

DOI: <https://doi.org/10.60797/IRJ.2024.145.67>

ОСОБЕННОСТИ ПОСТРОЕНИЯ РЕГРЕССИОННОЙ МОДЕЛИ ДЛЯ ОЦЕНКИ И ПРОГНОЗА УРОВНЯ ОЗЕРА ЛОВОЗЕРО

Научная статья

Батмазова А.А.¹, Гайдукова Е.В.^{2,*}, Дрегваль М.С.³¹ ORCID : 0000-0002-9135-4925;² ORCID : 0000-0002-3547-5538;³ ORCID : 0000-0002-2144-5928;^{1,2,3} Российский государственный гидрометеорологический университет, Санкт-Петербург, Российская Федерация

* Корреспондирующий автор (oderiut[at]mail.ru)

Аннотация

В статье приведен обзор методов многомерного статистического анализа и примеров их применения в гидрометеорологии. Исследовался уровенный режим озера Ловозеро с применением корреляционного и регрессионного анализа. Описаны используемые исходные данные, обоснована выбранная месячная дискретизация. Проведен корреляционный анализ гидрометеорологических характеристик с поэтапным исключением предикторов, создающих эффект мультиколлинеарности. Проведен нестандартизированный регрессионный анализ с исключением предикторов, оказывающих слабое влияние на зависимую переменную (уровень озера) и не удовлетворяющих критерию значимости коэффициента регрессии. По полученной регрессионной модели выполнен расчет уровней воды на зависимом материале, также выведено уравнение для прогноза с заблаговременностью один месяц. Для уравнений рассчитаны и оценены критерии эффективности. Выполнена стандартизация параметров уравнения регрессии, которая позволила оценить степень влияния параметров уравнения на функцию отклика – уровень озера. Сделан вывод о нецелесообразности проведения процедуры стандартизации в конкретном случае. По итогу всего исследования сделан вывод о влиянии на уровенный режим Ловозера метеорологических параметров – общая облачность и облачность нижнего яруса. Приведены комплексные графики, позволяющие проанализировать синхронность хода гидрометеорологических параметров: уровень воды, атмосферные осадки, общая облачность и облачность нижнего яруса. Проведена процедура фильтрации методом скользящего среднего для значений облачности с целью исключения высокочастотной составляющей для более удобного анализа. Сделан вывод о выполнимости критериев эффективности полученных регрессионных уравнений и, как следствие, возможности их применения для прогноза уровней воды оз. Ловозеро.

Ключевые слова: озерно-речные системы, численное моделирование, уровень воды, облачность, регрессия, заблаговременность.

A REGRESSION MODEL CONSTRUCTING FEATURES FOR ASSESSING AND FORECASTING THE LEVEL OF LAKE LOVOZERO

Research article

Batmazova A.A.¹, Gaidukova E.V.^{2,*}, Dregval M.S.³¹ ORCID : 0000-0002-9135-4925;² ORCID : 0000-0002-3547-5538;³ ORCID : 0000-0002-2144-5928;^{1,2,3} Russian State Hydrometeorological University, Saint-Petersburg, Russian Federation

* Corresponding author (oderiut[at]mail.ru)

Abstract

The article provides an overview of methods of multivariate statistical analysis and examples of their application in hydrometeorology. The level regime of Lake Lovozero was studied using correlation and regression analysis. The initial data used is described, and the selected monthly discretization is justified. A correlation analysis of hydrometeorological characteristics was carried out with the step-by-step exception of predictors that create the effect of multicollinearity. An unstandardized regression analysis was carried out with the exclusion of predictors that have a weak effect on the dependent variable (lake level) and do not satisfy the criterion of significance of the regression coefficient. Using the resulting regression model, water levels on the dependent material were calculated, and an equation for forecasting with a lead time of one month was also derived. Efficiency criteria were calculated and evaluated for the equations. The parameters of the regression equation were standardized, which made it possible to assess the degree of influence of the equation parameters on the response function – lake level. It is concluded that it is inappropriate to carry out the standardization procedure in a particular case. Based on the results of the entire study, a conclusion was made about the influence of meteorological parameters on the level regime of Lovozero – total cloudiness and low-level cloudiness. Complex graphs are presented that allow one to analyze the synchronicity of hydrometeorological parameters: water level, precipitation, total cloudiness and lower cloudiness. A filtering procedure was carried out using the moving average method for cloudiness values in order to exclude the high-frequency component for more convenient analysis. A conclusion is made about the feasibility of the efficiency criteria of the obtained regression equations and, as a consequence, the possibility of their use for forecasting water levels in Lake Lovozero.

Keywords: lake-river systems, numerical modeling, water level, cloudiness, regression, lead time.

Введение

Многомерный статистический анализ используется для оценки гидрометеорологических процессов и полей и включает различные методы, которые можно разделить на два крупных класса: методы, направленные на снижение размерности пространства переменных (например, регрессионный анализ, факторный анализ, метод главных компонент), методы, позволяющие объединять объекты в единообразные группы (например, методы дискриминантного и кластерного анализа). Методы широко применяются для решения практических и теоретических задач в гидрометеорологии.

Так, например, в работе Воронцова А. А. и др. [1] применен метод главных компонент, с помощью которого были выявлены синхронные и асинхронные колебания в полях среднемесячных значений температур воздуха. Был сделан вывод, что поля синхронных колебаний подобны среднемноголетнему полю. Главная синхронная компонента была аппроксимирована аналитическим выражением и спрогнозирована с помощью метода авторегрессий.

Методы спектрального анализа активно применяются в гидрометеорологии для оценки разномасштабной изменчивости тех или иных параметров рассматриваемой системы. Например, в работе Долгоносова Б. М. и Корчагина К. А [2] был получен спектр мощности речного стока, который может быть использован для анализа процессов переноса примесей с водосбора, что важно для оценки экологических последствий химического и бактериального загрязнения речного бассейна.

Авторы Рожков В. А. и Сухих Н. А. в своей работе [3] использовали векторный дисперсионный анализ для изучения изменчивости течений в Печорском море. Статистический анализ изменчивости по времени, пространству и глубине скоростей течений в Печорском море выявил наличие годовой ритмики и приливных колебаний, параметры которых также изменяются по глубине и в пространстве. Для количественной характеристики различий этих параметров необходимо использовать векторный дисперсионный анализ.

