

УДК 627.512

В. Е. ГВОЗДЕВ, А. Е. КОЛОДЕНКОВА**МЕТОД ЭКСТРАПОЛЯЦИОННОГО ПРОГНОЗИРОВАНИЯ
ПАВОДКОВОЙ СИТУАЦИИ
НА ОСНОВЕ РЕТРОСПЕКТИВНЫХ И ТЕКУЩИХ ДАННЫХ
(НА ПРИМЕРЕ РЕСПУБЛИКИ БАШКОРТОСТАН)**

Рассматривается метод краткосрочного прогнозирования паводковой ситуации в условиях малого объема и низкой точности исходных данных. Метод основан на построении одномерных непараметрических функциональных зависимостей. Построение функциональных зависимостей осуществляется на основе совместного анализа законов распределения уровней воды, отнесенных к различным временным сечениям. Рассматривается пример использования предлагаемого метода для краткосрочного прогнозирования паводковой ситуации на территории Республики Башкортостан. *Прогнозирование паводковой ситуации; временное сечение; малые по объему и низкие по точности исходные данные; непараметрические функциональные зависимости*

В «Повестке дня на XXI век», принятой в Рио-де-Жанейро (1992 год), подчеркивается, что в настоящее время, одним из необходимых условий принятия эффективных решений при управлении территориальными системами, является своевременное обеспечение лиц, задействованных на разных уровнях управления, достоверной, полной и качественной информацией о текущем и прогнозируемом состоянии объекта управления. Это создает основу для перехода от «реактивного» подхода, суть которого сводится к устранению уже имеющих место негативных явлений, к «превентивному» подходу, основанному на выработке опережающих управленческих решений, на базе прогнозных оценок состояния объекта управления.

Одной из задач управления сложными объектами, имеющей важное экономическое и социальное значение, является предупреждение негативных последствий паводков. Необходимым условием повышения эффективности действий по предупреждению негативных последствий в период прохождения паводка является обеспечение государственных органов, занятых снижением риска вредного воздействия вод, полной и достоверной информацией о прогнозируемой паводковой ситуации.

Паводок представляет собой сложное явление, протекает на больших территориях в условиях многообразия географической среды. Трудности, связанные с информацион-

ным обеспечением предупреждения негативных последствий паводков обусловлены: низким качеством и ограниченным объемом исходных данных; уникальностью протекания паводка на различных участках территории; недостаточной изученностью механизмов, определяющих динамику изменения уровней воды на исследуемой территории; высокой стоимостью и длительным сроком подготовки картографической основы и др. В силу вышеизложенного, совершенствование методов прогнозирования негативных последствий паводка является актуальной задачей.

**МЕСТО ЗАДАЧИ КРАТКОСРОЧНОГО
ПРОГНОЗИРОВАНИЯ ПАВОДКОВОЙ
СИТУАЦИИ В СИСТЕМЕ ПРЕДУПРЕЖДЕНИЯ
НЕГАТИВНЫХ ПОСЛЕДСТВИЙ ПАВОДКА**

Несмотря на значительное число работ, посвященных проблеме прогнозирования паводковой ситуации посредством различных моделей (в первую очередь математических и геоинформационных), задача решена далеко не полностью. Известные математические модели разработаны для конкретных территорий и их использование для других территорий оказывается малоэффективным; модели являются параметрическими, причем в методиках по их применению отсутствуют четкие рекомендации по выбору значений параметров (запасов воды в снежном покрове, характеристик осадков в период снеготаяния, в пе-

риод от схода снега до окончания паводка и т. д.) в зависимости от особенностей территорий. В силу того, что механизмы, определяющие характер протекания паводка недостаточно изучены и на каждой территории различны, для решения задач прогнозирования целесообразно использовать подходы, основанные на функциональной парадигме.

Необходимым компонентом создания информационной системы является установление места задачи краткосрочного прогнозирования в системе предупреждения негативных последствий паводка. Это, в свою очередь, требует разработки модели, описывающих структуру функций, реализуемых системой.

На рис. 1 и рис. 2 приводятся фрагменты структурно-функциональной модели, разработанной с использованием методологии IDEF0, на основе методических рекомендаций для председателей противопаводковых комиссий муниципальных образований Республики Башкортостан по выполнению комплекса противопаводковых мероприятий [1].

Анализ содержания функций и состава информационных потоков между ними является основанием для разработки архитектуры информационной подсистемы, реализующей информационную поддержку действий по предупреждению негативных последствий паводков (рис. 3). Подробно архитектура подсистемы описана в [2].

ОБОБЩЕННАЯ СХЕМА ЭКСТРАПОЛЯЦИОННОГО ПРОГНОЗИРОВАНИЯ УРОВНЕЙ ВОДЫ НА ПОСТАХ КОНТРОЛЯ

Одной из функциональных задач, реализуемых подсистемой информационной поддержки действий по предупреждению негативных последствий паводков, является задача краткосрочного прогнозирования уровней воды на постах контроля.

