 1 — исходные данные наблюдений (в метрах от поверхности земли)

Год	Среднегодовой уровень $y$	Предвесенний минимальный среднемесячный уровень $x$	Год	Среднегодовой уровень $y$	Предвесенний минимальный среднемесячный уровень $x$
1952	17,57	19,51	1961	14,97	16,10
1953	17,66	19,02	1962	14,46	16,84
1954	18,74	19,04	1963	15,75	16,43
1955	14,48	15,80	1964	16,37	17,13
1956	15,44	17,52	1965	16,87	18,41
1957	13,92	15,96	1966	14,08	18,10
1958	15,22	16,31	1967	15,30	16,58
1959	15,76	16,99	1968	15,52	17,00
1960	16,48	18,28			

Из равенства (20) следует, что

$$\bar{y} = (\bar{y} - b\bar{x}) + bx. \quad (22)$$

Отсюда свободный член  $a$  уравнения регрессии (7) находится по формуле

$$a = \bar{y} - b\bar{x}. \quad (23)$$

Используя формулы (11), (13), (14), найдем уравнение регрессии и коэффициент корреляции между положением среднегодового и предвесеннего минимального уровня грунтовых карстовых вод для скв. 11-11 на Ордовикском плато в Ленинградской области. Исходные данные наблюдений приведены в табл. 1.

$$b = \frac{n \sum_{i=1}^n x_i y_i - \sum_{i=1}^n x_i \sum_{i=1}^n y_i}{n \sum_{i=1}^n x_i^2 - \left( \sum_{i=1}^n x_i \right)^2} = \frac{17 \cdot 4680,84 - 295,02 \cdot 268,59}{17 \cdot 5141,68 - 295,02^2} = \frac{334,86}{371,76} = 0,90;$$

$$a = \frac{\sum_{i=1}^n y_i - b \sum_{i=1}^n x_i}{n} = \frac{268,59 - 0,90 \cdot 295,02}{17} = 0,18;$$

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}} = \frac{19,71}{\sqrt{28,69 \cdot 21,87}} = 0,79,$$

Таким образом, уравнение регрессии имеет вид

$$\bar{y} = 0,18 + 0,90x.$$

Поскольку коэффициент корреляции  $r$  можно рассматривать в качестве меры взаимозависимости в случае нормального или близкого к нему распределения переменных, существенный интерес представляет знание законов распределения уровней грунтовых вод. В настоящее время однозначно решить вопрос о распределении уровней грунтовых вод, как правило, весьма трудно из-за сравнительно небольшой продолжительности имеющихся рядов гидрогеологических наблюдений, которая редко превышает 15—20 лет. Следует отметить, что, говоря о распределении, мы имеем в виду не уровни (срочные, среднедекадные, среднемесячные), взятые в хронологической последовательности\*, а фазовооднородные уровни, т. е. уровни, относящиеся к одной и той же фазе годового цикла и формирующиеся под влиянием одного и того же комплекса факторов. К последним, в частности, относятся экстремальные (предвесенние минимальные, весенние максимальные и др.) уровни. Кроме того, во многих случаях необходимо знать распределение среднегодовых уровней.

В литературе имеется очень мало сведений по этим вопросам. Так, в работе [46] исследуется распределение среднегодовых уровней грунтовых вод на территории Воронежской области и устанавливается, что оно незначительно отличается от нормального. Распределение среднегодовых и максимальных уровней грунтовых вод, а также годовых амплитуд колебаний уровней рассматривается в работе [40]. Авторы совершенно справедливо указывают, что «продолжительность наблюдений по большинству скважин недостаточна для обоснованного суждения о характере распределения параметров режима, так как характеризует только среднюю часть кривых распределений, не отражая ее крайние участки. С этой точки зрения выводы о законе распределения носят предварительный характер». В работе сделано предварительное заключение о том, что среднегодовые и максимальные уровни грунтовых вод в зоне умеренного увлажнения подчиняются нормальному закону распределения, в зоне избыточного увлажнения — асимметричному закону распределения. По нашему мнению, в зоне избыточного увлажнения распределение уровней грунтовых вод также является в большинстве случаев весьма близким к нормальному.

Так, в работе [25] рассмотрено распределение среднегодовых уровней грунтовых вод на территории г. Ленинграда и его пригородов и сделан вывод о том, что оно является нормальным. Помимо этого нами исследовалось распределение рядов среднегодовых и экстремальных уровней грунтовых вод продолжительностью от 20 до 38 лет по скважинам, расположенным в Ленинградской (Ордовикское плато) и Московской областях. Почти во

---

\* Анализ распределения подобных рядов посвящены работы чехословацких исследователей Р. Нетопла [74], Х. Кржижа [73] и Р. Музикаржа.

Таблица 2 — статистические характеристики рядов режимообразующих факторов

Режимообразующие факторы, $x$	$C_s$	$\sigma_{c_s}$	$\frac{C_s}{\sigma_{c_s}}$	$E$	$\sigma_E$	$\frac{E}{\sigma_E}$
Сумма осадков за XII—III месяцы	0,09	0,29	0,31	-0,8	-0,58	1,38
Сумма осадков за IX—XI месяцы	0,27	0,29	0,93	-0,3	-0,58	0,52
Средняя температура воздуха за I—III месяцы	0,29	0,29	1,00	0,35	0,58	0,60

всех случаях распределение этих рядов практически не отличалось от нормального.

Поскольку наблюдения за уровнями являются, как правило, весьма непродолжительными, в ряде случаев целесообразно исследовать распределение не самих уровней, а тех или иных режимообразующих факторов, определяющих их положение. Например, на основе данных метеостанции Ленинград за 70-летний период наблюдений (1900—1969 гг.) были вычислены коэффициенты асимметрии  $C_s$ , эксцесса  $E$  и их стандартные ошибки  $\sigma_{c_s}$  и  $\sigma_E$  для некоторых рядов режимообразующих факторов суммы осадков за XII—III и IX—XI месяцы, средней температуры воздуха за I—III месяцы. Результаты расчетов представлены в табл. 2.

Указанные коэффициенты и их стандартные ошибки вычислялись по следующим формулам:

$$C_s = \frac{m_3}{\sigma_x^3}; \quad m_3 = \frac{\sum_{i=1}^n (x_i - \bar{x})^3}{n}; \quad \sigma_{c_s} \approx \sqrt{\frac{6}{n}}; \quad (24)$$

$$E = \frac{m_4}{\sigma_x^4} - 3; \quad m_4 = \frac{\sum_{i=1}^n (x_i - \bar{x})^4}{n}; \quad \sigma_E \approx \sqrt{\frac{24}{n}}, \quad (25)$$

где  $n$  — число наблюдений.

Так как отношения коэффициентов асимметрии и эксцесса к их стандартным ошибкам оказались меньше 1,5, можно считать, что исследуемые ряды практически подчиняются нормальному закону распределения. Дополнительная проверка распределения этих рядов с помощью критерия  $\chi^2$  показала, что гипотеза об их нормальности является вполне реальной. Поскольку распределение режимообразующих факторов практически нормальное, можно утверждать, что таковым же является распределение

и уровней грунтовых вод при условии, что между изменениями уровней и факторов наблюдаются линейные либо близкие к ним зависимости.

Таким образом, как в зоне избыточного, так и в зоне переменного увлажнения распределение уровней грунтовых вод в большинстве случаев либо нормальное, либо близкое к нему.

Косвенным свидетельством нормальности распределения уровней грунтовых вод на значительных территориях являются данные о распределении уровней воды в реках, приведенные в работе [65]. Поскольку между уровнями воды в реках и уровнями грунтовых вод наблюдается, как правило, весьма тесная линейная корреляционная зависимость, можно предположить, что в тех случаях, когда распределение уровней воды в реках является нормальным, распределение уровней грунтовых вод также должно быть близким к нормальному. Если распределения уровней грунтовых вод ( $y$ ) и тех факторов, с которыми отыскивается корреляционная зависимость ( $x$ ), существенно отличаются от нормального, причем эти различия подтверждаются применением тех или иных статистических критериев ( $\chi^2$ ,  $\lambda$  Колмогорова и др.), следует путем соответствующих преобразований привести распределение величин  $y$  и  $x$  к нормальному и лишь после этого находить корреляционные зависимости между ними.

## 2.2 — оценка достоверности коэффициента корреляции и коэффициентов уравнения регрессии

Поскольку и коэффициенты корреляции, и коэффициенты уравнений регрессии определяются для выборок ограниченного объема, а не для генеральной совокупности, необходимо определить их достоверность и значимость. В случае, когда распределение величин  $y$  и  $x$  не сильно отличается от нормального, число членов выборки сравнительно велико ( $n > 30$ ), а коэффициент корреляции имеет относительно небольшое абсолютное значение, его стандартную ошибку можно определить по формуле

$$\sigma_r = \frac{1-r^2}{\sqrt{n}} \quad (26)$$

где  $\sigma_r$  — стандартная ошибка выборочного коэффициента корреляции;  $r$  — выборочное значение коэффициента корреляции;  $n$  — число наблюдений.

Связь между величинами  $y$  и  $x$  обычно считается установленной в случае, если  $r_{yx}/\sigma_r \geq 2 \div 3$ .

При числе членов ряда  $n < 30$  стандартную ошибку коэффициента корреляции можно определять по формуле

$$\sigma_r = \frac{1-r^2}{\sqrt{n-1}} \quad (27)$$

Так, если коэффициент корреляции между колебаниями среднегодовых уровней в двух скважинах за 20 лет совместных наблюдений составил 0,7 (т. е.  $r = 0,7$ ,  $n = 20$ ), то

$$\sigma_r = \frac{1-r^2}{\sqrt{n-1}} = \frac{0,51}{\sqrt{19}} = 0,12;$$

$$\frac{r}{\sigma_r} = \frac{0,70}{0,12} = 5,8 > 3.$$

Следовательно, связь между колебаниями уровней в исследуемых скважинах является реальной и с вероятностью 95% можно утверждать, что истинное значение коэффициента находится в интервале  $\pm 2\sigma_r$ , т. е.  $0,94 \geq r \geq 0,46$ . Значения 95%-ных

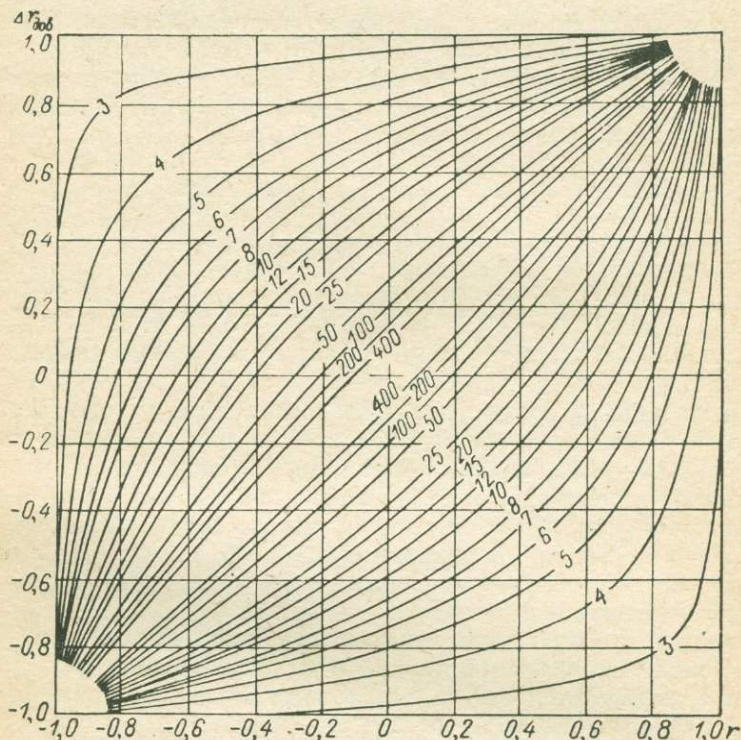


Рис. 2. 95%-ные доверительные пределы  $\Delta r_{\text{доп}}$  истинного значения выборочного коэффициента корреляции  $r$ , по И. С. Комарову [35].

доверительных пределов для выборочного значения коэффициента корреляции можно также снять с графика, приведенного на рис. 2. Цифры у кривых указывают объем использованной выборки. Так, при выборочном значении коэффициента корреляции  $r = 0,7$  для выборки из 20 членов истинное значение коэффициента с вероятностью 95% находится в интервале  $0,4 \leq r \leq 0,87$ .

При малых  $n$  и больших  $r$  для оценки истинного значения коэффициента корреляции можно использовать преобразование Р. А. Фишера:

$$z = \frac{1}{2} \ln \frac{1+r}{1-r} = 1,151 \lg \frac{1+r}{1-r}. \quad (28)$$

Распределение  $z$  близко к нормальному даже для небольших выборок и для значений  $r$ , близких к 1. Стандартная ошибка  $z$  не зависит от истинного значения  $r$  и равна:

$$\sigma_z = \frac{1}{\sqrt{n-3}}. \quad (29)$$

Таблицы для определения  $z$  по  $r$  для заданных доверительных вероятностей приведены в ряде работ [35, 56, 68 и др.]. Поясним порядок расчета, используя приведенные ранее данные ( $r = 0,7$ ,  $n = 20$ ). В указанной выше таблице, связывающей значения  $z$  и  $r$ , находим, что выборочному значению  $r = 0,7$  соответствует  $z = 0,87$ . Стандартная ошибка  $z$  равна

$$\sigma_z = \frac{1}{\sqrt{n-3}} = \frac{1}{\sqrt{17}} = 0,24; \quad \frac{z}{\sigma_z} = \frac{0,87}{0,24} = 3,6 > 3.$$

Доверительной вероятности 95% соответствует отклонение  $1,96 \cdot \sigma_z = 1,96 \cdot 0,24 = 0,47$ . Таким образом, предельные значения  $z$  равны  $0,87 + 0,47 \geq z \geq 0,87 - 0,47$ , т. е.  $1,34 \geq z \geq 0,40$ . Из этой же таблицы находим, что найденным значениям  $z$  соответствует  $0,87 \geq r \geq 0,39$ .

Таким образом, с вероятностью 95% можно утверждать, что истинное значение коэффициента корреляции заключено в интервале  $0,87-0,39$ . Приведенные примеры показывают, что расхождения в оценках коэффициента корреляции, полученных различными способами, сравнительно невелики. Поэтому при  $n \geq 10 \div 15$  и  $r \leq 0,9$  практически можно использовать любой способ для нахождения доверительных интервалов  $r$ , отдавая предпочтение наиболее простым из них.

Следует отметить, что приведенные выше оценки коэффициентов парной корреляции получены для случаев, когда в сравниваемых рядах отсутствует внутрирядная корреляция. Однако ряды гидрогеологических наблюдений нередко характеризуются той или иной внутрирядной связностью. Так, положительная корреляция между смежными членами ряда наблюдается при анализе среднегодовых уровней грунтовых вод, причем эта связность, как правило, возрастает с увеличением мощности зоны аэрации. Практически повсеместно обнаруживается та или иная положительная корреляция между уровнями грунтовых вод (суточными, среднедекадными, среднемесячными), взятыми в хронологической последовательности. Это же свойство характерно в той или иной степени для рядов режимобразующих факторов.

Наличие внутрирядной корреляции, по существу, уменьшает объемы использованных выборок в тем большей степени, чем теснее внутрирядные связи. В этом случае, имея  $n$  пар наблюдений, мы фактически располагаем информацией, соответствующей  $r$  парам независимых наблюдений, при этом  $n' < n$ . Поэтому следует иметь в виду, что в тех случаях, когда коэффициенты парной корреляции вычисляются для рядов, обладающих существенной внутрирядной связностью, их значения оказываются в той или иной степени завышенными, а значения доверительных пределов — заниженными.

Так как коэффициенты  $a$  и  $b$  уравнения регрессии вида  $\bar{y} = a + bx$ , так же как и коэффициенты корреляции  $r$ , получены не для генеральной совокупности, а для выборки данных ограниченного объема, необходимо установить доверительные пределы их истинных значений. Для вычисления двусторонних пределов можно использовать выражения

$$\begin{aligned} G_{в.н} &= b \pm t_{q, n-2} \sigma_b; \\ G_{в.н} &= a \pm t_{q, n-2} \sigma_a, \end{aligned} \quad (30)$$

где  $G_{в.н}$  — верхний и нижний доверительные интервалы;  $t_{q, n-2}$  — нормированное отклонение распределения Стьюдента;  $\sigma_b$  — стандартная ошибка коэффициента  $b$ ;  $\sigma_a$  — стандартная ошибка коэффициента  $a$ .

В свою очередь  $\sigma_a$  и  $\sigma_b$  находятся из следующих выражений

$$\begin{aligned} \sigma_a &= \frac{S_y}{\sqrt{n-2}}; \\ \sigma_b &= \frac{\sqrt{n} S_y}{\sqrt{(n-2) \sum_{i=1}^n (x_i - \bar{x})^2}}, \end{aligned} \quad (31)$$

где  $\sigma_a$  — стандартная ошибка коэффициента  $a = \bar{y}$  при  $x = \bar{x}$ ;  $n$  — число наблюдений; остальные обозначения приведены выше.

Вычислив доверительные интервалы для  $a$  и  $b$  или найдя отношения  $a/\sigma_a$  и  $b/\sigma_b$ , можно судить об устойчивости найденных коэффициентов уравнения регрессии.

С помощью формул (31) оценим устойчивость коэффициентов уравнения регрессии, полученного в разделе 2.1:

$$\begin{aligned} S_y &= \sqrt{\frac{1}{15} \left( 28,69 - \frac{19,71^2}{21,87} \right)} = \sqrt{\frac{10,93}{15}} = 0,85 \text{ м}; \\ \sigma_a &= \frac{S_y}{\sqrt{n-2}} = \frac{0,85}{3,87} = 0,22 \text{ м}; \\ \sigma_b &= \frac{\sqrt{n} S_y}{\sqrt{(n-2) \sum_{i=1}^n (x_i - \bar{x})^2}} = \frac{\sqrt{17} \cdot 0,85}{\sqrt{15 \cdot 21,87}} = \frac{3,49}{18,11} = 0,19; \\ \frac{a}{\sigma_a} &= \frac{15,80}{0,22} = 71,8; \quad \frac{b}{\sigma_b} = \frac{0,90}{0,19} = 4,74. \end{aligned}$$

Таким образом, отношения рассчитанных значений коэффициентов  $a$  и  $b$  к их стандартным ошибкам оказались больше 2,0, что позволяет считать их весьма устойчивыми.

Используя выражение (30), найдем доверительные интервалы для истинных значений  $a$  и  $b$  при 5%-ном уровне значимости и числе степеней свободы  $n = 17 - 2 = 15$ . Соответствующее значение  $t_q = 2,13$ . Отсюда для углового коэффициента  $b$  получаем следующие пределы:

$$G_{\text{верхн}} = 0,90 + 2,13 \cdot 0,19 = 1,30;$$

$$G_{\text{нижн}} = 0,90 - 2,13 \cdot 0,19 = 0,50.$$

Аналогично для свободного члена  $a$

$$G_{\text{верхн}} = 15,80 + 2,13 \cdot 0,22 = 16,22;$$

$$G_{\text{нижн}} = 15,80 - 2,13 \cdot 0,22 = 15,33.$$

Приведенные выше формулы позволяют оценить устойчивость коэффициентов уравнений регрессии лишь внутри использованных периодов наблюдений. В идеальном случае для оценки устойчивости этих коэффициентов следовало бы определить их для нескольких независимых периодов наблюдений и сравнить между собой. Например, если по какой-то скважине рассчитаны два уравнения, связывающие величину весенней амплитуды с суммой зимних осадков, при этом первое из них получено за период наблюдений с 1940 по 1955 г., второе — за период с 1956 по 1970 г., то сравнение знаков и абсолютных величин коэффициентов  $a$  и  $b$  этих уравнений покажет степень их устойчивости. Однако применение этого приема для режимных гидрогеологических наблюдений оказывается, как правило, практически невозможным из-за их сравнительной непродолжительности.

Общим правилом при выведении любых уравнений регрессии должна быть их проверка на независимом (по отношению к использованному периоду наблюдений) фактическом материале. Так, например, прогнозы по уравнению регрессии, рассчитанному по результатам наблюдений в 1950—1965 гг., необходимо сравнить с данными фактических замеров в 1966 и последующих годах.

В табл. 3 приводится проверка на независимом фактическом материале за 1969—1972 гг. уравнения регрессии среднегодового уровня, рассчитанного по результатам наблюдений за 1952—1968 гг. в разделе 2.1.

Проверка на независимом материале (правда, весьма ограниченного объема) подтвердила устойчивость и прогнозную ценность найденной зависимости среднегодового уровня от предвесеннего минимального.

Таблица 3 — проверка прогнозов среднегодового уровня (в метрах от поверхности земли)

Год	Предвесенний минимальный среднемесячный уровень	Среднегодовой уровень		Ошибка прогноза $\Delta$ , м
		Прогнозный	Фактический	
1969	17,23	15,69	15,62	0,07
1970	17,31	15,76	15,58	0,18
1971	15,84	14,44	15,07	0,63
1972	17,57	15,99	16,65	0,66

Примечание. Вероятная ошибка прогноза  $\Delta_{\text{вер}}$  составляет 0,92 м.

Используя формулу (19), найдем доверительные интервалы для прогнозного значения среднегодового уровня в 1972 г.

$$G_{\text{в. н}} = y_k \pm \frac{t_{q, n-2} S_y}{\sqrt{n-2}} \sqrt{1 + \frac{n(x_k - \bar{x})^2}{\sum_{i=1}^n (x_i - \bar{x})^2}} =$$

$$= 15,99 \pm \frac{2,13 \cdot 0,85}{3,87} \sqrt{1 + \frac{17(17,57 - 17,36)^2}{21,87}} = 15,99 \pm 0,48 \text{ м.}$$

Таким образом, с вероятностью 95% прогножное значение среднегодового уровня находится в интервале 16,47—15,51 м.

### 2.3 — выбор факторов-предсказателей положения уровней

Метод парной корреляции для прогнозов режима уровня грунтовых вод использовался многими авторами [16, 21, 36, 37, 41, 49 и др.].

В качестве независимого переменного — предиктора — можно использовать: 1) уровни за предшествующий период; 2) уровни в наблюдательной скважине-аналоге; 3) один из таких режимобразующих факторов, как осадки, температура, влажность воздуха, давление за тот или иной промежуток времени и др.; 4) индексы атмосферной циркуляции; 5) индексы солнечной активности и т. д.

В случае применения парной корреляции для заполнения пропусков в рядах гидрогеологических наблюдений или для их удлинения в качестве предикторов могут также использоваться уровни поверхностных водотоков (морей, рек, озер).

При применении парной корреляции необходимо, чтобы выбранный предиктор оказывал на формирование прогнозируемого уровня наиболее существенное, преобладающее влияние. Применительно к имеющимся сравнительно непродолжительным

рядом гидрогеологических наблюдений ( $n \leq 20$ ) следует, как правило, использовать для составления прогнозов корреляционные связи, характеризующиеся коэффициентами корреляции  $r \geq \geq 0,70 \div 0,75$  и величиной отношения  $S/\sigma \leq 0,70 \div 0,75$ .

При использовании в качестве предикторов уровней за предшествующий период следует иметь в виду, что прогноз последующих экстремальных уровней грунтовых вод до предшествующим (например, весеннего максимального по предвесеннему минимальному, летне-осеннего минимального по весеннему максимальному и т. д.) возможен лишь в том случае, если сезонные колебания сглажены либо в результате регулирующего влияния

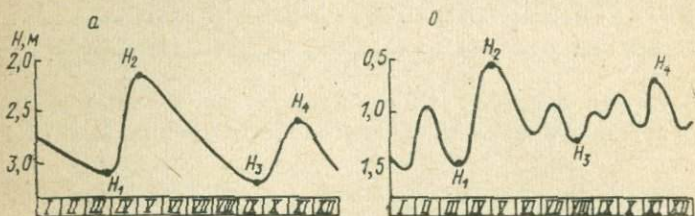


Рис. 3. Типичные графики колебаний уровней грунтовых вод в провинции сезонного (весенне-осеннего) питания при значительном (а) и незначительном (б) регулирующем влиянии зоны аэрации.

$H_1$  — предвесенний минимум;  $H_2$  — весенний максимум;  $H_3$  — летне-осенний минимум;  $H_4$  — осенний максимум.

зоны аэрации (в районах с междуречным видом режима), либо в результате постоянного притока грунтовых вод с водоразделов (в районах со склоновым и террасовым видами режима). Так, например, в гидрогеологической провинции сезонного (весенне-осеннего) питания, в зонах обильного и умеренного питания\* выравнивание колебаний уровней грунтовых вод, приуроченных преимущественно к песчаным и супесчаным отложениям, отмечается при глубине залегания среднесезонного среднегодового уровня не менее 2,5—3,5 м от поверхности земли (рис. 3).

В этих условиях, как правило, наблюдаются устойчивые зависимости между последующими и предыдущими экстремальными уровнями, которые можно использовать для составления прогнозов с различной заблаговременностью. Некоторые из подобных зависимостей представлены на рис. 4.

При залегании уровней грунтовых вод в этой же гидрогеологической провинции и тех же зонах на глубине менее 2,5—3,5 м от поверхности земли то или иное пополнение запасов грунтовых вод может происходить не только весной и осенью, но и летом, во время выпадения летних осадков, а на западе провинции — зимой, во время зимних оттепелей (см. рис. 3, б). В этих условиях связи между смежными экстремальными уровнями

\* Указанные гидрогеологические провинции и зоны выделены А. А. Коноплянцевым и соавторами [36].

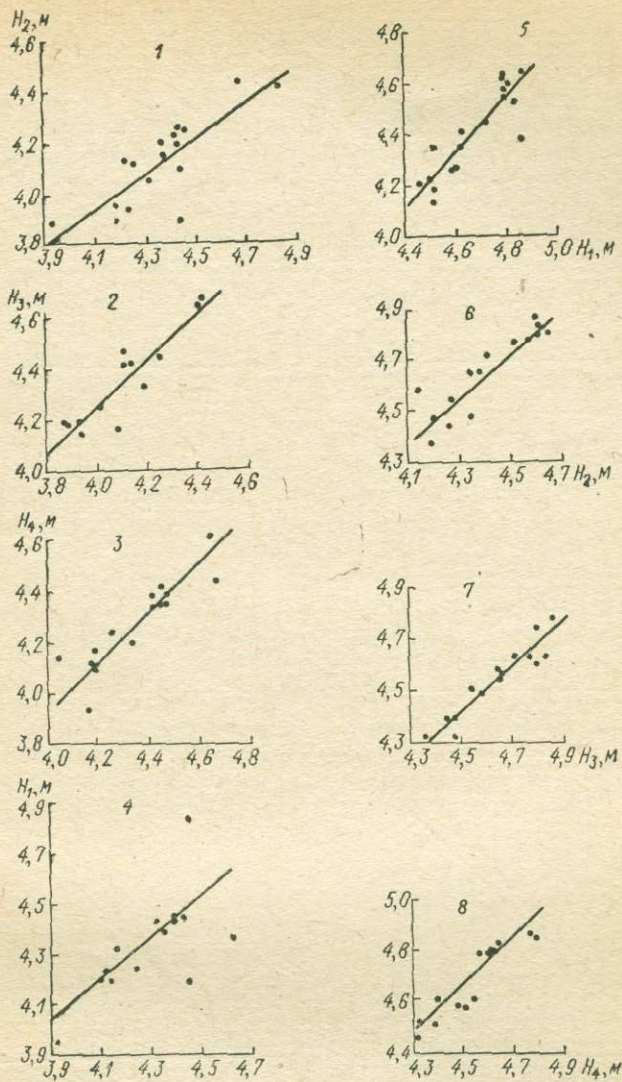


Рис. 4. Графики зависимости смежных экстремальных среднесесячных уровней грунтовых вод (в метрах от поверхности земли) в песках, Карельский перешеек.

1—4 — кол. 1; 5—8 — скв. 544;  $H_2$  — весенний максимальный;  $H_3$  — летне-осенний минимальный;  $H_4$  — осенний максимальный;  $H_1$  — предвесенний минимальный уровень следующего года.

Уравнения регрессии:

1) $H_2 = 1,05 + 0,7H_1$ ;	$r = 0,81 \pm 0,09$ ;	$S/\sigma = 0,56$ ;
2) $H_3 = 0,65 + 0,9H_2$ ;	$r = 0,92 \pm 0,04$ ;	$S/\sigma = 0,39$ ;
3) $H_4 = 0,14 + 0,95H_3$ ;	$r = 0,92 \pm 0,04$ ;	$S/\sigma = 0,39$ ;
4) $H_1 = 0,86 + 0,82H_4$ ;	$r = 0,74 \pm 0,10$ ;	$S/\sigma = 0,67$ ;
5) $H_2 = 0,38 + 1,02H_1$ ;	$r = 0,87 \pm 0,06$ ;	$S/\sigma = 0,43$ ;
6) $H_3 = 1,16 + 0,79H_2$ ;	$r = 0,88 \pm 0,06$ ;	$S/\sigma = 0,43$ ;
7) $H_4 = 0,51 + 0,87H_3$ ;	$r = 0,95 \pm 0,03$ ;	$S/\sigma = 0,32$ ;
8) $H_1 = 0,52 + 0,92H_4$ ;	$r = 0,93 \pm 0,04$ ;	$S/\sigma = 0,36$ .