Применение регрессионного анализа в области наук о Земле хорошо отражено в работе Пяткина Ф. В. и др. [4]. Ученые рассмотрели регрессионный метод выбора каналов гиперспектральных сенсоров для определения характеристик атмосферы. Метод применялся для расчетов концентрации углекислого газа для инфракрасного Фурье-спектрометра ИКФС-2 и интегрального содержания водяного пара в атмосфере. В работе использовался универсальный подход для определения предикторов и коэффициентов уравнения. Разработанная авторами методика интерпретации спутниковых данных для определения запасов водяного пара в атмосфере над безоблачной поверхностью океана дала хорошие результаты со стабильным среднеквадратическим отклонением линии регрессии от фактических данных.

Также содержательным примером возможности и необходимости применения регрессионного анализа в гидрометеорологии является исследование коллектива авторов из Российского государственного гидрометеорологического университета. В статье [5] авторы рассматривают методику долгосрочного прогнозирования стока рек Арктической зоны Российской Федерации, основанную на построении регрессионных зависимостей с различным набором предикторов. В результате расчетного эксперимента получен удовлетворительный результат прогнозов при учете взаимосвязей слоя стока весеннего половодья с наивысшим уровнем воды, максимальными снегозапасами, влажностью воздуха и температурой почвы на глубине 0,2 м. Метод регрессионного анализа позволил вывести прогнозное уравнение с минимальной заблаговременностью в один месяц с перспективой увеличения до двух месяцев.

В настоящей же статье представлены результаты использования регрессионного анализа для обоснования набора предикторов, с оценкой их статистической значимости, с целью получения достоверной регрессионной модели для описания и прогноза процесса изменения уровня воды в водоеме.

Исходные данные и методы исследования

2.1. Исходные данные

В качестве объекта исследования выбрано озеро Ловозеро с озерным гидрологическим постом (ОГП) с. Ловозеро с рядом наблюдений за уровнем воды (H) с 1932 года по настоящее время, но только с 2008 по 2021 гг. данные наблюдений имеются в открытом доступе на ресурсе АИС ГМВО [6].

Влияющими на уровеньный режим озера факторами являются расходы воды впадающих в озеро рек (Q) и метеорологические параметры, которые в совокупности своей являются приходными и расходными характеристиками уравнения водного баланса, позволяющего рассчитать коэффициент водообмена (K_B), который, в свою очередь, является наиболее показательным при изучении озер и используется при оценке структуры уровеньных рядов и режима стока вытекающих из озер рек [7], [8].

Гидрологическая информация представлена среднесуточными и среднемесячными расходами воды по притоку р. Сергевань в оз. Ловозеро.

В качестве метеорологических факторов были взяты среднемесячные значения: атмосферных осадков, температуры воздуха, общей облачности и облачности нижнего яруса, упругости водяного пара, продолжительности солнечного сияния, относительной влажности воздуха и атмосферного давления на уровне метеостанции (МС) Краснощелье.

В статье рассматривалась месячная дискретизация, то есть диапазон сезонной изменчивости в многолетнем разрезе, с переходом, в перспективе, на анализ синоптической изменчивости. Такой масштаб был выбран в связи с необходимостью изучения внутригодовых колебаний оз. Ловозеро и гидрометеорологических параметров, формирующих его уровеньный режим и имеющих свою уникальную ритмику годовой цикличности [7]. Таким образом, можно отследить ежегодные периоды максимального и минимального стояния уровней воды в Ловозере.

Для исследования был выбран весь имеющийся в свободном доступе период наблюдений. Метеорологическая информация бралась за этот же период с 2008 по 2021 гг. Однако в рядах гидрологической информации были обнаружены большие пропуски, поэтому был взят совместный период наблюдений за среднемесячными гидрометеорологическими параметрами с 2015 по 2021 год.

Хронологический график изменения рассматриваемых гидрометеорологических параметров представлен на рисунке 1.

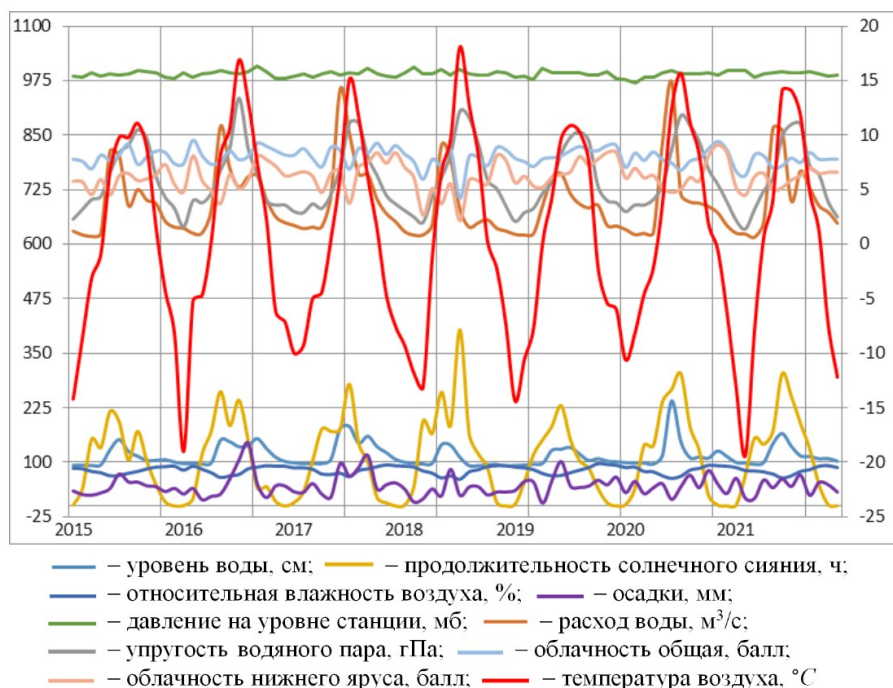


Рисунок 1 - Комплексный график изменения гидрометеорологических параметров
DOI: <https://doi.org/10.60797/IRJ.2024.145.67.1>

2.2. Методы исследования

Для описания и анализа связей гидрометеорологических процессов и явлений используется аппарат корреляционного и регрессионного анализа.

