

Интеллектуальные системы управления, анализ данных

© 2023 г. Ю.А. ДУБНОВ (yury.dubnov@phystech.edu)

(Федеральный исследовательский центр

“Информатика и управление” РАН, Москва;

Национальный исследовательский университет

“Высшая школа экономики”, Москва),

А.Ю. ПОПКОВ, канд. техн. наук (aropkov@isa.ru)

(Федеральный исследовательский центр

“Информатика и управление” РАН, Москва),

В.Ю. ПОЛИЩУК, канд. техн. наук (liquid_metal@mail.ru)

(Институт мониторинга климатических и экологических систем, Томск),

Е.С. СОКОЛ (sokoles@uriit.ru)

(Югорский НИИ Информационных технологий, Ханты-Мансийск),

А.В. МЕЛЬНИКОВ, д-р техн. наук (melnikovav@uriit.ru)

(Югорский НИИ Информационных технологий, Ханты-Мансийск),

Ю.М. ПОЛИЩУК, д-р техн. наук (yupolishchuk@gmail.com)

(Югорский НИИ Информационных технологий, Ханты-Мансийск),

Ю.С. ПОПКОВ, д-р техн. наук (popkov@isa.ru)

(Федеральный исследовательский центр

“Информатика и управление” РАН, Москва)

АЛГОРИТМЫ РАНДОМИЗИРОВАННОГО МАШИННОГО ОБУЧЕНИЯ ДЛЯ ПРОГНОЗИРОВАНИЯ ЭВОЛЮЦИИ ПЛОЩАДИ ТЕРМОКАРСТОВЫХ ОЗЕР В ЗОНАХ ВЕЧНОЙ МЕРЗЛОТЫ¹

Рандомизированное машинное обучение ориентировано на задачи, сопровождаемые значительной неопределенностью в данных и моделях. Алгоритмы машинного обучения формулируются в терминах функциональной задачи энтропийно-линейного программирования. Рассматривается методика их адаптации к задачам прогнозирования на примере временной эволюции площади термокарстовых озер в зонах вечной мерзлоты, которые являются генераторами метана — одного из парниковых газов, влияющих на изменения климата. Предлагаются процедуры рандомизированного машинного обучения, использующие модели динамической регрессии со случайными параметрами, и ретроспективные данные климатических параметров и дистанционного зондирования земной поверхности. Развивается алгоритм рандомизированного машинного обучения, позволяющий вычислять оценки функций плотности распределения вероятностей параметров модели и измерительных шумов. Рандомизированное прогнозирование реализовано в виде алгоритмов transforma-

¹ Работа выполнена при поддержке Российского научного фонда (проект 22-11-20023).

ции оптимальных распределений в соответствующие им случайные последовательности (алгоритмы сэмплирования). Развиваемые процедуры и технологии рандомизированного прогнозирования применены для обучения, тестирования и прогнозирования эволюции площади термокарстовых озер Западной Сибири.

Ключевые слова: термокарстовые озера, дистанционное зондирование, информационная энтропия, балансовые уравнения, динамическая регрессия, оптимизация, ляпуновская задача, сэмплирование, рандомизированное прогнозирование, рандомизированное машинное обучение.

DOI: 10.31857/S0005231023010051, **EDN:** LUKHYY

1. Введение

Проблемам, технологиям и алгоритмам машинного обучения посвящено огромное количество работ. Общее представление могут дать монографии [1–3]. Рандомизированное машинное обучение (РМО) фокусируется на восстановлении параметров зависимостей *в условиях неопределенности* в данных (неполнота, ошибки с неизвестными характеристиками, различные временные шкалы и др.) и моделях (недостаточность знаний о процессах, структурная неоднозначность, определяющие параметры и др.) [4]. В отличие от машинного обучения рандомизированная версия основана на оценках функций плотностей распределения вероятностей (ПРВ) параметров модели и измерительных шумов, соответствующих максимальной неопределенности в терминах информационной энтропии.

Термокарстовые озера являются объектом с высоким уровнем неопределенности. Заметим, что процессы образования и эволюции площади термокарстовых озер изучены недостаточно [5, 6] и исторические данные о них, в особенности получаемые со спутников, сопровождаются значительными погрешностями [7–9].

В данной статье предлагается технология рандомизированного прогнозирования эволюции площади термокарстовых озер, суть которой состоит в генерации оптимизированных ансамблей прогнозных траекторий путем сэмплирования энтропийно-оптимальных функций ПРВ параметров рандомизированных динамических моделей. Оптимальные характеристики указанных моделей определяются с использованием реальных исторических данных на основе оценок условно максимальной информационной энтропии.

2. Общая структура процедуры энтропийно-рандомизированного моделирования и прогнозирования

Рандомизация как средство придания искусственных, целесообразно организованных случайных свойств событиям, индикаторам, методам, не являющихся по своей природе случайными, является довольно распространенным приемом достижения позитивного эффекта. Тому существует много примеров

в самых разных областях науки, управления, экономики: рандомизированные численные методы оптимизации [10, 11], смешанные (случайные) стратегии в биржевой торговле [12], рандомизированное прогнозирование динамики населения [13], вибрационное управление технологическими процессами [14]. При этом рандомизация предполагает придание неслучайным объектам искусственных стохастических свойств с оптимальными, в выбранном смысле, вероятностными характеристиками. Вопрос о выборе количественных характеристик оптимальности всегда оказывается дискуссионным и неоднозначным. Он требует аргументов, которые бы как-то отражали важную специфику рандомизированного объекта. В частности, принципиальной особенностью процедур моделирования и прогнозирования является сопровождающая их неопределенность в используемых данных, предсказывающих моделях, методах генерации прогнозов и др.

Здесь будет использоваться в качестве характеристики неопределенности *информационная энтропия* [15]. В [16–18] показано, что энтропия согласно первому закону термодинамики является естественным функционалом, характеризующим процессы всеобщей эволюции. Согласно второму закону термодинамики ее максимизация определяет наилучшее состояние эволюционного процесса при наихудших воздействиях на него (максимальной неопределенности). Стоит упомянуть еще одно качество информационной энтропии, связанное с измерительными и иными погрешностями, являющимися важными характеристиками данных. Использование информационной энтропии для учета влияния указанных погрешностей позволяет оценить вероятностные характеристики шумов, оказывающих наихудшее воздействие на процедуры прогнозирования [19].

2.1. Рандомизированное моделирование

В основе рандомизированного моделирования лежат математические модели «вход-выход» со случайными параметрами. На данном этапе технологии формируются структуры базовой модели (*РПМ-Б*), ориентированной на имеющиеся исторические данные о входе и выходе, и вспомогательной модели такого же класса (*РПМ-В*), предназначенной для воспроизведения входа базовой модели, с использованием «подходящего» входного процесса для *РПМ-Б*. В качестве такового применяется случайная последовательность с оптимизированными свойствами.

Рассмотрим рандомизированную параметризованную базовую модель (*РПМ-Б*) и согласованные с ней массивы исторических данных. *РПМ-Б* преобразует массив исторических входных данных $X = [\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(s)}]$, где $\mathbf{x}^{(j)} \in \mathbb{R}^n$, в модельный выход, характеризуемый матрицей $\hat{Z} = [\hat{\mathbf{z}}^{(1)}, \dots, \hat{\mathbf{z}}^{(s)}]$, где $\mathbf{z}^{(j)} \in \mathbb{R}^m$.

В общем случае это преобразование предполагается динамическим, т.е. модельный выход, наблюдаемый в момент времени j , зависит от входа, наблюдаемого на некотором *историческом* интервале $j - \varrho, \dots, j$, т.е. от мат-

рицы $X_\varrho^{(j)} = [\mathbf{x}^{(j-\varrho)}, \dots, \mathbf{x}^{(j)}]$. Математическим образом этой связи является векторный функционал $\hat{\Omega}(X_\varrho^{(j)}, \mathbf{a} | P(\mathbf{a}))$ со случайными параметрами $\mathbf{a} \in R^d$ интервального типа

$$(2.1) \quad \mathbf{a} \in \mathcal{A} = [\mathbf{a}^-, \mathbf{a}^+].$$

Вероятностные свойства параметров характеризуются функцией ПРВ $P(\mathbf{a})$, которая предполагается непрерывно-дифференцируемой.

Выход *РПМ-Б* в j -й момент времени (момент измерения) представляет собой ансамбль $\hat{\mathbb{Z}}^{(j)}(P(\mathbf{a}))$ случайных векторов

$$(2.2) \quad \hat{\mathbf{z}}^{(j)}(\mathbf{a}) = \hat{\Omega}(X_\varrho^{(j)}, \mathbf{a} | P(\mathbf{a})), \quad j = \overline{1, s}.$$

Наблюдаемый выход *РПМ-Б* представим в виде

$$(2.3) \quad \mathbf{v}^{(j)}(\mathbf{a}, \xi^{(j)}) = \hat{\mathbf{z}}^{(j)}(\mathbf{a}) + \xi^{(j)}, \quad j = \overline{1, s},$$

где добавлен измерительный шум $\xi^{(j)} \in R^m$ интервального типа:

$$(2.4) \quad \xi^{(j)} \in \Xi_j = [\xi_-^{(j)}, \xi_+^{(j)}], \quad j = \overline{1, s}$$

с непрерывно-дифференцируемыми функциями ПРВ $Q_j(\xi^{(j)})$, $j = \overline{1, s}$, в соответствии с которыми генерируется для каждого момента измерения выхода объекта ансамбль $\mathbb{K}^{(j)}(Q_j(\xi^{(j)}))$.