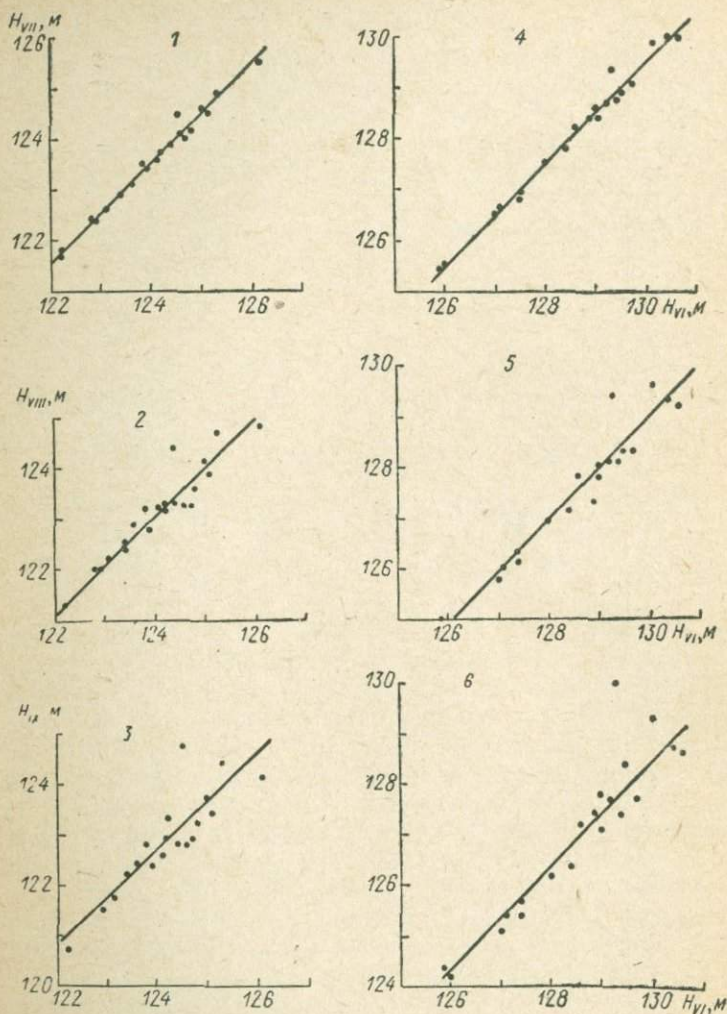


Рис. 5. Графики зависимости среднемесячных уровней грунтовых карстовых вод (в абс. отметках) по скв. 1-1 (1, 2, 3) и 3-3 (4, 5, 6) на Ордовикском плаге в Ленинградской обл.

Уравнения регрессии:

- 1)  $H_{VII} = -1,2 + 1,006H_{VI}$ ;  $r = 0,99 \pm 0,01$ ;  $S/\sigma = 0,13$ ;
- 2)  $H_{VIII} = 4,74 + 0,954H_{VI}$ ;  $r = 0,96 \pm 0,02$ ;  $S/\sigma = 0,29$ ;
- 3)  $H_{IX} = 2,97 + 0,965H_{VI}$ ;  $r = 0,91 \pm 0,04$ ;  $S/\sigma = 0,14$ ;
- 4)  $H_{VII} = -2,32 + 1,014H_{VI}$ ;  $r = 0,99 \pm 0,01$ ;  $S/\sigma = 0,14$ ;
- 5)  $H_{VIII} = -1,96 + 1,007H_{VI}$ ;  $r = 0,97 \pm 0,02$ ;  $S/\sigma = 0,24$ ;
- 6)  $H_{IX} = -6,50 + 1,038H_{VI}$ ;  $r = 0,92 \pm 0,04$ ;  $S/\sigma = 0,39$ .

становятся значительно менее тесными ( $r = 0,3 \div 0,6$ ), а использование их для прогнозов методом парной корреляции — весьма затруднительным.

Для составления прогнозов можно использовать не только уравнения регрессии между смежными экстремальными уровнями, но и между любыми срѳчными, среднемесячными и среднегодовыми уровнями, находящимися на ветви сезонного или многолетнего спада или подъема. Например, можно находить зависимости между среднемесячными уровнями за VI и VII, VI и VIII, VI и IX месяцы в тех случаях, когда в летне-осенний период наблюдается режим снижения уровней, близкий к «независимому»\* (рис. 5).

Для составления прогнозов среднегодовых уровней можно использовать зависимости между среднегодовыми уровнями в  $i$  и  $i + 1$ ,  $i$  и  $i + 2$ ,  $i$  и  $i + n$ ,  $i + 1$  и  $i + 2$  и другие годы в том случае, если  $i$ -й год является началом многолетнего (на протяжении  $n$  лет) снижения (или подъема) уровней в течение нескольких циклов колебаний примерно одинаковой продолжительности\*\*.

В зонах обильного и умеренного сезонного питания при глубине залегания среднемноголетнего среднегодового уровня не менее 2,5—3,5 м от поверхности земли в большинстве случаев наблюдаются четкие связи между весенними максимальными и среднегодовыми уровнями, которые можно использовать для прогноза последних (рис. 6). При меньшей глубине залегания уровней эти связи значительно ослабевают ( $r = 0,3 \div 0,5$ ) и, как правило, не могут быть использованы для прогнозов методом парной корреляции. Для этих условий нередко удается установить зависимость среднегодового уровня от предвесеннего минимального (рис. 7) или, как это рекомендовано в работе [33], от полусуммы предвесеннего минимума и весеннего максимума. Весьма тесные связи между весенними максимальными и среднегодовыми уровнями наблюдаются также в зоне скудного питания провинции сезонного питания грунтовых вод.

В провинции круглогодичного (преимущественно зимнего) питания грунтовых вод существует тесная зависимость между зимне-весенними и среднегодовыми уровнями, которую можно использовать для прогноза последних (рис. 8).

В провинции сезонного весенне-осеннего питания, в зонах обильного и умеренного питания, при глубине залегания среднемноголетних среднегодовых уровней до 2,5—3,5 м от поверхности земли наблюдаются тесные связи ( $r = 0,8 \div 0,9$ ) между летне-осенними минимальными и среднегодовыми уровнями грунтовых

\* Под «независимым» подразумевается режим снижения, определяемый положением уровня в начальный момент времени и не зависящий от метеорологических условий на поверхности земли.

\*\* Этот метод, названный методом регрессий, подробно рассмотрен в работе [37].

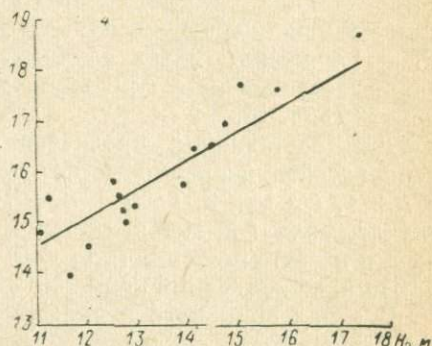
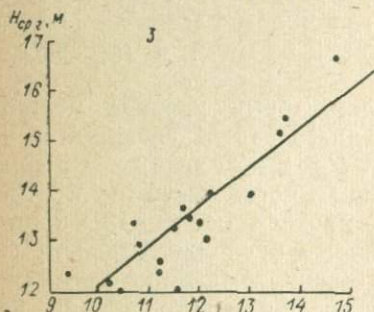
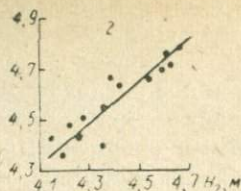
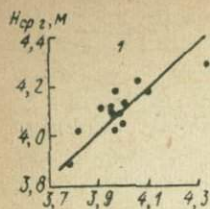


Рис. 6. Графики зависимости среднегодовых уровней грунтовых вод (в метрах от поверхности земли) в песках (1 — скв. 710, 2 — скв. 545) и в известняках (3 — скв. 3-3, 4 — скв. 11-11) от весенних максимальных среднемесячных уровней  $H_2$  (северо-запад РСФСР, Ленинградская обл.).

Уравнения регрессии:

- 1)  $H_{ср.г.} = 1,58 + 0,63 H_2$ ;  $r = 0,84 \pm 0,08$ ;  $S/\sigma = 0,55$ ;
- 2)  $H_{ср.г.} = 1,06 + 0,80 H_2$ ;  $r = 0,92 \pm 0,05$ ;  $S/\sigma = 0,37$ ;
- 3)  $H_{ср.г.} = 3,58 + 0,83 H_2$ ;  $r = 0,97 \pm 0,02$ ;  $S/\sigma = 0,24$ ;
- 4)  $H_{ср.г.} = 7,63 + 0,61 H_2$ ;  $r = 0,94 \pm 0,03$ ;  $S/\sigma = 0,35$ .

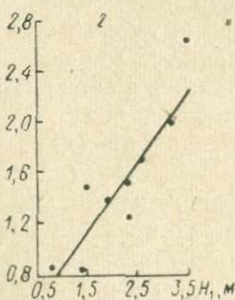
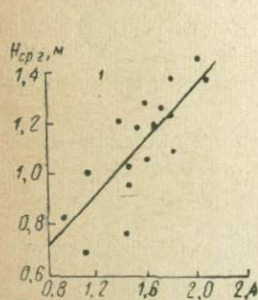
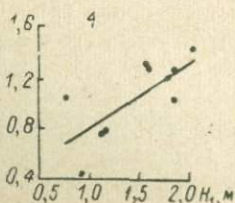
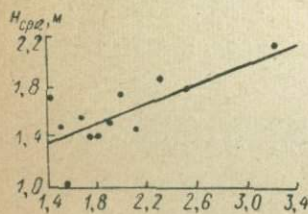


Рис. 7. Графики зависимости среднегодовых уровней грунтовых вод (в метрах от поверхности земли) в песках супесях (1 — скв. 11, 2 — скв. 153) и в суглинках (3 — скв. 100, 4 — скв. 170) от предвесенних минимальных среднемесячных уровней грунтовых вод  $H_1$  (г. Ленинград).

Уравнения регрессии:

- 1)  $H_{ср.г.} = 0,26 + 0,55 H_1$ ;  
 $r = 0,79 \pm 0,09$ ;  $S/\sigma = 0,62$ ;
- 2)  $H_{ср.г.} = 0,26 + 0,57 H_1$ ;  
 $r = 0,88 \pm 0,06$ ;  $S/\sigma = 0,48$ ;
- 3)  $H_{ср.г.} = 0,77 + 0,41 H_1$ ;  
 $r = 0,74 \pm 0,14$ ;  $S/\sigma = 0,67$ ;
- 4)  $H_{ср.г.} = 0,31 + 0,5 H_1$ ;  
 $r = 0,73 \pm 0,15$ ;  $S/\sigma = 0,68$ .



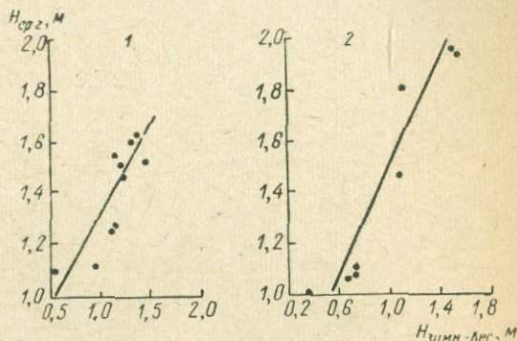
вод (рис. 9)\*. Из-за низкой заблаговременности их нельзя использовать для прогнозов, однако соответствующие уравнения регрессии можно применять для реконструкции среднегодовых уровней в тех случаях, когда имелись существенные пропуски в наблюдениях в периоды, предшествующие или последующие за летне-осенним спадом уровней.

В ряде случаев удается обнаружить весьма тесную зависимость того или иного уровня (или амплитуды) от одного режимообразующего фактора и использовать ее для составления прогнозов. Так, на территории северо-запада РСФСР часто наблюдаются весьма тесные связи между предвесенним минимальным уровнем

Рис. 8. Графики зависимости среднегодовых уровней грунтовых вод (в метрах от поверхности земли) от зимне-весенних максимальных среднемесячных уровней (1 — скв. 133, 2 — скв. 421, Калининградская обл.).

Уравнения регрессии:

- 1)  $H_{\text{ср.г}} = 0,74 + 0,57 H_{\text{зимн.-весн.}}$   
 $r = 0,88 \pm 0,07; S/\sigma = 0,44;$   
 2)  $H_{\text{ср.г}} = 0,42 + 1,08 H_{\text{зимн.-весн.}}$   
 $r = 0,90 \pm 0,07; S/\sigma = 0,43.$



и зимней температурой, которые ослабевают с увеличением глубины залегания уровней. При среднемноголетней глубине залегания предвесеннего минимума примерно до 2,0 м от поверхности земли соответствующие уравнения можно использовать для составления прогнозов этого уровня (рис. 10). На обширных территориях наблюдаются весьма тесные корреляционные связи между амплитудой весеннего (или зимне-весеннего) подъема уровней и суммой зимних осадков, между суммой зимне-весенних осадков и положением среднегодового уровня и т. д.

Ряд исследователей режима [30, 38, 41 и др.] находили зависимости между изменениями уровней грунтовых вод (главным образом среднегодовых) и изменениями солнечной активности. Поскольку некоторые показатели солнечной активности прогнозируются с достаточно высокой заблаговременностью, их можно использовать в качестве предсказателей положения уровней. Связи между уровнями и солнечной активностью могут быть двух видов: 1) внутри циклов, средняя продолжительность которых для годовых значений чисел Вольфа  $W$  составляет 11 лет; 2) внутривековые.

\* На рис. 9 и далее в тексте индексом «в» обозначены наблюдательные скважины Валдайской научно-исследовательской гидрологической лаборатории им. В. А. Урываева в Новгородской области.

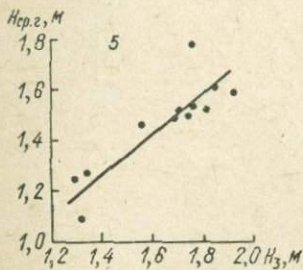
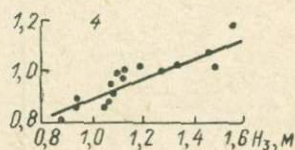
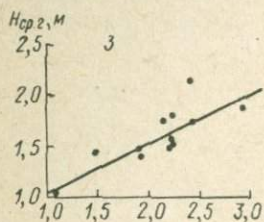
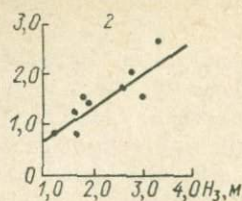
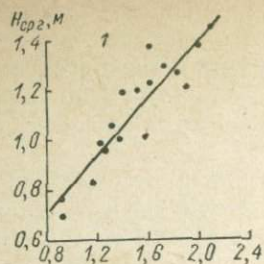


Рис. 9. Графики зависимости среднегодовых уровней грунтовых вод (в метрах от поверхности земли) в песках, супесях (1 — скв. 11, 2 — скв. 153, 5 — скв. 133) и в суглинках (3 — скв. 100, 4 — скв. 14в) от летне-осенних минимальных среднемесячных уровней  $H_3$ . Скв. 11, 100, 153 — г. Ленинград, скв. 133 — Калининградская обл., скв. 14в — Новгородская обл.

Уравнения регрессии:

- 1)  $H_{ср.г} = 0,25 + 0,57H_3$ ;  $r = 0,91 \pm 0,04$ ;  $S/\sigma = 0,27$ ;
- 2)  $H_{ср.г} = 0,07 + 0,65H_3$ ;  $r = 0,86 \pm 0,08$ ;  $S/\sigma = 0,39$ ;
- 3)  $H_{ср.г} = 0,57 + 0,48H_3$ ;  $r = 0,81 \pm 0,10$ ;  $S/\sigma = 0,60$ ;
- 4)  $H_{ср.г} = 0,47 + 0,42H_3$ ;  $r = 0,90 \pm 0,04$ ;  $S/\sigma = 0,44$ ;
- 5)  $H_{ср.г} = 0,10 + 0,82H_3$ ;  $r = 0,87 \pm 0,07$ ;  $S/\sigma = 0,50$ .

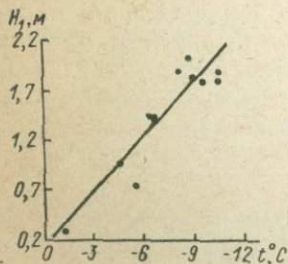


Рис. 10. График зависимости предвесеннего минимального среднемесячного уровня грунтовых вод  $H_1$  (в метрах от поверхности земли) от средней температуры воздуха за декабрь — февраль (скв. 170, г. Ленинград).

Уравнение регрессии:

$$H_1 = 0,15 - 0,18t_{XII-II}; r = 0,91 \pm 0,06; S/\sigma = 0,41.$$

Связи первого вида обнаруживаются лишь в том случае, если многолетние колебания уровней грунтовых вод имеют отчетливо выраженный циклический характер, при этом продолжительности

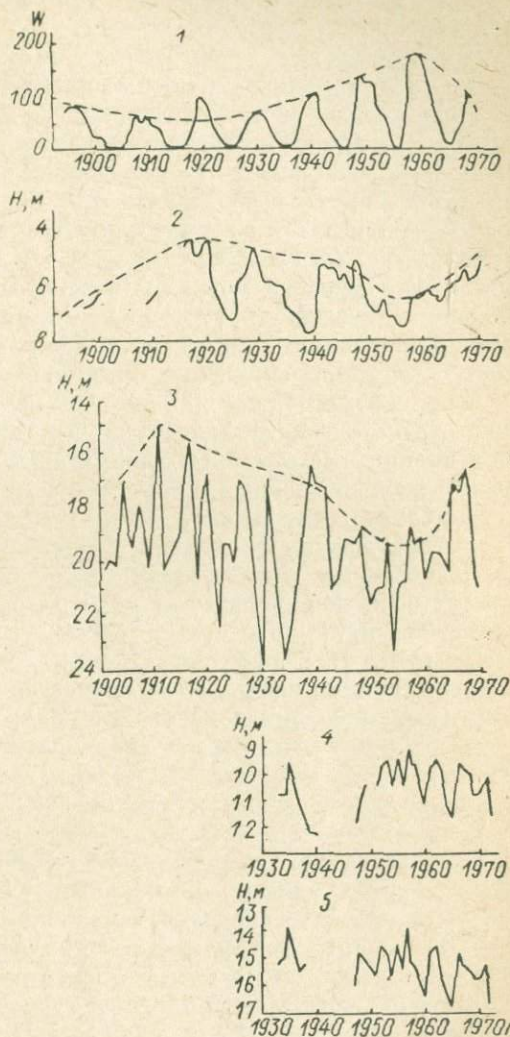


Рис. 11. Графики колебаний среднегодовых значений чисел Вольфа W (1) и среднегодовых уровней грунтовых вод (в метрах от поверхности земли) по шурфу 1 (2), по скв. В-12 (3, по данным [73]), 5-5 (4) и 11-11 (5). Шурф 1 — Воронежская обл.; скв В-12 — Чехословакия; скв. 5-5, 11-11 — Ленинградская обл.

циклов, выявленных в колебаниях уровней и солнечной активности, сопоставимы друг с другом (рис. 11, 1, 2). Существенной трудностью при составлении прогнозов, учитывающих эти связи, является их знакопеременный характер, т. е. положительная корреляция между изменениями солнечной активности и положением уровней, наблюдавшаяся на протяжении нескольких циклов, может в последующих циклах смениться отрицательной,

и наоборот. Если продолжительности циклов колебаний уровней и солнечной активности существенно отличаются друг от друга (рис. 11, 1, 3, 4, 5), то использование показателей солнечной активности (например, годовых значений  $W$ ) в качестве предикторов становится затруднительным.

Внутривековые связи обнаруживаются лишь при анализе достаточно длительных рядов наблюдений за уровнями. Так, в течение XX в. по некоторым скважинам (рис. 11, 1, 2, 3) отмечается обратная зависимость между вековыми изменениями солнечной активности и положением уровней грунтовых вод, которую можно использовать для составления сверхдолгосрочных качественных, а в ряде случаев и количественных прогнозов режима уровня. Следует отметить, что в качестве предсказателей положения уровней могут быть использованы не только годовые значения чисел Вольфа, но и некоторые другие показатели солнечной активности, например те из них, которые характеризуют корпускулярное излучение Солнца.

При использовании для прогнозов режима любого вероятностно-статистического метода, в том числе парной корреляции, для оценки достоверности полученных результатов необходимо иметь хотя бы общее представление о физическом характере устанавливаемых связей. Так, например, зависимость между величиной весеннего подъема уровней и суммой зимних осадков должна быть при прочих равных условиях прямой. Следовательно, коэффициент  $b$  в уравнении вида  $y = a + bx$ , где  $y$  — весенний подъем уровней,  $x$  — сумма зимних осадков, должен быть положительным. В провинции сезонного весенне-осеннего питания корреляционная связь между предвесенними минимальными уровнями, выраженными в метрах от поверхности земли, и зимними температурами должна быть обратной, т. е. чем ниже зимние температуры, тем значительнее вертикальный отток грунтовых вод к фронту промерзания и тем глубже положение предвесеннего минимума.

Следовательно, в уравнениях регрессии вида  $y = a + bx$ , где  $y$  — предвесенний минимальный уровень в метрах от поверхности земли,  $x$  — среднемесячная температура за XII—I или XII—II месяцы, коэффициент  $b$  должен быть со знаком минус. Таким образом, знак коэффициента  $b$  качественно определяет характер связи между прогнозируемым уровнем и соответствующим фактором. Если знак коэффициента  $b$  противоречит представлениям о физической сущности связи между прогнозируемым уровнем и тем или иным фактором, то даже в том случае, когда будет получено сравнительно высокое значение коэффициента корреляции  $r$ , следует признать, что данная связь является случайной, а соответствующие уравнения нельзя использовать для составления прогнозов. Чаще всего подобные явления наблюдаются в тех случаях, когда для расчетов уравнений используются сравнительно короткие ряды гидрогеологических наблюдений.

Таблица 4 — связь между колебаниями уровней грунтовых вод и уровней воды в реках

Сопоставляемые скважины и реки	Продолжительность совместных наблюдений, годы	Уравнение регрессии	Коэффициент корреляции и его стандартная ошибка	S/σ
Скв. 170 — р. Мга	12	$y = 7,52 + 2,454x$	$0,90 \pm 0,06$	0,41
Скв. 77 — р. Черная	9	$y = 50,74 + 1,409x$	$0,92 \pm 0,05$	0,39
Скв. 85 — р. Сестра	9	$y = 36,80 + 2,346x$	$0,83 \pm 0,10$	0,58
Скв. 160 — р. Мга	16	$y = 32,89 + 1,568x$	$0,79 \pm 0,12$	0,62

Помимо линейных связей между прогнозируемым уровнем и тем или иным фактором могут наблюдаться и нелинейные связи. Однако, располагая гидрогеологическими наблюдениями весьма небольшой продолжительности, следует в большинстве случаев ограничиться установлением линейных связей и нахождением линейных уравнений регрессии. В случае, если связи между уровнем и каким-либо фактором носят явно выраженный нелинейный характер, можно путем соответствующих преобразований привести их к линейным.

Парная корреляция может применяться не только для прогнозов, но также для удлинения и восстановления пропусков в рядах наблюдений за уровнями грунтовых вод. Чаще всего для этого используется связь между колебаниями уровней грунтовых вод и уровней воды в реках. Ниже, в табл. 4, приведены данные, характеризующие зависимость между колебаниями среднегодовых уровней грунтовых вод в скважинах, расположенных на междуречьях, и среднегодовых уровней воды в реках на территории г. Ленинграда и Ленинградской области.

Приведенные в табл. 4 уравнения использовались для удлинения рядов среднегодовых уровней грунтовых вод по гидрологическим данным.

## 2.4 — расчет линейного уравнения множественной регрессии

Поскольку естественный режим уровня грунтовых вод формируется, как правило, при одновременном воздействии нескольких факторов, для его прогноза во многих случаях оказывается целесообразным применение множественной корреляции. В области гидрологии множественная корреляция уже довольно

давно используется для прогнозов речного стока. Одним из первых для гидрогеологических прогнозов применил ее венгерский исследователь режима подземных вод И. Богарди [70].

В отечественной гидрогеологической литературе первое описание успешного применения линейной множественной корреляции содержится в работе А. А. Коноплянцева [37]. За последние годы этот метод нашел сравнительно широкое применение для анализа и прогноза режима уровня грунтовых вод [22, 58 и др.].

Уравнение множественной линейной регрессии имеет вид

$$\bar{y} = a + \sum_1^j b_j x_j, \quad (32)$$

где  $\bar{y}$  — прогнозируемый уровень — предиктант;  $x_j$  — режимообразующий фактор, или уровень за предшествующий период — предиктор;  $a$  — свободный член;  $b$  — коэффициенты уравнения регрессии.

Задача сводится к нахождению величин  $a$  и  $b_j$  по наблюдаемым значениям  $y$  и  $x$ . Рассмотрим простейший случай, когда колебания уровней определяются воздействием двух факторов, т. е.

$$\bar{y} = a + b_1 x_1 + b_2 x_2. \quad (33)$$

Согласно методу наименьших квадратов коэффициенты  $a$  и  $b$  находятся из условия

$$u = \sum_{i=1}^n (y_i - a - b_1 x_{i1} - b_2 x_{i2})^2 = \min. \quad (34)$$

Функция  $u$  обращается в минимум при условии, если  $du/da = 0$ ,  $du/db_1 = 0$ ,  $du/db_2 = 0$ . Найдя частные производные и приравняв их 0, получим систему нормальных уравнений:

$$na + b_1 \sum_{i=1}^n x_{i1} + b_2 \sum_{i=1}^n x_{i2} = \sum_{i=1}^n y_i; \quad (35)$$

$$a \sum_{i=1}^n x_{i1} + b_1 \sum_{i=1}^n x_{i1}^2 + b_2 \sum_{i=1}^n x_{i1} x_{i2} = \sum_{i=1}^n x_{i1} y_i; \quad (36)$$

$$a \sum_{i=1}^n x_{i2} + b_1 \sum_{i=1}^n x_{i1} x_{i2} + b_2 \sum_{i=1}^n x_{i2}^2 = \sum_{i=1}^n x_{i2} y_i. \quad (37)$$

Решив эту систему относительно  $b_1$  и  $b_2$  и проведя необходимые преобразования, получим

$$b_1 = \frac{\sigma_y}{\sigma_{x_1}} \frac{r_{yx_1} - r_{yx_2} r_{x_1 x_2}}{1 - r_{x_1 x_2}^2}; \quad (38)$$

$$b_2 = \frac{\sigma_y}{\sigma_{x_2}} \frac{r_{yx_2} - r_{yx_1} r_{x_1 x_2}}{1 - r_{x_1 x_2}^2}, \quad (39)$$

где  $r$  — коэффициенты парной корреляции между соответствующими переменными;  $\sigma$  — стандартные отклонения переменных.

Подставив величины  $b_1$  и  $b_2$  в одно из уравнений (35)—(37), можно найти величину  $a$ .

Коэффициент множественной корреляции для случая двух независимых переменных вычисляется по формуле

$$R = \sqrt{\frac{r_{yx_1}^2 + r_{yx_2}^2 - 2r_{yx_1}r_{yx_2}r_{x_1x_2}}{1 - r_{x_1x_2}^2}}. \quad (40)$$

Стандартная ошибка прогноза, осредненная для всей выборки данных, вычисляется по приближенной формуле

$$S_y = \sigma_y \sqrt{1 - R^2}. \quad (41)$$

Для вычисления стандартной ошибки прогноза при фиксированных значениях предикторов применяются весьма громоздкие формулы, которые здесь не приводятся. Отметим лишь, что так же, как и в случае парной корреляции, ошибки прогноза возрастают с увеличением отклонений предикторов от нормы.

Формула для вычисления частного коэффициента корреляции  $y$  по  $x_1$  при фиксированном  $x_2$ , т. е. при исключении влияния  $x_2$  на  $y$ , имеет вид

$$r_{yx_1 \cdot x_2} = \frac{r_{yx_1} - r_{yx_2}r_{x_1x_2}}{\sqrt{(1 - r_{yx_2}^2)(1 - r_{x_1x_2}^2)}}. \quad (42)$$

Аналогично для вычисления частного коэффициента корреляции  $y$  по  $x_2$  при фиксированном  $x_1$  имеем:

$$r_{yx_2 \cdot x_1} = \frac{r_{yx_2} - r_{yx_1}r_{x_1x_2}}{\sqrt{(1 - r_{yx_1}^2)(1 - r_{x_1x_2}^2)}}. \quad (43)$$

Доля детерминированного вклада каждого из  $j$  предикторов может быть определена по формуле

$$k_{x_j} = r_{yx_j} b_j \frac{\sigma_{x_j}}{\sigma_y}. \quad (44)$$

В большинстве случаев вклад каждого предиктора является положительным, т. е. знаки коэффициентов парной корреляции  $r_{yx}$  и коэффициентов регрессии  $b_j$  совпадают. Вклады могут стать отрицательными лишь в случае значительной корреляционной связи между предикторами.