Задача корреляционного анализа – выявление связи между переменными и оценка ее тесноты. Тесноту статистических связей между двумя рядами характеризует коэффициент корреляции, который является параметрическим показателем и изменяется в пределах $-1 \leq r \leq 1$.

Задача регрессионного анализа – изучение формы и установление зависимости между переменными. Статистическое уравнение множественной линейной регрессии имеет вид:

$$y_i = a_0 + a_1x_1 + \dots + a_nx_n + \varepsilon_n = \sum_i^n a_ix_i + a_0 + \varepsilon_i, \quad (1)$$

где:

x_i – значения предикторов; a_i – коэффициенты регрессии; a_0 – свободный член уравнения; ε_i – остатки, не описываемые уравнением. Для определения неизвестных коэффициентов a_i – применяется метод наименьших квадратов. Величина коэффициента регрессии a_i показывает, насколько в среднем изменится зависимая переменная y при изменении переменной x на единицу своего измерения.

Уравнение регрессии можно получить по стандартизированным данным [9]:

$$z_y = \beta_1z_1 + \beta_2z_2 + \dots + \beta_nz_n + \varepsilon_n = \sum_i^n \beta_iz_i + \varepsilon_i \quad (2)$$

где:

z_y, β_i, z_i – стандартизированные значения зависимой переменной, коэффициента регрессии и независимой переменной соответственно; ε – вектор остатков. Преимущество стандартизированного коэффициента регрессии β в том, что он показывает относительную роль переменной x в описании изменчивости функции отклика – y . Например, если $|\beta_1| > |\beta_2|$, то x_1 оказывает на y большее влияние, чем x_2 . Нестандартизированные коэффициенты a_i таким свойством не обладают.

Результаты исследования

3.1. Множественная линейная регрессия по нестандартизированным данным

Построение регрессионной модели начинается с анализа взаимосвязей между предикторами и предиктантом, который можно осуществить по корреляционной матрице (см. табл. 1) и геометрическому ее аналогу – графу корреляции (см. рис. 2).

Таблица 1 - Корреляционная матрица среднемесячных гидрометеорологических характеристик
DOI: <https://doi.org/10.60797/IRJ.2024.145.67.2>

	H , см	Q , м ³ /с	Упругость водяного пара, гПа	Прод-ть солнечного сияния, ч	Относ. влажность воздуха, %	P на уровне ст., мб	Облачность общая, балл	Облачность нижнего яруса, балл	Осадки, мм	Температура воздуха, °С
H , см	1,00	-	-	-	-	-	-	-	-	-
Q , м ³ /с	0,89	1,00	-	-	-	-	-	-	-	-
Упругость водяного пара, гПа	0,51	0,51	1,00	-	-	-	-	-	-	-
Прод-ть солнечного сияния, ч	0,47	0,59	0,62	1,00	-	-	-	-	-	-
Относ. Влажность воздуха, %	-0,48	-0,61	-0,47	-0,92	1,00	-	-	-	-	-
P на уровне ст., мб	0,35	0,38	0,34	0,34	-0,30	1,00	-	-	-	-
Облачность общая, балл	0,05	-0,01	-0,10	-0,58	0,49	-0,08	1,00	-	-	-
Облачность нижнего яруса, балл	-0,08	-0,22	-0,16	-0,71	0,71	-0,14	0,87	1,00	-	-
Осадки, мм	0,37	0,36	0,47	0,12	-0,03	-0,06	0,23	0,22	1,00	-
Температура воздуха, °С	0,62	0,66	0,95	0,72	-0,59	0,42	-0,08	-0,19	0,43	1,00

Анализ корреляционной матрицы и графа корреляции выявил следующее:

1. Наличие высокой связи между предиктантом (уровень воды в озере) и многими предикторами: упругость водяного пара, температура воздуха, относительная влажность воздуха, продолжительность солнечного сияния, атмосферное давление на уровне метеостанции.

2. Выявлены предикторы, которые не будут участвовать в дальнейшей проверке ввиду наличия относительно высокого коэффициента корреляции с предиктантом и отсутствия мультиколлинеарности, а именно: расходы воды, атмосферные осадки и температура воздуха (отмечены в табл. 1 зеленым цветом).

3. Оценивался мультиколлинеарный эффект между переменными, так называемая ложная корреляция, которая может заключаться, во-первых, в физически необоснованной корреляции между предикторами; во-вторых, в опосредующем или обуславливающем влиянии двух и более предикторов на прогнозируемый элемент [11], [12]. Пример ложной корреляции отмечен в табл. 1 красным цветом. В рассматриваемом случае один из предикторов (упругость водяного пара и температура воздуха) можно исключить, предварительно оценив степень влияния каждого на предиктант. Теоретически логично, что можно наблюдать сильную связь между упругостью водяного пара и температурой воздуха, так как, чем выше температура, тем больше воздух может содержать водяного пара [13]. В таком случае имеет смысл исключить из числа предикторов упругость водяного пара, так как этот показатель является следствием от изменения температуры, которая оказывает большее влияние на уровень воды в озере с соответственно сопутствующими процессами.