Случайные векторы (2.3) образуют ансамбль $\mathbb{V}^{(j)}(P(\mathbf{a}), Q_j(\xi^{(j)}))$, математическое ожидание которого имеет вид

$$(2.5) \quad \begin{aligned} \mathcal{M} \left\{ \mathbf{v}^{(j)}(\mathbf{a}, \xi^{(j)}) \right\} &= \int_{\mathcal{A}} \hat{\mathbf{z}}^{(j)}(\mathbf{a}, P(\mathbf{a})) d\mathbf{a} + \int_{\Xi_j} Q_j(\xi^{(j)}) \xi^{(j)} d\xi^{(j)} = \\ &= \varphi^{(j)} \left[P(\mathbf{a}), Q_j(\xi^{(j)}) \right], \quad j = \overline{1, s}. \end{aligned}$$

2.2. Обучение РПМ

Обучение РПМ состоит в *оценивании ПРВ* ее параметров и шумов измерений по имеющимся данным. Он реализуется с помощью алгоритмов рандомизированного машинного обучения (*РМО-А*) [19, 20].

Для *РПМ-Б* он имеет вид

$$(2.6) \quad [P^*(\mathbf{a}), Q^*(\xi)] = \arg \max_{P(\mathbf{a}), Q(\xi)} \mathcal{H}[P(\mathbf{a}), Q(\xi)]$$

на множестве нормированных функций $P^*(\mathbf{a}), Q(\xi)$, для которых выполняются условия эмпирических балансов средних (2.5):

$$(2.7) \quad \varphi_j \left[P(\mathbf{a}), Q_j(\xi^{(j)}) \right] = \mathbf{y}^{(j)}, \quad j = \overline{1, s},$$

где $\mathbf{y}^{(j)} \in R^m$ — вектор реальных измерений выхода объекта, $\xi = \{\xi^{(1)}, \dots, \xi^{(s)}\}$.

Качество оценивания указанных *ПРВ* характеризуется энтропийным функционалом

$$(2.8) \quad \mathcal{H}[P(\mathbf{a}), Q(\bar{\xi})] = - \int_{\mathcal{A}} P(\mathbf{a}) \ln P(\mathbf{a}) d\mathbf{a} - \sum_{j=1}^s \int_{\Xi_j} Q_j(\xi^{(j)}) \ln Q_j(\xi^{(j)}) d\xi^{(j)}.$$

Задача (2.6), (2.7) относится к классу функциональных энтропийно-линейных задач ляпуновского типа [21], которые имеют аналитическое решение, получаемое с помощью множителей Лагранжа $\Theta = [\theta^j, j = \overline{1, s}]$ (векторы $\theta^j \in R^m$):

$$(2.9) \quad P^*(\mathbf{a}) = \frac{\exp \left(- \sum_{j=1}^s \langle \theta^{(j)}, \hat{\mathbf{z}}^{(j)}(\mathbf{a}) \rangle \right)}{\mathcal{P}(\Theta)},$$

$$Q_j^*(\xi^{(j)}) = \frac{\exp(-\langle \theta^{(j)}, \xi^{(j)} \rangle)}{Q_j(\theta^{(j)})}, \quad j = \overline{1, s}; \quad Q(\xi) = \prod_{j=1}^s Q_j^*(\xi^{(j)}).$$

В знаменателях этих выражений стоят нормировочные константы

$$(2.10) \quad \mathcal{P}(\Theta) = \int_{\mathcal{A}} \exp \left(- \sum_{j=1}^s \langle \theta^{(j)}, \hat{\mathbf{z}}^{(j)}(\mathbf{a}) \rangle \right) d\mathbf{a},$$

$$Q_j(\theta^{(j)}) = \int_{\Xi_j} \exp \left(- \langle \theta^{(j)}, \xi^{(j)} \rangle \right) d\xi^{(j)}, \quad j = \overline{1, s}.$$

Оптимальные ПРВ и нормировочные константы параметризованы множителями Лагранжа, которые определяются решением следующих балансовых уравнений:

$$(2.11) \quad \frac{\mathcal{U}_j(\Theta)}{\mathcal{P}(\Theta)} + \frac{\mathcal{T}_j(\theta^{(j)})}{Q_j(\theta^{(j)})} = \mathbf{y}^{(j)}, \quad j = \overline{1, s},$$

где

$$(2.12) \quad \mathcal{U}_j(\Theta) = \int_{\mathcal{A}} \hat{\mathbf{z}}^{(j)}(\mathbf{a}) \exp \left(- \sum_{j=1}^s \langle \theta^{(j)}, \hat{\mathbf{z}}^{(j)}(\mathbf{a}) \rangle \right) d\mathbf{a},$$

$$\mathcal{T}_j(\theta^{(j)}) = \int_{\Xi_j} \xi^{(j)} \exp \left(- \langle \theta^{(j)}, \xi^{(j)} \rangle \right) d\xi^{(j)}, \quad j = \overline{1, s}.$$

2.3. Тестирование обученных РПМ

Тестирование оптимизированных РПМ осуществляется с использованием тестовых последовательностей исторических данных путем сэмплирования энтропийно-оптимизированных ПРВ параметров РПМ и измерительных шумов. Сэмплирование предполагает трансформацию функции ПРВ в соответствующую последовательность случайных векторов.

Общий метод генерации последовательностей случайных векторов с заданной функцией ПРВ изложен в [22]. В результате с помощью метода Монте-Карло генерируется ансамбль случайных траекторий наблюдаемого выхода РПМ (с учетом наихудших в энтропийных терминах измерительных шумов).

В процедурах тестирования используются эмпирические траектории средних по ансамблю (*mean*) (M — количество траекторий в ансамбле)

$$(2.13) \quad \bar{\mathbf{v}}[k] = \frac{1}{M} \sum_{i=1}^M \hat{\mathbf{v}}^{(i)}[k], \quad k = \overline{0, N}.$$

Качество обученных РПМ характеризуется абсолютной ошибкой

$$(2.14) \quad \Delta = \sqrt{\sum_{k=1}^M (\bar{\mathbf{v}}[k] - \mathbf{v}^r[k])^2}$$

и (или) относительной ошибкой

$$(2.15) \quad \delta = \frac{\sum_{k=1}^M (\bar{\mathbf{v}}[k] - \mathbf{v}^r[k])^2}{\sqrt{\sum_{k=1}^M \bar{\mathbf{v}}^2[k]} + \sqrt{\sum_{k=1}^M (\mathbf{v}^r[k])^2}},$$

где $\mathbf{v}^r[k]$ — исторические тестовые данные.

2.4. Рандомизированное прогнозирование

Применение динамических моделей «вход-выход» для прогнозирования требует некоторой их модификации. Дело в том, что в этом случае необходимо иметь прогноз входного процесса. Для формирования такого прогноза можно использовать идею энтропийно-рандомизированного машинного обучения.

Рассмотрим рандомизированную параметризованную вспомогательную модель (РПМ-В). В отличие от РПМ-Б входом в этой модели является случайная матрица $\Gamma = [\gamma^{(1)}, \dots, \gamma^{(s)}]$, где случайные векторы $\gamma^{(j)} \in R^n$ интервального типа, т.е.

$$(2.16) \quad \gamma^{(j)} \in \mathcal{G}_j = [\gamma_-^{(j)}, \gamma_+^{(j)}], \quad \Gamma \in \mathcal{G} = \bigcup_{j=1}^s \mathcal{G}_j.$$

РПМ-В преобразует случайную матрицу Γ в случайную матрицу $X = [\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(s)}]$. Математическим образом этой связи, так же как для РПМ-Б, является векторный функционал $\hat{\Psi}(\Gamma_{\varrho}^{(j)}, \mathbf{b} | W(\Gamma_{\varrho}^{(j)}, \mathbf{b}))$, в котором матрица $\Gamma_{\varrho}(j) = [\gamma^{(j-\varrho)}, \dots, \gamma^{(j)}]$ и параметры $\mathbf{b} \in R^p$ интервального типа:

$$(2.17) \quad \mathbf{b} \in \mathcal{B} = [\mathbf{b}^-, \mathbf{b}^+].$$

Вероятностные свойства функционала $\hat{\Psi}$ будем характеризовать функцией плотности совместного распределения вероятностей параметров $\mathbf{b}$ и матрицы $\Gamma_{\varrho}^{(j)}$ — ПРВ $W(\Gamma_{\varrho}^{(j)}, \mathbf{b})$, которая предполагается непрерывно-дифференцируемой.