Постановка задачи краткосрочного прогнозирования на основе ретроспективных и текущих данных имеет следующий вид.

Дано: архив измеренных значений уровней воды на r -м посту контроля $\{\Pi_r^{(j)}(t_i)\}$, где $i = \overline{1; N}$, $j = \overline{1; M}$. Здесь N — число временных сечений, в которых зарегистрированы архивные данные $\{\Pi_r^{(j)}(t_i)\}$; M — число процессов, по которым формировались архивные данные. Предполагается что, в моменты времени t_i , $i = \overline{1; N}$, в которых осуществлялись замеры параметра, являются оди-

наковыми для всех процессов; известно значение $\Pi_r(t_q)^{(M+1)}$ $(M+1)$ -го процесса в k -й момент времени $k \in \{t_1, t_2, \dots, t_N\}$.

Требуется: Рассчитать значение $\Pi_r(t_q)^{(M+1)}$, $k < q < N$.

Предлагаемый подход к решению задачи основан на построении функциональных зависимостей, устанавливающих взаимосвязь между уровнями воды в k -м и q -м временных сечениях, $\Pi_r(t_q)^{(M+1)} = \varphi_{kq}(\Pi_r(t_k)^{(M+1)})$. Эти зависимости строятся на основе обработки архивных данных $\{\Pi_r(t_k)\}$, $\{\Pi_r(t_q)\}$, зарегистрированных в k -м и q -м временных сечениях.

В основе решения задачи лежит известный подход, основанный на определении функции распределения функции случайного аргумента, описанный в [3]. Из известного соотношения (1)

$$F_2(y) = \int_{\varphi(x) < y} f_1(x) dx, \quad (1)$$

следует, что если зависимость $y = \phi(x)$ является непрерывной, монотонной, безынерционной и однозначной и известны законы распределения $F_1(x)$ и $F_2(y)$ входной и выходной случайных величин X и Y , то для заданного значения x можно подобрать такое y , при котором $F_1(x) = F_2(y)$. В случае положительной зависимости будет справедливо равенство (2)

$$\int_{a(x)}^{x_p} f_1(x) dx = \int_{a(y)}^{y_p} f_2(y) dy. \quad (2)$$

Обобщенная схема решения задачи прогнозирования представлена на основе ретроспективных и текущих данных на рис. 4.

В центре предлагаемого подхода к решению задачи краткосрочного прогнозирования лежит оценивание законов распределения контролируемого параметра в различных временных сечениях.

ТЕОРЕТИЧЕСКАЯ ОСНОВА МЕТОДА ОЦЕНИВАНИЯ ФУНКЦИОНАЛЬНЫХ ЗАВИСИМОСТЕЙ В УСЛОВИЯХ МАЛОГО ОБЪЕМА И НИЗКОЙ ТОЧНОСТИ ИЗМЕРИТЕЛЬНЫХ ДАННЫХ

Постановка задачи

Допустим, известны выборочные значения входного X и выходного Y параметров.

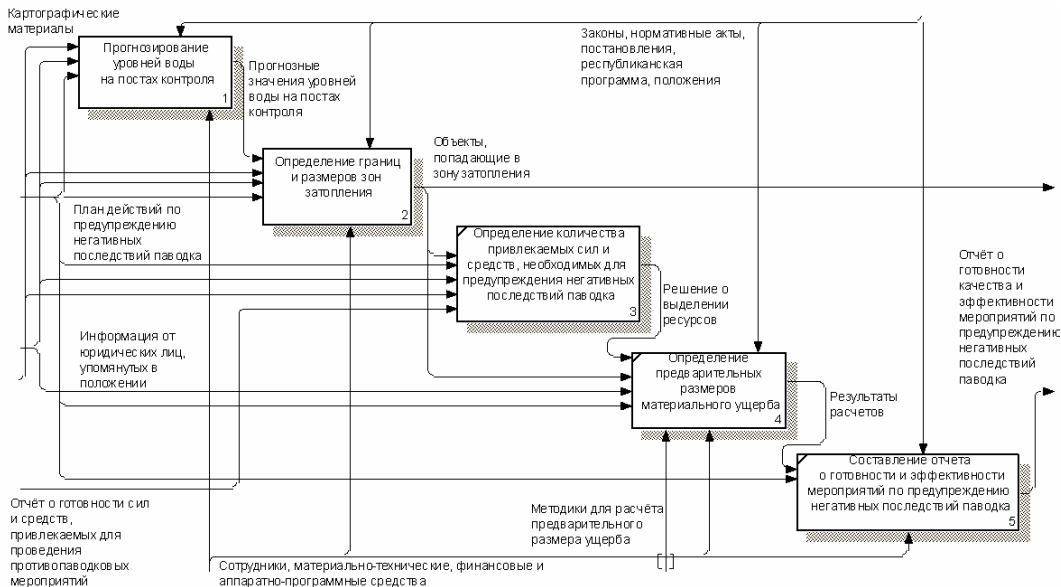


Рис. 1. Декомпозиция третьего уровня «Проведение превентивных мероприятий по предупреждению негативных последствий паводка»