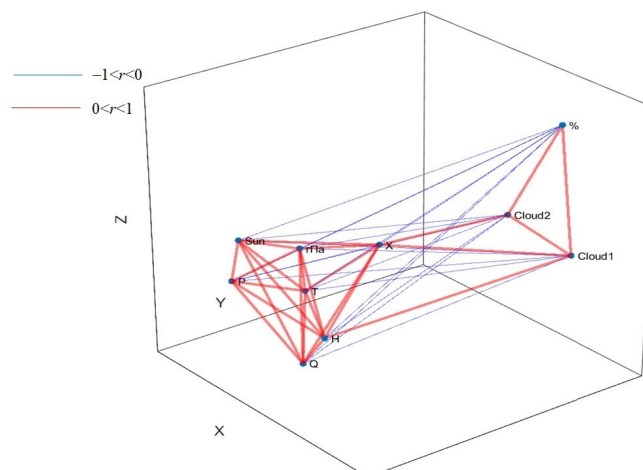


Рисунок 2 - Граф корреляции среднемесячных гидрометеорологических характеристик
DOI: <https://doi.org/10.60797/IRJ.2024.145.67.3>

Примечание: Q – расход воды; H – уровень воды; X – атмосферные осадки; T – температура воздуха; P – давление на уровне станции; % – относительная влажность воздуха; $Cloud1$ – общая облачность; $Cloud2$ – облачность нижнего яруса; Sun – продолжительность солнечного сияния; zPa – упругость водяного пара

Для дальнейшего отбора предикторов выполняется регрессионный анализ с последовательным исключением и/или добавлением переменных.

При первом «прогоне» использовались все имеющиеся в наличии предикторы. Потребовалось пять итераций с различными наборами параметров, в результате которых было получено регрессионное уравнение, давшее наилучший результат, из пяти предикторов, а именно: среднемесячный расход воды (Q , м³/с), атмосферные осадки (мм), температура воздуха (°C), а также общая облачность и облачность нижнего яруса в баллах. Соответствующие процессу перебора коэффициенты множественной регрессии R и стандартные ошибки представлены в таблице 2.

Таблица 2 - Результаты регрессионного анализа

DOI: <https://doi.org/10.60797/IRJ.2024.145.67.4>

	Количество предикторов					
	9	8	7	7	6	5
Множественный R	0,910	0,909	0,901	0,907	0,907	0,907

R^2	0,828	0,827	0,812	0,823	0,823	0,823
Стандартная ошибка, см	11,72	11,66	12,05	11,70	11,61	11,52

В гидрологии полученная регрессионная зависимость может использоваться для практических расчетов, если регрессия объясняет более 50 % разброса относительно предиктанта, т. е. $|R| \geq 0,7 \Rightarrow R^2 \approx 0,5$ [14]. Поскольку указанные два условия соблюдались для каждого набора предикторов, то в качестве критерия отбора параметров уравнения использовалась стандартная ошибка.

Полученное регрессионное уравнение имеет вид ($R = 0,907$):

$$H = 7,81X_1 - 0,015X_2 + 0,19X_3 - 7,07X_4 + 6,39X_5 + 104,1 \quad (3)$$

где:

где H – среднемесячный уровень воды оз. Ловозеро; X_1 – среднемесячные расходы воды р. Сергевань (Q , м³/с); X_2 – среднемесячные атмосферные осадки (мм); X_3 – температура воздуха (°C); X_4 – общая облачность в баллах; X_5 – облачность нижнего яруса в баллах.

Полученное уравнение регрессии может использоваться в качестве прогностической зависимости, однако это корректно только в том случае, если уравнение является надежным. Поэтому для уравнения (3) оценивались следующие статистические характеристики с целью проверки соблюдения требований, предъявляемых в гидрологии к уравнению линейной регрессии:

- σ_R – стандартная ошибка коэффициента парной корреляции;
- σ_a – стандартная ошибка коэффициента регрессии (далее будут приведены результаты проверки, где a_1 – среднемесячные расходы воды р. Сергевань; a_2 – среднемесячные атмосферные осадки; a_3 – температура воздуха; a_4 – общая облачность; a_5 – облачность нижнего яруса);
- σ_b – стандартная ошибка свободного члена.

Регрессионное уравнение (3) и его параметры проверялись на соблюдение следующих условий: $n > 10$, $|R| > 0,7$, $|R|/\sigma_R \geq 2$, $|a|/\sigma_a \geq 2$, $|b|/\sigma_b \geq 2$. В результате проверки получены следующие выводы: длина исследуемого ряда $n = 72$; $|R| = 0,907$; $|R|/\sigma_R = 43$; $|a_1|/\sigma_{a1} = 12,39$; $|a_2|/\sigma_{a2} = 0,23$; $|a_3|/\sigma_{a3} = 0,87$; $|a_4|/\sigma_{a4} = 1,99$; $|a_5|/\sigma_{a5} = 2,89$; $|b|/\sigma_b = 6,17$.

Для выведенного уравнения регрессии (3) соблюдаются практически все условия для того, чтобы его можно было использовать в прогностических целях. Однако, несмотря на такой хороший результат, встает вопрос, что делать с предикторами, неудовлетворяющими требованиям: среднемесячные атмосферные осадки, температура воздуха и общая облачность.

Было принято решение выполнить регрессионный анализ, исключив предиктор X_2 – среднемесячные атмосферные осадки (мм) и X_3 – температура воздуха (°C), оставив при этом параметр X_4 – общая облачность в баллах, поскольку он близок к выполнению требования $|a|/\sigma_a \geq 2$. Результаты анализа и проверки условий приведены в таблице 3.

Таблица 3 - Оценка требований уравнения линейной регрессии с тремя предикторами

DOI: <https://doi.org/10.60797/IRJ.2024.145.67.5>

	$n > 10$	$ R > 0,7$	$ R /\sigma_R \geq 2$	$ a /\sigma_a \geq 2$	$ b /\sigma_b \geq 2$
	72	0,906	42		
a_1				16,96	
a_2		–		2,07	–
a_3				2,89	
b				–	6,17

Примечание: среднемесячные расходы воды a_1 , общая облачность a_2 , облачность нижнего яруса a_3

Таким образом, исключив из регрессионного анализа два неудовлетворяющих условиям надежности предиктора (среднемесячные атмосферные осадки и температуру воздуха) – весомо повысили значимость остальных параметров.

Сокращенная регрессионная модель имеет вид:

$$H = 8,11X_1 - 7,27X_2 + 6,36X_3 + 104,2, \quad (4)$$

где:

X_1 – среднемесячные расходы воды (Q , м³/с); X_2 – общая облачность в баллах; X_3 – облачность нижнего яруса в баллах.

Итоговое уравнение проверялось на независимом материале. Эффективность методики оценивалась с помощью критерия S/σ , где S – средняя квадратическая погрешность проверочных прогнозов, σ – среднее квадратическое отклонение прогнозируемого значения элемента от среднего [15].