Выход РПМ-В в j -й момент времени (момент измерения) представляет собой ансамбль $\hat{\mathbf{X}}^{(j)}(W(\Gamma_{\varrho}(j), \mathbf{b}))$ случайных векторов

$$(2.18) \quad \hat{\mathbf{x}}^{(j)}(\Gamma_{\varrho}^{(j)}, \mathbf{b}) = \hat{\Psi}(\Gamma_{\varrho}^{(j)}, \mathbf{b} | W(\Gamma_{\varrho}^{(j)}, \mathbf{b})), \quad j = \overline{1, s}.$$

Наблюдаемый выход РПМ-В представим в виде

$$(2.19) \quad \mathbf{f}^{(j)}(\Gamma_{\varrho}^{(j)}, \mathbf{b}, \eta^{(j)}) = \hat{\mathbf{x}}^{(j)}(\Gamma_{\varrho}^{(j)}, \mathbf{b}) + \eta^{(j)}, \quad j = \overline{1, s},$$

где добавлен измерительный шум $\eta^{(j)} \in R^m$ интервального типа

$$(2.20) \quad \eta^{(j)} \in \mathcal{E}_j = [\eta_-^{(j)}, \eta_+^{(j)}], \quad j = \overline{1, s}$$

с непрерывно-дифференцируемыми функциями ПРВ $E_j(\eta^{(j)})$, $j = \overline{1, s}$, в соответствии с которыми генерируется для каждого момента измерения выхода объекта ансамбль $\mathbb{E}^{(j)}(E_j(\eta^{(j)}))$.

Случайные векторы (2.19) образуют ансамбль $\mathbb{F}^{(j)}(W(\Gamma_{\varrho}^{(j)}, \mathbf{b}), E_j(\eta^{(j)}))$, математическое ожидание которого имеет вид:

$$(2.21) \quad \begin{aligned} \mathcal{M} \left\{ \mathbf{f}^{(j)}(\Gamma_{\varrho}^{(j)}, \mathbf{b}, \eta^{(j)}) \right\} &= \int_{\mathcal{B} \cap \mathcal{G}} W(\Gamma_{\varrho}^{(j)}, \mathbf{b}) \hat{\mathbf{x}}^{(j)}(\Gamma_{\varrho}^{(j)}, \mathbf{b}) d\Gamma_{\varrho}^{(j)} d\mathbf{b} + \\ &+ \int_{\mathcal{E}_j} E_j(\eta^{(j)}) \eta^{(j)} d\eta^{(j)} = \psi^{(j)} \left[W(\Gamma_{\varrho}^{(j)}, \mathbf{b}), E_j(\eta^{(j)}) \right], \quad j = \overline{1, s}. \end{aligned}$$

Для РПМ-В алгоритм рандомизированного машинного обучения имеет вид

$$(2.22) \quad \left[W^*(\Gamma_{\varrho}^{(j)}, \mathbf{b}), E^*(\eta) \right] = \arg \max_{W(\Gamma_{\varrho}^{(j)}, \mathbf{b}), E(\eta)} \mathcal{H} [W(\Gamma_{\varrho}, \mathbf{b}), E(\eta)],$$

при условии, что функции $W(\Gamma_{\varrho}^{(j)}, \mathbf{b}), E(\eta)$ нормированы и выполняются эмпирические балансы

$$(2.23) \quad \psi_j \left[W(\Gamma_{\varrho}^{(j)}, \mathbf{b}), E_j(\eta^{(j)}) \right] = \mathbf{y}^{(j)}, \quad j = \overline{1, s}.$$

В (2.22) энтропийный функционал имеет вид:

$$(2.24) \quad \begin{aligned} \mathcal{H}[W(\Gamma_{\varrho}, \mathbf{b}), E(\eta)] = & - \sum_{j=1}^s \int_{\mathcal{B} \cap \mathcal{G}_j} W(\Gamma_{\varrho}^{(j)}, \mathbf{b}) \ln W(\Gamma_{\varrho}^{(j)}, \mathbf{b}) d\Gamma_{\varrho}^{(j)} d\mathbf{b} - \\ & - \sum_{j=1}^s \int_{\mathcal{E}_j} E_j(\eta^{(j)}) \ln E_j(\eta^{(j)}) d\eta^{(j)}. \end{aligned}$$

Обозначим через $\Lambda \{\lambda^{(1)}, \dots, \lambda^{(s)}\}$, $\lambda^{(j)} \in R^m$ множители Лагранжа для задачи (2.22, 2.23). Тогда ее решение можно представить в виде:

$$(2.25) \quad \begin{aligned} W^*(\Gamma_{\varrho}^{(j)}, \mathbf{b}) = & \frac{\exp\left(-\langle \lambda^{(j)}, \hat{\mathbf{x}}^{(j)}(\Gamma_{\varrho}^{(j)}, \mathbf{b}) \rangle\right)}{\mathcal{W}_j(\lambda^{(j)})}, \quad W^*(\Gamma_{\varrho}, \mathbf{b}) = \prod_{j=1}^s W^*(\Gamma_{\varrho}^{(j)}, \mathbf{b}), \\ E_j^*(\eta^{(j)}) = & \frac{\exp\left(-\langle \lambda^{(j)}, \eta^{(j)} \rangle\right)}{\mathcal{E}_j(\lambda^{(j)})}, \quad j = \overline{1, s}; \quad E(\eta) = \prod_{j=1}^s E_j^*(\eta^{(j)}). \end{aligned}$$

В знаменателях этих выражений стоят нормировочные константы

$$(2.26) \quad \begin{aligned} \mathcal{W}_j(\lambda^{(j)}) = & \int_{\mathcal{B} \cap \mathcal{G}} \exp\left(-\langle \lambda^{(j)}, \hat{\mathbf{x}}^{(j)}(\Gamma_{\varrho}^{(j)}, \mathbf{b}) \rangle\right) d\Gamma_{\varrho}^{(j)} d\mathbf{b}, \\ \mathcal{E}_j(\lambda^{(j)}) = & \int_{\mathcal{E}_j} \exp\left(-\langle \lambda^{(j)}, \eta^{(j)} \rangle\right) d\eta^{(j)}, \quad j = \overline{1, s}. \end{aligned}$$

Из этих равенств следует, что вероятностные свойства входного вспомогательного шума и параметров модели взаимосвязаны для нелинейного функционала $\hat{\Psi}$.

Из (2.7) видно, что оптимальные ПРВ параметризованы множителями Лагранжа Λ , которые определяются решением следующих балансовых уравнений:

$$(2.27) \quad \frac{\mathcal{N}_j(\lambda^{(j)})}{\mathcal{W}_j(\lambda^{(j)})} + \frac{\mathcal{S}_j(\lambda^{(j)})}{\mathcal{E}_j(\lambda^{(j)})} = \mathbf{y}^{(j)}, \quad j = \overline{1, s},$$

где

$$(2.28) \quad \begin{aligned} \mathcal{N}_j(\lambda^{(j)}) = & \int_{\mathcal{B} \cap \mathcal{G}_j} \hat{\mathbf{z}}(\Gamma_{\varrho}^{(j)}, \mathbf{b}) \exp\left(-\langle \theta^{(j)}, \hat{\mathbf{z}}(\Gamma_{\varrho}^{(j)}, \mathbf{b}) \rangle\right) d\Gamma_{\varrho}^{(j)} d\mathbf{b}, \\ \mathcal{S}_j(\lambda^{(j)}) = & \int_{\mathcal{E}_j} \eta^{(j)} \exp\left(-\langle \lambda^{(j)}, \eta^{(j)} \rangle\right) d\eta^{(j)}, \quad j = \overline{1, s}. \end{aligned}$$

После того как определена функция ПРВ параметров модели и входной последовательности производится генерация ансамблей прогнозных траекторий с помощью их сэмплирования [22].

Сэмплирование предполагает трансформацию функции ПРВ в соответствующую последовательность случайных векторов. В процедурах рандомизированного прогнозирования используются эмпирические вероятностные и числовые характеристики ансамблей:

- эмпирические функции плотности распределения вероятностей (*эПРВ*) $\mathcal{P}_k(\hat{\mathbf{v}}[k])$;
- эмпирические функции распределения вероятностей (*эРВ*) $\mathfrak{P}_k(\hat{\mathbf{v}}[k])$;
- траектории максимумов *эПРВ* параметров и шумов (*max-pn*)

$$(2.29) \quad \begin{aligned} \hat{\mathbf{y}}[k] &= \mathbb{B}(\mathbf{a}^*, X_{(k-\rho)}), \quad \mathbf{a}^* = \arg \max \mathbf{a} P^*(\mathbf{a}), \\ \xi[k] &= \arg \max \xi Q_k^*(\xi[k]), \\ \hat{\mathbf{v}}[k] &= \hat{\mathbf{y}}[k] + \xi[k], \quad k = \overline{0, N}; \end{aligned}$$

- траектории максимумов *эПРВ* наблюдаемых выходов РПМ для $k = \overline{0, N}$ (*max-ePDF*)

$$(2.30) \quad \check{\mathbf{v}}[k] = \arg \max \mathcal{P}_k(\hat{\mathbf{v}}[k]), \quad k = \overline{0, N};$$

- траектории средних по ансамблю (*mean*) (M — количество траекторий в ансамбле)

$$(2.31) \quad \bar{\mathbf{v}}[k] = \frac{1}{M} \sum_{i=1}^M \hat{\mathbf{v}}^{(i)}[k], \quad k = \overline{0, N};$$