Относительный вклад каждого предиктора в ту часть дисперсии предиктанта, которая определяется всеми  $j$  использованными предикторами, находится из выражения, приведенного в работе Г. А. Алексева [1].

$$\delta_{x_j} = \frac{k_{x_j}}{R^2} = \frac{r_{yx_j} b_j \frac{\sigma_{x_j}}{\sigma_y}}{R^2}. \quad (45)$$

В случае использования более двух предикторов формулы для вычисления уравнений регрессии становятся весьма громоздкими. Так, для трех предикторов уравнение регрессии, выраженное через  $r$  и  $\sigma$ , имеет следующий вид:

$$\begin{aligned} & \bar{y} - \bar{y} = \\ & = \frac{\sigma_y}{\sigma_{x_1}} \left[ \frac{r_{yx_1}(1 - r_{x_2x_3}^2) - r_{yx_2}r_{x_1x_2} - r_{yx_3}r_{x_1x_3} + r_{x_2x_3}(r_{yx_2}r_{x_1x_3} + r_{yx_3}r_{x_1x_2})}{1 - r_{x_2x_3}^2 - r_{x_1x_2}^2 - r_{x_1x_3}^2 + 2r_{x_2x_3}r_{x_1x_2}r_{x_1x_3}} \right] \times \\ & \times (x_1 - \bar{x}_1) + \frac{\sigma_y}{\sigma_{x_2}} \left[ \frac{r_{yx_2}(1 - r_{x_1x_3}^2) - r_{yx_1}r_{x_1x_2} - r_{yx_3}r_{x_2x_3} + r_{x_1x_3}(r_{yx_1}r_{x_2x_3} + r_{yx_2}r_{x_1x_3})}{1 - r_{x_2x_3}^2 - r_{x_1x_2}^2 - r_{x_1x_3}^2 + 2r_{x_2x_3}r_{x_1x_2}r_{x_1x_3}} \right] (x_2 - \bar{x}_2) + \\ & + \frac{\sigma_y}{\sigma_{x_3}} \left[ \frac{r_{yx_3}(1 - r_{x_1x_2}^2) - r_{yx_1}r_{x_1x_3} - r_{yx_2}r_{x_2x_3} + r_{x_1x_2}(r_{yx_1}r_{x_2x_3} + r_{yx_2}r_{x_1x_3})}{1 - r_{x_2x_3}^2 - r_{x_1x_2}^2 - r_{x_1x_3}^2 + 2r_{x_2x_3}r_{x_1x_2}r_{x_1x_3}} \right] \times \\ & \times (x_3 - \bar{x}_3), \end{aligned} \quad (46)$$

где  $r$  — коэффициенты парной корреляции между переменными  $y, x_1, x_2, x_3$ ;  $\sigma$  — стандартные отклонения переменных;  $\bar{y}, \bar{x}_1, \bar{x}_2, \bar{x}_3$  — средние значения переменных.

Раскрыв все скобки и решив это уравнение относительно  $y$ , получим уравнение вида

$$\bar{y} = a + b_1x_1 + b_2x_2 + b_3x_3. \quad (47)$$

Коэффициент множественной корреляции для случая трех предикторов находится с помощью определителей по формуле

$$R = \sqrt{1 - \frac{D_0}{D_{00}}}; \quad (48)$$

$$\begin{aligned} D_0 = & 1 - r_{yx_1}^2 - r_{yx_2}^2 - r_{yx_3}^2 - r_{x_1x_2}^2 - r_{x_1x_3}^2 - r_{x_2x_3}^2 + r_{yx_1}^2 r_{x_2x_3}^2 + \\ & + r_{x_1x_2}^2 r_{yx_3}^2 + r_{yx_2}^2 r_{x_1x_3}^2 + 2(r_{x_1x_2}r_{x_1x_3}r_{x_2x_3} + r_{x_2x_3}r_{yx_2}r_{yx_3} + r_{yx_1}r_{yx_2}r_{x_1x_3} + \\ & + r_{yx_1}r_{yx_3}r_{x_1x_2}) - 2(r_{yx_1}r_{yx_2}r_{x_1x_3} + r_{yx_2}r_{yx_3}r_{x_1x_2} + r_{yx_1}r_{yx_3}r_{x_2x_3}); \end{aligned} \quad (49)$$

$$D_{00} = 1 - r_{x_1x_2}^2 - r_{x_2x_3}^2 - r_{x_1x_3}^2 + 2r_{x_1x_2}r_{x_1x_3}r_{x_2x_3}. \quad (50)$$

Таким образом, для определения величин  $a, b_1, b_2, b_3$  и  $R$  необходимо вычислить шесть коэффициентов парной корреляции и четыре стандартных отклонения. При объемах выборок около 15—25 членов и с помощью настольных клавишных машин данная задача может быть решена за 1,0—1,5 дня. В настоящее время имеется большое число программ для ЭВМ различных типов, с помощью которых расчет уравнений регрессии не представляет существенных трудностей.

Таблица 5—исходные данные наблюдения

Год	Предвесенний минимальный уровень $y$	Среднемесячный уровень за декабрь месяца предыдущего года $x_1$	Средняя температура за декабрь $x_2$ , °С	Год	Предвесенний минимальный уровень $y$	Среднемесячный уровень за декабрь месяца предыдущего года $x_1$	Средняя температура за декабрь $x_2$ , °С
1958	1,17	0,89	-6,7	1964	1,80	1,49	-7,2
1959	1,52	1,39	-6,3	1965	2,02	1,86	-6,5
1960	1,94	1,63	-9,5	1966	2,42	2,19	-12,0
1961	1,68	1,54	-2,3	1967	1,88	1,74	-10,0
1962	1,93	1,86	-5,6	1968	2,46	1,69	-12,0
1963	1,42	1,05	-10,7	1969	2,34	1,45	-9,9

Ниже приводится пример расчета линейного уравнения регрессии для прогноза предвесеннего минимального среднемесячного уровня грунтовых вод в метрах от поверхности земли для скв. 2 на Северо-Западе РСФСР. В качестве предсказателей-предикторов использованы среднемесячный уровень за декабрь месяца предыдущего года и среднемесячная температура за XII—II месяцы. Исходные данные для расчета приведены в табл. 5 (уровни — в метрах от поверхности земли).

Коэффициенты парной корреляции между величинами  $y$ ,  $x_1$ ,  $x_2$  и соответствующие стандартные отклонения оказались следующими:  $r_{yx_1} = 0,77$ ;  $r_{yx_2} = -0,51$ ;  $r_{x_1x_2} = -0,18$ ;  $\sigma_y = 0,38$  м;  $\sigma_{x_1} = 0,34$  м;  $\sigma_{x_2} = -2,8^\circ$  С.

Используя значения  $r$  и  $\sigma$ , вычислим коэффициенты уравнения регрессии  $a$ ,  $b_1$ ,  $b_2$ , коэффициент множественной корреляции  $R$ , коэффициенты частной корреляции и вклад каждого предиктора в дисперсию предиктанта.

$$b_1 = \frac{0,38}{0,34} \cdot \frac{0,77 - (-0,51)(-0,18)}{1 - (-0,18)^2} = 0,784;$$

$$b_2 = \frac{0,38}{2,80} \cdot \frac{-0,51 - 0,77(-0,18)}{1 - (-0,18)^2} = -0,053;$$

$$a = \frac{22,58 - 14,72 - 5,23}{12} = 0,219;$$

$$R = \sqrt{\frac{0,77^2 + (-0,51)^2 - 2 \cdot 0,77(-0,51)(-0,18)}{1 - (-0,18)^2}} = 0,858;$$

$$S_y = 0,38 \quad 0,51 = 0,19 \text{ м};$$

$$r_{yx_1, x_2} = \frac{-0,51 - 0,77(-0,18)}{\sqrt{(1 - 0,77^2)[1 - (-0,18)^2]}} = -0,60;$$

$$r_{yx_2, x_1} = \frac{0,77 - (-0,51) \cdot (-0,18)}{\sqrt{[1 - (-0,51)^2][1 - (-0,18)^2]}} = 0,80;$$

$$k_{x_1} = \frac{0,77 \cdot 0,784 \cdot 0,34}{0,38} = 0,54; *$$

$$k_{x_2} = \frac{-0,51 \cdot (-0,053) \cdot 2,8}{0,38} = 0,20;$$

$$k_{x_1} + k_{x_2} = 0,54 + 0,20 = 0,74 = R^2.$$

Таким образом, уравнение регрессии предвесеннего минимального уровня имело вид:

$$\tilde{y} = 0,219 + 0,784x_1 - 0,053x_2.$$

Коэффициент множественной корреляции этого уравнения оказался равным 0,858. Величина отношения  $S/\sigma = 0,19 : 0,38 = 0,50$ , что позволяет считать принятую методику прогноза предвесеннего минимума удовлетворительной.

Вклад первого предиктора — среднемесячного уровня за декабрь — в общую дисперсию предвесеннего минимума составил 54%, вклад второго предиктора — средней температуры за декабрь — февраль — 20%. Иными словами, первый предиктор определяет 54% дисперсии предиктанта, второй — 20%. Коэффициенты частной корреляции оказались больше по абсолютной величине, чем соответствующие парные коэффициенты корреляции, т. е.  $|r_{yx_1 \cdot x_2}| > |r_{yx_1}|$  и  $|r_{yx_2 \cdot x_1}| > |r_{yx_2}|$ .

## 2.5 — расчет уравнения нормализованной множественной регрессии

Как указывалось выше, в зонах избыточного и переменного увлажнения распределения прогнозируемых уровней и используемых для прогнозов режимобразующих факторов во многих случаях близки к нормальному, а корреляционные связи между ними являются линейными либо близкими к ним. В тех случаях, когда эти условия не выполняются, исходные переменные необходимо предварительно нормализовать.

Ниже кратко излагается способ нормализации переменных, предложенный Г. А. Алексеевым [1].

Пусть между  $j+1$  переменными  $x_0, x_1, x_2, \dots, x_j$  существуют попарные монотонно возрастающие или убывающие нелинейные корреляционные связи\*. Тогда уравнение регрессии можно представить в виде

$$\bar{u}_0(x_0) = b_{01}u_1(x_1) + b_{02}u_2(x_2) + \dots + b_{0j}u_j(x_j), \quad (51)$$

где  $u_j(x_j)$  — нелинейные монотонные функции;  $b_{0j}$  — коэффициенты уравнения регрессии.

\* Здесь  $x_0$  — предиктанта (прогнозируемый уровень),  $x_j$  — предикторы (режимобразующие факторы).

Для нормализации переменных последовательно выполняются следующие преобразования. Все переменные ранжируются, т. е. располагаются в возрастающем (или убывающем) порядке. Так, уровни грунтовых вод в абсолютных отметках следует располагать в порядке их возрастания, а замеренные в метрах от поверхности земли — в порядке убывания. Каждому ранговому номеру переменной соответствует эмпирическая вероятность неперевышения  $P_j$  наблюдаемых значений  $x_{ji}$  ( $i = 1, 2, \dots, n$ ), которая находится по формуле

$$P_j(x_{ji}) \approx \frac{m(x_{ji}) - 0,25}{n + 0,5}, \quad (52)$$

где  $m(x_{ji}) = 1, 2, \dots, n$  — ранговые номера переменных  $x_{ji}$ ;  $n$  — число наблюдений переменных.

Полученные эмпирические вероятности неперевышения приравниваются к значениям теоретических вероятностей неперевышения. Последние определяются по известной таблице значений нормальной нормированной интегральной функции распределения:

$$P_j(x_{ji}) = P_m = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{u_m} e^{-\frac{u^2}{2}} du = \Phi(u_m). \quad (53)$$

Соответствующие эмпирические значения  $u_j(x_{ji}) = u_m$  нормализованной переменной  $u_j(x_j)$  определяются как обратные функции, или квантили, т. е.

$$u_m = F(P_m) = F[P_j(x_{ji})] = u_j(x_{ji}). \quad (54)$$

Для упрощения вычислений в упомянутой работе [1] приведены затабулированные эмпирические вероятности неперевышения  $P_m = P_j(x_{ji})$  для  $m = 1, 2, \dots, n$  и соответствующие им значения нормализованных переменных  $u_m = F[P_m]$  для выборок объемом  $n = 10 \div 50$ .

Не останавливаясь подробно на свойствах нормализованных переменных, отметим, что эмпирический коэффициент линейной (нормализованной) корреляции  $r_{jk}$  между двумя нормализованными переменными определяется по формуле

$$r_{jk} = r_{u_j u_k} = \frac{\sum_{i=1}^n u_{ji} u_{ki}}{\sum_{i=1}^n u_{ji}^2} = \frac{\mu_{jk}(n)}{\sigma_u^2(n)}. \quad (55)$$

Эмпирический коэффициент ковариации между двумя переменными находится из выражения

$$\mu_{jk}(n) = \frac{1}{n-1} \sum_{i=1}^n u_{ji} u_{ki}. \quad (56)$$



Коэффициент множественной нормализованной корреляции определяется по формуле

$$R_0^2 = r_{01}b_{01} + r_{02}b_{02} + \dots + r_{0j}b_{0j} = \sum_1^j r_{0j}b_{0j}. \quad (63)$$

Стандартная ошибка прогноза вычисляется как

$$S_0 = \sqrt{1 - R_0^2}. \quad (64)$$

Ниже приводится пример расчета уравнения нормализованной регрессии весеннего максимума для скв. 3-3 на Ордовикском плато в Ленинградской области. Исходные данные наблюдений, значения нормализованных переменных и их попарные произведения представлены в табл. 6. В качестве независимых переменных — предикторов — использованы среднемесячный уровень за январь  $x_1$  и сумма осадков за декабрь — март  $x_2$ . Зависимым переменным — предиктантом — является весенний максимальный среднемесячный уровень грунтовых карстовых вод  $x_0$ . Все уровни замерены в метрах от поверхности земли.

По суммам попарных произведений нормализованных переменных вычисляем эмпирические коэффициенты ковариации и нормализованной корреляции:

$$\begin{aligned} \mu_{01} &= \frac{5,7572}{17} = 0,3387; & r_{01} &= \frac{0,3387}{0,869} = 0,390; \\ \mu_{02} &= -\frac{9,0899}{17} = -0,5347; & r_{02} &= -\frac{0,5347}{0,869} = -0,615; \\ \mu_{12} &= \frac{1,8996}{17} = 0,1117; & r_{12} &= \frac{0,1117}{0,869} = 0,128. \end{aligned}$$

Находим коэффициенты уравнения регрессии и коэффициент множественной корреляции:

$$\begin{aligned} b_{01} &= \frac{0,390 - (-0,615) \cdot 0,128}{1 - 0,128^2} = \frac{0,469}{0,984} = 0,476. \\ b_{02} &= \frac{-0,615 - 0,390 \cdot 0,128}{1 - 0,128^2} = \frac{-0,665}{0,984} = -0,675; \end{aligned}$$

$$R_0^2 = 0,390 \cdot 0,476 + (-0,615) \cdot (-0,675) = 0,186 + 0,415 = 0,601;$$

$$R_0 = \sqrt{0,601} = 0,78.$$

Таким образом, уравнение нормализованной регрессии имеет следующий вид:

$$\bar{u}_0(x_0) = 0,476u_1(x_1) - 0,675u_2(x_2).$$

Используя зависимость  $x_0 = f[u_0]$  (рис. 12), получаем искомого уравнение регрессии, которое можно применять для составления прогнозов весеннего максимума:

$$\bar{x}_0 = f_0[0,476u_1(x_1) - 0,675u_2(x_2)].$$

Таблица 6 — наблюдаемые и нормализованные значения исходных переменных и произведения нормализованных величин

Год	Исходные переменные			Ранговые номера			Эмпирические значения нормализованных переменных*			Попарные произведения нормализованных переменных		
	$x_{0i}$	$x_{1i}$	$x_{2i}$	$m(x_{0i})$	$m(x_{1i})$	$m(x_{2i})$	$u_0(x_{0i})$	$u_1(x_{1i})$	$u_2(x_{2i})$	$u_0 u_{1j}$	$u_0 u_{2i}$	$u_1 u_{2i}$
1952	11,90	14,48	229	12	11	16	0,345	0,204	1,045	0,0704	0,3605	0,2132
1953	11,16	12,87	181	6	5	9	-0,493	-0,653	-0,068	0,3219	0,0335	0,0444
1954	12,93	12,00	120	15	3	3	0,831	-1,045	-1,045	-0,8684	-0,8684	1,0920
1955	9,33	13,65	232	1	8	18	-1,744	-0,204	1,744	0,3558	-3,0415	-0,3558
1956	11,66	14,08	159	10	10	6	0,068	0,068	-0,493	0,0046	-0,0335	-0,0335
1957	10,12	13,25	202	2	6	12	-1,315	-0,493	0,345	0,6483	-0,4537	-0,1701
1958	10,55	11,84	215	3	2	15	-1,045	-1,315	0,831	1,3742	-0,8684	-1,0928
1959	10,90	13,46	205	5	7	13	-0,653	-0,345	0,493	0,2253	-0,3219	-0,1701
1960	13,65	15,11	126	16	16	4	1,045	1,045	-0,831	1,0920	-0,8684	-0,8684
1961	11,85	15,04	167	11	15	8	0,204	0,831	-0,204	0,1695	-0,0416	-0,1695
1962	11,57	14,65	229	8	13	17	-0,204	0,493	1,315	-0,1006	-0,2683	0,6483
1963	12,08	10,99	105	13	1	2	0,493	-1,744	-1,315	-0,8598	-0,6483	2,2934
1964	13,65	14,81	141	17	14	5	1,315	0,653	-0,653	0,8587	-0,8587	-0,4264
1965	14,81	17,69	187	18	18	11	1,744	1,744	0,204	3,0415	0,3558	0,3558
1966	10,81	17,64	209	4	17	14	-0,831	1,315	0,653	-1,0928	-0,5426	0,8587
1967	11,30	12,31	166	7	4	7	-0,345	-0,831	-0,345	0,2867	0,1190	0,2867
1968	11,63	13,66	185	9	9	10	-0,068	-0,068	0,068	0,0046	-0,0046	-0,0046
1969	12,85	14,63	102	14	12	1	0,653	0,345	-1,744	0,2253	-1,1388	-0,6017
Сумма										5,7572	-9,0899	1,8996

\* Определяются по приложению 1 из работы Г. А. Алексеева [11].

Таблица 7 — проверка прогнозов весеннего максимума (в метрах от поверхности земли)

Год	Прогноз весеннего максимума	Весенний максимум по данным фактических замеров	Ошибка прогноза $\Delta$ , м
1970	13,90	13,16	0
1971	10,70	11,31	0,61
1972	14,80	14,62	0,18

Примечание. Вероятная ошибка прогноза  $\Delta_{\text{вер}}$  составляет 0,89 м.

С помощью этого уравнения и графиков  $u_j = f[x_j]$  (рис. 12) составим прогноз весеннего максимума на 1970 г., для которого  $x_1 = 15,51$  м,  $x_2 = 117$  мм.

$$\tilde{x}_0 = f_0 [0,476 \cdot 1,12 - 0,675 \cdot (-1,12)] = f_0 [1,29] = 13,90 \text{ м.}$$

В табл. 7 приводятся прогнозы весеннего максимума на 1970—1972 гг. и их сравнение с фактическими уровнями.

Интересно отметить, что коэффициенты парной корреляции между ненормализованными переменными (см. графы 2—4 табл. 6),

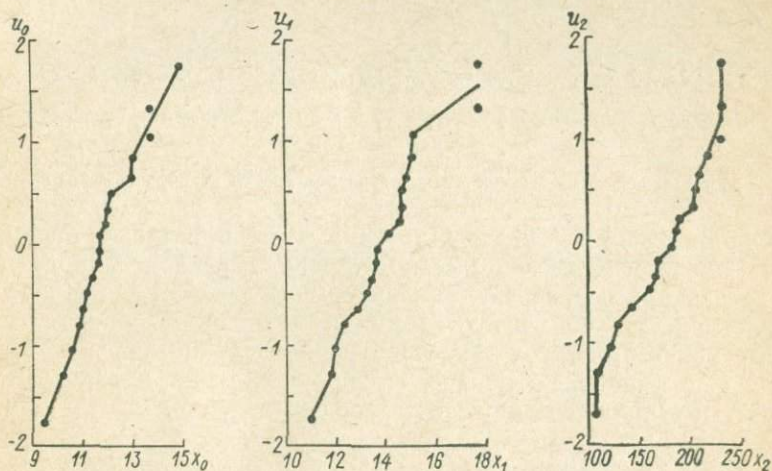


Рис. 12. Графики связи между исходными  $x_j$  и нормализованными  $u_j$  переменными по данным табл. 6.

а также коэффициент множественной корреляции ненормализованного уравнения регрессии составили:  $r_{x_0x_1} = 0,377$ ;  $r_{x_0x_2} = -0,587$ ;  $r_{x_1x_2} = 0,21$ ;  $R = 0,78$ . Эти значения практически равны приведенным выше значениям соответствующих коэффициентов для нормализованных переменных, т. е. нормализация не дала ощутимого эффекта. В данном случае это естественно,

поскольку почти прямолинейный вид графиков на рис. 12 показывает, что распределение исходных данных уже, по существу, было близким к нормальному. Предварительная нормализация дала наибольший эффект в том случае, если распределение исходных данных существенно отличается от нормального, а форма связи между переменными — от прямолинейной.

В качестве еще одного примера на рис. 13 изображены графики связи между нормализованными и исходными переменными, построенные по данным табл. 5. Почти прямолинейный вид

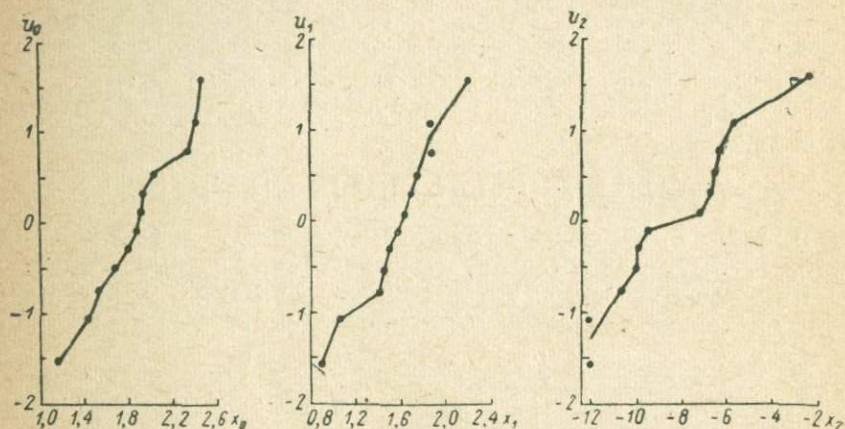


Рис. 13. Графики связи между исходными  $x_i$  и нормализованными  $u_i$  переменными по данным табл. 5.

$x_0$  — предвесенний минимальный среднемесячный уровень;  $x_1$  — среднемесячный уровень за декабрь;  $x_2$  — средняя температура воздуха за декабрь — февраль.

графиков показывает, что и в этом случае распределение переменных — предвесеннего минимального уровня, среднемесячного уровня за декабрь и средней температуры за декабрь — февраль — практически можно считать нормальным. Таким образом, сделанные ранее выводы о характере распределений уровней грунтовых вод в зоне избыточного увлажнения на Северо-Западе РСФСР получили дополнительное подтверждение.

## 2.6 — оценка достоверности коэффициента множественной корреляции и коэффициентов уравнения множественной регрессии

Поскольку коэффициент множественной корреляции  $R$ , а также величины  $a$  и  $b_j$  уравнений множественной регрессии определяются не для генеральной совокупности, а для выборок, ограниченных определенными периодами наблюдений, необходимо выяснить их достоверность и значимость.

Стандартную ошибку  $\sigma_R$  коэффициента множественной корреляции  $R$ , а также доверительные интервалы  $R$  можно вычислить

по формулам, аналогичным приведенным в разделе 2.2. Так, для вычисления стандартной ошибки  $R$  используется формула

$$\sigma_R = \frac{1-R^2}{\sqrt{n-j}}. \quad (65)$$

Стандартная ошибка квадрата коэффициента множественной корреляции находится по следующей формуле, приведенной в работе Г. А. Алексеева [1]:

$$\sigma_{R^2} = 2R\sigma_R + \sigma_R^2. \quad (66)$$

Отсюда относительная стандартная ошибка квадрата коэффициента корреляции равна

$$\frac{\sigma_{R^2}}{R^2} = 2 \frac{\sigma_R}{R} + \left(\frac{\sigma_R}{R}\right)^2 \approx 2 \frac{\sigma_R}{R}. \quad (67)$$

Для выборок сравнительно небольшого объема полезно вычислять скорректированное значение  $R$ , учитывающее продолжительность использованного периода наблюдений и число предикторов. Соответствующая формула имеет вид

$$R' = \sqrt{1 - \frac{n-1}{n-j}(1-R^2)}, \quad (68)$$

где  $R'$  — скорректированный коэффициент множественной корреляции;  $R$  — коэффициент множественной корреляции, вычисленный по данным наблюдений;  $n$  — длина использованного ряда или число наблюдений каждой из переменных;  $j$  — число факторов-предикторов.

В работе К. Брукса и Н. Карузера [8] приведена весьма удобная для практического использования таблица значимых величин коэффициента множественной корреляции  $R$  в зависимости от величин  $n$  и  $j$  (см. приложение).

С помощью приведенных выше формул оценим значимость коэффициента множественной корреляции  $R$ , вычисленного в разделе 2.4.

$$\sigma_R = \frac{1-0,86^2}{\sqrt{10}} = 0,08; \quad \frac{R}{\sigma_R} = 10,7.$$

Скорректированное значение коэффициента множественной корреляции оказалось равным

$$R' = \sqrt{1 - \frac{11}{10}(1-0,86^2)} = 0,85.$$

По приложению находим, что при  $n = 12$ ,  $j = 2$  и уровне значимости  $q = 5\%$  значимой величиной коэффициента множественной корреляции является  $R = 0,697$ , при  $q = 1\%$   $R = 0,80$ . Поскольку найденная величина  $R = 0,86 > R = 0,80$ , она является достоверной даже при однопроцентном уровне значимости.

Задача оценки значимости коэффициентов  $a$  и  $b_j$  в уравнении множественной регрессии является несколько более сложной. Представим уравнение регрессии (32) в виде

$$\bar{y} = a + \sum_{i=1}^j b_i (x_i - \bar{x}_i), \quad (69)$$

где  $\bar{x}_i$  — среднее значение предиктора  $x_i$ ,  $a = \bar{y}$ , т. е. среднему значению предиктанта.

Для вычисления стандартных ошибок  $a$  и  $b_j$  используются следующие выражения:

$$\sigma_a = \frac{S_y}{\sqrt{n-j-1}}; \quad (70)$$

$$\sigma_{b_j} = \frac{S_y}{\sqrt{n-j-1}} \sqrt{\frac{D_{ss}}{D_{00}}}, \quad (71)$$

(100 -  $q$ )%-ные доверительные границы для  $a$  и  $b_j$  находятся по формулам

$$a \pm t_{q, n-j-1} \frac{S_y}{\sqrt{n-j-1}} \quad (72)$$

$$b_j \pm t_{q, n-j-1} \frac{S_y}{\sqrt{n-j-1}} \sqrt{\frac{D_{ss}}{D_{00}}} \quad (73)$$

где  $S_y$  — стандартная ошибка прогноза;  $n$  — число наблюдений каждой из переменных;  $j$  — число факторов-предикторов;  $D_{00}$  — определитель матрицы, элементами которой являются ковариации предикторов;  $D_{ss}$  — главный минор определителя  $D_{00}$ , получаемый исключением из  $D_{00}$   $s$ -й строки и  $s$ -го столбца;  $t_q$  — значения  $q$ -процентных пределов для распределения Стьюдента.

Каждый элемент матрицы ковариаций предикторов можно представить как

$$l_{rs} = \frac{1}{n} \sum_{i=1}^n (x_{ri} - \bar{x}_r)(x_{si} - \bar{x}_s), \quad (1 \leq r \leq s \leq j). \quad (74)$$

При  $r = s$  получаем выражение для дисперсий предиктора:

$$l_{rr} = \frac{1}{n} \sum_{i=1}^n (x_{ri} - \bar{x}_r)^2. \quad (75)$$

Таким образом, определитель матрицы ковариаций предикторов будет иметь вид

$$D_{00} = \begin{vmatrix} l_{11} & l_{12} & \dots & l_{1j} \\ l_{21} & l_{22} & \dots & l_{2j} \\ \dots & \dots & \dots & \dots \\ l_{j1} & l_{j2} & \dots & l_{jj} \end{vmatrix} \quad (76)$$

На главной диагонали матрицы ( $l_{11}, l_{22}, \dots, l_{jj}$ ) помещены дисперсии всех предикторов. С помощью формулы (73) найдем доверительные интервалы для коэффициентов  $b_1$  и  $b_2$  уравнения, рассчитанного в разделе 2.4. Определим дисперсии и ковариации предикторов:

$$l_{11} = \sigma_{x_1}^2 = 0,34^2 = 0,1156 \text{ м}^2;$$

$$l_{22} = \sigma_{x_2}^2 = -2,8^2 = 7,84^\circ \text{ C};$$

$$l_{12} = l_{21} = \frac{1}{n} \sum_{i=1}^n (x_{1i} - \bar{x}_1)(x_{2i} - \bar{x}_2) = -0,172 \text{ м} \cdot ^\circ \text{C}.$$

Определитель матрицы ковариаций предикторов будет иметь следующий вид:

$$D_{00} = \begin{vmatrix} l_{11} & l_{12} \\ l_{21} & l_{22} \end{vmatrix} = \begin{vmatrix} 0,1156 & -0,172 \\ -0,172 & 7,84 \end{vmatrix} = \\ = 7,84 \cdot 0,1156 - (-0,172) \cdot (-0,172) = 0,876.$$

Миноры определителя соответственно равны  $D_{11} = 7,84$ ;  $D_{22} = 0,1156$ , отсюда

$$\sigma_{b_1} = \frac{0,19}{\sqrt{9}} \sqrt{\frac{7,84}{0,876}} = 0,19; \quad \frac{b_1}{\sigma_{b_1}} = \frac{0,784}{0,19} = 4,1;$$

$$\sigma_{b_2} = \frac{0,19}{\sqrt{9}} \sqrt{\frac{0,1156}{0,876}} = -0,024; \quad \frac{b_2}{\sigma_{b_2}} = \frac{-0,053}{-0,024} = 2,2.$$

При  $x_1 = \bar{x}_1$  и  $x_2 = \bar{x}_2$   $\sigma_a = \frac{0,19}{\sqrt{9}} = 0,065$ ;  $\frac{a}{\sigma_a} = \frac{1,88}{0,065} = 29$ .