Поверочный период составил один год – 2021. На рисунке 3 представлен совместный график хода фактических и поверочных (рассчитанных по уравнению регрессии) среднемесячных уровней воды оз. Ловозеро за 2021 год. Длина

ряда $n < 15$, отношение $S/\sigma < 0,70$, применяемую методику (регрессионную модель) можно считать удовлетворительной.

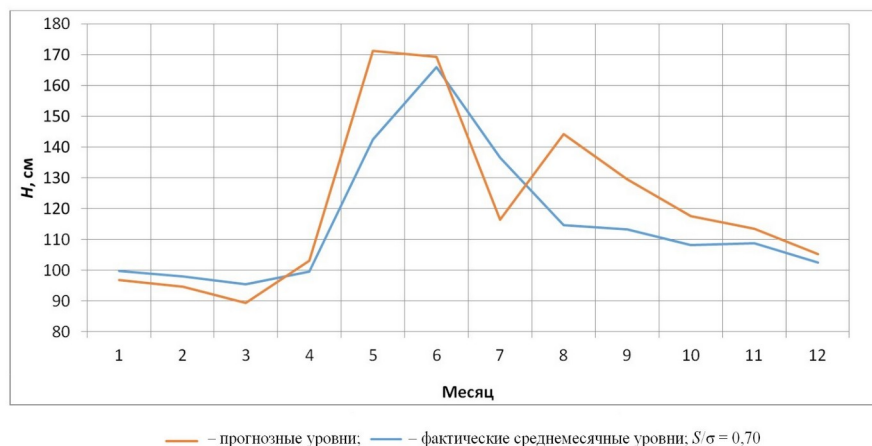


Рисунок 3 - График изменения фактических и прогнозных уровней воды оз. Ловозеро, 2021 год
DOI: <https://doi.org/10.60797/IRJ.2024.145.67.6>

Также была разработана регрессионная модель с учетом заблаговременности в 1 месяц (т.е. в уравнение подставлялись предикторы за предыдущий месяц):

$$H = 5,54X_1 + 11,6X_2 - 6,59X_3 + 42,35 \quad (5)$$

На рисунке 4 и 5 представлен график изменения фактических и расчетных уровней воды оз. Ловозеро на зависимом материале и график поверочного прогноза с оценками эффективности методики (на независимом материале) с заблаговременностью 1 месяц.

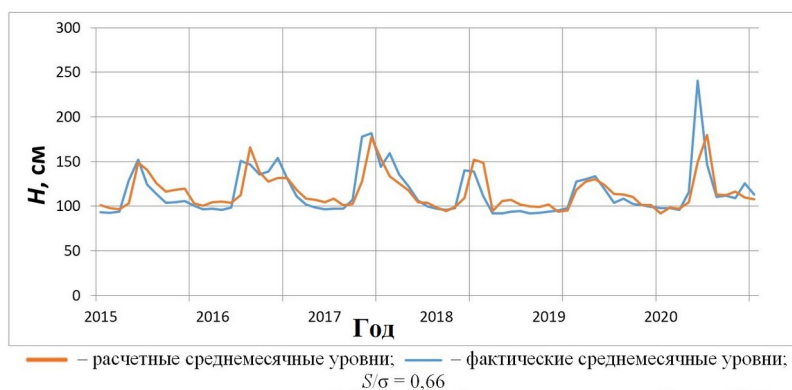


Рисунок 4 - График изменений фактических и расчетных уровней воды оз. Ловозеро с заблаговременностью 1 месяц
DOI: <https://doi.org/10.60797/IRJ.2024.145.67.7>

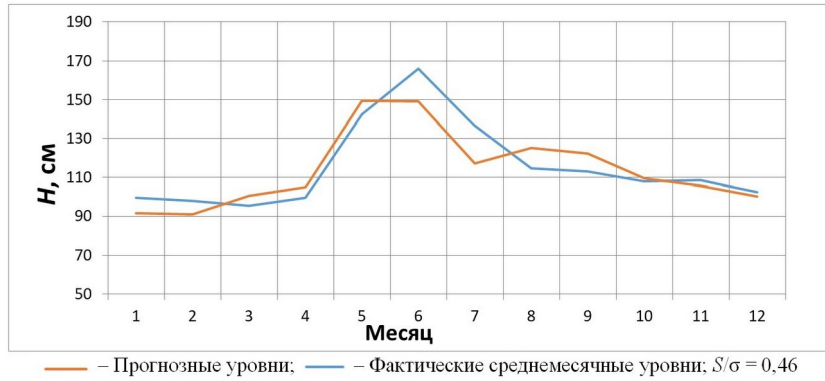


Рисунок 5 - График изменений фактических и прогнозных уровней воды оз. Ловозеро с заблаговременностью 1 месяц, 2021 год

DOI: <https://doi.org/10.60797/IRJ.2024.145.67.8>

Как видно из приведенных на графиках критериев S/σ , методика дает удовлетворительный результат как на зависимом, так и на независимом материале, при заблаговременности месяц с учетом выбранных предикторов.

3.2. Множественная линейная регрессия по стандартизированным данным

Для сравнения имеющегося набора данных и удобства их анализа применяется процесс стандартизации данных, в результате которого выводятся стандартизированные коэффициенты регрессии. Эти коэффициенты β_i используются для оценки силы влияния независимой переменной на зависимую.

Стандартизация рядов проводилась для отобранных предикторов и предиктанта. После анализа результатов сделан вывод, что процедура стандартизации не влияет на конечный результат и ей можно пренебречь. Однако, стандартизация выбранных параметров уравнения позволила выяснить следующее: при наборе таких гидрометеорологических характеристик как среднемесячный расход воды, общая облачность и облачность нижнего яруса влияние предиктора X_3 (облачность нижнего яруса) выражено сильнее, чем X_2 (общая облачность), при составлении прогнозного уравнения без учета заблаговременности. Обратная ситуация наблюдается при наличии заблаговременности в 1 месяц. Тогда влияние общей облачности X_2 несколько больше, чем облачности нижнего яруса X_3 .