- медианные траектории (*med*) (M — количество траекторий в ансамбле)

$$(2.32) \quad \hat{\mathbf{v}}^{(i^*)}[k] \Rightarrow \sum_{i=1}^{i^*} \mathfrak{P}_k(\hat{\mathbf{v}}^{(i)}[k]) = \sum_{i=i^*+1}^M \mathfrak{P}_k(\hat{\mathbf{v}}^{(i)}[k]).$$

Важными характеристиками ансамблей являются так называемые *доверительные* множества, в которых с определенным отклонением или с определенной вероятностью находятся траектории ансамбля. Наиболее информативными являются:

- дисперсионная трубка $\mathcal{D}(\bar{\mathbf{v}}[k])$, $k = \overline{0, N}$

$$(2.33) \quad d(\bar{v}_j[k]) = \sqrt{\frac{1}{M-1} \sum_{i=1}^M \left(\hat{v}_j^{(i)}[k] - \bar{v}_j[k] \right)^2}, \quad j = \overline{1, m},$$

$$\mathcal{D}(\bar{\mathbf{v}}[k]) = \left\{ \hat{\mathbf{v}}[k] : \bar{v}_j[k] - d(\hat{v}_j[k]) \leq \hat{v}_j[k] \leq \bar{v}_j[k] + d(\hat{v}_j[k]), \right. \\ \left. j = \overline{1, m}; k = \overline{0, N} \right\};$$

- интерквартильное множество $\mathcal{I}_{(\kappa_1, \kappa_2)}(\hat{\mathbf{v}}[k])$, $k = \overline{0, N}$

$$(2.34) \quad \mathcal{I}_{(\kappa_1, \kappa_2)}(\hat{\mathbf{v}}[k]) = \mathcal{I}_{\kappa_2}(\hat{\mathbf{v}}[k]) \setminus \mathcal{I}_{\kappa_1}(\hat{\mathbf{v}}[k]), \quad \kappa_1 < \kappa_2 < 1,$$

$$\mathcal{I}_{\kappa_2}(\hat{\mathbf{v}}[k]) = \left\{ \hat{\mathbf{v}}[k] : \hat{v}_j[k] < \hat{v}_j^{\kappa_2}[k] \right\}, \quad \hat{v}_j^{\kappa_2}[k] \Rightarrow \mathfrak{P}_k(\hat{v}_j[k]) = \kappa_2,$$

$$\mathcal{I}_{\kappa_1}(\hat{\mathbf{v}}[k]) = \left\{ \hat{\mathbf{v}}[k] : \hat{v}_j[k] < \hat{v}_j^{\kappa_1}[k] \right\}, \quad \hat{v}_j^{\kappa_1}[k] \Rightarrow \mathfrak{P}_k(\hat{v}_j[k]) = \kappa_1,$$

$$(2.35) \quad k = \overline{0, N}, \quad j = \overline{1, m}.$$

Указанные вероятностные и числовые характеристики энтропийно-оптимальных ансамблей используются для характеристики процедур рандомизированного моделирования и прогнозирования площади термокарстовых озёр.

3. Структуры рандомизированных моделей состояния термокарстовых озёр

Важным этапом технологии рандомизированного прогнозирования является формирование рандомизированных моделей эволюции площади $S[n]$ термокарстовых озёр и влияющих на нее климатических параметров: среднегодовой температуры $T[n]$ и средней суммы осадков $R[n]$. Для моделирования указанных переменных используются математические модели линейной динамической регрессии [23], но *со случайными параметрами (ЛДРР)*.

Феноменологические основы моделирования исследуемого объекта построены на исследованиях, изложенных в [8]. В ней показана зависимость площади озёр от двух климатических факторов (среднегодовой температуры и среднегодовой суммы осадков), а также наличие временной «памяти» прошлых состояний объекта.

Согласно феноменологии исследуемого объекта структурная схема ЛДРР имеет вид, изображенный на рис. 1. Она состоит из рандомизированных моделей площади ЛДРР-П, температуры ЛДРР-Т и осадков ЛДРР-О. Измерительные ошибки в данных имитируются шумами ξ , η , ζ соответственно.

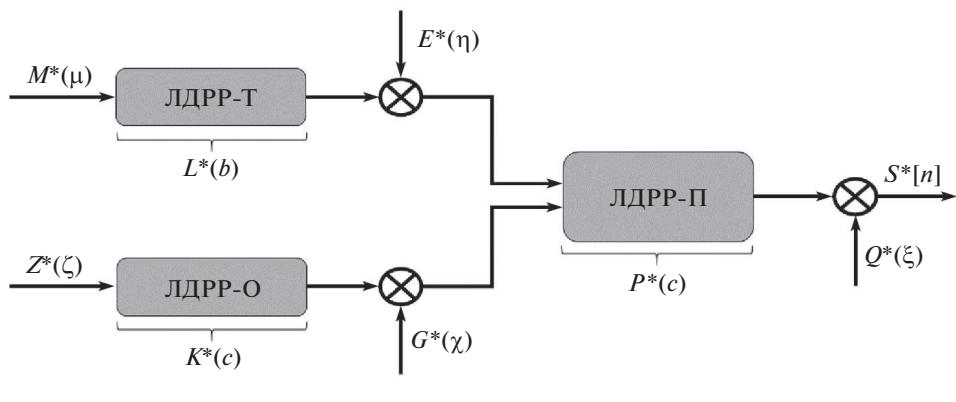


Рис. 1. Блок-схема модели ЛДРР.

Временная эволюция площади $S[n]$ описывается следующим уравнением динамической рандомизированной регрессии с двумя влияющими факторами (среднегодовой температурой $T[n]$ и годовой суммой осадков $R[n]$):

$$(3.1) \quad S[n] = a_0 + \sum_{k=1}^p a_k S[n-k] + a_{(p+1)} T[n] + a_{(p+2)} R[n],$$

где p — глубина «памяти» модели. Случайные параметры интервального типа

$$(3.2) \quad a_k \in \mathcal{A}_k = [a^-_k, a^+_k], \quad k = \overline{0, (p+2)}; \quad \mathbf{a} = \{a_0, \dots, a_{p+2}\} \in \mathcal{A} = \bigcup_{k=0}^{p+2} \mathcal{A}_k$$

характеризуются функцией ПРВ $P(\mathbf{a})$.

Наблюдаемый выход модели, искаженный измерительным $\xi[n]$ шумом, имеет вид

$$(3.3) \quad v[n] = S[n] + \xi[n],$$

где

$$(3.4) \quad \xi[n] \in \Xi_n = [\xi^-[n], \xi^+[n]].$$

Вероятностные свойства характеризуются функциями ПРВ $Q_n(\xi[n])$.

Рандомизированная модель температуры (ЛДРР-Т) $T[n]$ и ее наблюдаемый выход $t[n]$ имеют вид

$$(3.5) \quad T[n] = b_0 + \sum_{k=1}^p b_k T[n-k] + \mu[n], \quad t[n] = T[n] + \eta[n].$$

Параметры этой модели случайные, интервальные

$$(3.6) \quad b_k \in \mathcal{B}_k = [b^-_k, b^+_k], \quad k = \overline{0, p}, \quad \mathbf{b} = \{b_0, \dots, b_p\}, \quad \mathbf{b} \in \mathcal{B} = \bigcup_{k=0}^p \mathcal{B}_k.$$

Входная случайная последовательность также интервального типа:

$$(3.7) \quad \mu[n] \in \mathcal{M}_n = [\mu^-[n], \mu^+[n]], \quad n = \overline{0, N};$$

$$\mu = \{\mu[0], \dots, \mu[N]\}, \quad \mu \in \mathcal{M} = \bigcup_{n=0}^N \mathcal{M}_n.$$

Случайные параметры модели и входная случайная последовательность характеризуются совместной функцией ПРВ $W(\mathbf{b}, \mu)$.

Измерительный шум в (3.5) $\eta[n] \in \mathcal{E}_n = [\eta^-[n], \eta^+[n]]$ — интервальный, с функциями ПРВ $E_n(\eta[n])$.

Модель осадков (ЛДРР-О) $R[n]$ и ее наблюдаемый выход $r[n]$ имеют вид:

$$(3.8) \quad R[n] = c_0 + \sum_{k=1}^p c_k R[n-k] + \zeta[n], \quad r[n] = R[n] + \chi[n].$$

Параметры этой модели случайные, независимые, интервальные

$$(3.9) \quad c_k \in \mathcal{C}_k = [c^-_k, c^+_k], \quad k = \overline{0, p}; \quad \mathbf{c} = \{c_0, \dots, c_p\}, \quad \mathbf{c} \in \mathcal{C} = \bigcup_{k=0}^p \mathcal{C}_k$$

и случайная последовательность

$$(3.10) \quad \zeta[n] \in \mathcal{Z}_n = [\zeta^-[n], \zeta^+[n]], \quad \zeta = \{\zeta[0], \dots, \zeta[N]\}, \quad \zeta \in \mathcal{Z} = \bigcup_{n=0}^N \mathcal{Z}_n$$

характеризуются совместной функцией ПРВ $F(\mathbf{c}, \zeta)$. Измерительный шум $\chi[n] \in \mathcal{G}_n = [\chi^-[n], \chi^+[n]]$ с функцией ПРВ $G_n(\chi[n])$.