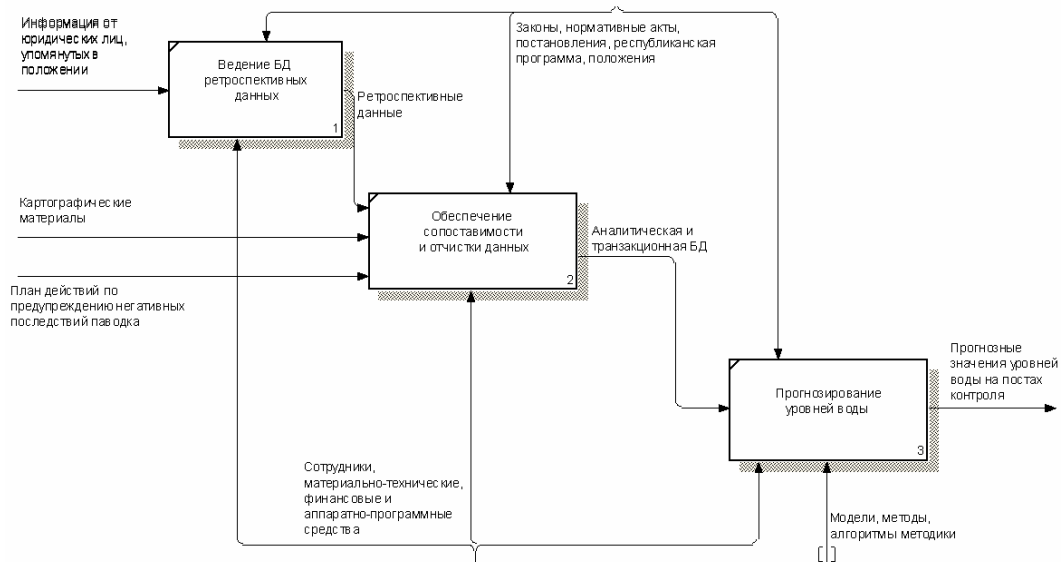


Рис. 2. Декомпозиция третьего уровня «Прогнозирование уровней воды на постах контроля»

Границы (хотя бы одна) интервалов физически возможных значений параметров полагаются известными из физических соображений $x \in [a^{(x)}, b^{(x)}]; y \in [a^{(y)}, b^{(y)}]$. Зависимость $y = \phi(x)$ полагается непрерывной, однозначной, монотонной, безынерционной и априорно неизвестной.

Требуется построить эмпирическую оценку $y = \hat{\phi}(x)$ на основе выборочных значений параметров X и Y .

Теоретическая основа метода

В основе метода лежит решение обратной, по отношению к (1), задачи, представлен-

ной в виде (2). Из (2) следует, что, задаваясь различными возрастающими значениями $\lambda_p \in [0;1]$ при известных $F_1(x), F_2(y)$, из соотношения (3)

$$F_1(x_p) = F_2(y_p) = \lambda_p, \quad p = 1, 2, \dots \quad (3)$$

можно получить серию пар чисел $\{x_p, y_p\}$, которые представляют значения функции $y = \phi(x)$ в точках x_p .

На рис. 5 приведена графическая иллюстрация метода для $a^{(x)} = a^{(y)} = 0; b^{(x)} = b^{(y)} = \infty$.