Если же рассматривать уравнение, состоящее из пяти предикторов (среднемесячные расходы воды, среднемесячные атмосферные осадки и температура воздуха, общая облачность и облачность нижнего яруса), стандартизация показывает в обоих случаях (с заблаговременностью и без) доминирующее влияние облачности нижнего яруса над общей облачностью в описании процесса изменения уровня оз. Ловозеро.

Для удобства анализа критериев S/σ , которые послужили индексом качества выведенных уравнений регрессии, приведена в таблице 4.

Таблица 4 - Сводная таблица критериев S/σ для уравнений (4), (5)

DOI: <https://doi.org/10.60797/IRJ.2024.145.67.9>

$n \leq 15 \Rightarrow S/\sigma \leq 0,7$		S/σ
$n > 25 \Rightarrow S/\sigma \leq 0,85$		
Без заблаговременности	зависимый материал ($n=72$)	0,43
	поверочный прогноз ($n=12$)	0,70
Заблаговременность 1 месяц	зависимый материал ($n=72$)	0,66
	поверочный прогноз ($n=12$)	0,46

Дискуссия (обсуждение)

Анализируя полученные результаты (см. табл. 1), можно заметить, что «численный» вклад общей облачности и облачности нижнего яруса в изменение уровня оз. Ловозеро совершенно невелик. Однако в процессе перебора предикторов было установлено, что отсутствие этих характеристик сильно ухудшает ситуацию с оправдываемостью прогноза и другими критериями, оценивающими тесноту связи.

Ниже представлены графики хода гидрометеорологических величин, которые немного прояснили вопрос появления в регрессионных уравнениях таких предикторов как облачность нижнего яруса и общая облачность (см. рис. 6).

На рисунке 6 можно заметить, что реакция характеристик идет с некоторым запаздыванием. Сначала наблюдается формирование облачности, причем ход облачности нижнего яруса выражен лучше, чем общая облачность, а затем выпадают осадки, вызывая подъем уровня воды.

Однако, ход облачности достаточно «зашумлен». Чтобы понять, действительно ли есть цикличность в периодах высокой и низкой балльности облачности, была проведена статистическая фильтрация. Применялось, для сравнения, два вида фильтрации: скользящее среднее (СС) и экспоненциальное сглаживание.

Был выбран метод скользящего среднего из следующих соображений: данные по облачности предоставляются ВНИИГМИ-МЦД [16] в срочном формате, в сутках восемь сроков, и для получения месячного значения облачности эти данные осредняются. Также в пользу фильтра скользящего среднего по сравнению с экспоненциальным сглаживанием идет факт величины отклонений фактических значений облачности от спрогнозированных. В методе СС при различных комбинациях ширины окна сглаживания значения отклонений σ находятся в интервале $0 \leq \sigma \leq 2$. Тогда как при применении экспоненциального сглаживания при различных вариантах значения константы сглаживания отклонения σ варьируются от 0 до 3 ($0 \leq \sigma \leq 3$) и выше.

Для исследуемой метеорологической характеристики (облачность) экспериментальным путем [14] выбрана ширина окна $n=4$, поскольку показала наилучший результат сглаживания и дала наименьшие отклонения от фактических значений облачности.

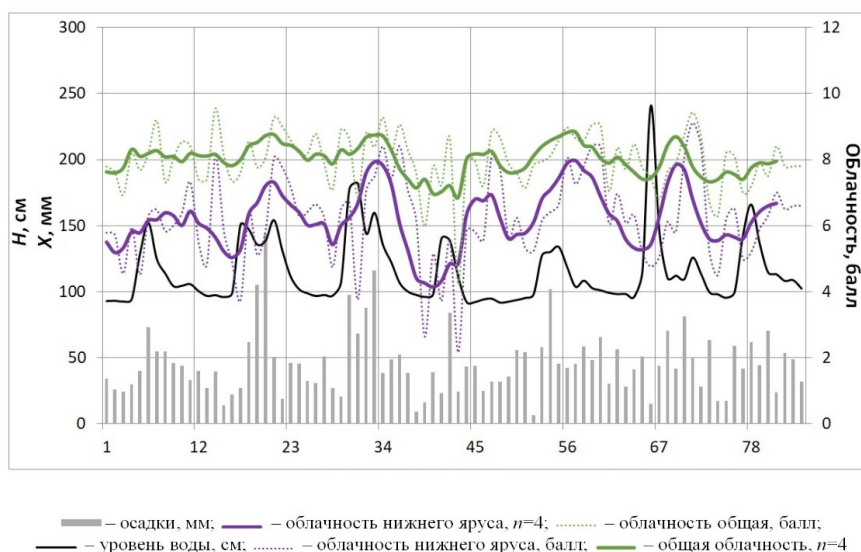


Рисунок 6 - Комплексный график хода гидрометеорологических характеристик
DOI: <https://doi.org/10.60797/IRJ.2024.145.67.10>

Из графиков на рисунке 6 видно, что наиболее выраженный ход и влияние на уровень воды оз. Ловозеро имеет облачность нижнего яруса, это же доказывает и проведенный выше регрессионный анализ.

Заключение

В результате использования регрессионного анализа для обоснования набора предикторов с целью получения достоверной регрессионной модели для описания и прогноза уровня режима водоема сделан следующий вывод: для оценки процесса изменения уровня озера Ловозеро с месячной дискретизацией будет достаточен учет трех нестандартизированных предикторов – среднемесячные расходы воды притока, общая облачность и облачность нижнего яруса.

Полученные регрессионные модели показывают удовлетворительные результаты с заблаговременностью один месяц. Для увеличения заблаговременности прогнозов необходимо учитывать дополнительные предикторы, описывающие циркуляцию атмосферы и другие характеристики водосборов [17].

Планируется расширенное исследование по комплексной оценке роли облачности при прогнозе изменений гидрологических характеристик на озерно-речных системах Карелии и Северо-Западного Федерального округа.

Финансирование

Работа выполнена в рамках госзадания Минобрнауки РФ № FSZU-2023-0002.

Конфликт интересов

Не указан.