Следует заметить, что на этапе обучения модели ЛДРР-П имеются данные о площади, температуре и осадкам. Но на этапе прогнозирования данные о температуре и осадкам отсутствуют. Модели ЛДРР-Т и ЛДРР-О предназначены для прогнозирования среднегодовой температуры и среднегодовой суммы осадков, для чего используются вспомогательные случайные последовательности $\mu[n]$, $\zeta[n]$, совместные вероятностные характеристики которых определяются в результате обучения ЛДРР-Т и ЛДРР-О.

4. Алгоритмы РМО для оценивания ПРВ

Для обучения модели используются данные из обучающего массива [24], структурированные по трем географическим зонам (индекс r) и временным интервалам 1973–2007 гг. По историческим данным были определены методом корреляционного анализа параметры «памяти» p соответствующих моделей и сформированы следующие массивы (матрицы) данных:

$$(4.1) \quad S_p^{(r)} = \begin{pmatrix} \mathbf{S}_{(0)}^{(r)} \\ \dots \\ \mathbf{S}_{(24-p)}^{(r)} \end{pmatrix}, \quad T_p^{(r)} = \begin{pmatrix} \mathbf{T}_{(0)}^{(r)} \\ \dots \\ \mathbf{T}_{(24-p)}^{(r)} \end{pmatrix}, \quad R_p^{(r)} = \begin{pmatrix} \mathbf{R}_{(0)}^{(r)} \\ \dots \\ \mathbf{R}_{(24-p)}^{(r)} \end{pmatrix},$$

где

$$(4.2) \quad \mathbf{S}_{(n-p)}^{(r)} = \{1, S^{(r)}[n-1], \dots, S^{(r)}[n-p]\}, \quad \mathbf{S}_{(p)}^{(r)} = \{S^{(r)}[p], \dots, S^{(r)}[24]\},$$

$$\mathbf{T}_{(n-p)}^{(r)} = \{1, T^{(r)}[n-1], \dots, T^{(r)}[n-p]\}, \quad \mathbf{T}_{(p)}^{(r)} = \{T^{(r)}[p], \dots, T^{(r)}[24]\},$$

$$\mathbf{R}_{(n-p)}^{(r)} = \{1, R^{(r)}[n-1], \dots, R^{(r)}[n-p]\}, \quad \mathbf{R}_{(p)}^{(r)} = \{R^{(r)}[p], \dots, R^{(r)}[24]\},$$

$$n = \overline{p, 24}.$$

Эти массивы использовались в алгоритмах РМО для оценивания функций ПРВ соответствующих моделей.

1. Алгоритм РМО для ЛДРР-П имеет вид

$$(4.3) \quad \mathcal{H}[P(\mathbf{a}, Q(\xi))] = - \int_{\mathcal{A}} P(\mathbf{a}) \ln P(\mathbf{a}) d\mathbf{a} - \\ - \sum_{n=p}^{24} \int_{\Xi_n} Q_n(\xi[n]) \ln Q_n(\xi[n]) d\xi[n] \Rightarrow \max$$

при ограничениях

$$(4.4) \quad \int_{\mathcal{A}} P(\mathbf{a}) d\mathbf{a} = 1, \quad \int_{\Xi_n} Q_n(\xi[n]) d\xi[n] = 1, \\ \int_{\mathcal{A}} P(\mathbf{a}) \mathbf{D}_{(n-p)}^{(r)} \mathbf{a} d\mathbf{a} + \int_{\Xi_n} Q(\xi[n]) \xi[n] d\xi[n] = S^{(r)}[n], \quad n = \overline{p, 24},$$

где блочный вектор-строка равна

$$(4.5) \quad \mathbf{D}_{(n-p)}^{(r)} = \left[\mathbf{S}_{(n-p)}^{(r)}, T^{(r)}[n], R^{(r)}[n] \right].$$

Решение этой задачи, параметризованное множителями Лагранжа $\theta = \{\theta_p, \dots, \theta_{24}\}$, имеет вид

$$(4.6) \quad P^*(\mathbf{a}, \theta) = \frac{\exp\left(-\langle \theta, D_p^{(r)} \mathbf{a} \rangle\right)}{\mathcal{P}(\theta)}, \quad \mathcal{P}(\theta) = \int_{\mathcal{A}} \exp\left(-\langle \theta, D_p^{(r)} \mathbf{a} \rangle\right) d\mathbf{a},$$

где блочная матрица равна

$$(4.7) \quad D_p^{(r)} = \begin{pmatrix} S_p^{(r)} & \mathbf{T}_{(p)}^{(r)} & \mathbf{R}_{(p)}^{(r)} \end{pmatrix}.$$

Измерительный шум для ЛДРР-П имеет энтропийно-оптимальные ПРВ следующего вида:

$$(4.8) \quad Q_n^*(\xi[n], \theta_n) = \frac{\exp(-\xi[n] \theta_n)}{\mathcal{Q}_n(\theta_n)}, \quad \mathcal{Q}_n(\theta_n) = \int_{\Xi_n} \exp(-\xi[n] \theta_n) d\xi[n].$$

Множители Лагранжа θ определяются из следующей системы уравнений:

$$(4.9) \quad \mathcal{P}^{-1}(\theta) \int_{\mathcal{A}} \exp\left(-\langle \theta, D_p^{(r)} \mathbf{a} \rangle\right) \mathbf{D}_{(n-p)}^{(r)} \mathbf{a} d\mathbf{a} + \\ + \mathcal{Q}_n^{-1}(\theta) \int_{\Xi} \exp(-\xi[n] \theta_n) d\xi[n] = S^{(r)}[n], \\ n = \overline{p, 24}.$$

Для ЛДРР-Т и ЛДРР-О алгоритм РМО отличается от (4.3)–(4.4) тем, что нужно оценивать совместные функции *ПРВ* $W(\mathbf{b}, \mu)$ и $F(\mathbf{c}, \zeta)$ параметров и входных случайных последовательностей, а также функции *ПРВ* измерительных шумов $E(\eta)$ и $G(\chi)$.

2. Адаптируя алгоритм РМО для ЛДРР-Т, будем иметь:

$$(4.10) \quad \mathcal{H}[W(\mathbf{b}, \mu), E(\eta)] = - \int_{\mathcal{B} \cap \mathcal{M}} W(\mathbf{b}, \mu) \ln W(\mathbf{b}, \mu) d\mathbf{b} d\mu - \\ - \sum_{n=p}^{24} \int_{\mathcal{E}_n} E_n(\eta[n]) \ln E_n(\eta[n]) d\eta[n] \Rightarrow \max$$

при ограничениях

$$(4.11) \quad \int_{\mathcal{B} \cap \mathcal{M}} W(\mathbf{b}, \mu) d\mathbf{b} d\mu = 1, \quad \int_{\mathcal{E}_n} E_n(\eta[n]) d\eta[n] = 1, \\ \int_{\mathcal{B} \cap \mathcal{M}} W(\mathbf{b}, \mu) \left[\mathbf{T}_{(n-p)}^{(r)} \mathbf{b} + \mu[n] \right] d\mathbf{b} d\mu[n] + \int_{\mathcal{E}_n} E_n(\eta[n]) \eta[n] d\eta[n] = T^{(r)}[n], \\ n = \overline{p, 24}.$$

В выражениях (4.10)–(4.11) вектор $\mu = \{\mu[p], \dots, \mu[24]\}$.

Обозначим множители Лагранжа для указанной задачи через $\vartheta = \{\vartheta_p, \dots, \vartheta_{24}\}$. Тогда ее решение — оптимальные функции *ПРВ* — можно представить, в силу линейности модели и независимости параметров и элементов вспомогательной случайной последовательности, в следующем виде:

$$(4.12) \quad W^*(\mathbf{b}, \mu, \vartheta) = L^*(\mathbf{b}, \vartheta) M^*(\mu, \vartheta), \quad M^*(\mu, \vartheta) = \prod_{n=p}^{24} M_n^*(\mu[n], \vartheta_n), \\ L^*(\mathbf{b}, \vartheta) = \frac{\exp\left(-\langle \vartheta, T_p^{(r)} \mathbf{b} \rangle\right)}{\mathcal{L}(\vartheta)}, \quad \mathcal{L}(\vartheta) = \int_{\mathcal{B}} \exp\left(-\langle \vartheta, T_p^{(r)} \mathbf{b} \rangle\right) d\mathbf{b}, \\ M_n^*(\mu[n], \vartheta_n) = \mathbb{M}_n^{-1}(\vartheta_n) \exp(-\vartheta_n \mu[n]), \quad \mathbb{M}_n(\vartheta_n) = \int_{\mathcal{M}_n} \exp(-\vartheta_n \mu[n]) d\mu[n], \\ E_n^*(\eta[n], \vartheta_n) = \mathbb{E}_n^{-1}(\vartheta_n) \exp(-\eta[n] \vartheta_n), \quad \mathbb{E}_n(\vartheta_n) = \int_{\mathcal{E}_n} \exp(-\eta[n] \vartheta_n) d\eta[n],$$

