

Непараметрические байесовские сети как инструмент комплексирования данных мультимасштабного анализа временных рядов и дистанционного зондирования

Н. С. Пыко^{1✉}, Д. В. Тишин^{1,2}, П. Ю. Искандиров^{1,2},
А. М. Гафуров², Б. М. Усманов², М. И. Богачев¹

¹Санкт-Петербургский государственный электротехнический университет "ЛЭТИ" им. В. И. Ульянова (Ленина), Санкт-Петербург, Россия

²Казанский (Приволжский) федеральный университет, Казань, Россия

✉goststalker13@gmail.com

Аннотация

Введение. Непараметрические байесовские сети представляют собой перспективный инструмент для анализа, визуализации, интерпретации и прогнозирования структурных и динамических характеристик сложных систем. Современные междисциплинарные исследования подразумевают комплексную обработку разнородных данных, получаемых с помощью датчиков различной физической природы. При исследовании лесного фонда широко применяются методы как непосредственных дендрологических измерений, так и дистанционного наблюдения с использованием беспилотных летательных аппаратов. Информацию, полученную с помощью этих методов, необходимо анализировать во взаимосвязи с данными гидрометеорологического мониторинга.

Цель работы. Исследование возможности автоматизации мониторинга благополучия лесного фонда на основе комплексирования данных наземных исследований, дистанционных мультиспектральных измерений и гидрометеорологических наблюдений с использованием математического аппарата непараметрических байесовских сетей.

Материалы и методы. Для оценки долговременной совместной динамики природно-климатических показателей и радиального прироста деревьев использован модифицированный метод мультимасштабного взаимного корреляционного анализа с удалением фонового тренда, описываемого моделью скользящего среднего. Взаимосвязи между различными показателями оценивались на основе безусловных и условных непараметрических коэффициентов корреляции Спирмена, которые использовались для реконструкции и параметризации непараметрической байесовской сети.

Результаты. Построена мультимасштабная непараметрическая байесовская сеть, характеризующая безусловные и условные статистические взаимосвязи между параметрами, полученными в результате дистанционного зондирования, гидроклиматических и дендрологических измерений. Предложенная модель показала хорошее качество прогнозирования состояния растительного фонда. Коэффициенты корреляции между наблюдаемыми и предсказываемыми показателями превышают значения 0.6, а при предсказании тренда прироста годичных колец деревьев коэффициент корреляции составляет 0.77.

Заключение. Предложенная непараметрическая байесовская сетевая модель отражает взаимосвязи между различными факторами, влияющими на лесную экосистему. Байесовская сеть может использоваться для оценки рисков и улучшения планирования экологического управления.

Ключевые слова: дистанционное зондирование, мультиспектральные вегетационные индексы, корреляционный анализ с удалением тренда, частные корреляции, направленный ациклический граф, непараметрическая байесовская сеть

Для цитирования: Непараметрические байесовские сети как инструмент комплексирования данных мультимасштабного анализа временных рядов и дистанционного зондирования / Н. С. Пыко, Д. В. Тишин, П. Ю. Искандиров, А. М. Гафуров, Б. М. Усманов, М. И. Богачев // Изв. вузов России. Радиоэлектроника. 2023. Т. 26, № 3. С. 32–47. doi: 10.32603/1993-8985-2023-26-3-32-37

Конфликт интересов. Авторы заявляют об отсутствии конфликта интересов.

Источник финансирования. Исследование выполнено за счет гранта Российского научного фонда № 22-76-10042, <https://rscf.ru/project/22-76-10042/>.