Рецензия

Все статьи проходят рецензирование. Но рецензент или автор статьи предпочли не публиковать рецензию к этой статье в открытом доступе. Рецензия может быть предоставлена компетентным органам по запросу.

Funding

The work was carried out within the framework of the state assignment of the Ministry of Education and Science of the Russian Federation № FSZU-2023-0002.

Conflict of Interest

None declared.

Review

All articles are peer-reviewed. But the reviewer or the author of the article chose not to publish a review of this article in the public domain. The review can be provided to the competent authorities upon request.

Список литературы / References

1. Воронцов А. А. Применение физико-статистического метода определения главных компонент полей температуры воздуха в прогнозах погоды / А. А. Воронцов, С. Р. Степаненко, В. Н. Яхрюшин // Наука России: Цели и задачи. — Екатеринбург, 2017. — С. 51–56. DOI: 10.18411/sr-10-02-2017-3-13
2. Долгоносов Б. М. Спектры мощности речного стока в анализе эколого-гидрологических проблем / Б. М. Долгоносов, К. А. Корчагин // Институт водных проблем РАН. — URL: <http://www.iwp.ru/sites/files/iwp.ru/pub/18/2010-02-19/Publ%20Dolg-Korchl.pdf> (дата обращения 28.04.2024).
3. Рожков В. А. Векторный дисперсионный анализ изменчивости течений в Печорском море / В. А. Рожков, Н. А. Сухих // Проблемы Арктики и Антарктики. — 2017. — № 1(111). — С. 30–38.
4. Пяткин Ф. В. Применение регрессионного анализа для расчетов по данным приборов ИКФС-2 и МТВЗА спутника «Метеор-2» / Ф. В. Пяткин, В. В. Голомолзин, А. А. Косторная // Интерэкспо ГЕО Сибирь. — 2019. — Т. 4. — № 1. — С. 3–10. DOI: 10.33764/2618-981X-2019-4-1-3-10
5. Гайдукова Е. В. Долгосрочный прогноз стока рек АЗРФ с использованием регрессионных зависимостей / Е. В. Гайдукова, А. О. Яновская, М. С. Дрегваль [и др.] // Современные тенденции и перспективы развития гидрометеорологии в России. — Иркутск, 2023. — С. 59–67.
6. Состояние поверхностных вод // АИС ГМВО Автома-тизированная информационная система государственного мониторинга водных объектов. — URL: <https://gmvo.skniivh.ru/index.php?id=296> (дата обращения: 28.04.2024).
7. Мякишева Н. В. Многокритериальная классификация озер / Н. В. Мякишева. — СПб.: Гидрометеиздат, 2009. — 159 с.
8. Догановский А. М. Уровенный режим озер – интегральный показатель динамических биоценозов / А. М. Догановский. — СПб., 1994. — 348 с.
9. Малинин В. Н. Статистические методы анализа гидрометеорологической информации. Том 1 / В. Н. Малинин // Первичный анализ и построение эмпирических зависимостей. — СПб.: РГГМУ, 2020. — 256 с.
10. Малинин В. Н. Статистические методы анализа гидрометеорологической информации. Том 2 / В. Н. Малинин // Первичный анализ и построение эмпирических зависимостей. — СПб.: РГГМУ, 2020. — 196 с.
11. Базилевский М. П. Метод выпрямления искаженных из-за мультиколлинеарности коэффициентов в регрессионных моделях / М. П. Базилевский // Информатика и ее применения. — 2021. — Т. 15. — С. 60–65. DOI: 10.14357/19922264210209
12. Мокшина С. И. Метод построения содержательно интерпретируемых регрессионных моделей в условиях мультиколлинеарности / С. И. Мокшина, Г. В. Шуршикова, С. С. Щекунских // Современная экономика: проблемы и решения. — 2017. — № 5(89). — С. 81–94. DOI: 10.17308/meps.2017.5/1700
13. Гицба Я. В. Характеристика гидрометеорологических параметров акватории Сухумской бухты / Я. В. Гицба, Я. А. Экба // Доклады российского научно-технического общества радиотехники, электроники и связи им. А. С. Попова. — 2020. — С. 50–55.
14. Сикан А. В. Методы статистической обработки гидрометеорологической информации / А. В. Сикан. — СПб.: РГГМУ, 2007. — 279 с.
15. Георгиевский Ю. М. Гидрологические прогнозы / Ю. М. Георгиевский, С. В. Шаночкин. — СПб.: РГГМУ, 2007. — 436 с.
16. Федеральная служба по гидрометеорологии и мониторингу окружающей среды Всероссийский научно-исследовательский институт гидрометеорологической информации // Мировой центр данных. — URL: <http://meteo.ru/data/> (дата обращения: 28.04.2024).
17. Гайдукова Е. В. Учет испарения при математическом моделировании речного стока / Е. В. Гайдукова, С. В. Шаночкин, М. А. Москалюк // Ученые записки Российского государственного гидрометеорологического университета. — 2018. — № 52. — С. 79–87.