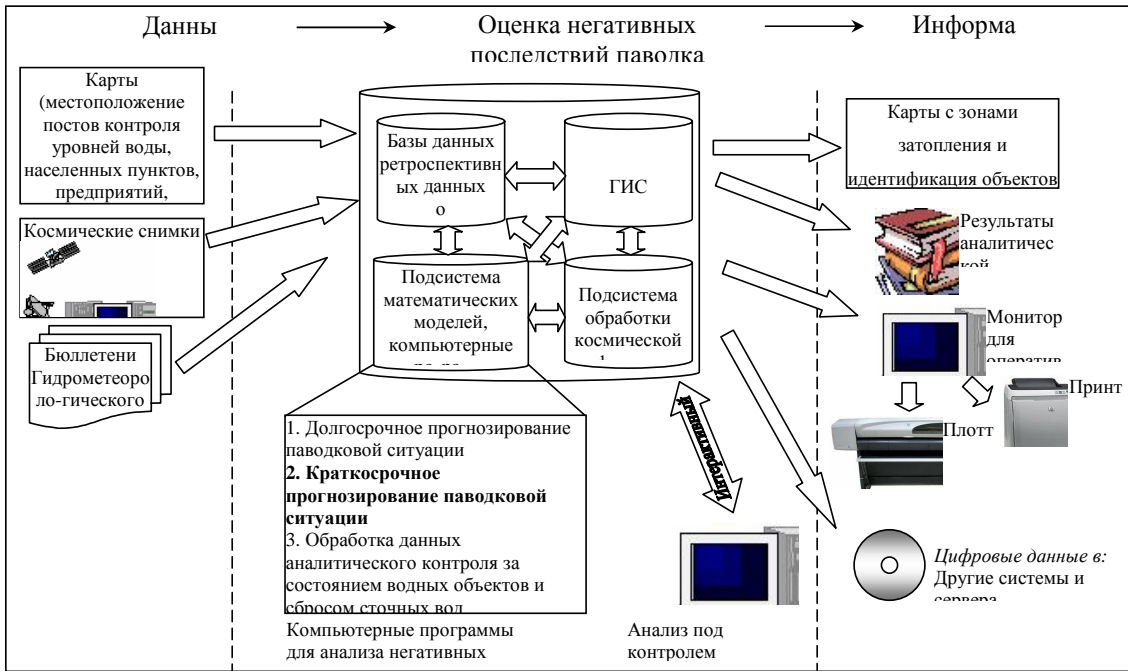


Рис. 3. Архитектура системы информационной поддержки мероприятий по предупреждению негативных последствий паводка

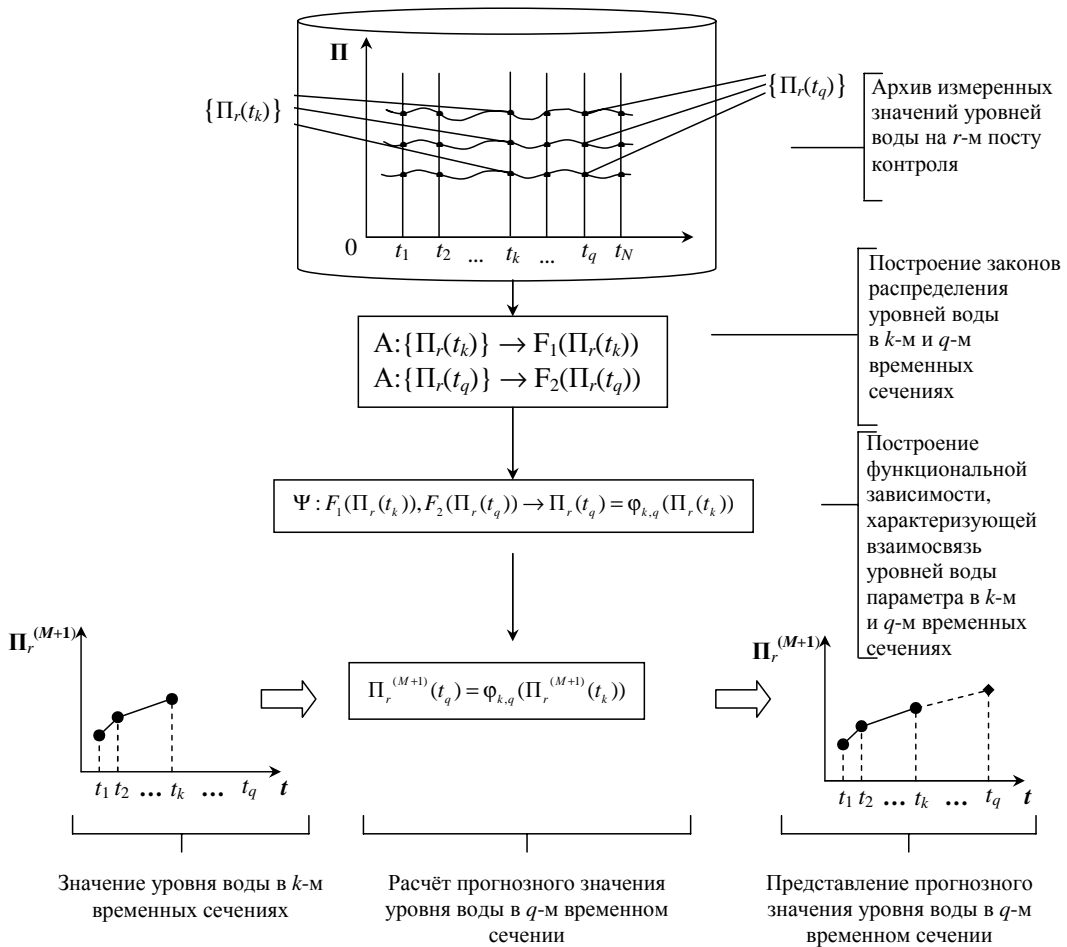


Рис. 4. Обобщенная схема решения задачи прогнозирования уровней воды на основе ретроспективных и текущих данных