где множители Лагранжа ϑ определяются из следующей системы уравнений:

$$\begin{aligned}
& \mathcal{L}^{-1}(\vartheta) \int_{\mathcal{B}} \exp \left(-\langle \vartheta, T_p^{(r)} \mathbf{b} \rangle \right) \mathbf{T}_{(n-p)}^{(r)} \mathbf{b} d\mathbf{b} + \\
& + \mathbb{M}_n^{-1}(\vartheta_n) \int_{\mathcal{M}_n} M_n^*(\mu[n], \vartheta_n) \mu[n] d\mu[n] + \\
(4.13) \quad & + \mathbb{E}_n^{-1}(\vartheta_n) \int_{\tilde{\mathcal{E}}_n} E_n^*(\eta[n], \vartheta_n) \eta[n] d\eta[n] = T^{(r)}[n], \quad n = \overline{p, 24}.
\end{aligned}$$

3. Для ЛДРР-О алгоритм РМО, аналогичный (4.12)–(4.13), имеет вид:

$$\begin{aligned}
F^*(\mathbf{c}, \zeta, \lambda) &= V^*(\mathbf{c}, \lambda) Z^*(\zeta, \lambda), \quad Z^*(\zeta, \lambda) = \prod_{n=p}^{24} Z_n^*(\zeta[n], \lambda_n), \\
(4.14) \quad V^*(\mathbf{c}, \lambda) &= \frac{\exp \left(-\langle \lambda, R_p^{(r)} \mathbf{c} \rangle \right)}{\mathcal{V}(\lambda)}, \quad \mathcal{V}(\lambda) = \int_{\mathcal{C}} \exp \left(-\langle \lambda, R_p^{(r)} \mathbf{c} \rangle \right) d\mathbf{c}, \\
Z_n^*(\zeta[n], \lambda_n) &= \mathbb{Z}_n^{-1}(\lambda_n) \exp \left(-\lambda_n \zeta[n] \right), \quad \mathbb{Z}_n(\lambda_n) = \int_{\tilde{\mathcal{Z}}_n} \exp \left(-\lambda_n \zeta[n] \right) d\zeta[n], \\
G_n^*(\chi[n], \lambda_n) &= \mathbb{G}_n^{-1}(\lambda_n) \exp \left(-\lambda_n \chi[n] \right), \quad \mathbb{G}_n(\lambda_n) = \int_{\mathcal{G}_n} \exp \left(-\lambda_n \chi[n] \right) d\chi[n],
\end{aligned}$$

где множители Лагранжа $\lambda = \{\lambda_p, \dots, \lambda_{24}\}$ определяются из следующей системы уравнений:

$$\begin{aligned}
(4.15) \quad & \mathcal{V}^{-1}(\lambda) \int_{\mathcal{C}} \exp \left(-\langle \lambda, R_p^{(r)} \mathbf{c} \rangle \right) \mathbf{R}_{(n-p)}^{(r)} \mathbf{c} d\mathbf{c} + \\
& + \mathbb{Z}_n^{-1}(\lambda_n) \int_{\tilde{\mathcal{Z}}_n} Z_n^*(\zeta[n], \lambda_n) \zeta[n] d\zeta[n] + \\
& + \mathbb{G}_n^{-1}(\lambda_n) \int_{\mathcal{G}_n} G_n^*(\chi[n], \lambda_n) \chi[n] d\chi[n] = R^{(r)}[n], \\
& n = \overline{p, 24}.
\end{aligned}$$

Из равенств (4.6), (4.8), (4.12), (4.14) видно, что энтропийно-оптимальные функции ПРВ — экспоненциального класса, параметризованные соответствующими множителями Лагранжа, значения которых определяются балансовыми уравнениями (4.9), (4.13), (4.15).

5. Результаты обучения, тестирования и прогнозирования временной эволюции площади термокарстовых озер Западной Сибири

1. *Рандомизированное обучение (1973–1997 гг.)*. На предварительном этапе обучения определялись порядки p соответствующих моделей, с использованием исторических данных из обучающей коллекции. Для этого вычислялись автокорреляционные функции $r[k]$ и определялось $k_{\max} = p$, для которого $r[k_{\max}] \leq \delta$, где $\delta = 0,1$.

В результате РМО определены энтропийно-оптимальные ПРВ параметров моделей, вспомогательных случайных последовательностей и измерительных шумов.

ЛДРР-П ($p = 4$). Аналитические выражения для соответствующих параметризованных ПРВ представлены в (4.6), (4.8). Поскольку ЛДРР-П линейные, то они все экспоненциального типа:

$$\begin{aligned}
 P^*(\mathbf{a}, \theta) &= \prod_{k=0}^{(p+2)} P_k^*(a_k), \quad P_k^*(a_k) = \frac{\exp(-q_k a_k)}{\mathcal{P}_k}, \\
 \mathcal{P}_k &= \int_{\mathcal{A}_k} \exp(-q_k a_k) da_k, \\
 (5.1) \quad q_0 &= \sum_{n=p}^{24} \theta_n, \quad q_k = \sum_{n=p}^{24} \theta_n S^{(r)}[n-k], \quad k = \overline{1, p}, \\
 q_{p+1} &= \sum_{n=p}^{24} \theta_n T^{(r)}[n], \quad q_{p+2} = \sum_{n=p}^{24} \theta_n R^{(r)}[n], \\
 Q^*(\xi, \bar{\theta}) &= \frac{\exp(-\bar{\theta} \xi)}{\mathcal{Q}}, \quad \mathcal{Q} = \int_{\Xi} \exp(-\bar{\theta} \xi) d\xi, \quad \bar{\theta} = \frac{q_0}{24-p}.
 \end{aligned}$$

На рис. 2 представлены примеры энтропийно-оптимальных функций ПРВ для параметров $P^*(a_2, a_4)$ (a) и шума $Q^*(\xi)$ ($\bar{\theta}$), восстановленные в резуль-

Таблица 1. *ЛДРР-П*: оценки параметров модели

Северная зона термокарстовых озер				
Модель	Порядок	Параметры	Оценки МНК	Оценки РМО
S	4	a_0	-0,2888	-0,2750
		a_1	0,1069	0,1126
		a_2	-0,2224	-0,2212
		a_3	-0,1289	-0,1333
		a_4	0,0535	0,0533
		a_5	0,8330	0,8322
		a_6	0,6245	0,6080

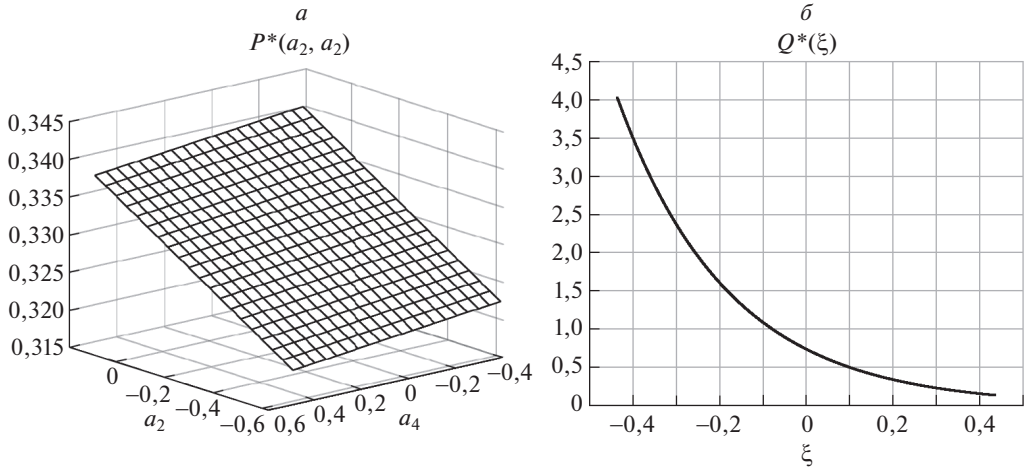


Рис. 2. Примеры восстановленных ПРВ.

тате обучения с использованием исторических данных северной зоны термокарстовых озер, а в табл. 1 приведены средние значения оценок параметров модели площади, полученные на основе РМО и оценками методом наименьших квадратов (МНК). Как видно из данной таблицы, оценки МНК и средние оценки РМО практически совпадают, что объяснимо линейностью используемых РПМ. Однако даже в этом случае процедура РМО позволяет генерировать ансамбль траекторий выхода модели и определять его числовые характеристики.