Список литературы на английском языке / References in English

1. Vorontsov A. A. Primenenie fiziko-statisticheskogo metoda opredeleniia glavnykh komponent polei temperatury vozdukh v prognozakh pogody [Application of a physical-statistical method for determining the main components of air temperature fields in weather forecasts] / A. A. Vorontsov, S. R. Stepanenko, V. N. Yakhryushin // Nauka Rossii: Celi i zadachi [Russian Science: Goals and objectives]. — Yekaterinburg, 2017. — P. 51–56. DOI: 10.18411/sr-10-02-2017-3-13 [in Russian]
2. Dolgonosov B. M. Spektry moshhnosti rechnogo stoka v analize jekologo-gidrologicheskikh problem [Power spectra of river flow in the analysis of ecological and hydrological problems] / B. M. Dolgonosov, K. A. Korchagin // Institut vodnyh problem RAN [Institute of Water Problems of the Russian Academy of Sciences]. — URL: <http://www.iwp.ru/sites/files/iwp.ru/pub/18/2010-02-19/Publ%20Dolg-Korchl.pdf> (accessed: 28.04.2024). [in Russian]
3. Rozhkov V. A. Vektornyj dispersionnyj analiz izmenchivosti techenij v Pechorskom more [Factorial dispersion analysis of the variability of currents in the Pechora Sea] / V. A. Rozhkov, N. A. Sukhykh // Problemy Arktiki i Antarktiki [Problems of the Arctic and Antarctic]. — 2017. — № 1(111). — P. 30–38. [in Russian]
4. Pyatkin F. V. Primenenie regressionnogo analiza dlja raschetov po dannym priborov IKFS-2 i MTVZA sputnika «Meteor-2» [Application of regression analysis for calculations based on the data of the IKFS-2 and MTVZA devices of the Meteor-2 satellite] / F. V. Pyatkin, V. V. Golomolzin, A. A. Kostornaya // Interjekspo GEO Sibir' [Interexpo GEO Siberia]. — 2019. — Vol. 4. — № 1. — P. 3–10. DOI: 10.33764/2618-981X-2019-4-1-3-10 [in Russian]
5. Gaidukova E. V. Dolgosrochnyj prognoz stoka rek AZRF s ispol'zovaniem regressionnykh zavisimostej [Long-term forecast of the flow of the rivers of the Russian Arctic using regression dependencies] / E. V. Gaidukova, A. O. Yanovskaya, M. S. Dregval [et al.] // Sovremennye tendencii i perspektivy razvitiya gidrometeorologii v Rossii [Current trends and prospects for the development of hydrometeorology in Russia]. — Irkutsk, 2023. — P. 59–67. [in Russian]

6. Sostojanie poverhnostnyh vod [Composition of ground waters] // AIS GMVO Avtoma-tizirovannaja informacionnaja sistema gosudarstvennogo monitoringa vodnyh objektov [AIS GMVO Automated information system for state monitoring of water bodies]. — URL: <https://gmvo.skniivh.ru/index.php?id=296> (accessed: 28.04.2024). [in Russian]
7. Myakisheva N. V. Mnogokriterial'naja klassifikacija ozer [Multicriteria classification of lakes] / N. V. Myakisheva. — St. Petersburg: Hydrometeoizdat, 2009. — 159 p. [in Russian]
8. Doganovsky A. M. Urovennyj rezhim ozer – integral'nyj pokazatel' dinamicheskikh biocenozov [The level regime of lakes is an integral indicator of dynamic biocenoses] / A. M. Doganovsky. — St. Petersburg, 1994. — 348 p. [in Russian]
9. Malinin V. N. Statisticheskie metody analiza gidrometeorologicheskoy informacii. Tom 1 [Statistical methods of analysis of hydrometeorological information. Vol. 1] / V. N. Malinin // Pervichnyj analiz i postroenie jempiricheskikh zavisimostej [Primary analysis and construction of empirical dependencies]. — St. Petersburg: RGGMU, 2020. — 256 p. [in Russian]
10. Malinin V. N. Statisticheskie metody analiza gidrometeorologicheskoy informacii. Tom 2 [Statistical methods of analysis of hydrometeorological information. Vol. 2] / V. N. Malinin // Pervichnyj analiz i postroenie jempiricheskikh zavisimostej [Primary analysis and construction of empirical dependencies]. — St. Petersburg: RGGMU, 2020. — 196 p. [in Russian]
11. Bazilevsky M. P. Metod vyprjamenija iskazhennyh iz-za mul'tikollinearosti koeficientov v regressionnyh modeljah [Method of rectification of coefficients distorted due to multicollinearity in regression models] / M. P. Bazilevsky // Informatika i ee primenija [Informatics and its applications]. — 2021. — Vol. 15. — P. 60–65. DOI: 10.14357/19922264210209 [in Russian]
12. Mokshina S. I. Metod postroenija sodержatel'no interpretiruemyh regressionnyh modelej v uslovijah mul'tikollinearosti [Method of constructing meaningfully interpreted regression models in multicollinearity conditions] / S. I. Mokshina, G. V. Shurshikova, S. S. Shchekunskikh // Sovremennaja jekonomika: problemy i reshenija [Modern Economics: problems and solutions]. — 2017. — № 5(89). — P. 81–94. DOI: 10.17308/meps.2017.5/1700 [in Russian]
13. Gitsba Ya. V. Harakteristika gidrometeorologicheskikh parametrov akvatorii Suhumskoj buhty [Characteristics of hydrometeorological parameters of the Sukhumi Bay water area] / Ya. V. Gitsba, Ya. A. Ekba // Doklady rossijskogo nauchno-tehnicheskogo obshhestva radiotekhniki, jelektroniki i svjazi im. A. S. Popova [Reports of the Russian Scientific and Technical Society of Radio Engineering, Electronics and Communications named after A. S. Popov]. — 2020. — P. 50–55. [in Russian]
14. Sikan A. V. Metody statisticheskoy obrabotki gidrometeorologicheskoy informacii [Methods of statistical processing of hydrometeorological information] / A. V. Sikan. — St. Petersburg: RGGMU, 2007. — 279 p. [in Russian]
15. Georgievsky Y. M. Gidrologicheskie prognozy [Hydrological forecasts] / Y. M. Georgievsky, S. V. Shanochkin. — St. Petersburg: RGGMU, 2007. — 436 p. [in Russian]
16. Federal'naja sluzhba po gidrometeorologii i monitoringu okruzhajushhej sredy Vserossijskij nauchno-issledovatel'skij institut gidrometeorologicheskoy informacii [Federal Service for Hydrometeorology and Environmental Monitoring All-Russian Scientific Research Institute of Hydrometeorological Information] // Mirovoj centr dannyh [World Data Center]. — URL: <http://meteo.ru/data/> (accessed: 28.04.2024). [in Russian]
17. Gaidukova E. V. Uchet isparenija pri matematicheskom modelirovanii rechnogo stoka [Accounting for evaporation in mathematical modeling of river flow] / E. V. Gaidukova, S. V. Shanochkin, M. A. Moskalyuk // Uchenye zapiski Rossijskogo gosudarstvennogo gidrometeorologicheskogo universiteta [Scientific notes of the Russian State Hydrometeorological University]. — 2018. — № 52. — P. 79–87. [in Russian]