Статья поступила в редакцию 19.03.2023; принята к публикации после рецензирования 21.04.2023; опубликована онлайн 29.06.2023

Nonparametric Bayesian Networks as a Tool of Multiscale Time Series Analysis and Remote Sensing Data Integration

Nikita S. Pyko¹✉, Denis V. Tishin^{1,2}, Pavel Yu. Iskandirov^{1,2},
Artur M. Gafurov², Bulat M. Usmanov², Mikhail I. Bogachev¹

¹ Saint Petersburg Electrotechnical University, St Petersburg, Russia

² Kazan Federal University, Kazan, Russia

✉ goststalker13@gmail.com

Abstract

Introduction. Nonparametric Bayesian networks are a promising tool for analyzing, visualizing, interpreting and predicting the structural and dynamic characteristics of complex systems. Modern interdisciplinary research involves the complex processing of heterogeneous data obtained using sensors of various physical nature. In the study of the forest fund, both methods of direct dendrological measurements and methods of remote observation using unmanned aerial vehicles are widely used. Information obtained using these methods must be analyzed in conjunction with hydrometeorological monitoring data.

Aim. Investigation of the possibility of automating the monitoring of the well-being of the forest fund based on the integration of ground survey data, remote multispectral measurements and hydrometeorological observations using the mathematical apparatus of nonparametric Bayesian networks.

Materials and methods. To assess the long-term joint dynamics of natural and climatic indicators and the radial growth of trees, a modified method of multiscale cross-correlation analysis was used with the removal of the background trend described by the moving average model. Relationships between various indicators were estimated based on the unconditional and conditional nonparametric Spearman correlation coefficients, which were used to reconstruct and parameterize the nonparametric Bayesian network.

Results. A multiscale nonparametric Bayesian network was constructed to characterize both unconditional and conditional statistical relationships between parameters obtained from remote sensing, hydroclimatic and dendrological measurements. The proposed model showed a good quality of the plant fund state forecasting. The correlation coefficients between the observed and predicted indicators exceed 0.6, with the correlation coefficient comprising 0.77 when predicting the growth trend of annual tree rings.

Conclusion. The proposed nonparametric Bayesian network model reflects the relationship between various factors that affect the forest ecosystem. The Bayesian network can be used to assess risks and improve environmental management planning.

Keywords: remote sensing, multispectral vegetation indices, detrended cross-correlation analysis, partial correlations, directed acyclic graph, nonparametric Bayesian network

For citation: Pyko N. S., Tishin D. V., Iskandirov P. Yu., Gafurov A. M., Usmanov B. M., Bogachev M. I. Nonparametric Bayesian Networks as a Tool of Multiscale Time Series Analysis and Remote Sensing Data Integration. Journal of the Russian Universities. Radioelectronics. 2023, vol. 26, no. 3, pp. 32–47. doi: 10.32603/1993-8985-2023-26-3-32-47

Conflict of interest. The authors declare no conflicts of interest.

Acknowledgements. The study was supported by the Russian Science Foundation grant № 22-76-10042, <https://rscf.ru/project/22-76-10042/>.

Submitted 19.03.2023; accepted 21.04.2023; published online 29.06.2023

Введение. В последние десятилетия возрастает интерес к применению байесовских сетей (БС) для анализа и визуализации сложных взаимосвязей между случайными величинами. БС применяются в климатологии, сейсмологии, гидрологии, медицине, при определении

безопасности сложных инженерных систем и надежности инфраструктуры, а также во многих других областях знаний [1–5]. Подробный обзор особенностей построения и приложений БС приведен в [6, 7]. БС представляет собой направленный ациклический граф, узлы

Непараметрические байесовские сети как инструмент комплексирования данных

мультимасштабного анализа временных рядов и дистанционного зондирования

Nonparametric Bayesian Networks as a Tool of Multiscale Time Series Analysis

and Remote Sensing Data Integration

которого соответствуют анализируемым случайным величинам, а дуги (ребра графа) отражают зависимость между величинами. Между узлами сети на основании проведенных заранее исследований или экспертных оценок устанавливаются отношения "родитель – потомок". Для каждого узла без родителей задается безусловное распределение, а сила зависимости между узлами определяется условными распределениями. Отсутствие дуг между узлами соответствует ситуации условной независимости случайных величин, соответствующих узлам. Отражая причинно-следственные связи между переменными, БС позволяет прогнозировать состояние узлов сети, обеспечивая тем самым поддержку принятия решений.