ЛДРР-Т. Аналитические выражения для соответствующих ПРВ имеют вид:

$$\begin{aligned}
 W^*(\mathbf{b}, \mu, \vartheta) &= L^*(\mathbf{b}, \vartheta) M^*(\mu, \vartheta), \quad M^*(\mu, \vartheta) = \prod_{n=p}^{(24)} M_n^*(\mu[n], \vartheta_n), \\
 L^*(\mathbf{b}, \vartheta) &= \prod_{k=0}^p L_k^*(b_k), \quad L_k^*(b_k) = \frac{\exp(-w_k b_k)}{\mathcal{L}_k}, \\
 \mathcal{L}_k &= \int_{\mathcal{B}_k} \exp(-w_k b_k) db_k, \\
 (5.2) \quad w_0 &= \sum_{n=p}^{24} \vartheta_n, \quad w_k = \sum_{n=p}^{24} \vartheta_n T^{(r)}[n-k], \quad k = \overline{1, p}, \\
 M_n^*(\mu[n], \vartheta_n) &= \frac{\exp(-\vartheta_n \mu[n])}{\mathfrak{M}_n}, \quad \mathfrak{M}_n = \int_{\mathcal{M}_n} \exp(-\vartheta_n \mu[n]) d\mu[n], \\
 E^*(\eta, \bar{\vartheta}) &= \frac{\exp(-\bar{\vartheta} \eta)}{\mathfrak{E}}, \quad \mathfrak{E} = \int_{\mathcal{E}} \exp(-\bar{\vartheta} \eta) d\eta, \quad \bar{\vartheta} = \frac{w_0}{24-p}.
 \end{aligned}$$

ЛДРР-О. Аналитические выражения для соответствующих ПРВ имеют вид:

$$\begin{aligned}
 F^*(\mathbf{c}, \zeta, \chi) &= V^*(\mathbf{c}, \lambda) Z^*(\zeta, \lambda), \quad Z^*(\zeta, \lambda) = \prod_{n=p}^{(24)} Z_n^*(\zeta[n], \lambda_n), \\
 V^*(\mathbf{c}, \lambda) &= \prod_{k=0}^p V_k^*(c_k), \quad V_k^*(b_k) = \frac{\exp(-s_k c_k)}{\mathcal{V}_k}, \\
 \mathcal{V}_k &= \int_{\mathcal{C}_k} \exp(-s_k c_k) dc_k, \\
 s_0 &= \sum_{n=p}^{24} \lambda_n, \quad s_k = \sum_{n=p}^{24} \lambda_n R^{(r)}[n-k], \quad k = \overline{1, p}, \\
 Z_n^*(\zeta[n], \lambda_n) &= \frac{\exp(-\lambda_n \zeta[n])}{\mathfrak{Z}_n}, \quad \mathfrak{Z}_n = \int_{\mathcal{Z}_n} \exp(-\lambda_n \zeta[n]) d\zeta[n], \\
 G^*(\chi, \bar{\lambda}) &= \frac{\exp(-\bar{\lambda} \chi)}{\mathfrak{G}}, \quad \mathfrak{G} = \int_{\mathcal{G}} \exp(-\bar{\lambda} \chi) d\chi, \quad \bar{\lambda} = \frac{w_0}{24-p}.
 \end{aligned}
 \tag{5.3}$$

2. *Тестирование (1998–2007 гг.).* Процедура тестирования применяется к комбинации обученных моделей, описанной в разделе 2, с использованием тестовых коллекций данных. Основу процедуры составляют сэмплирование



Рис. 3. Динамика площади.

Результаты тестирования для области С
 $AbsErr = 1,2863$; $RelErr = 0,0801$

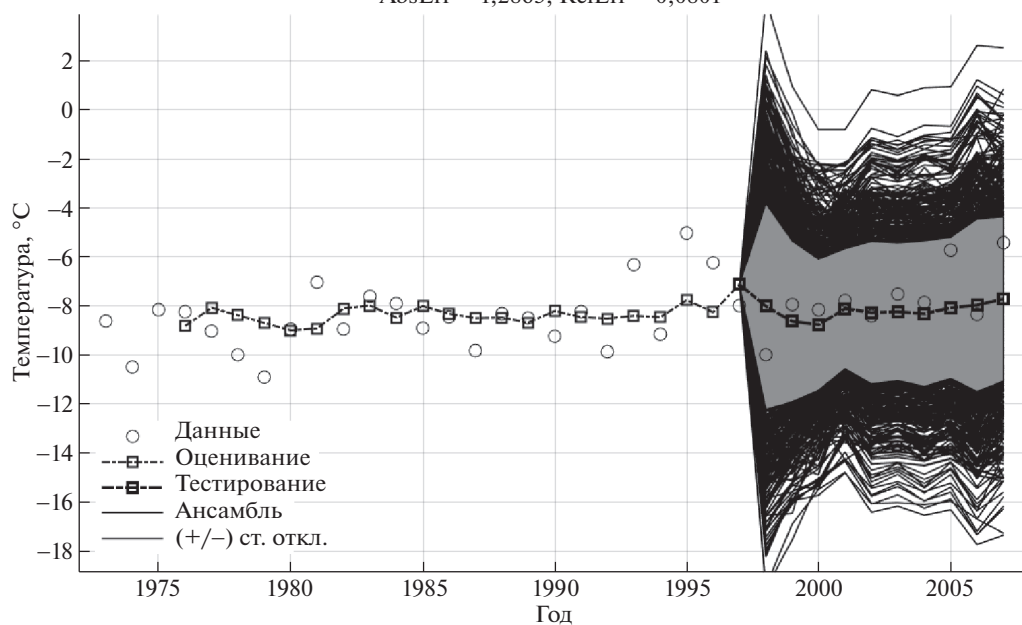


Рис. 4. Динамика среднесуточной температуры.

Результаты тестирования для области С
 $AbsErr = 107,6399$; $RelErr = 0,1675$

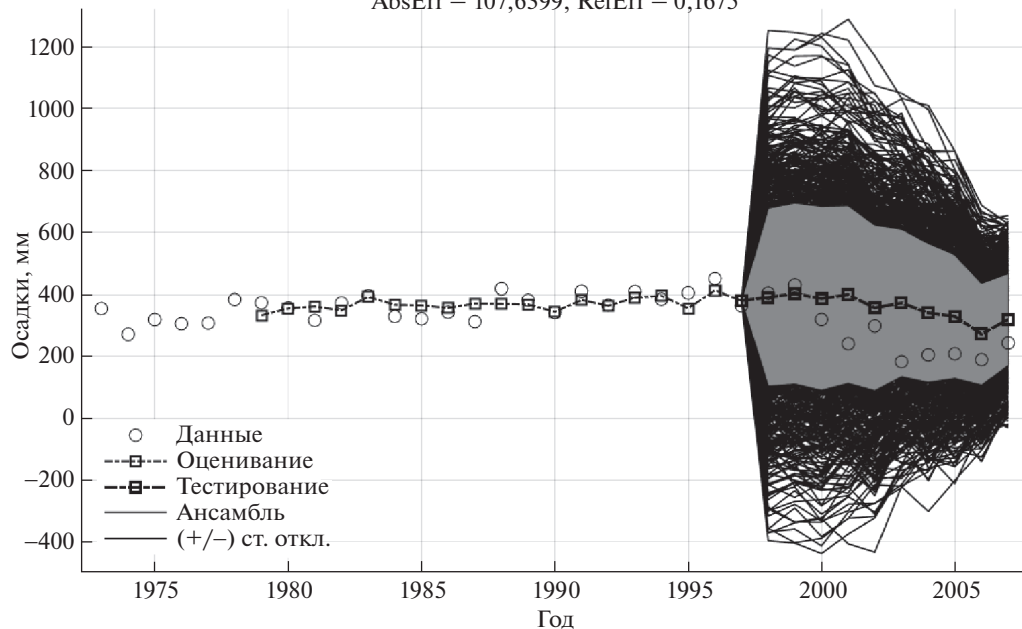


Рис. 5. Динамика годовой суммы осадков.

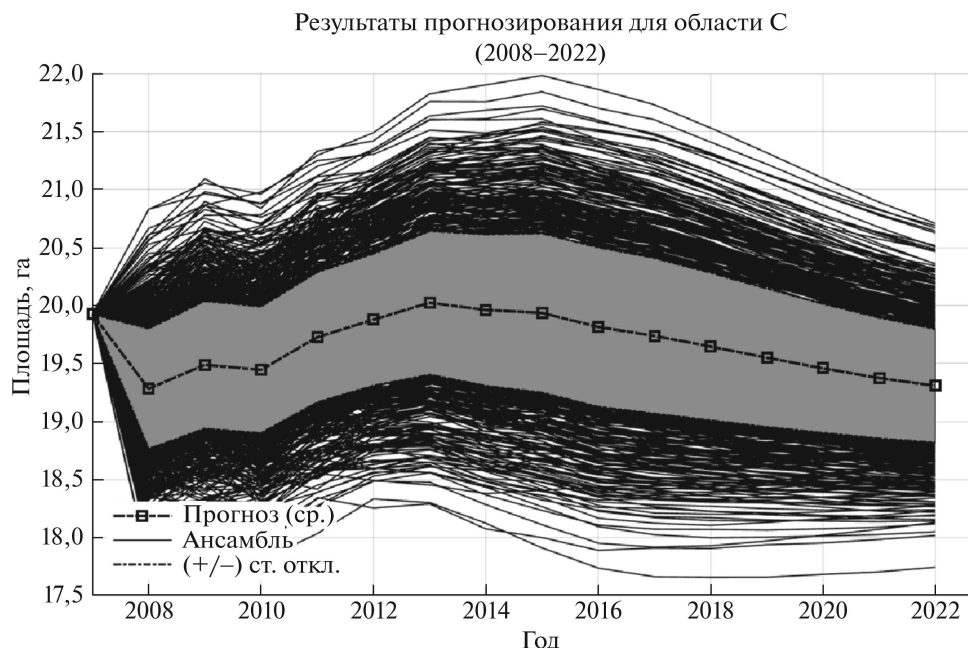


Рис. 6. Прогнозные траектории площади озер: ансамбль, дисперсионные трубки, средние траектории.