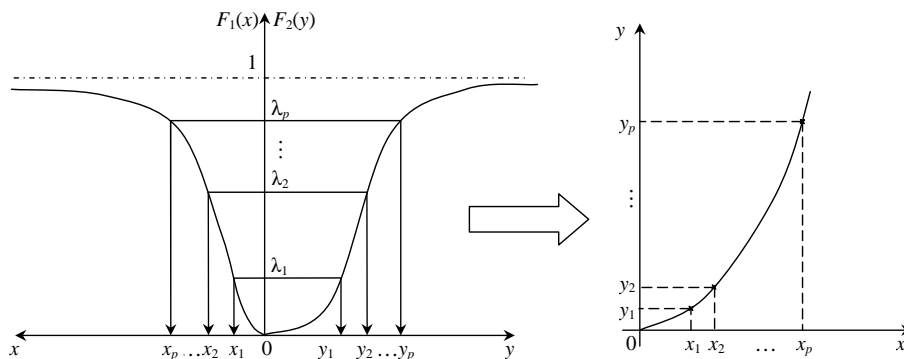


Рис. 5. Графическая иллюстрация метода непараметрического оценивания функциональных зависимостей на основе законов распределения входного и выходного параметров

Если шаг изменения λ_p от минимального до максимального значения выбрать достаточно малым, можно получить большое количество пар чисел $\{x_p, y_p\}$, так что визуально график $y = \hat{\varphi}(x)$ окажется практически непрерывным. При этом функциональная зависимость будет тем более нелинейной, чем сильнее различаются типы законов распределения $F_1(x)$ и $F_2(y)$.

Из описания метода следует, что точность непараметрического метода оценивания функциональных зависимостей будет тем выше, чем ближе к истинным окажутся оценки $\hat{F}_1(x)$ и $\hat{F}_2(y)$, определяемые по выборочным значениям X и Y . В связи с этим отдельного рассмотрения требует задача оценивания законов распределения случайных величин по выборочным данным.

ОЦЕНИВАНИЕ ЗАКОНОВ РАСПРЕДЕЛЕНИЯ СЛУЧАЙНЫХ ВЕЛИЧИН ПО ВЫБОРОЧНЫМ ДАННЫМ

Решению задачи оценивания законов распределения случайных величин по выборочным данным, в том числе по выборкам малого объема, посвящено достаточно много публикаций. В работе [4] проведен анализ известных решений и обоснована целесообразность применения энтропийного подхода. Это обусловлено тем, что:

1) Для оценивания различных законов распределения используется унифицированная параметрическая модель, что исключает необходимость выдвижения проверки гипотез о типе функции распределения.

2) Параметры модели, получаемые по значениям начальных моментов, совпадают с оценками, получаемыми методом максимального правдоподобия, из чего следует, что они являются состоятельными, несмещенными и эффективными; задача построения оценки закона распределения имеет единственное ре-

шение; получаемые оценки законов распределения являются наиболее вероятными при имеющихся результатах наблюдений.

3) Оценки функций распределения обеспечивают минимум максимального отклонения от истинной плотности распределения, а также удовлетворяют условию минимума критерия Хи-квадрат.

4) Эмпирическим путем было установлено, что точность аппроксимации, получаемая при учете первых четырех моментов, практически не изменяется по сравнению с результатами, получаемыми при использовании лишь первых двух моментов, что на первый взгляд противоречило известному методу моментов [5].

В работе [6] было показано, что противоречие является кажущимся и связано с тем, что при учете границ интервала физически возможных значений случайной величины практически вся информация о законе распределения заключается в первых моментах, причем идентификатором формы закона распределения является отношение математического ожидания к среднеквадратическому отклонению. Это позволило разработать унифицированную непараметрическую схему оценивания законов распределения по значениям двух первых моментов для случая, когда случайная величина принадлежит интервалу $[0; \infty)$.

Непараметрическая схема оценивания интегральных законов распределения случайных величин имеет следующий вид:

$$[\hat{v}_1, \hat{\sigma}] \rightarrow [\hat{v}_1^{(n)}] \rightarrow \{F_1^{(n)}(z^*)\} \rightarrow \{\hat{F}^{(n)}(z^*)\} \rightarrow \hat{F}(z), \quad (4)$$

где $\hat{v}_1, \hat{\sigma}$ — эмпирические оценки математического ожидания и среднеквадратического отклонения, определяемые по выборочным данным; $\hat{v}_1^{(n)}$ — отношение \hat{v}_1 к $\hat{\sigma}$

(идентификатор формы закона распределения); $\hat{F}^{(n)}(z^*)$ — значения нормированной интегральной функции распределения, рассчитанные в заранее определенных точках $\{z^*\}$ на основе предварительно сформированного набора $\{F_3^{(n)}(z^*)\}$ нормированных интегральных функций распределения; $F(z)$ — оценка интегральной функции распределения, получаемая из нормированной оценки с учетом характеристики масштаба $\hat{\sigma}$.