Таким образом, отношения коэффициентов уравнения регрессии к их стандартным ошибкам оказалось больше 2, что свидетельствует об их значительной устойчивости. Используя выражение (73), определим доверительные интервалы для коэффициента  $b_1$  при 5%-ном уровне значимости ( $q_{5,9} = 2,26$ ):

$$G_{\text{верхн}} = 0,19 \cdot 2,26 + 0,784 = 1,21;$$

$$G_{\text{нижн}} = 0,784 - 0,19 \cdot 2,26 = 0,35.$$

Следовательно, с вероятностью 95% можно утверждать, что  $1,21 > b_1 > 0,35$ .

Аналогичные расчеты можно провести и для остальных коэффициентов уравнения регрессии. Если в результате подобных расчетов выяснится, что коэффициент  $b_j$  при каком-либо предикторе оказался неустойчивым ( $b_j < 2\sigma_{b_j}$ ), то в общем случае следует исключить этот предиктор из уравнения регрессии, которое рассчитывается вновь, уже без его участия.

Эффективность вклада каждого предиктора в общую дисперсию предиктанта определяется путем сравнения относительного вклада этого предиктора  $\delta_{x_j}$  с величиной относительной стандартной ошибки квадрата коэффициента множественной корреляции [1]. В случае, если

$$\delta_{x_j} < (1,5 \div 2) \frac{\sigma_{R^2}}{R^2} \approx (3 \div 4) \frac{\sigma_R}{R}, \quad (77)$$

то предиктор  $x_j$ , как недостаточно эффективный, можно из дальнейшего рассмотрения исключить.

Возвращаясь к приведенному выше примеру (см. табл. 5), находим

$$\begin{aligned} \delta_{x_1} &= \frac{0,54}{0,735} = 0,73; & \delta_{x_2} &= \frac{0,20}{0,735} = 0,27; \\ \frac{\sigma_{R^2}}{R^2} &\approx \frac{2\sigma_R}{R} = \frac{2,0 \cdot 0,08}{0,858} = 0,186; \\ \frac{\delta_{x_1}}{2 \frac{\sigma_R}{R}} &= \frac{0,73}{0,186} = 3,9; & \frac{\delta_{x_2}}{2 \frac{\sigma_R}{R}} &= \frac{0,27}{0,186} \approx 1,5. \end{aligned}$$

Таким образом, оба предиктора эффективны, а использование их в уравнении регрессии целесообразно.

Следует отметить, что условие (77) является, вообще говоря, весьма жестким, так как зачастую даже при  $\frac{b_j}{\sigma_{b_j}} > 2 \div 3$  оказывается, что  $\delta_{x_j} < (1,5 \div 2) \frac{\sigma_{R^2}}{R^2}$ . В связи с этим его выполнение в ряде случаев скорее желательно, чем необходимо.

## 2.7 — выбор факторов-предикторов для использования в уравнениях множественной регрессии

Возможности современной счетной техники, в частности ЭВМ, позволяют находить уравнения регрессии для прогнозов уровней с использованием большого числа предикторов. Казалось бы, что увеличение числа предикторов ведет к наиболее полному учету комплекса режимобразующих факторов, влияющих на формирование того или иного прогнозируемого уровня, и должно существенно повысить точность прогнозов. Однако в действительности это не так. Увеличение числа предикторов при неизменном объеме использованной выборки приводит к тому, что уравнения регрессии все более точно аппроксимируют прогнозируемый уровень внутри использованного периода наблюдений, однако при этом коэффициенты уравнений  $a$  и  $b_j$  становятся все более неустойчивыми и по абсолютной величине и по знаку. Неустой-

чивость коэффициентов приводит к тому, что эти уравнения искажают не только количественный характер связи прогнозируемого уровня с теми или иными факторами, но даже и ее качественный характер. В связи с этим проверка подобных уравнений на независимом фактическом материале дает неудовлетворительный результат, т. е. прогнозные уровни существенно отличаются от фактически измеренных, и применение таких уравнений для целей прогноза становится невозможным.

В литературе нет четких указаний о рациональном соотношении числа использованных предикторов  $j$  с числом наблюдений  $n$  или с продолжительностью периода наблюдений [1, 53, 64, 68 и др.]. Известно, что при  $n/j \rightarrow 1$  коэффициент множественной корреляции  $R \rightarrow 1$ , при этом устойчивость коэффициентов уравнения становится минимальной. Таким образом, чем больше величина отношения  $n/j$ , т. е. чем меньше предикторов используется в уравнении регрессии, тем более устойчивыми будут коэффициенты уравнения регрессии при неизменном объеме выборки  $n$ .

С другой стороны, режимобразующие факторы, используемые для прогнозов уровней, в той или иной степени коррелируют друг с другом. Давление, температура, осадки, влажность воздуха, взятые за одни и те же или смежные промежутки времени, оказываются связанными между собой. Так, например, согласно работе [15], на европейской территории СССР наблюдается сравнительно слабая, но отчетливая положительная корреляция между осадками и температурой в холодное время года (X—III месяцы) и более сильная, но уже отрицательная корреляция между этими же факторами в теплое время года. Как указывалось ранее, на Северо-Западе РСФСР часто наблюдаются существенные корреляционные связи между зимними и предвесенними минимальными уровнями и зимними температурами и т. д.

Использование в уравнении регрессии скоррелированных предикторов уменьшает устойчивость его коэффициентов в тем большей степени, чем теснее предикторы связаны друг с другом и чем слабее их связь с предиктантом. Использование сильно скоррелированных предикторов может привести к тому, что знаки некоторых коэффициентов регрессии изменятся на противоположные и им невозможно будет дать сколько-нибудь удовлетворительное физическое объяснение.

Таким образом, задача сводится к выбору минимального числа наиболее информативных предикторов, которые были бы наименее зависимы друг от друга. Корреляционные связи между предиктантом и предикторами, а также между самими предикторами можно учесть наиболее полно, вычислив матрицы нормированных коэффициентов корреляции. В качестве примера в табл. 8 приводится подобная матрица, вычисленная для одной из скважин на Северо-Западе РСФСР, характеризующей междуречный вид режима грунтовых вод в сусясах.

Таблица 8—матрица коэффициентов корреляции

	$y$	$x_1$	$x_2$	$x_3$	$x_4$	$x_5$
$y$	1	$0,72 \pm 0,16$	$-0,83 \pm 0,10$	$-0,50 \pm 0,25$	$0,05 \pm 0,28$	$-0,44 \pm 0,22$
$x_1$		1	$-0,30 \pm 0,23$	$-0,45 \pm 0,22$	$0,16 \pm 0,24$	$-0,68 \pm 0,15$
$x_2$			1	$0,91 \pm 0,05$	$0,03 \pm 0,28$	$0,45 \pm 0,22$
$x_3$				1	$0,01 \pm 0,28$	$0,48 \pm 0,21$
$x_4$					1	$-0,43 \pm 0,23$
$x_5$						1

В таблице приняты следующие условные обозначения:  $y$  — весенний максимальный среднемесячный уровень от поверхности земли, м;  $x_1$  — среднемесячный уровень за февраль;  $x_2$  — сумма осадков за декабрь — февраль;  $x_3$  — сумма осадков за декабрь — март;  $x_4$  — среднемесячная температура воздуха за сентябрь — октябрь предыдущего года;  $x_5$  — среднемесячная температура за декабрь — февраль. На пересечении столбца и строки помещены соответствующий коэффициент корреляции и его стандартная ошибка.

Анализ корреляционной матрицы показывает, что наиболее тесные связи наблюдаются между прогнозируемым уровнем — весенним максимумом — и уровнем за февраль месяца ( $r_{yx_1} = 0,72 \pm 0,16$ ), а также суммой осадков за декабрь — февраль ( $r_{yx_2} = -0,83 \pm 0,10$ ). В то же время связность между самими предикторами оказалась незначительной ( $r_{x_1x_2} = -0,30 \pm 0,23$ ). Таким образом, эти факторы можно использовать при выведении уравнения регрессии весеннего максимума.

Один из предикторов ( $x_4$ ) практически не связан с предиктантом ( $r_{yx_4} = 0,05 \pm 0,28$ ), в силу чего он не несет о нем никакой прогнозной информации и включение его в уравнение регрессии не рационально.

Несмотря на то что наблюдается корреляционная связь между весенним максимумом и зимними температурами ( $r_{yx_5} = -0,44 \pm 0,22$ ), включение этого предиктора в уравнение регрессии нежелательно по следующим соображениям. Зимние температуры тесно связаны с уровнями за февраль месяц ( $r_{x_1x_5} = -0,68 \pm 0,15$ ), и использование в уравнении регрессии двух скоррелированных предикторов существенно уменьшает устойчивость его коэффициентов. С другой стороны, этому же будет способствовать и то обстоятельство, что увеличение числа предикторов с 2 до 3 уменьшит величину отношения  $n/j$ , где  $n$  — продолжительность периода наблюдений;  $j$  — число предикторов. Изменение устойчивости коэффициентов уравнения регрессии после введения каждого дополнительного предиктора можно оценить по приведенным выше формулам (70), (71), найдя отношение  $b_j/\sigma_{b_j}$ .

Неустойчивость коэффициентов уравнений регрессии может наглядно проявляться в следующем. Пусть для одной скважины (например, для той, данные о которой приведены в табл. 8) получено уравнение регрессии весеннего максимума вида

$$y = a + b_1x_1 - b_2x_2 - b_5x_5. \quad (78)$$

В то же время для другой скважины, находящейся в аналогичных условиях формирования режима, в которой колебания уровней практически почти синхронны с колебаниями уровня в первой скважине, по данным за тот же период наблюдений получено уравнение регрессии вида

$$y = a + b_1x_1 - b_2x_2 + b_5x_5. \quad (79)$$

В этом случае невозможно найти какое-либо физическое объяснение изменению знака коэффициента  $b_5$  при зимней температуре на противоположный, поскольку оно связано исключительно с его малой устойчивостью.

При оценке правильности полученных решений необходимо исходить из общих физических предпосылок о характере влияния каждого использованного фактора на прогнозируемый уровень. Например, если при выведении уравнения регрессии любого экстремального уровня в качестве одного из предикторов был использован уровень за предшествующий период, то знак коэффициента уравнения регрессии при этом предикторе должен быть положительным. Точно так же в уравнении регрессии весеннего максимума знаки коэффициентов при предикторах — осадках за тот или иной промежуток времени должны быть отрицательными\* и т. д. Если же знак коэффициента при каком-либо предикторе не соответствует представлению о физическом характере связи, то, несмотря на то что коэффициент множественной корреляции  $R$  может быть высоким, полученное уравнение нельзя использовать на независимом материале, т. е. для целей прогноза.

Проведенные многочисленные расчеты коэффициентов уравнений регрессии и их устойчивости показали, что, как правило, при продолжительности режимных наблюдений, не превышающей 15—25 лет, следует использовать в уравнениях регрессии не более 2—3 предикторов. Весьма желательно при этом использование линейных уравнений регрессии вида (32), поскольку применение нелинейных зависимостей по сравнению с линейными существенно уменьшает устойчивость коэффициентов  $b_j$  при неизменных объемах использованных выборок.

Корреляционный анализ, проведенный по данным наблюдений в опорных и представительных скважинах, позволил установить, что при прогнозировании весеннего максимального среднемесячного уровня грунтовых вод на Северо-Западе РСФСР

\* Весенние максимумы измерены в метрах от поверхности земли.

Таблица 9 — эффективность и точность прогнозов весеннего максимального среднемесячного уровня грунтовых вод

Номер скважин	Исполь- зованный период наблюдений, годы	Скорректирован- ный коэффициент множественной корреляции $R'$ и его стандартная ошибка $\sigma R'$	Стандарт- ная ошибка проверочных прогнозов $S, м$	Стандарт- ное отклонение весеннего максимума $\sigma, м$	$S/\sigma$
11	17	$0,73 \pm 0,12$	0,47	0,27	0,63
22	17	$0,85 \pm 0,07$	0,08	0,16	0,50
153	11	$0,79 \pm 0,12$	0,21	0,36	0,59
3—3	18	$0,78 \pm 0,09$	0,83	1,32	0,63
42—1	14	$0,81 \pm 0,10$	0,79	1,34	0,60
2	13	$0,86 \pm 0,08$	0,15	0,32	0,47
3	15	$0,75 \pm 0,11$	0,17	0,27	0,63
Среднее	15	0,80			0,58

в качестве предикторов следует использовать среднемесячный уровень за январь или февраль и сумму осадков за декабрь — февраль, в которую в отдельных случаях вводилась со знаком минус величина зимнего испарения за этот же период. Для некоторых районов (например, для грунтовых карстовых вод на ордовикском плато в Ленинградской области) целесообразно учитывать сумму осадков за декабрь — март.

Скорректированные коэффициенты множественной корреляции  $R'$  уравнений регрессии весеннего максимума, полученных для опорных и представительных скважин на территории Северо-Запада РСФСР, а также величины отношений  $S/\sigma$ , характеризующие эффективность использованной методики, представлены в табл. 9.

При прогнозировании предвесеннего минимального среднемесячного уровня грунтовых вод на Северо-Западе РСФСР в качестве предикторов следует использовать среднемесячный уровень за декабрь предыдущего года (или за январь) и среднюю температуру за декабрь — январь или декабрь — февраль.

Данные об уравнениях регрессии предвесеннего минимума для опорных и представительных скважин на этой территории представлены в табл. 10.

Таким образом, методику прогноза весеннего максимума и предвесеннего минимума, оцененную по величине отношения  $S/\sigma$ , следует признать удовлетворительной. Однако высокие значения коэффициента множественной корреляции  $R$  и сравнительно низкие значения отношения  $S/\sigma$  свидетельствуют лишь о том, что найденные уравнения хорошо аппроксимируют фактические данные в пределах использованных периодов наблюдений. Для того чтобы убедиться в прогнозной ценности полученных

Таблица 10 — эффективность и точность прогнозов предвещенного минимального среднемесячного уровня грунтовых вод

Номер скважин	Использованный период наблюдений, годы	Скорректированный коэффициент множественной корреляции $R'$ и его стандартная ошибка $\sigma_{R'}$	Стандартная ошибка проверочных прогнозов $S$ , м	Стандартное отклонение зимнего минимума $\sigma$ , м	$S/\sigma$
11	18	$0,80 \pm 0,09$	0,19	0,32	0,60
160	14	$0,94 \pm 0,04$	0,07	0,22	0,32
251	10	$0,80 \pm 0,12$	0,39	0,71	0,55
2	12	$0,89 \pm 0,08$	0,19	0,38	0,50
1в	15	$0,86 \pm 0,07$	0,13	0,25	0,52
11в	11	$0,81 \pm 0,11$	0,13	0,22	0,59
26в	15	$0,82 \pm 0,08$	0,04	0,08	0,50
27в	14	$0,81 \pm 0,09$	0,10	0,18	0,55
Среднее	13	0,84			0,51

уравнений регрессии, их необходимо проверить на независимом фактическом материале. Подобная проверка показала, что в большинстве случаев ошибки прогнозов были меньше вероятных или допустимых ошибок, что позволяет использовать найденные уравнения для составления оперативных прогнозов.

Применяя корреляционный анализ, аналогичный описанному выше, можно подобрать предикторы для прогнозов экстремальных уровней не только в зоне избыточного увлажнения, но и в других природных условиях.

Следует отметить, что множественную корреляцию, как правило, наиболее рационально применять для прогнозов естественного режима уровня грунтовых вод при глубине его залегания до 3—4,0 м от поверхности земли. В этих условиях наблюдается наиболее активное воздействие на режим уровня разнообразных метеорологических факторов, прежде всего осадков и температуры. При большей глубине залегания уровней воздействие на их колебания метеорологических факторов ослабевает, поэтому для прогнозов естественного режима уровня в этих условиях можно во многих случаях использовать парную корреляцию, в частности между смежными экстремальными уровнями грунтовых вод.

## 2.8 — графическая множественная корреляция

Для прогнозов режима уровня грунтовых вод наряду с численной успешно может быть применена графическая множественная корреляция. Ее преимуществами по сравнению с численной корреляцией является сравнительная простота, меньшая

трудоемкость и большая наглядность. При использовании графической корреляции не требуется применения сложной счетной техники, в частности ЭВМ. К недостаткам этих методов относятся определенная субъективность при проведении изолиний, а также невозможность контролировать устойчивость полученных решений. Описание различных способов графической множественной корреляции (последовательной регрессии, метода поправок, соосных графиков) применительно к задачам метеорологических

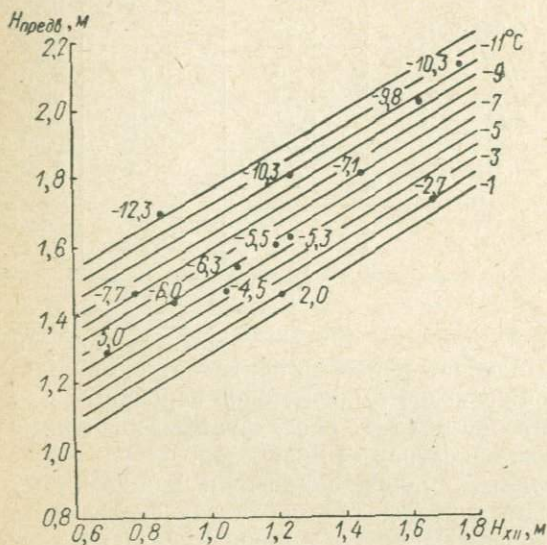


Рис. 14. Графики зависимости предвесеннего минимального уровня грунтовых вод  $H_{предв}$  (в метрах от поверхности земли) от среднемесячного уровня за декабрь  $H_{XII}$  и средней температуры воздуха за декабрь — февраль для одной из скважин на Северо-Западе РСФСР.

и гидрологических прогнозов приводится в некоторых работах [47, 53]. Впервые в отечественной гидрологической практике графическая множественная корреляция (метод поправок) была использована В. П. Пиварелис [54], а для прогноза режима уровня подземных вод — В. Н. Озябкиным [37]. Позднее В. С. Ковалевский [32, 33] для прогнозов экстремальных уровней грунтовых вод применил хорошо известные в гидрологии и метеорологии соосные графики.

При использовании двух предикторов графическая множественная корреляция осуществляется следующим образом. В системе  $yx_1$  наносятся точки, соответствующие этим переменным, и около каждой из них указывается значение переменной  $x_2$ . Через поле полученных точек проводятся изолинии значений  $x_2$ . На рис. 14 построены графики для прогноза предвесеннего минимума по одной из скважин на Северо-Западе РСФСР. В качестве предикторов использованы среднемесячный уровень за декабрь и средняя температура за XII—II месяцы.

## Применение корреляционных функций и авторегрессионных моделей для составления прогнозов режима уровня

### 3.1 — вычисление корреляционных функций и их анализ

Внутреннюю структуру рядов уровней грунтовых вод, т. е. степень связи между значениями уровня в разные моменты времени, можно охарактеризовать с помощью нормированной автокорреляционной функции. Одна из формул для ее вычисления имеет вид

$$r(\tau) = \frac{\sum_{t=1}^{n-\tau} (x_t - \bar{x}_1) (x_{t+\tau} - \bar{x}_2)}{\left[ \sum_{t=1}^{n-\tau} (x_t - \bar{x}_1)^2 \sum_{t=1}^{n-\tau} (x_{t+\tau} - \bar{x}_2)^2 \right]^{\frac{1}{2}}}, \quad (80)$$

где  $r(\tau)$  — коэффициент корреляции между уровнями, разделенными интервалом времени  $\tau$ ;  $x_t, x_{t+\tau}$  — значения уровней (суточных, среднемесячных, среднегодовых), разделенных интервалом времени  $\tau$ ;  $\bar{x}_1, \bar{x}_2$  — средние значения первых  $n - \tau$  и последних  $n - \tau$  членов ряда соответственно;  $n$  — число членов ряда.

При  $\tau = 0$   $r(\tau) = 1$ . При  $\tau > 0$  корреляционная функция характеризует степень линейной зависимости между наблюдениями, разделенными интервалом времени  $\tau$ .

Корреляционные функции рекомендуется вычислять для рядов большой продолжительности, содержащих несколько десятков, а еще лучше сотен членов. В этих случаях существенно повышается точность и надежность значений  $r(\tau)$  даже при сравнительно большой величине сдвига  $\tau$ . При вычислении корреляционной функции по сравнительно непродолжительным рядам к полученным значениям  $r(\tau)$  следует относиться с известной осторожностью.

Точность определения  $r(\tau)$  при любом сдвиге  $\tau$  можно определить по приведенным ранее формулам (26), (28), (29), в которых  $n$  следует заменить на  $n - \tau$ . В литературе не существует четких

указаний о допустимой величине сдвига  $\tau$ . Естественно, что чем меньше  $\tau$  и чем больше разность  $n - \tau$ , тем более точно вычисляются значения  $r(\tau)$ , однако при этом значительно уменьшается объем информации об исходном ряде, которую мы получаем в результате анализа корреляционной функции. Так, например, в работе О. А. Дроздова и А. С. Григорьевой [15] используется

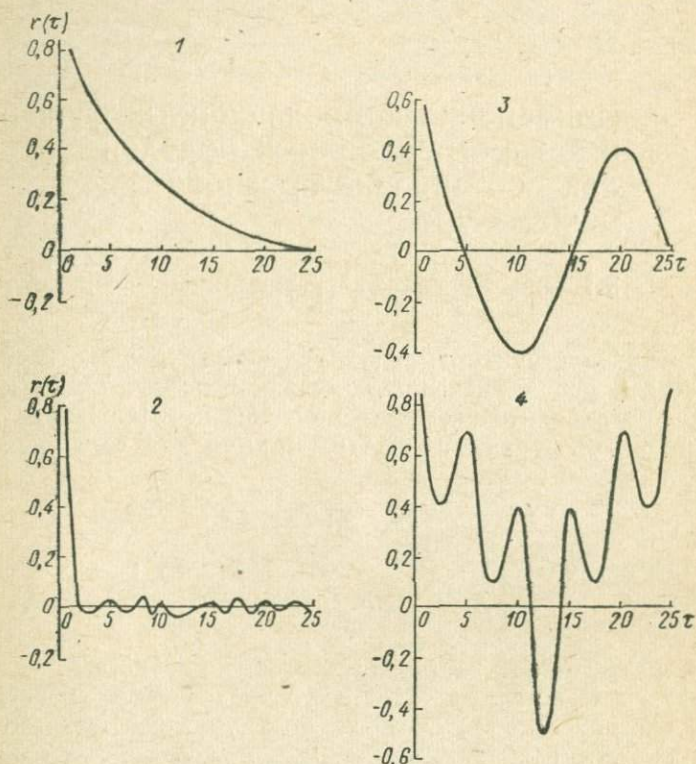


Рис. 15. Графики корреляционных функций процессов:

1 — с длиннопериодной составляющей ( $T > 25$ ); 2 — не содержащих периодических составляющих; 3 — с одной периодической составляющей ( $T = 20$ ); 4 — с двумя периодическими составляющими ( $T_1 = 25, T_2 = 5$ ).

сдвиг  $\tau$ , равный  $5/7$  длины исходного ряда. В работе И. П. Дружинина с соавторами [56] величина сдвига  $\tau$  превышает 0,5 и даже 0,75 длины исходного ряда. В Руководстве Всемирной метеорологической организации [72] рекомендуемая величина сдвига  $\tau$  составляет не более  $1/3$  длины ряда и т. д. При вычислении корреляционных функций рядов уровней грунтовых вод мы, как правило, использовали величину  $\tau$ , не превышающую 0,5—0,6 длины исходного ряда.

Корреляционные функции позволяют судить о временной структуре исследуемых рядов уровней. В случае, если корре-

корреляционная функция  $r(\tau)$  с возрастанием  $\tau$  уменьшается сравнительно медленно (рис. 15, 1), это свидетельствует о наличии существенной связности членов анализируемого ряда уровней. Если с возрастанием  $\tau$  корреляционная функция быстро убывает и затем колеблется около 0 (рис. 15, 2), то это свидетельствует об отсутствии существенной связности между членами ряда.

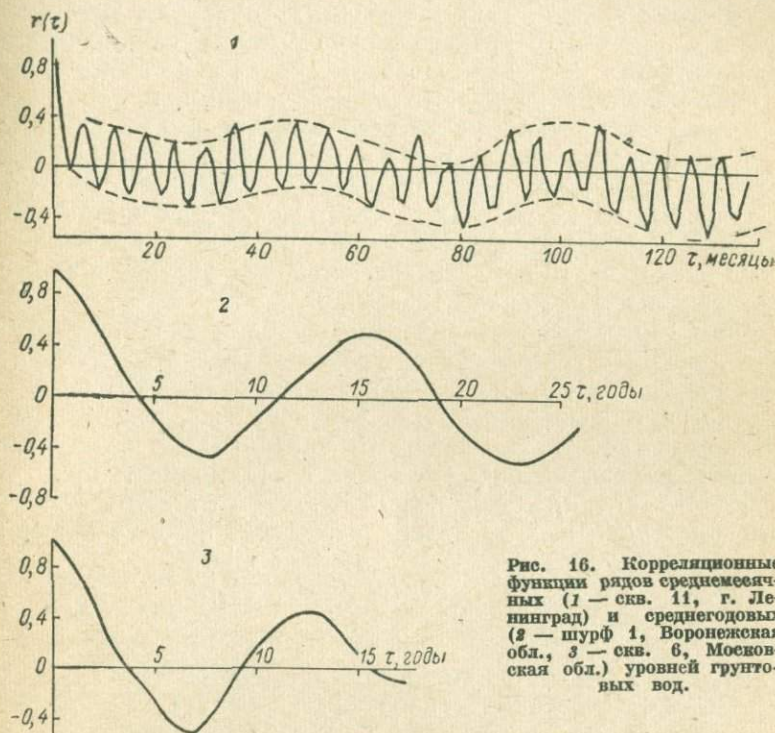


Рис. 16. Корреляционные функции рядов среднемесячных (1 — скв. 11, г. Ленинград) и среднегодовых (2 — шурф 1, Воронежская обл., 3 — скв. 6, Московская обл.) уровней грунтовых вод.

Если построить корреляционную функцию синусоидального временного ряда с периодом  $T$ , то она будет иметь вид косинусоиды с тем же периодом  $T$  (рис. 15, 3). Таким образом, по виду корреляционной функции можно судить о наличии одной или нескольких периодических составляющих в исходном ряде уровней (рис. 15, 4).

Проанализируем с этой точки зрения некоторые конкретные корреляционные функции. На рис. 16 представлена корреляционная функция ряда среднемесячных уровней грунтовых вод по данным наблюдений в скв. 11 (г. Ленинград, 1949—1969 гг.). Правильно чередующиеся максимумы и минимумы этой функции, разделенные интервалом 6 месяцев, фиксируют сезонную составляющую колебаний продолжительностью 0,5 года. Кроме того, по виду корреляционной функции, в частности, по поведению ее верхней и нижней огибающих, можно судить и о других

периодических составляющих, содержащихся в исходном ряде среднемесячных уровней. Так, например, минимальное значение корреляционной функции при  $\tau = 128$  свидетельствует о наличии периодической составляющей ориентировочной продолжительностью 256 месяцев, или 21 год. Четкие максимумы и минимумы огибающих, разделенные интервалом  $\tau = 44 \div 50$  месяцев, свидетельствуют о наличии 4-летней периодической составляющей в колебаниях уровня. Таким образом, по корреляционным функциям среднемесячных уровней можно выявить не только сезонные, но и более продолжительные периодические составляющие.

Представленные на рис. 16 корреляционные функции рядов среднегодовых уровней по шурфу 1 (Воронежская область) и скв. 6 (Московская область) имеют почти правильную косинусоидальную форму, что свидетельствует о наличии в этих рядах существенным периодических составляющих. Для среднегодовых уровней в шурфе 1 характерен период колебаний продолжительностью около 15—16 лет, для среднегодовых уровней в скв. 6 — продолжительностью около 13—14 лет.