оптимальных ПРВ [22] и генерация ансамблей случайных траекторий на интервале тестирования. Оценка качества моделей производилась по степени близости средних модельных траекторий к реальным данным на интервале наблюдений в терминах относительной ошибки (2.15). В табл. 2 приведены значения относительных ошибок тестирования для северной зоны. На рис. 3–5 показаны примеры тестируемых траекторий с указанием соответствующих ошибок.

3. *Рандомизированное прогнозирование (2008–2023 гг.).* Все прогнозы строятся с использованием модели (рис. 1), элементами которой являются ЛДРР-П, ЛДРР-Т и ЛДРР-О. Энтропийно-оптимальные ПРВ генерируют ансамбли траекторий, характеризующие временную эволюцию площади термокарстовых озер на соответствующем интервале прогноза. Отличие интервала прогноза от интервала тестирования заключается в том, что в данном случае для обучения моделей используются все точки данных.

Для каждой точки по ансамблю вычисляется среднее значение (*mean*) и стандартное отклонение (*std*). Результаты прогнозирования для северной зоны представлены на рис. 6, где показан ансамбль прогнозных траекторий по

Таблица 2. Относительные ошибки тестирования

Оценка	ЛДРР-П	ЛДРР-Т	ЛДРР-О
<i>RelErr</i>	0,0089	0,0801	0,1675

площади озер и отмечены траектория средних значений (темная кривая) и дисперсионная трубка (серая зона).

6. Обсуждение результатов

Процессы формирования и эволюции площади термокарстовых озер в зонах вечной мерзлоты изучены недостаточно как с точки зрения внутренней геологической динамики, так и внешних климатических факторов, влияющих на их формирование. Процедуры трансформации спутниковых снимков в количество и площадь озер сопровождаются весьма существенными погрешностями. Поэтому модельная информация и данные о состоянии исследуемого объекта являются неопределенными.

Алгоритмы РМО ориентированы на оценивание характеристик таких моделей с использованием данных с ошибками неизвестной природы. В результате восстанавливаются функции ПРВ параметров моделей и измерительных шумов, соответствующие максимальной неопределенности.

Сэмплирование оптимальных функций ПРВ и применение метода Монте-Карло позволяют генерировать ансамбли траекторий, характеризующих эволюцию площади термокарстовых озер. Статистическая обработка этих ансамблей позволяет получать различные их числовые характеристики, прежде всего средние траектории и множество траекторий, границами которого являются среднеквадратичные траектории (дисперсионные трубки). Сравнение средних траекторий с реальными данными в терминах относительных среднеквадратичных (интегральных) ошибок для различных участков поверхности составляет 8–17%.

Проведенное исследование данного объекта с применением, адаптацией и тестированием алгоритмов рандомизированного машинного обучения продемонстрировало достаточно высокую точность по воспроизведению средних траекторий (8–17%). Максимальные размеры дисперсионных трубок, вычисленных по генерируемым ансамблям, составляли $\pm 9\%$.

Во всех исследованиях использовались линейные рандомизированные модели динамической регрессии с памятью $p = 3 \div 5$. Для моделирования эволюции площади они дали вполне приемлемые результаты. Но для климатических параметров ошибки тестирования оказывались существенно выше. По-видимому, с переходом на нелинейные варианты моделей удастся повысить точность воспроизведения реальных данных.

Проблемной частью метода РМО остаются балансовые уравнения, в которые входят так называемые интегральные компоненты. Это — многомерные определенные интегралы с параметрами, вычисляемые на простых множествах (параллелепипедах). В данном исследовании, где использовались линейные модели, они вычислялись аналитически и можно было использовать программное интегрирование соответствующих аналитических выражений. Однако при переходе на нелинейные модели такой путь исключен.

7. Заключение

Данная работа является одной из посвященных исследованию работоспособности и эффективности метода рандомизированного машинного обучения. В ней сделан уклон на проблемы использования оптимальных рандомизированных моделей для прогнозирования. Их особенность состоит в том, что прогнозируемая входная последовательность для этих моделей, как правило, неизвестна. В этой ситуации предлагается ее генерировать с помощью вспомогательной случайной последовательности, энтропийно-оптимальные характеристики которой можно восстановить при машинном обучении соответствующей модели.

СПИСОК ЛИТЕРАТУРЫ

1. *Vapnik V.N.* Statistical Learning Theory. John Wiley & Sons, 1998.
2. *Bishop C.* Pattern Recognition and Machine Learning. N.Y. Springer, 2007.
3. *Friedman J., Hastie T., Tibshirani R.* The elements of statistical learning. Volume 1, Springer series in statistics, Berlin. Springer, 2009.
4. *Popkov Yu.S., Dubnov Yu.A., Popkov A.Yu.* Randomized Machine Learning: Statement, Solution, Applications // Proc. IEEE Int. Conf. on Intelligent Systems. 2016. P. 27–39.
5. *Zuidhoff F.S., Kolstrup E.* Changes in palsa distribution in relation to climate change in Laivadalen, Northern Sweden, especially 1960–1997 // Permafrost and Periglacial Processes. 2000. V. 11. P. 55–69.
6. *Kirpotin S., Polishchuk Y., Bruksina N.* Abrupt changes of thermokarst lakes in Western Siberia: impacts of climatic warming on permafrost melting // Int. J. Environmental Studies. 2009. V. 66. No. 4. P. 423–431.
7. *Karlson J.M., Lyon S.W., Destouni G.* Temporal behavior of lake size-distribution in a thawing permafrost landscape in Northwestern Siberia // Remote Sensing. 2014. No. 6. P. 621–636.
8. *Bryksina N.A., Polishchuk Yu.M.* Analysis of changes in the number of thermokarst lakes in permafrost of Western Siberia on the basis of satellite images // Cryosphere of Earth. 2015. V. 19. No. 2. P. 114–120.
9. *Liu Q., Rowe M.D., Anderson E.J., Stow C.A., Stumpf R.P.* Probabilistic forecast of microcystin toxin using satellite remote sensing, *in situ* observation and numerical modeling // Environment Modelling and Software. 2020. V. 128. P. 104705.
10. *Vidyasagar M.* Statistical Learning Theory and Randomized Algorithms for Control // IEEE Control System Magazine. 1998. V. 1. No. 17. P. 69–88.
11. *Граничин О.Н., Поляк Б.Т.* Рандомизированные алгоритмы оценивания и оптимизации при почти произвольных помехах. М.: Наука, 2002.
12. *Biondo A.E., Pluchino A., Rapisarda A., Helbing D.* Are random trading strategies more successful than technical ones? // PLoS ONE. 2013. V. 6. No. 7. P. e68344.
13. *Lutz W., Sandersen S., Scherbov S.* The end of world population growth // Nature. 2001. V. 412. No. 6846. P. 543–545.
14. *Цирлин А.М.* Методы усредненной оптимизации и их применение. М.: Физматлит, 1997.

15. *Shannon C.* Communication Theory of Secrecy Systems // Bell System Technical Journal. 1949. V. 28. No. 4. P. 656–715.
16. *Jaynes E.T.* Information Theory and Statistical Mechanics // Physics Review. 1957. V. 106. P. 620–630.
17. *Jaynes E.T.* Papers on probability, statistics and statistical physics. Dordrecht. Kluwer Academic Publisher, 1989.
18. *Jaynes E.T.* Probability Theory. The logic and science. Cambridge University Press, 2003.
19. *Попков Ю.С., Попков А.Ю., Дубнов Ю.А.* Рандомизированное машинное обучение при ограниченных объемах данных. М.: УРСС, 2019.
20. *Popkov Y., Popkov A.* New Method of Entropy-Robust Estimation for Randomized Models under Limited Data // Entropy. 2014. V. 16. P. 675–698.
21. *Иоффе А.Д., Тихомиров В.М.* Теория экстремальных задач. М.: Наука, 1984.
22. *Дарховский Б.С., Попков Ю.С., Попков А.Ю., Алиев А.С.* Метод генерации случайных векторов с заданной функцией плотности распределения вероятностей // АиТ. 2018. № 9. С. 31–45. <https://doi.org/10.31857/S000523100001408-2>
Darkhovsky B.S., Popkov Y.S., Popkov A.Y., Aliev A.S. A Method of Generating Random Vectors with a Given Probability Density Function // Autom. Remote Control. 2018. V. 79. No. 9. P. 1569–1581. <https://doi.org/10.1134/S0005117918090035>
23. *Айвазян С.А., Енюков И.С., Мешалкин Л.Д.* Прикладная статистика: Исследование зависимостей. М.: Финансы и статистика, 1985.
24. *Электронный ресурс:* <https://cloud.uriit.ru/index.php/s/0DOrxL9RmGqXsV0>.

Статья представлена к публикации членом редколлегии А.Н. Соболевским.

Поступила в редакцию 20.04.2022

После доработки 21.06.2022

Принята к публикации 29.09.2022