Набор нормированных интегрированных функций распределения $\{\hat{F}^{(n)}(z^*)\}$ формируется по следующей схеме.

Для заранее определенных одномерных параметрических законов распределения $y = f(x, \vec{\theta})$ непрерывной случайной величины $z \in [0; \infty)$ рассчитывались значения вектора параметров $\vec{\theta}^*$ таким образом, чтобы величина среднеквадратического отклонения σ^* равнялась единице. Для полученных значений $\vec{\theta}^*$ рассчитывалось значение математического ожидания ν_1^* . В ходе исследований в качестве $f(x, \vec{\theta})$ использовались следующие типы законов распределения: логнормальное, Хи-квадрат, F -распределения, распределение Эрланга, распределение Вейбулла.

Для полученных $\vec{\theta}^*$, соответствующих вышеречисленным типам законов распределения, рассчитывались значения интегральных функций распределения $F_3^{(n)}(z^*)$ в заранее выбранных точках z^* . Объединение $F_3^{(n)}(z^*)$ в результирующую таблицу $\{F_3^{(n)}(z^*)\}$ производилось в порядке возрастания ν_1^* .

Непараметрическая схема оценивания обладает лучшими аппроксимирующими воз-

можностями по сравнению с упоминавшейся выше моделью, получаемой на основе энтропийного подхода. Это обусловлено тем, что $\{F_3^{(n)}(z^*)\}$ содержит целый спектр параметрических моделей $f(x, \vec{\theta})$ в то время, как энтропийный подход во всех случаях предписывает использовать аналитическую зависимость одной и той же структуры. Особо следует подчеркнуть, что непараметрическая схема не требует выражения и проверки статистических гипотез и представляет собой формальную процедуру.

ПРИМЕР РЕШЕНИЯ ЗАДАЧИ КРАТКОСРОЧНОГО ПРОГНОЗИРОВАНИЯ УРОВНЕЙ ВОДЫ НА ПОСТАХ КОНТРОЛЯ

В качестве примера использования метода рассмотрим задачу краткосрочного прогнозирования паводковой ситуации по измерительным данным, зарегистрированным на одном из стационарных постов контроля на территории Республики Башкортостан.

Дано: архив, содержащий десять временных рядов, характеризующих уровни воды на посту контроля Глуховский. Значения уровней воды $\{y_i^{(j)}\}$ фиксировались в двадцати шести временных сечениях ($j = \overline{1; 10}, i = \overline{1; 26}$) (рис. 6). Задано значение одиннадцатого временного ряда $y_{14}^{(11)}$, не входящего в архив, отнесенное к четырнадцатому временному сечению.

Требуется: рассчитать прогнозные значения одиннадцатого временного ряда в q -х ($q=15, 16, \dots$) временных сечениях.