Таким образом, корреляционная функция характеризует наличие линейных связей между уровнями, разделенными интервалом времени  $\tau$ , и принимает максимальные значения в тех случаях, когда  $\tau$  равно или кратно периоду, содержащемуся в исходном ряде уровней. В связи с этим максимальные и достоверные значения корреляционной функции  $r(\tau)$  (например, при  $r(\tau) \geq 0,6 \div 0,7$  и  $r(\tau)/\sigma_{r(\tau)} > 2$ ) можно использовать для составления прогнозных линейных уравнений регрессии, которые имеют вид

$$x_{t+\tau} - \bar{x}_{t+\tau} = r(\tau) \frac{\sigma_{x_{t+\tau}}}{\sigma_{x_t}} (x_t - \bar{x}_t), \quad (81)$$

где  $x_{t+\tau}$  и  $x_t$  — уровни в моменты (интервалы) времени  $t + \tau$  и  $t$ ;  $\bar{x}_{t+\tau}$  и  $\bar{x}_t$  — соответствующие средние значения уровней;  $\sigma_{x_{t+\tau}}$ ,  $\sigma_{x_t}$  — стандартные отклонения рядов уровней  $x_{t+\tau}$  и  $x_t$ ;  $r(\tau)$  — значение корреляционной функции при сдвиге  $\tau$ . При сравнительно небольших значениях  $\tau$   $\frac{\sigma_{x_{t+\tau}}}{\sigma_{x_t}} \approx 1$ ,  $\bar{x}_{t+\tau} \approx \bar{x}_t$  и формула приобретает вид

$$x_{t+\tau} - \bar{x}_t = r(\tau) (x_t - \bar{x}_t). \quad (82)$$

В этом случае, зная положение уровня в момент времени  $t$ , можно с заблаговременностью  $\tau$  составить прогноз о его вероятном положении в момент времени  $t + \tau$ .

Например, корреляционная функция среднегодовых уровней грунтовых вод по скв. 6 (Московская область) имеет максимальное значение (за исключением  $r_{(0)} = 1$ ) при  $\tau = 1$  (см. рис. 16), при этом  $r_{(1)} = 0,78$ . Уравнение регрессии для прогноза среднегодовых уровней с заблаговременностью  $\tau = 1$  год имеет вид

$$\tilde{x}_{t+1} = 4,86 + 0,78 (x_t - 4,86) \pm 0,38.$$

В качестве пределов изменения прогнозируемых уровней с обеспеченностью 68% указана соответствующая стандартная ошибка проверочных прогнозов. Эффективность и точность приведенной зависимости характеризуется величиной отношения  $S/\sigma = 0,66$ .

При  $\tau > 1$  уравнения регрессии для прогнозов среднегодовых уровней в скв. 6 с заблаговременностью свыше 1 года становятся недостаточно точными и эффективными.

### 3.2 — простейшие авторегрессионные модели, описывающие колебания уровней

Поскольку значения корреляционной функции многих рядов уровней убывают до 0 не сразу, а постепенно, естественно задаться вопросом, нельзя ли предсказать будущие значения уровня по одному или нескольким его прошлым значениям, т. е. решить задачу экстраполяции уровня.

Проблеме экстраполяции временных рядов посвящены классические работы А. Н. Колмогорова [34], А. М. Яглома [69] и других. Линейная экстраполяция гидрологических и некоторых метеорологических рядов рассматривается в работе Ю. М. АLEXИНА [2]. Задача линейной экстраполяции (предсказания) данных режимных гидрогеологических наблюдений впервые была поставлена В. Н. Озябкиным [51] и позднее доведена до численного решения на ЭВМ С. П. Сидоркиной [61].

В связи с линейной экстраполяцией рядов уровней необходимо рассмотреть вопрос об их стационарности. Колебания уровней подземных вод (или любого другого элемента режима) называются стационарными в широком смысле в случае, если их математическое ожидание и дисперсия постоянны во времени, а значения корреляционной функции зависят лишь от интервала времени  $\tau$  и не зависят от начала отсчета. В настоящее время однозначно решить вопрос о стационарности естественных колебаний уровней грунтовых вод весьма трудно из-за непродолжительности имеющихся рядов наблюдений. Часто нестационарность связана с тем, что в колебаниях уровней отчетливо проявляются циклы, продолжительность которых значительно превосходит период наблюдений. С другой стороны, она может быть вызвана направленным воздействием на режим уровня, которое осуществляется с помощью разнообразных мелиоративных и инженерных мероприятий, таких как осушение, орошение территорий, сооружение плотин и водохранилищ, вызывающее подпор грунтовых вод, и др. (рис. 17).

В случае явной нестационарности рядов уровней с помощью известных математических приемов можно выделить в них тренд\*,

\* Под трендом подразумевается более или менее равномерное снижение или повышение уровня, которое отмечается в течение всего периода наблюдений.

Таблица 11 — средние значения и стандартные отклонения рядов среднегодовых уровней по шурфу 1

Номер ряда	$\bar{x} \pm \sigma_{\bar{x}}, \text{ м}$	$\sigma_{x_i}, \text{ м}$
1 (1915—1940 гг.)	$6,18 \pm 0,43$	0,99
2 (1941—1967 гг.)	$6,47 \pm 0,23$	0,66
3 (1915—1967 гг.)	$6,33 \pm 0,20$	0,80

исключить его из исходных данных и привести ряды к стационарным. Однако применение вероятностно-статистических методов для прогноза этих рядов оказывается крайне затруднительным, так как предсказать развитие тренда за пределами использованного периода наблюдений оказывается, как правило, весьма трудно. Нами сделана попытка оценить в первом приближении стационарность некоторых натуральных данных режимных гидрогеологических наблюдений, а именно: среднегодовых уровней грунтовых вод по шурфу 1 (Воронежская область, 1915—1967 гг.) и

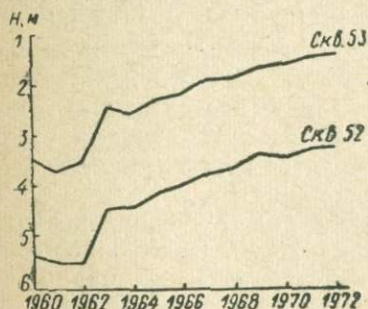
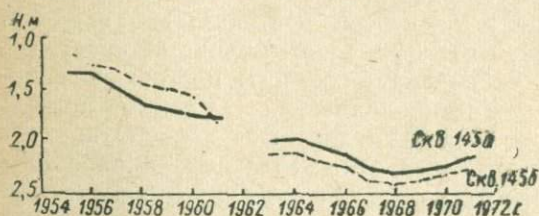


Рис. 17. Графики колебаний среднегодовых уровней грунтовых вод (в метрах от поверхности земли) в скв. 52, 53 (Молдавская ССР, пос. Рыбница, зона подпюра) и предвесенних минимальных уровней в скв. 145а, 145б (Ленинградская обл., м. олов осушения).



скв. 6 (Московская область, 1934—1967 гг.), а также среднемесячных уровней по скв. 160 (Ленинградская область, 1955—1968 гг.).

Для того чтобы оценить стационарность ряда, т. е. независимость его статистических характеристик от начала отсчета, можно вычислить эти характеристики отдельно для двух его неперекрывающихся частей и сравнить их между собой. В случае стационарности значения этих характеристик для каждого отрезка исходного ряда должны незначительно отличаться друг от друга. Так, например, ряд среднегодовых уровней по шурфу 1 был

Таблица 12 — значения корреляционной функции рядов среднегодовых уровней по шурфу 1

Номер ряда	$r(1)$	$r(2)$	$r(3)$	$r(4)$
1 (1915— 1940 гг.)	0,85 [0,93—0,69]	0,62 [0,81—0,31]	0,36 [0,70—0,02]	0,02 [0,42—(-0,38)]
2 (1941— 1967 гг.)	0,69 [0,81—0,42]	0,43 [0,69—0,02]	0,18 [0,56—(-0,20)]	0,12 [0,50—(-0,26)]
3 (1915— 1967 гг.)	0,85 [0,91—0,74]	0,62 [0,76—0,43]	0,35 [0,59—0,10]	0,09 [0,37—(-0,19)]

разбит на две примерно равные части: первая охватывала период наблюдений с 1915 по 1940 г., вторая — с 1941 по 1967 г. Для каждого ряда было вычислено среднее значение  $\bar{x}$  (в метрах от поверхности земли), стандартное отклонение  $\sigma_x$  (табл. 11) и первые четыре значения корреляционной функции  $r(r)$  (табл. 12).

При вычислении стандартной ошибки средних значений рядов  $\sigma_{\bar{x}}$  и стандартных отклонений самих рядов  $\sigma_x$  учитывалось наличие внутрирядной корреляции.

Поскольку, как указывалось выше, распределение среднегодовых уровней грунтовых вод в шурфе 1 близко к нормальному, для определения расхождений в значениях  $\bar{x}$  и  $\sigma_x$  рядов 1 и 2 можно использовать известные в математической статистике критерии. Так, для определения значимости расхождения средних следует найти величину

$$\sigma_{(\bar{x}_1 - \bar{x}_2)} = \sqrt{\frac{\sigma_{x_1}^2}{n_1} + \frac{\sigma_{x_2}^2}{n_2}}, \quad (83)$$

где  $\sigma_{x_1}$  и  $\sigma_{x_2}$  — стандартные отклонения рядов 1 и 2;  $n_1$  и  $n_2$  — число членов рядов 1 и 2.

Подставив соответствующие значения, получим

$$\sigma_{(\bar{x}_1 - \bar{x}_2)} = \sqrt{\frac{0,99^2}{26} + \frac{0,66^2}{27}} = 0,23 \text{ м.}$$

При уровне значимости  $q = 5\%$  расхождение между средними не должно превышать  $1,96\sigma_{\bar{x}_1 - \bar{x}_2} = 0,45$  м. Фактическая разность между  $\bar{x}_1$  и  $\bar{x}_2$  составила  $6,47 - 6,18 = 0,29$  м, т. е. практически ею можно пренебречь.

Различия дисперсий рядов 1 и 2 можно оценить с помощью критерия Р. Фишера

$$F = \frac{\sigma_1^2}{\sigma_2^2} = \frac{0,99^2}{0,66^2} = 2,25. \quad (84)$$

Таблица 13 — средние значения, стандартные отклонения и значения корреляционной функции рядов среднегодовых уровней по шурфу 1\*

Показатели	Номер ряда и периоды наблюдений	
	1 (1915—1940 гг.)	2 (1941—1967 гг.)
$\bar{x} \pm \sigma_x, \text{ м}$	6,35±0,33	6,11±0,30
$\sigma_x, \text{ м}$	0,94	0,84
$r_{(1)}$	0,79	0,82
$r_{(2)}$	0,49	0,43
$r_{(3)}$	0,18	0,18
$r_{(4)}$	0,20	0,15

\* После исключения периодической составляющей продолжительностью 67 лет.

При уровне значимости  $q = 1\%$  величина  $F$  не должна превышать 2,56\*. Следовательно, в первом приближении можно считать, что различия дисперсий рядов 1 и 2 являются несущественными.

В табл. 12 под каждым значением корреляционной функции указан интервал, соответствующий его 95%-ной вероятности. Анализ данных табл. 11, 12 показывает, что средние значения, стандартные отклонения и первые значения корреляционных функций рядов 1 и 2 мало отличаются друг от друга. Это позволяет в свою очередь считать ряд 3 близким к стационарному.

Устойчивость исследуемых рядов существенно возрастает после исключения из исходных данных периодической составляющей ориентировочной продолжительностью 67 лет, выделенной в результате проведенного частотного анализа\*\* (табл. 13).

Аналогичные вычисления были выполнены для рядов среднегодовых уровней по скв. 6 (1934—1967 гг.), а также для среднемесячных и среднемесечных экстремальных\*\*\* уровней грунтовых вод по скв. 160, расположенной в Ленинградской области, за период наблюдений с 1955 по 1968 г., при этом исходные ряды по скв. 160 также разбивались на две части: первая охватывала период наблюдений с 1955 по 1961 г., вторая — с 1962 по 1968 г. Найденные статистические характеристики этих рядов оказались весьма близкими. Так, например, среднее значение и дисперсия для

\* Таблицы для нахождения предельных значений  $F$  помещены в работах [35, 64] и многих других.

\*\* Методика выявления периодических составляющих в колебаниях уровня изложена в разделе 4.

\*\*\* Этот ряд составлен из предвесенних минимальных, весенних максимальных, летне-осенних минимальных и осенних максимальных среднемесечных уровней, взятых в хронологической последовательности за весь период наблюдений.

первой половины ряда экстремальных среднемесячных уровней по скв. 160 оказались равными  $1,83 \pm 0,05$  м и  $0,052$  м<sup>2</sup>, для второй половины —  $1,93 \pm 0,05$  м и  $0,046$  м<sup>2</sup>. Первые значения корреляционной функции этих же рядов, из которых предварительно исключена сезонная составляющая продолжительностью 0,5 года, показаны на рис. 18.

Хотя приведенные выше доказательства стационарности рядов уровней грунтовых вод являются недостаточно полными из-за относительной краткости периодов наблюдений, особенно при использовании отдельных частей рядов, по-видимому, можно считать исследованные ряды весьма близкими к стационарным. В связи с этим для их экстраполяции (прогнозирования) может быть использована авторегрессионная модель вида

$$x_t - \bar{x} = \sum_1^m a_m (x_{t-m} - \bar{x}) + z_t, \quad (85)$$

где  $x_t$  — уровень, прогнозируемый на момент времени  $t$ ;  $\bar{x}$  — среднее значение уровня за период наблюдений;  $a_m$  — коэффициенты авторегрессии;  $x_{t-m}$  — известные значения уровня в моменты времени  $t - m$ ;  $z_t$  — случайная компонента.

Задача прогнозирования сводится к нахождению  $\bar{x}$  и  $a_m$  из условия

$$u = \sum_{t=m+1}^n \left[ (x_t - \bar{x}) - \sum_1^m a_m (x_{t-m} - \bar{x}) \right]^2 = \min, \quad (86)$$

где  $n$  — использованный период наблюдений, остальные обозначения приведены выше.

Дифференцируя уравнение (86) по  $\bar{x}$  и  $a_m$  и приравнявая эти производные нулю, а также принимая допущения, что при сдвиге времени  $\tau = 1 \div 3$  среднее значение ряда и его ковариации существенно не меняются, в работах [13 и 71] авторы приводят сравнительно простые выражения для первых трех коэффициентов авторегрессии и остаточной дисперсии.

Так, для процесса авторегрессии 1-го порядка ( $m = 1$ ) имеем

$$a_1^* \approx r_{(1)}. \quad (87)$$

Нетрудно убедиться, что при  $a_1 = r_{(1)}$  выражение (85) превращается в приведенную ранее формулу (82), где  $r_{(\tau)} = r_{(1)}$ .

Для процесса авторегрессии 2-го порядка ( $m = 2$ )

$$a_1 \approx \frac{r_{(1)} [1 - r_{(2)}]}{1 - r_{(1)}^2}; \quad a_2 \approx \frac{r_{(2)} - r_{(1)}^2}{1 - r_{(1)}^2}, \quad (88)$$

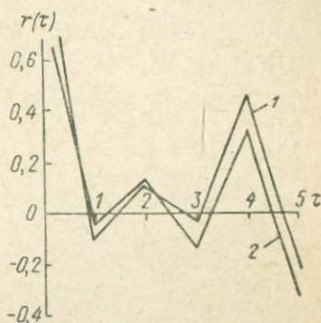


Рис. 18. Корреляционные функции рядов экстремальных среднемесячных уровней грунтовых вод по скв. 160 (Ленинградская обл.).  
1 — 1955—1961 гг., 2 — 1962—1968 гг.

где  $a_1, a_2$  — коэффициенты авторегрессии;  $r(\tau)$  — значения корреляционной функции при сдвиге  $\tau = 1, 2$ .

Аналогичным образом могут быть найдены коэффициенты авторегрессии более высоких порядков. Остаточная дисперсия для процесса авторегрессии любого порядка находится из выражения

$$\sigma_{z_t}^2 = \frac{n-m}{n-2m-1} \left[ C_{(0)} - \sum_1^m a_m C_m \right], \quad (89)$$

где  $n$  — число членов ряда;  $m$  — порядок авторегрессии;  $C_{(0)}$  — дисперсия исходного ряда;  $C_{(m)}$  — автоковариации при сдвиге  $\tau = m$ .

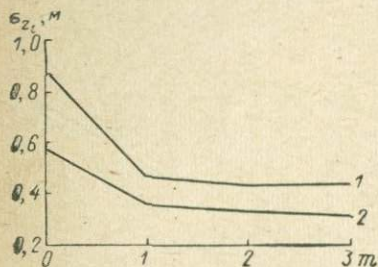


Рис. 19. Графики зависимости стандартного отклонения остатка  $\sigma_{z_t}$  от порядка авторегрессии  $m$ .

1 — шурф 1 (Воронежская обл.); 2 — скв. 6 (Московская обл.).

Имея первые значения корреляционной функции  $r(\tau)$ , можно найти коэффициенты авторегрессии  $a_m$  и стандартные отклонения остатка  $\sigma_{z_t}$ . Вычислив значения  $\sigma_{z_t}$  для авторегрессионных моделей 1, 2 и 3-го порядков, можно решить вопрос о порядке процесса  $m$ . Для построения графиков  $\sigma_{z_t} = f(m)$  (рис. 19) были использованы ряды среднегодовых уровней по шурфу 1 (1915—1967 гг.) и скв. 6 (1934—1967 гг.).

Анализ графиков показывает, что с увеличением порядка авторегрессии от 2 до 3 дисперсия остатка уменьшается очень незначительно, поэтому для прогнозов с заблаговременностью  $(1 \div 2) \Delta t$  следует ограничиться моделями 1-го и 2-го порядков. Для окончательного решения вопроса о порядке модели нужно анализ дисперсий остатка  $\sigma_{z_t}$  дополнить вычислением доверительных пределов для каждого вновь добавляемого коэффициента авторегрессии  $a_m$  [13]. Отметим, что с увеличением порядка модели  $m$  устойчивость коэффициентов  $a_m$  при неизменном объеме выборки  $n$  существенно уменьшается.

Проверка моделей 1—3-го порядков для шурфа 1 и скв. 6 показала, что наименьшие ошибки прогноза получаются при использовании модели 1-го порядка для шурфа 1 и 2-го порядка — для скв. 6. Соответствующие прогнозные уравнения имели следующий вид:

$$\bar{x}_t = 6,33 + 0,85(x_{t-1} - 6,33) \pm 0,47 \text{ (для шурфа 1);}$$

$$\bar{x}_t = 4,86 + 1,08(x_{t-1} - 4,86) - 0,37_1(x_{t-2} - 4,86) \pm 0,33 \text{ (для скв. 6).}$$

Результаты проверки этих уравнений на независимом фактическом материале представлены в табл. 14.

Таблица 14 — проверка прогнозов среднегодовых уровней грунтовых вод (в метрах от поверхности земли)

Номер шурфа и скважины	Прогнозный уровень	Фактический уровень	Ошибка прогноза $\Delta$ , м	Номер шурфа и скважины	Прогнозный уровень	Фактический уровень	Ошибка прогноза $\Delta$ , м
Заблаговременность 1 год							
1968 г.				1970 г.			
1	5,74 ± 0,47	5,25	0,49	1	5,75 ± 0,47	5,47	0,28
6	4,69 ± 0,33	4,67	0,02	6	4,96 ± 0,33	4,45	0,51
1969 г.				1971 г.			
1	5,42 ± 0,47	5,65	0,23	1	5,60 ± 0,47	4,95	0,65
6	4,82 ± 0,33	4,89	0,07	6	4,41 ± 0,33	4,80	0,39
Заблаговременность 2 года							
1969 г.				1971 г.			
1	5,83 ± 0,47	5,65	0,18	1	5,84 ± 0,47	4,95	0,89
6	4,84 ± 0,33	4,89	0,05	6	4,96 ± 0,33	4,80	0,16
1970 г.							
1	5,55 ± 0,47	5,47	0,08				
6	4,89 ± 0,33	4,45	0,44				

Примечание. Вероятная ошибка  $\Delta_{\text{вер}}$  для шурфа 1 составляет 0,58 м, для скв. 6—0,38 м.

В табл. 14 в качестве пределов изменения прогнозируемых уровней указана стандартная ошибка остатка  $\sigma_{z1}$ . При составлении прогнозов с заблаговременностью 1 год использованы только предшествующие фактические уровни. При большей заблаговременности наряду с фактическими используются и соответствующие прогнозные данные.

Таким образом, проверка на независимом фактическом материале показала, что ошибки прогнозов были, как правило, меньше вероятных ошибок, а найденные зависимости имели существенную прогнозную ценность.

Эффективность и точность прогнозных зависимостей внутри использованных периодов наблюдений может быть оценена величиной отношения  $S/\sigma$ , где  $\sigma$  — стандартное отклонение исходного ряда уровней. Для шурфа 1 это отношение оказалось равным 0,55, для скв. 6 — 0,58, что позволяет считать методику прогноза удовлетворительной.

В заключение следует отметить, что подобная методика применима для прогнозов уровней грунтовых вод в случае их сравнительно глубокого (не менее 3—4 м от поверхности) залегания на междуречных пространствах, сложенных песками, супесями и легкими суглинками. При этом многолетние колебания среднегодовых или любых фазовооднородных уровней в результате регулирующего влияния зоны аэрации становятся в той или иной степени связанными, а значения корреляционных функций  $r_{(\tau)}$  этих рядов при  $\tau = 1 \div 3$  достигают сравнительно больших величин. При неглубоком залегании уровней значения  $r_{(\tau)}$  существенно уменьшаются ( $r_{(\tau)} = 0,4 \div 0,2$  при  $\tau = 1 \div 3$ ), поэтому применение в этих условиях для целей прогноза авторегрессионных моделей, аналогичных описанному, становится, как правило, малоэффективным.

## 4

### Применение методов периодограммного анализа для прогнозов естественного режима уровня грунтовых вод

#### 4.1 — интегральные кривые и сглаживание

Несмотря на то что эти методы, вообще говоря, не являются методами периодограммного анализа, именно они наиболее часто используются для выявления периодов и циклов в рядах наблюдений.

Если обозначить члены исходного ряда через  $x_1, x_2, x_t$ , то его среднее значение будет

$$\bar{x} = \frac{1}{n} \sum_1^t x_t, \quad (90)$$

где  $n$  — число членов ряда, а отклонения каждого члена от среднего

$$d_t = x_t - \bar{x}. \quad (91)$$

Интегральным называется ряд, члены которого равны

$$x_1; x_1 + x_2; x_1 + x_2 + \dots + x_t, \text{ т. е. } \sum_1^t x_t, \quad (92)$$

а разностно-интегральным — члены которого равны

$$d_1; d_1 + d_2; d_1 + d_2 + \dots + d_t, \text{ т. е. } \sum_1^t d_t. \quad (93)$$

Разностно-интегральный ряд может быть представлен также в виде

$$\sum_1^t (k_t - 1), \quad (94)$$

где  $k_t = \frac{x_t}{\bar{x}}$  — модульный коэффициент для данного значения  $x_t$ .

Если исходный ряд содержит закономерные составляющие, осложненные случайной компонентой, то в преобразованном

разностно-интегральном ряду закономерные составляющие проявляются более четко, а действие случайных факторов сглаживается. На рис. 20 представлены разностно-интегральные кривые среднегодовых уровней грунтовых вод по скв. 458 (г. Горький) и 9-9 (Ленинградская область). На первой разностно-интегральной кривой отчетливо прослеживаются два цикла колебаний продолжительностью 11 и 12 лет, а также длительной многолетней подъем уровней, который, по-видимому, связан с влиянием

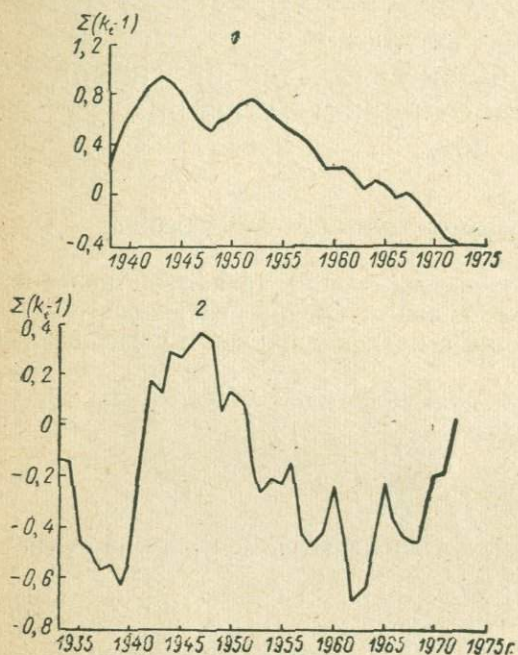


Рис. 20. Разностно-интегральные кривые среднегодовых уровней грунтовых вод:

1 — скв. 458 (г. Горький), 2 — скв. 9-9 (Ленинградская обл., ряд частично реконструирован).

искусственных факторов. На втором графике выделяется цикл колебаний 23 года, осложненный менее продолжительными 4- и 5-летними циклами.

Следует отметить, что при построении разностно-интегральных рядов начальные фазы и амплитуды циклических (или периодических) составляющих, содержащихся в анализируемом процессе, испытывают искажения, при этом амплитуда длиннопериодных составляющих увеличивается больше, чем короткопериодных. Кроме того, если члены исходного ряда не связаны друг с другом, то коэффициенты корреляции между соседними членами соответствующего разностно-интегрального ряда довольно значительны, особенно в середине ряда, в результате чего интегральная кривая становится плавной и кажется закономерно изменяющейся.

При анализе цикличности с помощью скользящего осреднения исходный ряд  $x_1, x_2, x_t$  преобразуется в ряд

$$\frac{1}{n} \sum_1^n x_i; \quad \frac{1}{n} \sum_2^{n+1} x_i; \quad \frac{1}{n} \sum_{z+1-n}^z x_i, \quad (95)$$

где  $z$  — число членов ряда;  $n$  — интервал осреднения, при этом  $z > n$ .

Если в исходном ряду  $x_1, x_2, x_t$  содержится периодическая составляющая с периодом  $T$ , то в преобразованном ряду ее амплитуда будет уменьшаться по сравнению с исходным рядом в отношении

$$C = \frac{\sin n \frac{z}{T}}{n \sin \frac{z}{T}}, \quad (96)$$

где  $T$  — длина периода;  $z$  — число членов ряда.

При  $T \leq n$  соответствующие периодические составляющие исходного ряда полностью гасятся. Если  $T > n$ , то чем больше  $T$ , тем менее убывает амплитуда периодической составляющей в сглаженном ряду, так как при малых углах

$$\sin n \frac{z}{T} \approx n \sin \frac{z}{T} \\ \text{и } C \approx 1. \quad (97)$$

Таким образом, при скользящем осреднении гасятся высокочастотные периодические составляющие и более отчетливо выделяются низкочастотные, при этом если  $T \gg n$ , то амплитуды и начальные фазы последних мало искажаются.

На рис. 21 представлены результаты последовательного скользящего осреднения частично реконструированного ряда среднегодовых уровней грунтовых вод в скв. 170 (г. Ленинград). После сглаживания по скользящим 5- и 11-леткам прослеживается один цикл продолжительностью не менее 26 лет.

Следует отметить, что при применении скользящего осреднения требуется известная осторожность, так как оно увеличивает

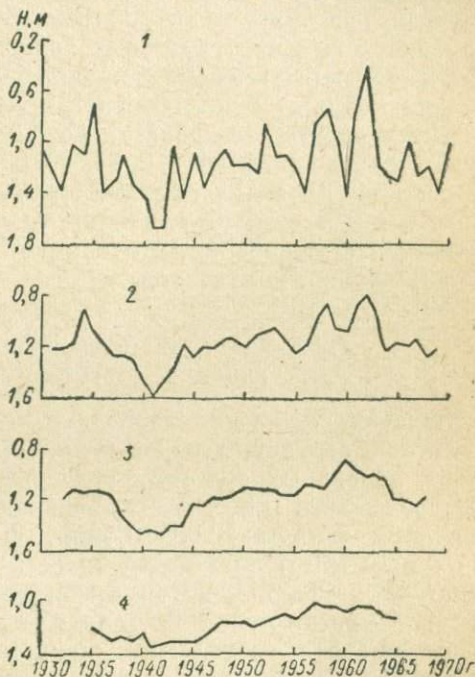


Рис. 21. Результаты сглаживания частично реконструированного ряда среднегодовых уровней грунтовых вод в скв. 170 (г. Ленинград) по скользящим  $n$ -леткам.

1 — исходный ряд; 2 —  $n = 3$ ; 3 —  $n = 5$ ; 4 —  $n = 11$  годам.

связность членов вновь полученного ряда в тем большей степени, чем больше интервал осреднения  $n$ . В результате этого возникают кажущиеся, или ложные, периодические составляющие, не содержащиеся в исходном ряду [63, 65].

Таким образом, интегральные кривые и скользящее осреднение применимы (с известными оговорками) для выявления скрытых цикличностей и периодичностей, на основе которого в ряде случаев оказывается возможным общий качественный прогноз колебаний уровней грунтовых вод. Так как оба эти метода являются в значительной степени визуальными, то применение их для количественных прогнозов весьма затруднительно.

Цикличность колебаний уровней грунтовых вод исследовалась рядом авторов [5, 16, 17, 19, 23, 39, 66, 67 и др.]. Оценка выявленных циклов проводилась этими авторами либо визуальными, либо близкими к ним методами, в том числе с помощью интегральных кривых и сглаживания, что в значительной степени затрудняло использование полученных данных для количественных прогнозов.

#### 4.2 — методика выявления периодических составляющих в колебаниях уровня

Задачей собственно периодограммного анализа является изучение частотной структуры рядов, т. е. нахождение периодов, амплитуд и начальных фаз гармоник, выделенных в исходном ряду. Классическими методами выявления скрытых периодичностей являются методы Бью — Балло, Лагранжа — Дэла, Шустера и их многочисленные модификации. Не останавливаясь на описании этих методов, которые можно найти в специальной литературе, отметим лишь, что в ряде случаев их применение затруднено тем, что исследуемый процесс оказывается существенно осложненной случайной компонентой. С помощью этих методов можно выделить лишь периоды, продолжительность которых значительно меньше длины исходного ряда, выделение же длиннопериодных составляющих практически невозможно.

В области изучения режима подземных вод периодограммный анализ, основанный на применении упрощенного метода Бью — Балло, впервые использовал М. Е. Альтовский [49]. Впоследствии этот метод применялся различными авторами. С его помощью можно найти продолжительность преобладающих в колебаниях уровней периодов, но невозможно дать их полное аналитическое выражение, поэтому полученные данные могут использоваться лишь для качественных прогнозов.

Таким образом, методы периодограммного анализа до настоящего времени практически не использовались для количественных прогнозов уровней подземных вод. Однако эта группа методов имеет существенные преимущества по сравнению с другими, а именно 1) значительная (от нескольких месяцев до нескольких лет) заблаговременность прогнозов, которую трудно получить

при использовании каких-либо других методов, 2) исходным материалом для составления прогнозов служат только данные режимных гидрогеологических наблюдений; метеорологические данные непосредственно для получения прогнозов не используются.

В настоящей работе сделана, по существу, первая попытка широкого применения периодограмм-анализа для сверхдолгосрочных (с заблаговременностью не менее 1 года) прогнозов уровней грунтовых вод.

Ниже приводятся определения ряда терминов, употребляемых в дальнейшем изложении.

1. Ритмические (циклические) колебания — колебания уровней (или других элементов режима), при которых последовательные максимумы или минимумы встречаются через неравные или приблизительно равные интервалы времени. Промежуток времени между двумя соседними максимумами (или минимумами) составляет ритм (цикл).

2. Периодические (гармонические) колебания — колебания уровней, в которых временные интервалы между соседними максимумами или минимумами и амплитуды колебаний постоянны для всего ряда наблюдений. Длина периода (периодической составляющей, гармоники)  $T$  — промежуток времени между двумя соседними равноотстоящими максимумами или минимумами.

3. Частота колебаний — величина, обратная периоду, т. е.  $\omega = 2\pi/T$ .

4. Частотная структура — совокупность периодических составляющих, выделенных в колебаниях анализируемого уровня.

Для выявления периодических составляющих в колебаниях уровней была использована стохастическая модель гидрогеологического процесса, которая может быть представлена суммой нескольких периодических составляющих, осложненных случайной компонентой:

$$x_t = \bar{x} + \sum_1^j A_j \cos\left(\frac{2\pi}{T_j} t - \varphi_j\right) + z_t, \quad (98)$$

где  $x_t$  — анализируемый ряд уровней (среднегодовых; средне-месячных и т. д.);  $\bar{x}$  — среднее значение уровня за период наблюдений;  $A_j$ ,  $T_j$ ,  $\varphi_j$  — амплитуда, период и начальная фаза каждой выделенной периодической составляющей;  $j$  — число периодических составляющих;  $z_t$  — случайная компонента.

Для реализации подобной модели Л. Д. Кнорингом и В. Н. Дечем был разработан алгоритм для БЭСМ-2М, который построен следующим образом. На вход ЭВМ поступает информация в виде равноотстоящих данных гидрогеологических наблюдений (например, суточные, среднемесячные, среднегодовые уровни). Для этого ряда вычисляется нормированная корреляционная функция по формуле (80).

Анализ этой функции, как известно, позволяет рассмотреть и принять одну из альтернативных гипотез.

1. Исследуемый процесс является случайным (содержит только компоненту  $z_t$ ), т. е. он определяется действием разнообразных факторов, не поддающихся в настоящее время аналитическому учету, и подчиняется нормальному или близкому к нему закону распределения. Для такого процесса характерна быстро убывающая нормированная корреляционная функция (см. рис. 15, 2).

2. Исследуемый процесс помимо случайной компоненты содержит одну или несколько периодических составляющих. Нормированная корреляционная функция такого ряда убывает значительно медленнее, чем в предыдущем случае, и обязательно содержит гармонические компоненты с тем или иным временем корреляции (см. рис. 15, 1, 3, 4). Иными словами, в первом случае исходный ряд наблюдений не содержит закономерных составляющих, т. е. определяется случайной компонентой  $z_t$ , во втором случае в исходном ряду имеются закономерные составляющие, которые можно прогнозировать за пределы использованного периода наблюдений.

Выполнение условия второй гипотезы позволяет подвергнуть анализируемый ряд преобразованию с целью определения всех  $3j + 1$  параметров, входящих в формулу (98).

Если предположить, что среднее значение случайной компоненты  $z_t$  равно 0, т. е.  $\sum_{t=0}^T \frac{z_t}{n} = 0$ , ее дисперсия равна  $\sigma_{z_t}^2$  и при каждом фиксированном значении времени  $t$  отклонения  $z_t$  подчинены одному и тому же нормальному закону распределения, то оценка параметров  $\bar{x}$  и  $T_j$  формулы (98) находится из условия существования минимума величины  $u$ :

$$u = \sum_{t=0}^T \left[ x_t - \bar{x} - \sum_1^j A_j \cos \left( \frac{2\pi}{T_j} t - \varphi_j \right) \right]^2 = \min, \quad (99)$$

где все обозначения приведены выше.

Величина  $u$ , как известно, является положительно определенной квадратичной формой переменных параметров  $\bar{x}$  и  $T_j$ . Это значит, что если последовательно производить перебор этих величин в выбранных диапазонах, то можно найти минимум величины  $u$ , который будет соответствовать искомому  $\bar{x}$  и  $T_j$ . Таким образом, задаваясь рядом значений  $\bar{x}$  и  $T_j$ , можно найти соответствующие им величины  $u$ , при этом два других неизвестных —  $A_j$  и  $\varphi_j$  — находятся методом наименьших квадратов. Так, используя известные тригонометрические соотношения, можно записать

$$A_j \cos \left( \frac{2\pi}{T_j} t - \varphi_j \right) = d_j \cos \frac{2\pi}{T_j} t + c_j \sin \frac{2\pi}{T_j} t, \quad (100)$$

$$\text{где } A_j = \sqrt{c_j^2 + d_j^2} \text{ и } \operatorname{tg} \varphi_j = \frac{c_j}{d_j}. \quad (101)$$

Подставляя (100) в (99) и дифференцируя по  $c_j$  и  $d_j$ , получаем

$$-\frac{1}{2} \frac{du}{dd_j} = \sum_{t=0}^T \left( x_t^* - d_j \cos \frac{2\pi}{T_j} t \right) \cos \frac{2\pi}{T_j} t;$$

$$-\frac{1}{2} \frac{du}{dc_j} = \sum_{t=0}^T \left( x_t^* - c_j \sin \frac{2\pi}{T_j} t \right) \sin \frac{2\pi}{T_j} t,$$
(102)

где  $x_t^* = x_t - \bar{x}$ .

Приравнявая левые части уравнений (102) нулю, можно найти  $c_j$  и  $d_j$ , а следовательно,  $A_j$  и  $\varphi_j$ . Таким образом, из ряда значений  $u$  выбирается его минимальное значение  $u_1 = \min$ , которому соответствуют искомые  $\bar{x}$  и  $T_1$ . Период  $T_1$  является одним из периодов, содержащихся в ряду  $x_t$ , а соответствующая ему периодическая составляющая имеет наибольшую амплитуду. При известных параметрах  $\bar{x}$  и  $T_1$  величины  $A_1$  и  $\varphi_1$  находятся по приведенным выше формулам (101) и (102). Определив все параметры первой периодической составляющей, ЭВМ выводит их на печать, формирует по ним саму периодическую составляющую и исключает (вычитает) ее из исходного ряда.

На печать выводятся в дискретных точках  $t$  найденная периодическая составляющая и остаток. Последний анализируется по описанному выше алгоритму, за исключением поиска  $\bar{x}$ -члена, что позволяет определить параметры следующей гармонической составляющей.

Эта процедура повторяется до тех пор, пока из исходного ряда  $x_t$  не будут исключены все содержащиеся в нем периодические составляющие, о чем свидетельствует определенный вид корреляционной функции (см. рис. 15, 2), вычисляемой каждый раз по соответствующему остатку. После этого на печать

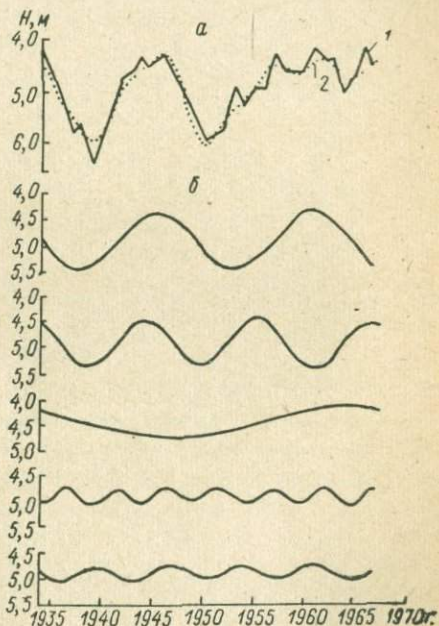


Рис. 22. Результаты выявления периодических составляющих в колебаниях среднегодовых уровней грунтовых вод по скв. 6 (Московская обл.).

$a$  — среднегодовые уровни (в метрах от поверхности земли) по данным фактических замеров (1) и вычисленные (2);  $b$  — периодические составляющие, выявленные в колебаниях уровней, продолжительностью 15, 11, 35, 5, 7 лет.

выводится сумма всех найденных периодических составляющих в точках  $t$  (рис. 22).

Следует отметить, что при фиксированном значении  $\bar{x}$  можно осуществить построение периодограммы в координатах  $u$  и  $T_j$ . Анализ периодограммы позволяет судить о частотной структуре исследуемого гидрогеологического процесса  $x_t$ , т. е. определить число содержащихся в нем периодических составляющих и оценить их амплитуды. На присутствие какой-либо периодической составляющей указывает минимум величины  $u$  при соответствующем  $T_j$ , при этом чем меньше величина  $u$ , тем больше амплитуда  $A_j$  периодической составляющей. Так, периодограмма на рис. 23

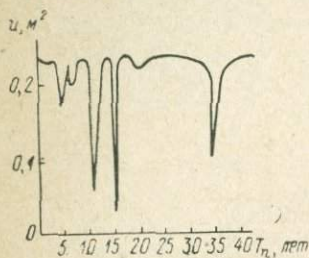


Рис. 23. Периодограмма ряда среднегодовых уровней грунтовых вод по скв. 6 (Московская обл.).

показывает, что в анализируемом ряде содержатся периодические составляющие с периодами (в порядке убывания амплитуд) 15, 11, 35, 5, 7 лет.

Для того чтобы решить вопрос о том, сколько гармоник следует выделять в исходном ряде уровней, нужно оценить в заданных доверительных интервалах все «пики» корреляционной функции, которая вычисляется после последовательного исключения из исходных данных каждой новой периодической составляющей.

Стандартные отклонения  $\sigma_{r(\tau)}$ , необходимые для вычисления доверительных интервалов значений корреляционной функции  $r(\tau)$ , определяются по формулам (26), (28), (29) либо, когда число членов ряда  $n$  достаточно велико, из выражения

$$\sigma_{r(\tau)} = \frac{1}{\sqrt{n}}. \quad (103)$$

Выделение периодических составляющих следует прекратить после того, как число отличающихся от 0 значений  $r(\tau)$  будет меньше количества этих значений, возможных при принятых доверительных интервалах. Например, при  $n = 100$   $\sigma_{r(\tau)} = 1/\sqrt{100} = 0,10$ . 95%-ный доверительный интервал для  $r(\tau)$  приблизительно равен  $r(\tau) \pm 0,10 \cdot 1,96 = r(\tau) \pm 0,20$ . Предположим, что корреляционная функция  $r(\tau)$  вычислялась для значений  $\tau$  от 0 до 25. Если после выделения очередной гармоники и вычисления корреляционной функции соответствующего остатка оказалось, что четыре значения  $r(\tau)$  превышают  $\pm 0,20$ , дальнейшее выявление периодических составляющих следует продолжить, так как возможное число таких значений для случайного временного ряда («белого шума») составляет  $25 \cdot 0,05 \approx 1$ .

Если после выделения очередной периодической составляющей лишь одно значение корреляционной функции будет превышать  $\pm 0,20$  или таких значений не будет вовсе, дальнейшее выявление гармоник следует прекратить. Проведенные расчеты такого рода

показали, что колебания уровней грунтовых вод описываются, как правило, первыми 3—4 гармониками, которые отличаются от «белого шума».

Любая модель или схема гидрогеологического процесса лишь в большей или меньшей степени учитывает все многообразие и взаимосвязь многочисленных природных явлений, определяющих этот процесс, в связи с чем каждая из этих схем имеет свои достоинства и недостатки. Достоинством предлагаемой модели колебаний уровней (или любых других элементов режима) подземных вод является то, что с ее помощью в исходном ряде можно дифференцированно выделить несколько неслучайных и независимых друг от друга периодических составляющих, каждая из которых получает вполне определенное аналитическое выражение, и случайную компоненту. Таким образом, выделенные периодические составляющие можно прогнозировать за пределы использованного интервала наблюдений.

Недостатком предлагаемой модели является ее «жесткость», т. е. постоянство амплитуд и продолжительностей выделяемых гармоник. Несмотря на этот недостаток, данную методику можно применять для различных гидрогеологических прогнозов, так как колебательный процесс (например, колебания уровней), независимо от причин, его вызывающих, обладает известной инерционностью, т. е. периоды, выявленные в прошлом, должны с той или иной степенью вероятности проявиться в ближайшем будущем. Например, если в течение 15—20-летнего периода наблюдений в колебаниях уровней были отмечены периоды 4 или 5 лет, то со значительной долей вероятности можно утверждать, что они так же будут проявляться в ближайшие 2—3 года.

Все многообразные закономерности, отмечаемые в системе Солнце — тропосфера — атмосфера — подземные воды, являются не строго периодическими, а квазипериодическими, т. е. для них характерны переменные во времени амплитуды и продолжительности циклов. Создание алгоритма, учитывающего это свойство природных явлений, связано с рядом значительных трудностей, главной из которых является отсутствие достаточно продолжительных рядов гидрогеологических наблюдений, по которым можно выявить закономерности изменений во времени продолжительности циклов и их амплитуд.

#### 4.3 — о соотношении гармонических составляющих и случайной компоненты в колебаниях уровней грунтовых вод

Исследуя периодичность колебаний уровней грунтовых вод, необходимо выяснить вопрос о соотношении периодических составляющих и случайной компоненты, иными словами, требуется найти ту долю в общих колебаниях уровня, которую можно строго детерминировать с помощью предлагаемой методики.

Для того чтобы исключить индивидуальные особенности колебаний уровня в каждой скважине и сделать их сопоставимыми, необходимо нормировать амплитуды выделенных гармоник, т. е. найти отношение

$$A_{j \text{ прив}} = \frac{A_j}{\sigma_{x_t}}, \quad (104)$$

где  $A_{j \text{ прив}}$  — приведенная амплитуда;  $A_j$  — амплитуда;  $\sigma_{x_t}$  — стандартное отклонение анализируемого уровня.

В этом случае, приведенную ранее формулу (98), описывающую изменения уровней на основе выявлений в их колебаниях периодических составляющих, можно представить в виде

$$x_t = \bar{x} + \bar{x}C_v \sum_1^j A_{j \text{ прив}} \cos \left( \frac{2\pi}{T_j} t - \varphi_j \right), \quad (105)$$

где  $A_{j \text{ прив}}$  — приведенная амплитуда каждой выделенной периодической составляющей;  $C_v = \frac{\sigma_{x_t}}{\bar{x}}$  — коэффициент вариации анализируемого уровня; остальные обозначения приведены выше.

Так как  $C_v = \frac{\sigma_{x_t}}{\bar{x}}$ , то  $\sigma_{x_t} = C_v \bar{x}$ . Из выражения (104) следует, что  $A_j = A_{j \text{ прив}} \sigma_{x_t}$ , откуда получаем  $A_j = C_v \bar{x} A_{j \text{ прив}}$ .

Подстановка выражения  $A_j$  в формулу (98) дает приведенную выше зависимость (105).

Величина, стоящая в выражении (105) под знаком суммы, достигнет своего максимального значения в том случае, когда экстремальные значения косинусов ( $\pm 1$ ) совпадут для всех выделенных гармоник. В этом случае выражение (105) примет вид

$$x_t = \bar{x} + \bar{x}C_v \sum_1^j A_{j \text{ прив}} = \bar{x} + l\sigma_{x_t}, \quad (106)$$

где  $l = \sum_1^j A_{j \text{ прив}}$ .

Если принять, что анализируемые уровни грунтовых вод подчиняются нормальному или близкому к нему закону распределения, то их колебания вокруг среднего значения составляют  $\pm 3\sigma$ . Таким образом, второе слагаемое в выражении (106) характеризует максимально возможную долю гармонических колебаний в общих колебаниях уровня. С применением этой зависимости были определены доли периодических составляющих, описываемых первыми 3—5 гармониками в колебаниях среднегодовых уровней грунтовых вод. Для рядов натуральных наблюдений на Северо-Западе РСФСР продолжительностью 15—19 лет при глубине залегания уровня 1,0—2,5 м от поверхности земли доля периодических колебаний в общих колебаниях среднегодовых уровней составила  $(1,8 \div 2,8)\sigma$ . При анализе частично

реконструированных\* рядов среднегодовых уровней грунтовых вод продолжительностью 23—26 лет доля периодических составляющих в общих колебаниях среднегодовых уровней была равна  $(1,8 \div 2,4)\sigma$ .

Аналогичные расчеты, проведенные для отдельных скважин в центральных областях РСФСР с продолжительными периодами натуральных наблюдений (скв. 6 в Московской области — 35 лет, шурф 1 в Воронежской области — 52 года), показали, что доля периодических составляющих в общих колебаниях среднегодовых уровней оставалась весьма высокой и составила для них соответственно 2,6 $\sigma$  и 2,7 $\sigma$ . Таким образом, в колебаниях среднегодовых уровней грунтовых вод на Северо-Западе и в центральных районах РСФСР, т. е. в зонах избыточного и переменного увлажнения, максимальная доля периодических составляющих, определяемая вкладом первых 3—5 выделенных в колебаниях гармоник, оказывается весьма высокой.

Близкие результаты были получены при анализе некоторых экстремальных, в частности, летне-осенних и предвесенних минимальных среднемесячных уровней грунтовых вод на Северо-Западе РСФСР. Доля периодических составляющих при анализе натуральных наблюдений продолжительностью 12—19 лет составила здесь  $(2,0 \div 2,5)\sigma$ . По приведенной выше методике определяется максимально возможная доля периодических составляющих в общих колебаниях уровней.

Для получения не максимальных, а усредненных величин, характеризующих вклад гармоник в колебания уровня, было использовано отношение стандартного отклонения случайной компоненты (напомним, что она получается в результате вычитания выделенных периодических составляющих из исходного ряда уровней) к стандартному отклонению исходного ряда, т. е.  $\lambda = \sigma_{z_t} / \sigma_{x_t}$ . В случае, если это отношение равно 1, в исследуемом ряду уровней не содержится периодических составляющих. Если это отношение равно 0, то исходный ряд полностью раскладывается на гармонические составляющие.

Таким образом, чем меньше значение  $\lambda$ , тем большую долю в колебаниях составляют периодические составляющие, которые можно прогнозировать за пределы использованного интервала наблюдений. Величина  $1 - \lambda$  является численным выражением доли периодических колебаний в общих колебаниях уровня. Нетрудно убедиться, что коэффициент  $\lambda$  в известном смысле аналогичен обычно применяемому критерию оценки точности и эффективности прогнозов, который выражается отношением  $S/\sigma$ , поскольку величина стандартного отклонения случайной компоненты  $\sigma_{z_t}$ , по существу, и является стандартной ошибкой проверочных прогнозов  $S$ . Таким образом, чем меньше величина  $\lambda$ ,

\* Реконструкция осуществлена по уравнениям регрессии, связывающим колебания уровней грунтовых вод в скважинах и уровней воды в реках. Глубина залегания уровней грунтовых вод не превышала 1,0—2,7 м.

Таблица 15 — соотношение периодических составляющих и случайной компоненты в колебаниях уровней грунтовых вод

Номера скважин	Местоположение	Использованный период наблюдений, годы	Номера вариантов	$\sigma_{z_t}$ , м	$\sigma_{x_t}$ , м	$\lambda = \frac{\sigma_{z_t}}{\sigma_{x_t}}$
11	г. Ленинград	1949—1970	1	0,32	0,46	0,70
			2	0,23	0,35	0,65
			3	0,05	0,21	0,24
			5	0,15	0,32	0,47
			6	0,13	0,30	0,43
153	То же	1957—1970	3	0,15	0,51	0,30
			5	0,42	0,81	0,52
			6	0,16	0,69	0,23
170	»	1957—1970	3	0,08	0,25	0,32
		1933—1970	4	0,16	0,25	0,64
85	Ленинградская область	1946—1970	4	0,08	0,17	0,47
860	То же	1946—1970	4	0,14	0,30	0,47
160	»	1953—1970	2	0,27	0,35	0,80
		1953—1970	3	0,06	0,19	0,31
		1946—1970	4	0,08	0,19	0,42
		1953—1970	5	0,10	0,19	0,53
		1953—1970	6	0,08	0,22	0,36
3—3	Ленинградская область Ордовикское плато	1950—1970	3	0,44	1,17	0,37
5—5	То же	1950—1970	3	0,30	0,64	0,47
42—1	»	1954—1970	3	0,45	1,16	0,39
6	Московская область	1934—1967	3	0,17	0,57	0,30
Шурф 1	Воронежская область	1915—1967	3	0,37	0,86	0,43

тем более эффективной оказывается методика прогноза, основанная на выявлении периодических составляющих в колебаниях урорня. По ряду наблюдательных скважин для вычисления  $\lambda$  использовались различные варианты осреднения исходных данных, а именно: среднемесячные уровни (вариант 1), среднесезонные\* уровни (вариант 2), среднегодовые уровни (вариант 3). Помимо этого использовались реконструированные ряды среднегодовых уровней грунтовых вод (вариант 4), а также ряды минимальных среднемесячных предвесенних (вариант 5) и летне-осенних (вариант 6) уровней. Данные о коэффициенте  $\lambda$  для различных вариантов приведены в табл. 15.

Анализ данных табл. 15 показывает, что с увеличением масштаба осреднения исходных данных, т. е. с переходом от средне-

\* Для территории Северо-Запада РСФСР под среднесезонными подразумеваются средние уровни за периоды XII-III, IV-V, VI-VIII, IX-XI месяцы.

месячных к среднесезонным и среднегодовым уровням величина  $\lambda$  уменьшается и, следовательно, в колебаниях уровней увеличивается доля периодических составляющих и уменьшается роль случайной компоненты. Коэффициент  $\lambda$  уменьшается с увеличением глубины залегания уровня (рис. 24). На график нанесены значения  $\lambda$ , полученные для некоторых рядов среднегодовых уровней грунтовых вод в четвертичных отложениях, при этом уровни по скв. 85, 160, 170, 860 частично реконструированы, так что продолжительность наблюдений составила 23—36 лет, а по скв. 6 по данным фактических замеров — 35 лет. Таким образом, были использованы сопоставимые (в первом приближении) по продолжительности ряды среднегодовых уровней, при этом уменьшение величины  $\lambda$  связано главным образом с увеличением сглаживающего влияния зоны аэрации на колебания уровней, которое наблюдается с возрастанием ее мощности. С другой стороны, с увеличением использованного периода наблюдений значение  $\lambda$  увеличивается (табл. 16, скв. 160, 170, варианты 3, 4). Это связано с тем, что, как отмечалось ранее, в действительности колебания уровней являются не строго периодическими, а циклическими, и с увеличением периода наблюдений неодинаковая продолжительность циклов сказывается все в большей степени.

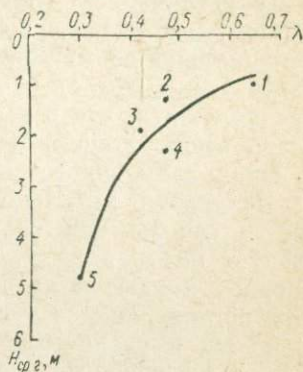


Рис. 24. График зависимости коэффициента  $\lambda$  от среднегодовой глубины залегания среднегодового уровня.

1 — скв. 170 (г. Ленинград);  
2 — скв. 860; 3 — скв. 160, 4 — скв. 85 (Ленинградская обл.);  
5 — скв. 6 (Московская обл.).

В настоящее время трудно решить вопрос о соотношении длины исходного ряда и заблаговременности прогноза. Очевидно, что если с увеличением длины анализируемого ряда уровней величина коэффициента  $\lambda$  значительно возрастает, то использование для прогнозов со сравнительно небольшой заблаговременностью длинного ряда не даст желаемых результатов. С другой стороны, если с возрастанием длины исходного ряда значения  $\lambda$  остаются практически неизменными, что свидетельствует об устойчивости выделенных периодических (циклических) составляющих во времени, то для прогнозов следует использовать длинные ряды уровней. Поскольку по подавляющему большинству наблюдательных точек продолжительность наблюдений не превышает 20—35 лет, в настоящее время вопрос о выборе оптимальной длины исходного ряда является в значительной мере преждевременным и для составления прогнозов следует использовать весь накопленный фактический материал.

Проведенный анализ значений коэффициента  $\lambda$ , характеризующего эффективность и точность методики прогноза, позволяет

сделать ряд методических выводов. Предлагаемую методику целесообразно применять для прогнозов уровней, осредненных за сравнительно продолжительный период времени: месяц, сезон, год. Минимальное значение коэффициента  $\lambda$  и соответственно максимальная эффективность и точность методики отмечены для рядов среднегодовых уровней грунтовых вод (см. табл. 16, вариант 3). Наряду с этим данную методику целесообразно использовать и для прогнозов экстремальных, в частности летне-осенних и предвесенних минимальных среднемесячных уровней.

С увеличением глубины залегания уровней эффективность данной методики существенно возрастает, поэтому наиболее рационально ее применять при глубине залегания уровней не менее 1,0—1,5 м от поверхности земли. В связи с этим можно предположить, что в зоне недостаточного увлажнения эффективность этой методики будет, по-видимому, более высокой, чем в зонах избыточного и переменного увлажнения.

Используя выделенные периодические составляющие в качестве основы для составления сверхдолгосрочных прогнозов, следует иметь в виду, что, как указывалось выше, колебания уровней являются не строго периодическими, а циклическими, что подтверждается уменьшением доли периодических составляющих с увеличением использованного периода наблюдений. Поэтому в общем случае целесообразно давать прогнозы с заблаговременностью не более 1—3 шагов  $\Delta t$ . При увеличении заблаговременности точность прогнозов может существенным образом понизиться в результате проявления цикличности колебаний уровней. По мере накопления фактических данных необходимо вновь проводить частотный анализ рядов уровней для уточнения характеристик выделяемых в них гармоник. Эти уточненные характеристики следует в дальнейшем использовать для составления новых прогнозов с указанной выше заблаговременностью.

#### 4.4 — составление прогнозов уровней и их проверка

Как указывалось в разделе 4.2, колебания уровней грунтовых вод можно представить в виде

$$x_t = \bar{x} + \sum_1^j A_j \cos \left( \frac{2\pi}{T_j} t - \varphi_j \right) + z_t. \quad (107)$$

Описанная методика позволяет находить  $3j + 1$  неизвестных параметров, входящих в уравнение (107), а именно:  $\bar{x}$ ,  $A_j$ ,  $T_j$  и  $\varphi_j$ . Характеристикой случайной компоненты  $z_t$  является величина ее стандартного отклонения  $\sigma_{z_t}$ . Таким образом, для составления прогнозов режима уровня выражение (107) следует записать в виде

$$\tilde{x}_t = \bar{x} + \sum_1^j A_j \cos \left( \frac{2\pi}{T_j} t - \varphi_j \right) \pm K \sigma_{z_t}, \quad (108)$$

где  $\bar{x}_t$  — прогнозируемое значение уровня;  $\sigma_{zt}$  — стандартное отклонение случайной компоненты  $z_t$ ;  $k=1 \div 3$ ; остальные обозначения приведены выше.