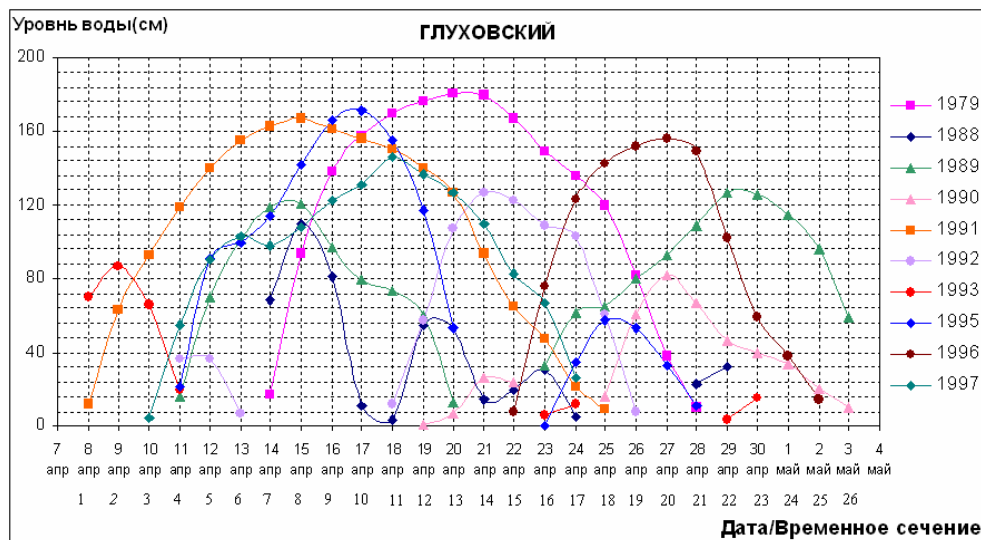


Рис. 6. Временные ряды уровней воды, зафиксированные на посту контроля

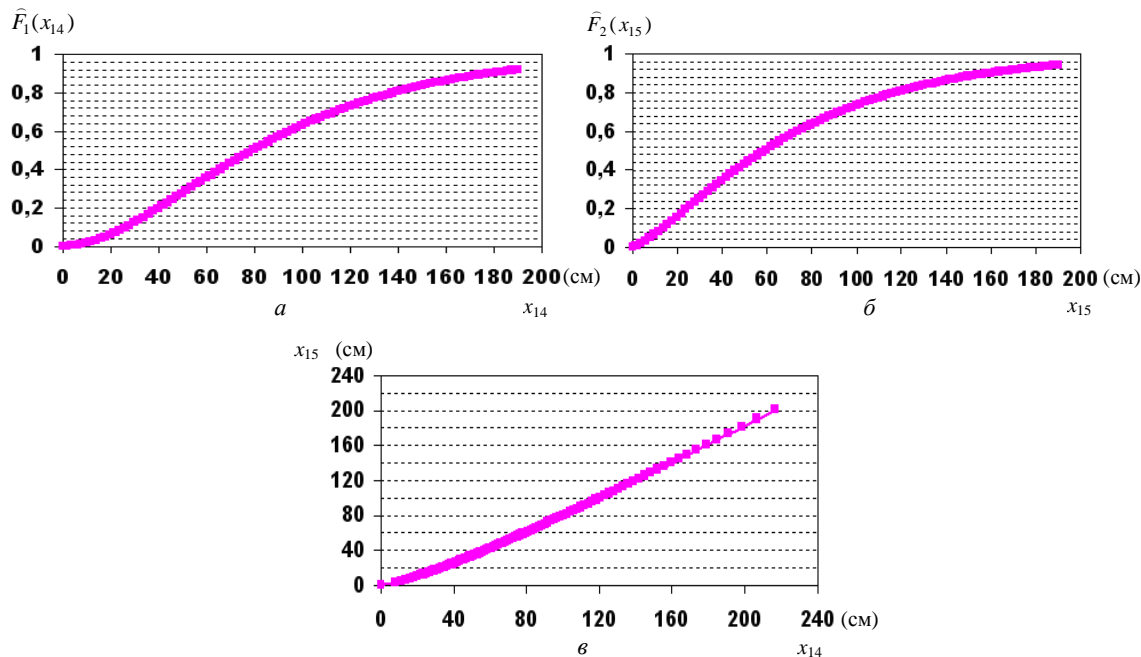


Рис. 7. Результаты обработки выборочных значений $\{x_{14}\}, \{x_{15}\}$

Решение: по зарегистрированным значениям j -го временного ряда ($j = \overline{1; 10}$) формируется выборка однородных случайных чисел $x_1^{(j)}, x_2^{(j)}, \dots, x_{N_j}^{(j)}$ по правилу: $x_i^{(j)} = y_i^{(j)} - y^{(кр)}$. Учитываются только те значения $x_i^{(j)}$, для которых $x_i^{(j)} \geq 0$. Здесь $y^{(кр)}$ — уровень воды, при котором происходит затопление поймы. Теоретически возможные значения $x_i^{(j)}$ в i -м временном сечении ($i = \overline{1; 26}$) составляют: $0 < x_i^{(j)} < \infty$. Значение $x_i^{(j)} = 0$ соответствует случаю $y_i^{(j)} = y^{(кр)}$. Объемы неотрицательных выборочных данных в i -м временном сечении $N_i (x_i^{(j)} \geq 0)$ удовлетворяют условию $N_i \leq 10$. Отметим, что числовые значения $N_i (i = \overline{1; 26})$ является различными.

Расчеты и построение непараметрических моделей велись относительно значения уровня воды, при котором происходит затопление поймы (для указанного поста выход на пойму происходит при уровне воды равном $y^{(кр)} = 550$ см).

Рассмотрим технологию расчета на примере двух временных сечений. В качестве k -го временного сечения выступает сечение с номером четырнадцать (21 апреля). Примем в качестве q -го временного сечения — временное сечение с номером пятнадцать (22 апреля). Значение уровня воды одиннадцатого временного ряда в четырнадцатом временном сечении (над уровнем $y^{(кр)}$) равно 62 см. Рассчитаем на основе сформированных вы-

борок случайных величин $\{x_{14}\}, \{x_{15}\}$ оценки математических ожиданий и среднеквадратических отклонений, которые для четырнадцатого и пятнадцатого временных сечений $M[x_{14}] = 93,8$ (см), $\sigma[x_{14}] = 63,7$ (см); $M[x_{15}] = 67,5$ (см), $\sigma[x_{15}] = 64,2$ (см).