Подставляя в выражение (108) соответствующие значения времени  $t$ , можно прогнозировать каждую периодическую составляющую и, следовательно, их сумму за пределы использованного периода наблюдений, при этом два первых слагаемых являются математическим ожиданием прогноза, а третье с определенной вероятностью указывает пределы его изменения. Так, на-

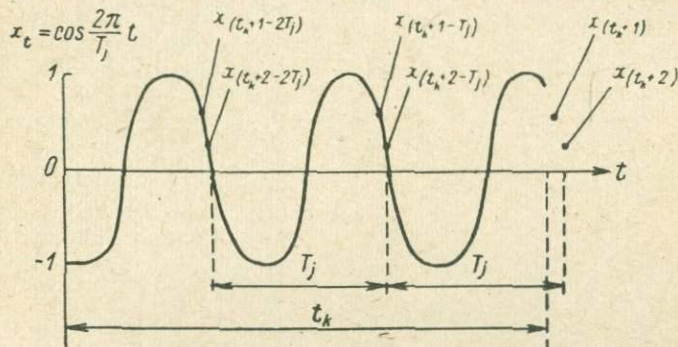


Рис. 25. Экстраполяция гармоники с периодом  $T_j$  и амплитудой  $A_j=1$  на два шага  $\Delta t$  ( $t_{k+1}$  и  $t_{k+2}$ ) за пределы использованного периода наблюдений  $t_k$ .

пример, расчетное уравнение для прогноза среднегодовых уровней по кв. 153 (г. Ленинград) имело вид

$$\bar{x}_t = 1,59 + 0,47 \cos \left( \frac{2\pi}{4} t - 2,89 \right) + 0,35 \cos \left( \frac{2\pi}{3} t - 4,49 \right) + 0,26 \cos \left( \frac{2\pi}{6} t - 0,31 \right) + 0,12 \cos \left( \frac{2\pi}{40} t - 2,32 \right) \pm 0,30. \quad (109)$$

Начальные фазы  $\varphi_j$  приведены к 1958 г. и даны в радианах. Подставляя в уравнение значения  $t$ , равные 14 и 15, что соответствует 1971 и 1972 гг., легко вычислить прогнозные значения среднегодовых уровней. В качестве пределов их изменения с вероятностью 95% указана величина  $\pm 2\sigma_{zt}$ . Поскольку ЭВМ выводит на печать значения каждой гармоники в точках  $t$ , то их экстраполяцию в большинстве случаев можно осуществить и без непосредственного вычисления косинусов, при этом прогнозные значения каждой гармоники с периодом  $T_j$  равны ее значениям внутри использованного периода наблюдений, кратным  $T_j$  (рис. 25), т. е.

$$\begin{aligned} x(t_{k+1}) &= \cos \frac{2\pi}{T_j} (t_k + 1) = \cos \left[ \frac{2\pi}{T_j} (t_k + 1) - mT_j \right]; \\ x(t_{k+2}) &= \cos \frac{2\pi}{T_j} (t_k + 2) = \cos \left[ \frac{2\pi}{T_j} (t_k + 2) - mT_j \right]. \end{aligned} \quad (110)$$

В качестве примера на рис. 26 показаны периодические составляющие, выявленные в колебаниях среднегодовых уровней грунтовых вод в скв. 153 (г. Ленинград), и составленный на этой основе прогноз.

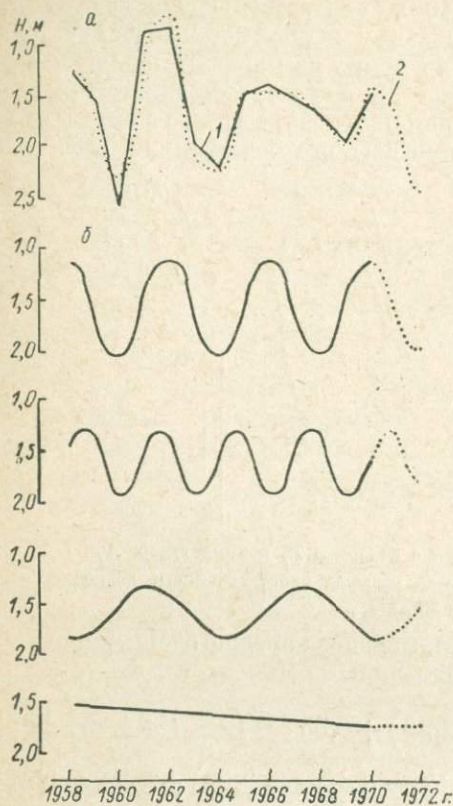


Рис. 26. Результаты выявления периодических составляющих в колебаниях среднегодовых уровней грунтовых вод по скв. 153 (г. Ленинград) и его прогноз.

а — среднегодовые уровни (в метрах от поверхности земли) по данным фактических замеров (1) и вычисленные (2); б — периодические составляющие, выявленные в колебаниях уровней, продолжительностью 4, 3, 6, 40 лет.

невозможным сравнить прогнозные и фактические значения уровней. Результаты подобной проверки для тех скважин, по которым наблюдения проводились без существенных перерывов, представлены в табл. 16.

Проверка прогнозов среднегодовых уровней на независимом материале показала, что в большинстве случаев ошибки прогнозов были меньше их вероятных значений. Данная методика может успешно применяться для прогнозов не только среднегодовых, но и экстремальных среднемесячных уровней, в частности, весьма

С помощью подобной методики, с заблаговременностью 1—2 года, были составлены прогнозы среднегодовых уровней грунтовых вод на 1968—1969 гг. для двух опорных точек (шурфа 1 и скв. 6), расположенных в центральных областях РСФСР, и с заблаговременностью 1 год (на 1969 г.) для двух опорных скважин (№ 11, 160) на территории Северо-Запада РСФСР. Проверка этих прогнозов показала, что лишь в одном случае из шести (для шурфа 1 в 1968 г.) ошибка прогноза превысила ее вероятное значение.

В дальнейшем по данным наблюдений, включающим 1969 г., был составлен прогноз о положении среднегодовых уровней грунтовых вод на Северо-Западе РСФСР в 1970 г., а по данным, включающим 1970 г. — в 1971 г. К сожалению, в 1970—1971 гг. по различным причинам наблюдения на некоторых опорных и представительных скважинах прерывались, поэтому для них оказалось невозможным

Таблица 16 — проверка прогнозов среднегодовых уровней грунтовых вод (в метрах от поверхности земли)

Номера скважин	Среднегодовые уровни		Ошибка прогноза $\Delta$ , м	Вероятная ошибка $\Delta_{\text{вер}}$ , м
	Прогнозный	Фактический		
1970 г.				
8а	0,81±0,40	0,80	0,01	0,11
77	1,27±0,15	1,43	0,16	0,11
84	0,93±0,10	0,96	0,03	0,08
153	1,35±0,30	1,49	0,14	0,34
170	0,90±0,20	1,06	0,16	0,18
1971 г.				
8а	1,00±0,10	1,01	0,01	0,11
77	1,11±0,15	1,21	0,10	0,11
84	0,84±0,10	0,86	0,02	0,08
153	1,76±0,30	1,72	0,04	0,34
160	2,06±0,20	2,00	0,06	0,13
170	1,02±0,20	1,06	0,04	0,18

Таблица 17 — проверка прогнозов летне-осенних минимальных среднемесячных уровней грунтовых вод (в метрах от поверхности земли)

Номера скважин	Летне-осенние минимальные среднемесячные уровни		Ошибка прогноза $\Delta$ , м	Вероятная ошибка $\Delta_{\text{вер}}$ , м
	Прогнозный	Фактический		
1970 г.				
11	2,05±0,25	2,18	0,13	0,20
153	2,23±0,30	2,44	0,21	0,46
160	2,13±0,15	2,25	0,12	0,18
1971 г.				
153	2,31±0,30	2,54	0,23	0,46
160	2,34±0,15	2,29	0,05	0,18

трудно прогнозируемых в зоне избыточного увлажнения летне-осенних минимальных уровней грунтовых вод. В табл. 17 представлены результаты проверки прогнозов этих уровней, составленных с заблаговременностью 1—2 года для некоторых опорных скважин на Северо-Западе РСФСР.

Отметим, что в настоящее время составить прогнозы экстремальных уровней с подобной заблаговременностью с помощью какого-либо иного метода практически крайне трудно. Таким образом, проведенная на независимом фактическом материале проверка показала, что методика прогноза уровня грунтовых вод, основанная на выявлении в его колебаниях периодических составляющих, является достаточно надежной и эффективной, что позволяет рекомендовать ее для дальнейшего использования в различных природных условиях.

Наряду с этим данная методика нуждается в дальнейшей проверке как в различных природных условиях, так и на независимом фактическом материале большого объема.

## О возможности распространения прогнозов по территории и их картированию

Несмотря на значительное расширение сети наблюдательных скважин, которое отмечается в последние годы, количество точек с периодами наблюдений более 15—20 лет все еще во многих районах СССР сравнительно невелико. В связи с этим с особой остротой встает вопрос о передаче или, образно говоря, «ретрансляции» прогнозов с точек с продолжительными периодами наблюдений на точки с непродолжительными наблюдениями, поскольку составить прогноз непосредственно по данным 2—3-летних наблюдений весьма затруднительно, а часто и вовсе невозможно. Следует отметить, что вопросы, связанные с передачей прогнозов с точки на точку и с распространением прогнозов по территории, практически никогда не потеряют своей актуальности, ибо скважины с продолжительными периодами наблюдений чаще всего расположены в том или ином удалении от интересующего нас пункта или района.

Главным условием, обеспечивающим подобную передачу прогнозов на расстояние, является максимальная сезонная и многолетняя синхронность хода уровней в сравниваемых скважинах. Основой для определения этой синхронности служит районирование территории по условиям формирования естественного режима уровня грунтовых вод. При исследовании режима уровня равнинных областей площадью от нескольких квадратных километров до нескольких десятков тысяч квадратных километров основное внимание следует обращать на весьма изменчивые в пространстве геологические факторы формирования режима, так как климатические факторы в пределах таких площадей изменяются сравнительно мало. Под геологическими факторами формирования режима подразумеваются: степень дренированности территории, геоморфологическое положение, литологический состав водовмещающих пород и пород зоны аэрации. В тех случаях, когда выделенные при районировании гидрогеологические участки и области характеризуются аналогичными геоло-

гических условиями формирования режима, наблюдается весьма высокая сезонная и многолетняя\* синхронность хода уровня при расстоянии между скважинами до 200—250 км. Если при этом одна из скважин имеет сравнительно короткий период наблюдений (не более 2—3 лет), то для передачи прогнозов на эту

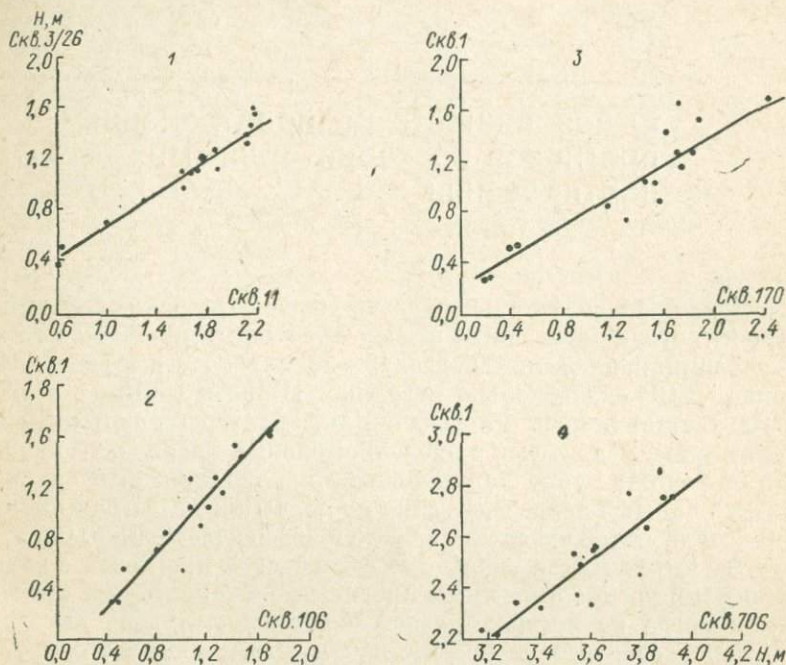


Рис. 27. Графики связи среднемесячных уровней грунтовых вод (в метрах от поверхности земли) по скважинам, находящимся в аналогичных условиях формирования режима уровня.

Расстояние между скважинами: 3/26 (Ленинградская обл., пос. Ефимовский) и 11 (г. Ленинград) — 240 км, 1 (Ленинградская обл., г. Волхов) и 170 (г. Ленинград) — 105 км, 1 (Ленинградская обл.) и 106 (г. Ленинград) — 110 км, 1 (Новгородская обл., пос. Крестцы) и 706 (г. Ленинград) — 240 км.

Уравнения регрессии:

- 1)  $y = 0,01 + 0,662x$ ;  $r = 0,98 \pm 0,01$ ;  $S/\sigma = 0,21$ ;
- 2)  $y = -0,18 + 1,09x$ ;  $r = 0,94 \pm 0,03$ ;  $S/\sigma = 0,35$ ;
- 3)  $y = 0,19 + 0,608x$ ;  $r = 0,93 \pm 0,04$ ;  $S/\sigma = 0,40$ ;
- 4)  $y = -0,14 + 0,732x$ ;  $r = 0,87 \pm 0,07$ ;  $S/\sigma = 0,50$ .

скважину со скважины с продолжительным периодом наблюдений можно использовать графики или уравнения связи, рассчитанные по среднемесячным или срочным уровням (рис. 27).

Следует иметь в виду, что пользоваться такими графиками можно лишь в тех случаях, если прогнозное значение уровня в скважине с длительными наблюдениями незначительно отли-

\* Под сезонной синхронностью подразумевается синхронность срочных или среднемесячных уровней, взятых в хронологической последовательности. Под многолетней подразумевается синхронность любых фазовооднородных или среднегодовых уровней.

чается от уровней в той же скважине, зафиксированных во время совместных наблюдений с другой скважиной, на которую передается прогноз.

В случае более продолжительных совместных наблюдений следует отдавать предпочтение отдельным графикам связи, построенным для фазооднородных и среднегодовых уровней (рис. 28),

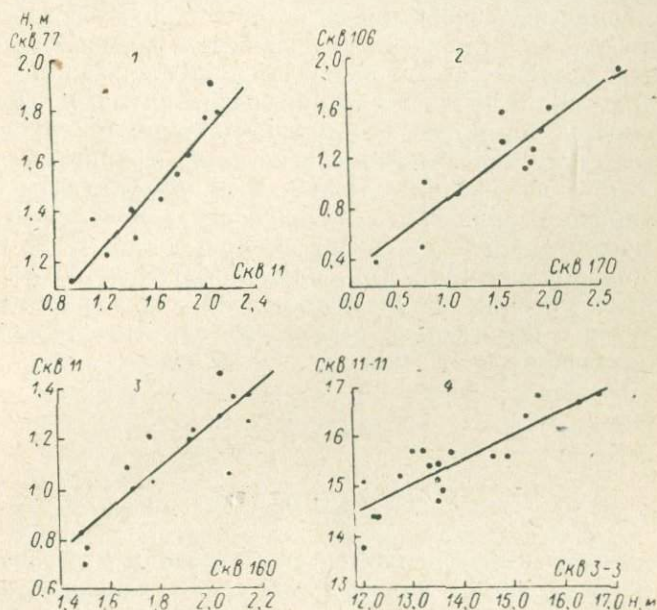


Рис. 28. Графики связи между предвесенними минимальными (1, 2) и среднегодовыми (3, 4) уровнями грунтовых вод (в метрах от поверхности земли) по скважинам, находящимся в аналогичных условиях формирования режима уровня.

Расстояния между скважинами: 77 (Ленинградская обл.) и 11 (г. Ленинград) — 60 км, 106 и 170 (г. Ленинград) — 20 км, 11 (г. Ленинград) и 160 (Ленинградская обл.) — 20 км, 11-11 и 3-3 (Ленинградская обл.) — 26 км.

Уравнения регрессии:

- 1)  $y = 0,58 + 0,563x$ ;  $r = 0,93 \pm 0,05$ ;  $S/\sigma = 0,37$ ;
- 2)  $y = 0,27 + 0,605x$ ;  $r = 0,90 \pm 0,06$ ;  $S/\sigma = 0,43$ ;
- 3)  $y = -0,42 + 0,833x$ ;  $r = 0,87 \pm 0,07$ ;  $S/\sigma = 0,50$ ;
- 4)  $y = 8,75 + 0,483x$ ;  $r = 0,87 \pm 0,06$ ;  $S/\sigma = 0,49$ .

которые можно использовать для передачи прогнозов с точки на точку.

При исследовании синхронности режима уровней грунтовых вод крупных территорий площадью в несколько сотен тысяч квадратных километров и больше основное внимание следует уделять районированию по климатическим режимообразующим факторам, определяющим особенности режима уровня. В этом случае роль геологических факторов формирования режима существенно уменьшается. Таким образом, по мере укрупнения масштаба исследований и уменьшения площади изучаемой территории возрастает роль аazonальных геологических

режимообразующих факторов и уменьшается роль зональных климатических факторов.

Важное значение для изучения региональных закономерностей режима уровня грунтовых вод и его прогноза имеют вопросы картирования режима. При картировании естественного режима уровня грунтовых вод крупных территорий возникает ряд методических вопросов, основными из которых являются следующие: выбор того или иного показателя положения уровня и расчетного периода наблюдений, выбор наблюдательных скважин.

Картирование и прогноз режима уровня крупных территорий осуществляются не в абсолютных показателях (в глубинах от поверхности или в абсолютных отметках), а в коэффициентах относительного положения уровней. Эти коэффициенты сравнительно широко используются в гидрологии; для гидрогеологического картирования их впервые применили С. М. Семенов и А. А. Коноплянцев [59]. Так, для определения коэффициента относительного положения уровней, подчиняющихся нормальному или близкому к нему закону распределения, эти авторы использовали следующие выражения:

$$\varepsilon = \frac{H_i - H_{\min}}{H_{\max} - H_{\min}}; \quad (111)$$

$$\varepsilon = \frac{H_i - H_{99\%}}{H_{1\%} - H_{99\%}}; \quad (112)$$

где  $H_i$  — прогнозный или фактический уровень  $i$ -го года;  $H_{\max}$ ,  $H_{\min}$  — максимальный и минимальный уровни за использованный период наблюдений;  $H_{1\%}$ ,  $H_{99\%}$  — уровни 1%-ной и 99%-ной обеспеченности;  $\varepsilon$  — коэффициент относительного положения уровня.

При  $H_i = H_{\max}$  или  $H_i = H_{1\%}$   $\varepsilon = 1$  и уровни занимают наиболее высокое положение. При  $H_i = H_{\min}$  или  $H_i = H_{99\%}$   $\varepsilon = 0$  и уровни занимают наиболее низкое положение за использованный период наблюдений.

В случае, если распределение уровней существенно асимметрично, В. С. Ковалевский [32] использовал следующие формулы для вычисления процента отклонения уровня от уровня 50%-ной обеспеченности:

$$\begin{aligned} +m &= \frac{(H_{50\%} - H_i) \cdot 100\%}{H_{50\%} - H_{1\%}}; \\ -m &= \frac{(H_{50\%} - H_i) \cdot 100\%}{H_{99\%} - H_{50\%}}; \end{aligned} \quad (113)$$

где  $+m$  — процентное отклонение уровня от уровня 50%-ной обеспеченности для случая, когда уровень  $H_i$  выше уровня  $H_{50\%}$ ;  $-m$  — процентное отклонение уровня от уровня 50%-ной обеспеченности для случая, когда уровень  $H_i$  ниже уровня  $H_{50\%}$ .

Коэффициенты относительного положения  $\varepsilon$ , так же как и процентные отклонения ( $\pm m$ , %), вычисляются только для фазовооднородных или среднегодовых уровней одного и того же года. Независимо от формулы, принятой для определения коэффициента относительного положения уровней, он должен вычисляться за один, общий для всех скважин, период наблюдений. Только в этом случае коэффициенты для разных скважин будут сопоставимыми друг с другом. Так, например, уровень 1%-ной обеспеченности, вычисленный для одной скважины за 10-летний период наблюдений, можно сопоставить с уровнем 1%-ной обеспеченности для другой скважины, вычисленным за тот же период наблюдений, но нельзя сопоставить с уровнем 1%-ной обеспеченности, вычисленным для третьей скважины за 30-летний период наблюдений.

При выборе продолжительности расчетного периода, используемого для вычисления коэффициентов  $\varepsilon$  или  $m$ , следует иметь в виду, что по длительным периодам наблюдений получаются наиболее достоверные значения этих коэффициентов, однако при этом значительно сокращается количество скважин, участвующих в построении карт режима. При выборе общего периода необходимо учитывать цикличность колебаний уровней. Следует стремиться к тому, чтобы использованный период был по крайней мере не меньше продолжительности основных циклов колебаний уровней.

Для составления карт режима необходимо использовать данные лишь по наиболее репрезентативным скважинам, которые с наибольшей полнотой отражают особенности режима уровня грунтовых вод тех или иных районов. Эти карты могут составляться как по фактическим, так и по прогнозным данным. В качестве примера на рис. 29, 30 представлены карты\* относительного положения фактически наблюдававшихся экстремальных и среднегодовых уровней грунтовых вод в 1971 г. на территории северо-западных и центральных областей РСФСР. Для вычисления коэффициентов относительного положения уровней был использован период наблюдений с 1958 г., несмотря на то что по отдельным скважинам имелись и более продолжительные ряды наблюдений. В колебаниях среднегодовых уровней в центральных районах РСФСР преобладают циклы продолжительностью 13—15 лет, в колебаниях среднегодовых и экстремальных уровней на Северо-Западе РСФСР — циклы продолжительностью 4—5 лет. Таким образом, использованный 14-летний период наблюдений почти полностью характеризует цикличность уровней, которая наблюдается на исследуемой территории. Поскольку распределение уровней грунтовых вод на этой территории, как правило, весьма близко к нормальному, для вычисления коэффициентов относительного положения уровней использовалась

\* Составлены автором совместно с Б. А. Бабкиным.

формула (111). На картах (см. рис. 29, 30) прослеживается смена в широтном направлении зон различной водности, связанная с особенностями циркуляции атмосферы в пределах исследуемой территории.

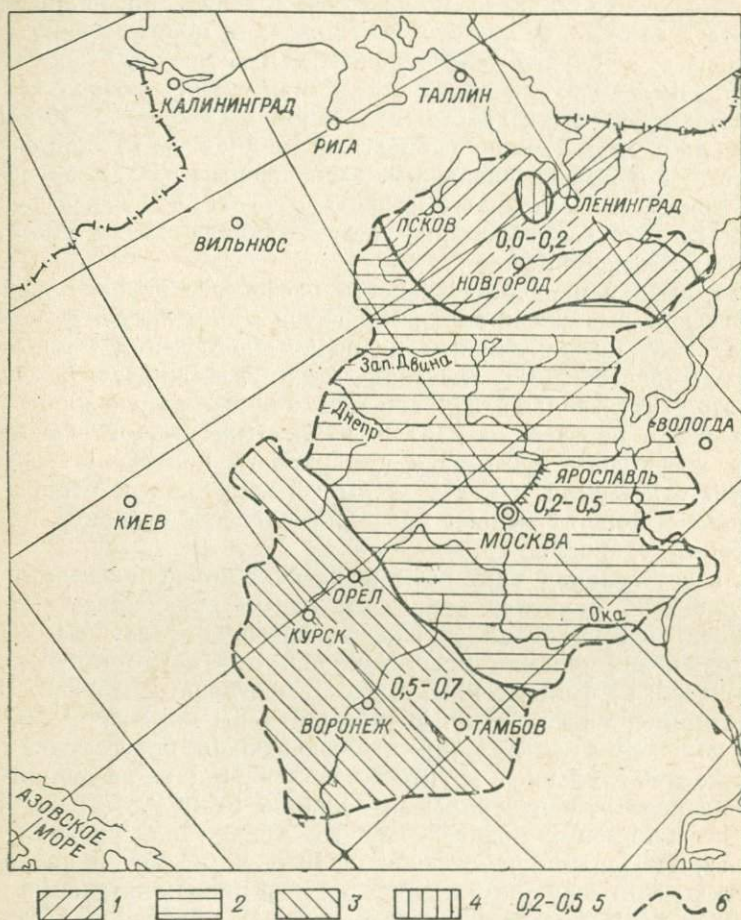


Рис. 29. Карта коэффициентов относительного положения летне-осенних минимальных уровней грунтовых вод на территории северо-западных и центральных областей РСФСР в 1971 г.

1 — коэффициенты относительного положения 0,0—0,2, т. е. уровни ниже многолетней нормы на 50—30% многолетней амплитуды их колебаний; 2 — коэффициенты относительного положения 0,2—0,5, т. е. уровни либо около нормы, либо ниже нее на величину до 30% многолетней амплитуды; 3 — коэффициенты относительного положения 0,5—0,7, т. е. уровни либо около нормы, либо выше нее на величину до 20% многолетней амплитуды; 4 — территория, в пределах которой данный уровень не фиксировался; 5 — коэффициенты относительного положения уровней; 6 — граница территории исследований.

При картировании режима уровня грунтовых вод сравнительно небольших территорий коэффициенты относительного положения будут определяться не столько климатической зональностью,

скольким распространением геологических факторов формирования режима. Так, например, предвесенний минимальный уровень грунтовых вод в 1972 г. на территории Ленинградской области в хорошо дренированных районах с междуречным видом режима

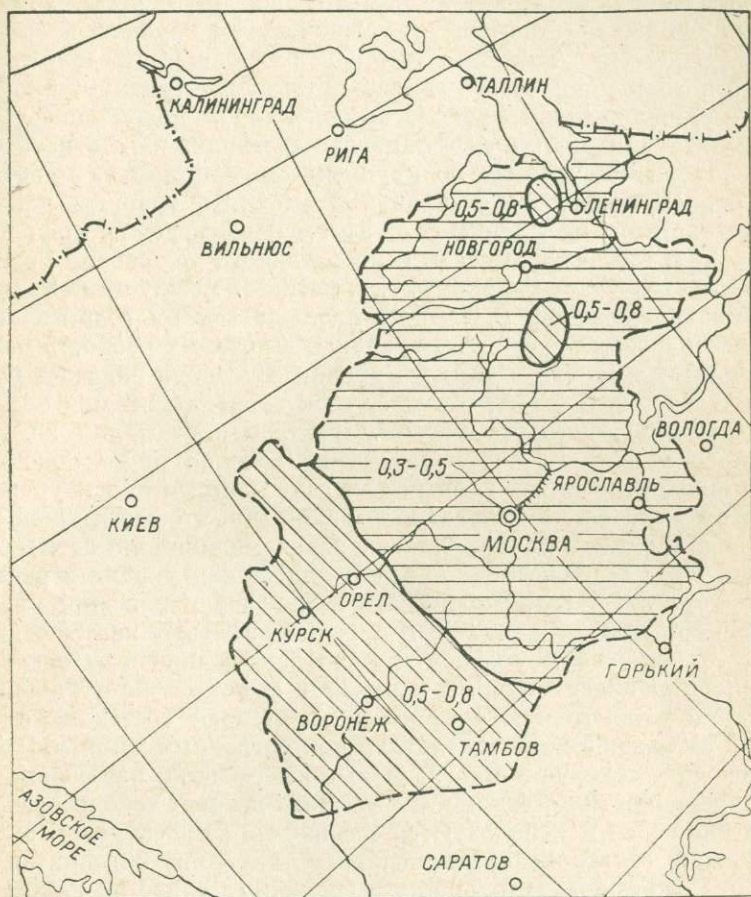


Рис. 30. Карта коэффициентов относительного положения среднегодовых уровней грунтовых вод на территории северо-западных и центральных областей РСФСР в 1971 г.

(в известняках) характеризовался коэффициентами  $\epsilon = 0,2 \div 0,4$ . В дренированных и слабо дренированных районах с междуречным видом режима, где грунтовые воды приурочены к пескам, супесям и суглинкам, предвесенний минимум характеризовался коэффициентом относительного положения  $\epsilon = 0,0 \div 0,1$ .

Это обстоятельство следует учитывать при составлении и картировании прогнозов для сравнительно небольших территорий, характеризующихся разнообразными геологическими условиями формирования режима уровня.

В настоящее время институтом ВСЕГИНГЕО и Гидрорежимной экспедицией МГ РСФСР начато составление по фактическим данным за ряд лет комплекса карт относительного положения уровней для европейской и частично азиатской части СССР и РСФСР. Анализ этих карт и сопутствующих им табличных и текстовых приложений поможет в решении ряда научных и практических вопросов. Так, по этим картам можно определить для каждого года площади с различным (высоким, низким, близким к норме) положением уровня, которое характеризуется соответствующими значениями коэффициента  $\epsilon$ . Анализ карт относительного положения уровней по крупным территориям за многолетний период позволит установить и уточнить границы районов с многолетним синхронным и асинхронным режимом уровня. В районах, где коэффициенты относительного положения уровней за ряд лет окажутся близкими, т. е. разность между ними не будет превышать в среднем 0,1, будет наблюдаться многолетняя синхронность режима уровней. Те районы, в которых разность между значениями  $\epsilon$  одного и того же уровня в течение ряда лет будет изменяться от 0 до 1, будут характеризоваться многолетней асинхронностью режима уровня. В связи с этим для каждого крупного района на территории СССР можно будет подобрать район-предиктор с синхронным или асинхронным ходом уровней, при этом желательно выбрать этот район так, чтобы синхронность (или асинхронность) проявлялась с запаздыванием в  $n$  лет. В этом случае, имея фактические данные о положении уровня в районе-предикторе в  $i$ -й год, можно дать прогноз о положении уровня в районе-предиктанте в  $(i + n)$ -й год. Используя карты относительного положения уровней, можно давать прогнозы о сроках наступления экстремальных уровней в том или ином районе.

Карты относительного положения уровней составляются не только по фактическим, но также и по прогнозным данным. В этом случае они содержат прогнозную информацию различной заблаговременности о тех или иных уровнях и могут использоваться, а частично уже и используются различными проектными, строительными, сельскохозяйственными и другими организациями в их практической деятельности в различных областях народного хозяйства.

---

## Заклучение

В работе рассмотрено применение некоторых вероятностно-статистических методов для прогнозов естественного режима уровня грунтовых вод, а именно: парной и множественной корреляции, гармонического анализа, а также корреляционные функции и авторегрессионные модели. Применение этих методов в практике режимных исследований стало возможным в результате накопления необходимого (хотя в большинстве случаев и минимального) объема фактических данных и все более широкого использования ЭВМ. Применение современной вычислительной техники наряду с несомненными преимуществами таит в себе и большие опасности, особенно для прогнозистов режима грунтовых вод, не обладающих значительным опытом и навыками в работе. Можно утверждать, что огромные технические возможности счетной техники в известном смысле находятся в противоречии с имеющимся объемом режимной гидрогеологической информации. Например, в настоящее время сравнительно просто получить уравнения регрессии с 5, 6 предикторами и более, однако их практическая ценность будет в большинстве случаев крайне ограниченной. Точно так же можно получить авторегрессионные модели весьма высоких порядков для прогнозов с заблаговременностью 5—10 шагов  $\Delta t$ , однако, как правило, и они не выдерживают проверки на независимом фактическом материале.

Следует иметь в виду, что вне зависимости от примененной методики наиболее простая прогнозная зависимость, т. е. та, в которой определяется минимальное число неизвестных коэффициентов, будет одновременно наиболее эффективной и надежной. Поэтому в общем случае, т. е. при наличии выборки весьма ограниченного объема около 15—25 членов, не следует для прогнозов уровней использовать сложные выражения, как-то: уравнения регрессии с большим числом предикторов, авторегрессии высокого порядка и т. д. Все найденные статистические характеристики (средние значения, дисперсии, коэффициенты корреляции и др.), а также коэффициенты прогнозных зависимостей следует оценивать с точки зрения их устойчивости, для чего

вычисляются их стандартные ошибки и доверительные интервалы, при этом, естественно, учитываются объемы использованных выборок. Те предикторы, коэффициенты при которых оказываются неустойчивыми, из прогнозных уравнений регрессии следует исключить.

Эффективность и точность всех использованных зависимостей оценивалась по величине отношения  $S/\sigma$ , где  $S$  — стандартная ошибка прогноза,  $\sigma$  — стандартное отклонение ряда уровней. Однако наиболее серьезным испытанием любого прогнозного уравнения является его проверка на независимом фактическом материале. В случае, если внутри использованного периода наблюдений удалось выявить наиболее важные, значимые связи между изменениями уровней и режимобразующими факторами или существенные закономерности в самих колебаниях уровней, то можно надеяться, что проверка на независимом материале даст удовлетворительные результаты.

В работе рассмотрены лишь некоторые методы прогноза естественного режима грунтовых вод. Уже теперь для анализа и прогноза режима уровня представляется целесообразным помимо описанных выше использовать методы компонентного и спектрального анализа.

В дальнейшем, по мере накопления новых данных и естественного удлинения имеющихся рядов наблюдений, возможности для применения разнообразных вероятностно-статистических методов прогноза будут все более и более расширяться, а сами методы совершенствоваться.

Приложение

Значимые величины коэффициента множественной корреляции  $R$

$n$	$j$								
	1	2	3	4	5	6	8	12	24
5%-ный уровень значимости									
10	0,632	0,758	0,839	0,898	0,942	0,973	1,000		
12	0,576	0,697	0,777	0,838	0,886	0,925	0,979	1,000	
14	0,532	0,648	0,726	0,786	0,835	0,877	0,941	1,000	
16	0,497	0,608	0,683	0,741	0,790	0,832	0,900	0,986	
18	0,468	0,574	0,646	0,703	0,751	0,792	0,861	0,958	
20	0,444	0,545	0,615	0,670	0,717	0,757	0,826	0,927	
22	0,423	0,520	0,587	0,641	0,686	0,726	0,794	0,896	
24	0,404	0,498	0,563	0,615	0,660	0,699	0,765	0,868	1,000
26	0,388	0,479	0,542	0,592	0,636	0,674	0,739	0,840	1,000
28	0,373	0,461	0,523	0,572	0,614	0,651	0,715	0,815	0,993
30	0,361	0,446	0,506	0,553	0,595	0,631	0,693	0,792	0,978
35	0,334	0,414	0,469	0,514	0,552	0,586	0,645	0,741	0,932
40	0,312	0,387	0,439	0,482	0,518	0,550	0,607	0,697	0,886
45	0,294	0,365	0,414	0,455	0,490	0,520	0,574	0,661	0,845
50	0,279	0,346	0,394	0,432	0,465	0,495	0,545	0,629	0,808
55	0,266	0,330	0,375	0,412	0,444	0,473	0,522	0,602	0,776
60	0,254	0,316	0,359	0,394	0,425	0,452	0,500	0,577	0,746

1%-ный уровень значимости

10	0,765	0,855	0,911	0,949	0,975	0,991	1,000	
12	0,708	0,800	0,860	0,904	0,938	0,963	0,993	1,000
14	0,661	0,753	0,814	0,861	0,898	0,928	0,971	1,000
16	0,623	0,712	0,773	0,821	0,859	0,891	0,942	0,996
18	0,590	0,677	0,737	0,785	0,824	0,857	0,911	0,980

Продолжение приложения

<i>n</i>	<i>j</i>								
	1	2	3	4	5	6	8	12	24
20	0,561	0,647	0,706	0,752	0,791	0,825	0,880	0,958	
22	0,537	0,620	0,677	0,724	0,762	0,796	0,852	0,934	
24	0,515	0,596	0,652	0,697	0,736	0,769	0,825	0,910	1,000
26	0,496	0,574	0,630	0,674	0,712	0,745	0,800	0,886	1,000
28	0,479	0,555	0,605	0,652	0,690	0,722	0,778	0,864	0,998
30	0,463	0,538	0,590	0,633	0,670	0,701	0,756	0,842	0,989
35	0,429	0,500	0,550	0,591	0,626	0,657	0,709	0,794	0,955
40	0,403	0,469	0,517	0,556	0,589	0,618	0,670	0,752	0,917
45	0,380	0,444	0,489	0,526	0,558	0,587	0,636	0,715	0,880
50	0,361	0,422	0,465	0,501	0,533	0,559	0,606	0,683	0,846
55	0,344	0,403	0,445	0,479	0,509	0,535	0,580	0,656	0,816
60	0,330	0,386	0,426	0,459	0,488	0,514	0,558	0,631	0,785

Примечание. *n* — число наблюдений каждой из переменных;  
*j* — число независимых переменных. При *j* = 1 коэффициент *R* обращается в обычный коэффициент парной корреляции.

---

## Список литературы

1. Алексеев Г. А. Объективные методы выравнивания и нормализации корреляционных связей. Л., Гидрометеиздат, 1971. 363 с.
2. Алехин Ю. М. Динамико-статистический метод прогноза геофизических макропроцессов (на примере прогнозов речного стока). — В кн.: Вопросы расчетов и прогноза речного стока. Л., 1961, с. 97—123. (Труды Ленингр. гидрометеорол. ин-та, вып. 11).
3. Багров Н. А. Статистические методы предсказания погоды. — «Метеорология и гидрология», 1964, № 2, с. 10—19.
4. Багров Н. А. О некоторых особенностях корреляционного анализа и их применении к прогнозам погоды. — «Метеорология и гидрология», 1968, № 1, с. 3—13.
5. Басов Г. Ф. Итоги шестидесятилетнего изучения гидрологической роли лесных полос и режима грунтовых вод Каменной Степи. — В кн.: Тр. III Всесоюз. гидрол. съезда, т. 2. Л., Гидрометеиздат, 1959, с. 105—112.
6. Богачев В. Г. Методы прогноза режима подземных вод Силурийского плато. — В кн.: Вопросы изучения и прогноза режима подземных вод. М., «Недра», 1964, с. 73—81.
7. Борович З. И. Определители и матрицы. Л., Изд-во Ленингр. гос. ун-та, 1965. 163 с.
8. Брукс К., Карузерс Н. Применение статистических методов в метеорологии. Л., Гидрометеиздат, 1963. 352 с.
9. Варзар С. М. Практическое применение метода Лагранжа — Дэла. «Зап. Гос. гидрол. ин-та», 1928, т. 2, с. 61—82.
10. Варзар С. М. К вопросу об отыскании скрытых периодичностей. — «Зап. Гос. гидрол. ин-та», 1934, т. 13, с. 39—77.
11. Вентцель Е. С. Теория вероятностей. М., «Наука», 1962. 560 с.
12. Гренджер К., Хатанака М. Спектральный анализ временных рядов в экономике. М., «Статистика», 1972. 312 с.
13. Дженкинс Г., Ваттс Д. Спектральный анализ и его приложения. Вып. 1. М., «Мир», 1971. 316 с.

14. Гвоздева В. Г. Опыт сверхдолгосрочного прогноза среднегодовых расходов р. Невы — «Труды Ленингр. гидрометеорол. ин-та», 1964, вып. 26, с. 95—112.

15. Дроздов О. А., Григорьева А. С. Многолетние циклические колебания осадков на территории СССР. Л., Гидрометеиздат, 1971. 158 с.

16. Дугинов В. И., Коробейников В. А. Многолетние колебания грунтовых вод Каменной Степи и их связь с колебаниями метеорологических элементов. — «Разведка и охрана недр», 1957, № 8, с. 43—49.

17. Дугинов В. И., Коробейников В. А. Ритмические колебания уровней грунтовых вод Каменной Степи и причины, вызывающие их. — В кн.: Тр. III Всесоюз. гидрол. съезда, т. 9. Л., Гидрометеиздат, 1959, с. 303—309.

18. Ефимов Н. В. Квадратичные формы и матрицы. М., «Наука», 1967. 160 с.

19. Ефремова Т. Н. Прогноз естественного режима уровня грунтовых вод. — «Разведка и охрана недр», 1960, № 8, с. 38—45.

20. Зальцберг Э. А. Определение элементов баланса грунтовых вод на отдельных участках территории г. Ленинграда. — В кн.: Мат-лы гидрорежимной экспедиции по изучению режима подземных вод и инж.-геол. процессов. М., 1970, с. 44—57. Труды Всесоюз. науч.-исслед. ин-та гидрогеологии и инж. геологии, вып. 27).

21. Зальцберг Э. А. О связях среднегодового и экстремальных уровней грунтовых вод. — «Информ. бюлл. Центр. треста инж.-строит. изысканий», 1970, № 3 (20), с. 13—20.

22. Зальцберг Э. А. О результатах множественной корреляции максимальных весенних и осенних уровней грунтовых вод. — В кн.: Вопросы изучения и прогноза режима подземных вод. М., 1970, с. 59—68. (Труды Всесоюз. науч.-исслед. ин-та гидрогеологии и инж. геологии, вып. 25).

23. Зальцберг Э. А. О цикличности колебаний уровней грунтовых вод. Там же, с. 69—79.

24. Зальцберг Э. А., Деч В. Н., Кноринг Л. Д. Долгосрочные прогнозы уровней подземных вод. — «Разведка и охрана недр». 1970, № 10, с. 43—47.

25. Зальцберг Э. А. О вероятностных прогнозах среднегодовых уровней грунтовых вод. — В кн.: Вопросы изучения режима подземных вод и инж.-геол. процессов. М., 1971, с. 41—48. (Труды Всесоюз. науч.-исслед. ин-та гидрогеологии и инж. геологии, вып. 40).

26. Зальцберг Э. А. О режиме и балансе подземных вод на низинном болоте в зоне избыточного увлажнения. — В кн.: Вопросы изучения режима подземных вод и инж.-геол. процессов. М., 1972, с. 14—21. (Труды Всесоюз. науч.-исслед. ин-та гидрогеологии и инж. геологии, вып. 49).

27. Зальцберг Э. А. Долгосрочные прогнозы естественного режима уровня грунтовых вод в условиях избыточного увлажнения. — В кн. Докл. на семинаре слушателей курсов. М., Изд-во Моск. ун-та, 1972, с. 15—19 (Междунар. высшие гидролог. курсы ЮНЕСКО).

28. Зальцберг Э. А. Применение множественной корреляции для прогнозов минимальных уровней грунтовых вод. — В кн.: Применение ста-

тист. методов в гидрологии. Л., Гидрометеониздат, 1973, с. 209—214. (Труды Гос. гидрол. ин-та, вып. 196).

29. Калинин Г. П. Проблемы глобальной гидрологии. Л., Гидрометеониздат, 1968. 377 с.

30. Кенесарин Н. А. Гидрогеологический прогноз земных проявлений спокойного солнца. — «Узб. геол. журн», 1963, № 2, с. 58—59.

31. Киселев П. А. Гидродинамические принципы количественной оценки питания подземных вод на основе анализа их режима на территории Белоруссии. — В кн.: Режим и баланс подземных вод. Минск, «Наука и техника», 1967, с. 3—92.

32. Ковалевский В. С. Долгосрочные прогнозы естественного режима грунтовых вод. М., 1972. 137 с. (Всесоюз. науч.-исслед. ин-т гидрогеологии и инж. геологии).

33. Ковалевский В. С. Прогнозы сезонного режима грунтовых вод. — В кн.: Подземный сток и методы его исследования. М., «Наука», 1972, с. 35—48.

34. Колмогоров А. Н. Интерполирование и экстраполирование стационарных случайных последовательностей. — «Изв. АН СССР. Сер. мат.», 1941, № 5, с. 3—14.

35. Комаров И. С. Накопление и обработка информации при инженерно-геологических исследованиях. М., «Недра», 1972. 295 с.

36. Коноплянцев А. А., Ковалевский В. С., Семенов С. М. Естественный режим подземных вод и его закономерности. М., Госгеолтехиздат, 1963. 231 с. (Труды Всесоюз. науч.-исслед. ин-та гидрогеологии и инж. геологии, № 2).

37. Коноплянцев А. А. Применение методов математической статистики для анализа и прогноза режима уровня подземных вод. М., 1967. 107 с. (Всесоюз. науч.-исслед. ин-т гидрогеологии и инж. геологии).

38. Коноплянцев А. А. Прогноз режима подземных вод методом аналогии. — «Разведка и охрана недр», 1968, № 2, с. 41—46.

39. Коноплянцев А. А. Вековые и многолетние колебания уровней грунтовых вод на территории СССР. — В кн.: Вопросы изучения и прогноза режима подземных вод. М., 1970, с. 34—58. (Труды Всесоюз. науч.-исслед. ин-та гидрогеологии и инж. геологии, вып. 25).

40. Коноплянцев А. А., Семенов С. М. Методические рекомендации по картированию естественного режима уровня грунтовых вод. М., 1971. 88 с. (Всесоюз. науч.-исслед. ин-т гидрогеологии и инж. геологии).

41. Коробейникова З. А., Коробейников В. А. Долгосрочный прогноз режима уровня грунтовых вод в Каменной Степи — «Разведка и охрана недр», 1959, № 10, с. 39—41.

42. Коробейников В. А. Составление сверхдолгосрочных прогнозов режима уровня грунтовых вод. — В кн.: Вопросы изучения и прогноза режима подземных вод. М., 1969, с. 68—83. (Труды Всесоюз. науч.-исслед. ин-та гидрогеологии и инж. геологии, вып. 23).

43. Крамбейн У., Кауфмен М., Мак-Кеммон Р. Модели геологических процессов. М., «Мир», 1973. 150 с.

44. Курош А. Г. Курс высшей алгебры. М., «Наука», 1965. 431 с.

45. Лебедев А. В. Методы изучения баланса грунтовых вод. М., Гостеолтехиздат, 1963. 192 с.
46. Лебедев А. В. Опыт статистической обработки результатов многолетних наблюдений за режимом грунтовых вод. — В кн.: Вопросы изучения и прогноза режима подземных вод. М., «Недра», 1964, с. 64—72.
47. Линслей Р. К., Колер М. А., Паулюс Д. Л. Х. Прикладная гидрология. Л., Гидрометеониздат, 1962. 759 с.
48. Максимов И. В., Смирнов Н. П. Опыт построения долгосрочного прогноза основных форм атмосферной циркуляции в северном полушарии компонентно-гармоническим методом. — В кн.: Увеличение заблаговременности и детализации долгосрочных метеорол. прогнозов для Арктики. Л., Гидрометеониздат, 1965, с. 231—259. (Труды Аркт. и антаркт. науч.-исслед. ин-та, т. 262).
49. Методическое руководство по изучению режима подземных вод. Ред. Альтовский М. Е., Коноплянцев А. А. М., Гостеолтехиздат, 1954. 196 с.
50. Монин А. С. Об использовании статистических методов в задачах прогноза погоды. — «Метеорология и гидрология», 1962, № 7, с. 3—9.
51. Озябкин В. Н. К методике долгосрочного прогноза уровня napopных вод в условиях их естественного режима. — В кн.: X конференция младших научных сотрудников, аспирантов и инженеров. М., 1966, с. 41—46. (Всесоюз. науч.-исслед. ин-т гидрогеологии и инж. геологии).
52. Опыт применения частотного анализа для изучения внутренней структуры рядов. — В кн.: Численные методы исследования гидрометеорол. условий в Арктике с использованием ЭЦВМ. Л., 1966, с. 147—157. Авт. Вольцингер Н. Е., Максимов И. В., Пясковский Р. В., Хаминов К. А. (Труды Аркт. и антаркт. науч.-исслед. ин-та, т. 277).
53. Пановский Г. А., Брайер Г. В. Статистические методы в метеорологии. Л., Гидрометеониздат, 1967. 209 с.
54. Пиварелис В. П. Способ графической корреляции в применении к гидрологии. — В кн.: Тр. III Всесоюз. гидрол. съезда, т. 2. Л., Гидрометеониздат, 1959, с. 507—512.
55. Попов Е. Г. Вопросы теории и практики прогнозов речного стока. М., Гидрометеониздат, 1963. 395 с.
56. Речной сток и геофизические процессы. М., «Наука», 1966. 295 с. Авт.: Дружинин И. П., Коноваленко З. П., Кукушкина В. П., Хамьянова Н. В.
57. Рождественский А. В. Анализ многолетних колебаний речного стока и методов их математического описания. — В кн.: Сб. работ по гидрологии, № 10. Л., Гидрометеониздат, 1970, с. 22—72.
58. Сакалаускене Д. Ю. Прогноз естественного режима уровня грунтовых вод в переувлажненных областях. — «Разведка и охрана недр», 1971, № 9, с. 43—48.
59. Семенов С. М., Коноплянцев А. А. Карты информации и прогнозов режима подземных вод. — «Разведка и охрана недр», 1969, № 2, с. 43—48.
60. Серебренников М. Г., Первозванский А. А. Выявление скрытых периодичностей. М., «Наука», 1965. 244 с.

61. Сидоркина С. П. О долгосрочном прогнозе гидрогеологических процессов методами теории стационарных случайных функций. — В кн.: Применение мат. методов при гидрогеол. и инж.-геол. исследованиях. М., 1970, с. 74—91. (Труды Всесоюз. науч.-исслед. ин-та гидрогеологии и инж. геологии, вып. 34).

62. Сиротенко О. Д. Многомерный регрессионный анализ, как метод исследования связи урожая сельскохозяйственных культур с гидрометеорологическими факторами. — «Метеорология и гидрология», 1969, № 12, с. 68—78.

63. Слуцкий Е. Е. Избранные труды. М., Изд-во АН СССР, 1960. 292 с.

64. Смирнов Н. В., Дунин-Барковский И. В. Курс теории вероятностей и математической статистики для технических приложений. М., «Наука», 1965. 511 с.

65. Соколовский Д. Л. Речной сток. Л., Гидрометеиздат, 1968. 539 с.

66. Субботин А. И. Об использовании наблюдений за подземными водами при исследованиях и прогнозах речного стока. — В кн.: Тр. III Всесоюз. гидрол. съезда, т. 9. Л., Гидрометеиздат, 1959, с. 103—111.

67. Токарев Н. С. Ритмические колебания климата и их влияние на режим поверхностных и подземных вод. — В кн.: Тр. лаб. гидрогеол. пробл. им. акад. Ф. П. Саваренского, т. 9. М., Изд-во АН СССР, 1950, с. 46—72.

68. Уланова Е. С., Сиротенко О. Д. Методы статистического анализа в агрометеорологии. Л., Гидрометеиздат, 1968. 198 с.

69. Яглом А. М. Введение в теорию стационарных случайных функций. — «Успехи мат. наук», 1952, т. 7, вып. 5 (51), с. 3—168.

70. Bogardi J. A varheto tavaszi maximalis havi közepes talajvi, zallasok előrejelzeze az Alföldön. — «Hidrologiai Közlöny», 1953, vol. 33, № 11—12, p. 415—422.

71. Box G. E., Jenkins G. M. Time series analysis. Forecasting and control. San Francisco, London, Cambridge, Amsterdam, Holden-Day, 1970. 553 p.

72. Climatic change. — «WMO Technical Note», 1966, № 79, 79 p.

73. Kříž H. Statistické zhodnocení nejdělsiho pozorování podzemních vod v ČSSR. — «Vodohospodársky časopis», 1971, r. 19, č. 6, s. 545—570.

74. Netopil R. K otázce stanovení charakteristických úrovní hladiny spodní vody. — «Vodohospodársky časopis», 1958, r. 6, č. 3, s. 209—220.

75. Rorabaugh M. Prediction of groundwater levels on basis of rainfall and temperature correlations. — «Trans. Amer. Geophys. Union», 1956, vol. 37, № 4, p. 436—441.

---

## оглавление

Предисловие . . . . .	3
1. Краткая классификация прогнозов естественного режима уровня грунтовых вод, оценка их точности и эффективности . . . . .	5
2. Применение парной и множественной корреляции для составления прогнозов естественного режима уровня грунтовых вод . . . . .	10
2.1 — расчет линейного уравнения парной регрессии . . . . .	—
2.2 — оценка достоверности коэффициента корреляции и коэффициентов уравнения регрессии . . . . .	16
2.3 — выбор факторов-предсказателей положения уровней . . . . .	21
2.4 — расчет линейного уравнения множественной регрессии . . . . .	31
2.5 — расчет уравнения нормализованной множественной регрессии . . . . .	36
2.6 — оценка достоверности коэффициента множественной корреляции и коэффициентов уравнения множественной регрессии . . . . .	42
2.7 — выбор факторов-предикторов для использования в уравнениях множественной регрессии . . . . .	46
2.8 — графическая множественная корреляция . . . . .	51
3. Применение корреляционных функций и авторегрессионных моделей для составления прогнозов режима уровня . . . . .	53
3.1 — вычисление корреляционных функций и их анализ . . . . .	—
3.2 — простейшие авторегрессионные модели, описывающие колебания уровней . . . . .	57

4. Применение методов периодограммного анализа для прогнозов естественного режима уровня грунтовых вод . . . . .	65
4.1 — интегральные кривые и сглаживание . . . . .	—
4.2 — методика выявления периодических составляющих в колебаниях уровня . . . . .	68
4.3 — о соотношении гармонических составляющих и случайной компоненты в колебаниях уровней грунтовых вод . . . . .	73
4.4 — составление прогнозов уровней и их проверка . . . . .	78
5. О возможности распространения прогнозов по территории и их картировании . . . . .	83
Заключение . . . . .	91
Приложение . . . . .	93
Список литературы . . . . .	95

*Эрнст Абрамович Зальцберг*

**Статистические методы прогноза  
естественного режима уровня грунтовых вод**

---

Редактор издательства Т. М. Пономарева  
Обложка художника Ю. И. Прошлецова  
Техн. редактор А. Б. Ящуржинская  
Корректор Р. С. Янбекова

Сдано в набор 29/VII 1975 г. Подписано в печать 8/IV 1976 г.  
М-37194. Формат 60 × 90<sup>1/16</sup>. Бумага тип. № 2. Печ. л. 6<sup>1/2</sup>.  
Уч.-изд. л. 6,55. Тираж 4000 экз. Заказ 461/333. Цена 34 коп.

Издательство «Недра», Ленинградское отделение.  
193171, Ленинград, С-171, ул. Фарфоровская, 12.

Ленинградская типография № 6 Союзполиграфпрома при Государственном  
комитете Совета Министров СССР по делам издательств, полиграфии и книж-  
ной торговли.  
196006, Ленинград, Московский пр., 91.

## УВАЖАЕМЫЕ ЧИТАТЕЛИ!

В 1976 г. Ленинградское отделение издательства «Недра» выпускает следующие книги.

### Геология

Курс общей геологии. Учебное пособие. Изд. 2, перераб. и доп. (I изд. — 1960 г.). Авт.: Серпухов В. И., Шалимов А. И., Мораховский В. Н. и др. 40 л., ц. 1 р. 80 к.

Бугаец А. Н., Дуденко Л. Н. Математические методы при прогнозировании месторождений полезных ископаемых. 20 л., ц. 1 р. 30 к.

Питулько В. М. Вторичные ореолы рассеяния в криолитозоне. 12 л., ц. 88 к.

Поротов Г. С., Веселов Е. В., Ларионов А. Г. Методика изучения комплексных магнетитовых руд. 12 л., ц. 88 к.

Басков Е. А. Палеогидрогеологический анализ при металлогенических исследованиях. 15 л., ц. 1 р. 74 к.

Вакье Виктор. Геомagnetизм в морской геологии. Амстердам, Лондон, Нью-Йорк, 1970. Пер. с англ. 15 л., ц. 1 р. 74 к.

Горбов А. Ф. Геохимия бора. 13 л., ц. 1 р. 54 к.

Дядькина И. Я., Орлова М. П. Месторождения флогопита. 15 л., ц. 1 р. 74 к.

Павлов А. Н. Геологический круговорот воды на Земле. 10 л., ц. 1 р. 24 к.

Периодические процессы в геологии. Авт.: Логвиненко Н. В., Айнемер А. И., Риттенберг М. И. и др. 20 л., ц. 2 р. 24 к.

Природные газы осадочной толщи. Авт.: Якуцени В. П., Несмелова З. Н., Тихомиров В. В. и др. 25 л., ц. 2 р. 90 к.

Раевский В. И., Медведев Н. С. Методы изучения изменчивости свойств пластовых месторождений. 10 л., ц. 1 р. 24 к.

Романовский С. И. Динамика формирования флиша. 13 л., ц. 1 р. 54 к.

Роненсон Г. В., Ортенберг Н. А. Методы расчленения метаморфических комплексов. 15 л., ц. 1 р. 74 к.

Синицын В. М. Климат латерита и боксита. 10 л., ц. 1 р. 24 к.

Шепард Ф. П. Морская геология. Изд. 3. Нью-Йорк, Ивенстон, Сан-Франциско, Лондон, 1973. Пер. с англ. 40 л., ц. 4 р. 40 к.

Эрнст Вернер. Геохимический анализ фаций. Амстердам, Лондон, Нью-Йорк, 1970. Пер. с англ. 10 л., ц. 1 р. 24 к.

Кузнецов С. С. Геологические экскурсии. Научно-популярная. 8 л., ц. 27 к.

Милашев В. А. Алмаз. Легенды и действительность. Научно-популярная. 8 л., ц. 27 к.

Новиков Э. А. Человек и литосфера. Научно-популярная. 8 л., ц. 27 к.

### Геодезия

Павлов В. И. Математическая обработка фотограмметрических измерений. 15 л., ц. 1 р. 74 к.

### Геофизика

Гамма-методы в рудной геологии. Под ред. А. П. Очкура. 25 л., ц. 1 р. 57 к.

Геофизическая аппаратура. Вып. 59. 13 л., ц. 84 к.

Руководство по применению метода переходных процессов в рудной геофизике. Под ред. Ф. М. Каменецкого. Авт.: Каменецкий Ф. М., Седов М. П., Макагонов П. П. и др., 10 л., ц. 54 к.

### Бурение

Морозов Ю. Т. Бурение направленных и многоствольных скважин малого диаметра. 13 л., ц. 93 к.

С аннотациями этих книг можно ознакомиться в «Плане выпуска литературы» на 1976 г. Своевременно оформляйте заказы в местных магазинах Книготорга или в магазинах № 17 «Недра» (199178, г. Ленинград, В-178, Средний пр., д. 61) и № 59 «Недра» (127412, г. Москва, И-412, Коровинское шоссе, д. 20). Тираж книги определяется собранными заказами.

*Ленинградское отделение издательства «Недра»*

34 коп.

1818