На рис. 7,а приведена оценка закона распределения превышения уровней воды над $y^{(кр)}$ в четырнадцатом временном сечении $\hat{F}_1(x_{14})$; на рис. 7,б приведена оценка закона распределения $\hat{F}_2(x_{15})$ превышения уровней воды над $y^{(кр)}$ в пятнадцатом временном сечении; на рис. 7,в приведена непараметрическая модель, соответствующая выборочным данным.

По полученной непараметрической модели уровней воды и зарегистрированному в четырнадцатом временном сечении значению $x_{14}^{(11)}$, равного 62 см, рассчитывалось ожидаемое значение $x_{15}^{(11)}$. Оно составило 43 см.

Таким образом, ожидаемый уровень воды равен $y_{(ф)15}^{(11)} = 550 + 43 = 593$ см, фактическое значение уровня воды оказалось равным 629 (см).

В табл. 1 приведены результаты, полученные по аналогичной расчетной схеме для других временных сечений, а в табл. 2 — результаты, полученные по тем же исходным данным посредством метода линейной экстраполяции, а также характеристики точности полученных оценок.

Таблица 1

Результаты прогнозирования уровней воды с помощью непараметрического метода при известном значении уровня воды во временном сечении номер 14

Номер временного сечения	14	15	16	17	18	19	20
Фактические значения уровня воды, см	612	629	595	572	603	667	694
Прогнозные значения уровня воды во временных сечениях, см		593	587	593	600	606	609
Относительная погрешность, %		5	7	8	9	10	12
Абсолютная погрешность, см		24	27	36	42	57	84

Таблица 2

Результаты прогнозирования уровней воды с помощью экстраполяционного метода при известном значении уровня воды во временном сечении номер 14

Номер временного сечения	14	15	16	17	18	19	20
Фактические значения уровня воды, см	612	629	595	572	603	667	694
Прогнозные значения уровня воды во временных сечениях, см		675	738	801	864	927	990
Относительная погрешность, %		8	24	39	40	43	46
Абсолютная погрешность, см		54	143	229	261	275	320

ЗАКЛЮЧЕНИЕ

Предлагаемый подход — экстраполяционное прогнозирование паводковой ситуации позволяет:

1) полностью формализовать процедуру прогнозирования, что позволяет реализовать ее в виде программной компоненты в системе информационной поддержки предупреждения негативных последствий паводка;

2) повысить точность прогнозирования уровней воды на постах контроля в условиях малых по объему и низким по точности исходных данных.

СПИСОК ЛИТЕРАТУРЫ

1. **Минприродресурсы РБ.** Методические рекомендации председателям противопаводковых комиссий муниципальных образований Республики Башкортостан по выполнению комплекса противопаводковых мероприятий / утв. министром природ. ресурсов РБ перв. зам. предс. противопаводковой комис. РБ А. А. Фаухутдиновым. 2006. 43 с.
2. **Ямалов, И. У.** Опыт создания системы информационной поддержки контроля и управления состоянием территориальных систем в Республике Башкортостан / И. У. Ямалов, С. В. Павлов, В. Е. Гвоздев [и др.] // Башкирский экологический вестник. 2003. № 2 (13). С. 5–11.
3. **Пугачев, В. С.** Теория вероятностей и математическая статистика / В. С. Пугачев. М. : Физматлит, 2002. 496 с.
4. **Гвоздев, В. Е.** Статистические методы оценки качества и надежности промышленных изделий по результатам малого числа испытаний : учебно-методич. разработка / В. Е. Гвоздев, Ю. Е. Альпов. Уфа : УАИ, 1988. 44 с.
5. **Вентцель, Е. С.** Теория вероятностей : учеб. для вузов. 7-е изд., стер. / Е. С. Вентцель. М. : Высш. шк., 2001. 575 с.
6. **Гвоздев, В. Е.** Влияние границ области возможных значений контролируемых параметров на точность оценивания законов распределения / В. Е. Гвоздев, Ю. Е. Альпов // Вопросы проектирования информационных кибернетических систем : межвуз. сб. Уфа : УАИ, 1987. С. 20–24.

ОБ АВТОРАХ



Гвоздев Владимир Ефимович, зав. каф. автоматиз. проектир. информ. систем. Дипл. инж. электрон. тех-ки (УАИ, 1978). Д-р техн. наук по АСУ (УГАТУ, 2000). Иссл. в обл. АСУ, открытых информ. систем, прикл. статистики, теории надежности, контроля и управ. сост. окр. среды.

Колоденкова Анна Евгеньевна, ассистент той же каф. Дипл. инж.-системотехн. по АСОИУ (УГАТУ, 2004). Иссл. в обл. прикладн. статистики, системн. анализа, контроля и управ-я состоянием территор. объектов, информац. систем, технологии системн. модел-я.