В случае отсутствия априорной информации о функциональном виде безусловных распределений переменных, соответствующих узлам сети, используют непараметрические байесовские сети (НПБС). Особенности работы с НПБС в значительной степени посвящена [6]. В общем случае сила связи между переменными, расположенными в узлах сети, может быть выражена в терминах различных метрик связности, но основная роль в данном контексте отводится корреляционным метрикам. При построении НПБС, как правило, используют ранговые корреляции Спирмена [6, 7], причем преимущество отдается частным взаимным корреляциям, отражающим "собственные" связи переменных в условиях, когда влияние других переменных исключается [8].

В [9, 10] исследуется роль частной корреляции и условной корреляции как меры условной независимости двух случайных величин, а также устанавливаются необходимые и достаточные условия совпадения частной корреляции с условной корреляцией. Авторы [9, 10] показывают, что совпадение условной и частной корреляций правомерно не только для нормально распределенных случайных величин, но и в широком классе других законов распределения, в частности эллиптических, полиномиальных, гипергеометрических, отрицательных гипергеометрических распределений и распределения Дирихле. Условная независимость переменных является ключевым понятием при построении БС. Узлы графа считаются

связанными тогда и только тогда, когда соответствующие им переменные не являются условно независимыми. Чтобы подтвердить условную независимость, особенно когда переменные непрерывны, как правило, проверяют степень близости к нулю частных коэффициентов корреляции. Только в случае гауссовского распределения из некоррелированности переменных следует их независимость. При построении НПБС понятия условной независимости и нулевой частной корреляции не являются эквивалентными. Однако на практике целесообразнее заменить условную независимость нулевой или достаточно малой частной корреляцией. Тогда несвязанные вершины в графической модели можно, по крайней мере, считать ортогональными друг другу после того, как влияние других переменных будет удалено при переходе к частным корреляциям [7].

Недостатком использования частных корреляций при построении НПБС является тот факт, что значения частных коэффициентов корреляции зачастую очень малы. В результате создается впечатление незначительности выявленных закономерностей. Поэтому при построении БС для визуализации связей между узлами сети в отдельных случаях возможно использование и полных коэффициентов корреляции.

Взаимный корреляционный анализ с удалением тренда. Исследование взаимного поведения временных рядов традиционно проводится с помощью методов корреляционного анализа, который в классическом понимании требует стационарности изучаемых данных. В случае нестационарных данных, включая ситуацию, связанную с общим внешним воздействием на анализируемые переменные, предпочтительным является взаимный корреляционный анализ с удалением тренда. Применяемые на практике процедуры удаления тренда носят мультимасштабный характер, поскольку в обрабатываемых данных могут присутствовать тренды, имеющие различный спектральный состав [11–14].

Методы оценивания взаимной корреляции с удалением тренда предполагают центрирование m исходных временных рядов x^1, x^2, \dots, x^m , содержащих каждый по N

отсчетов, и формирование на их основе рядов кумулятивных сумм, называемых профилями:

$$X^j = \sum_{k=1}^i x_k^j; j=1, 2, \dots, m; i=1, 2, \dots, N.$$

Следующей операцией является удаление тренда. Одной из возможных процедур удаления тренда является описанный в [15] метод центрального скользящего среднего (Central Moving Average – CMA). В соответствии с данным методом тренд оценивается по всей записи профиля в окне скользящего среднего с нечетной длиной s .

Ряд остатков после удаления тренда при использовании метода CMA имеет вид

$$Y_s(i) = X_s^j(i) - CM_s^j(i); i=1, 2, \dots, N,$$

$$\text{где } CM_s^j(i) = \frac{1}{s} \sum_{k=-(s-1)/2}^{(s-1)/2} X^j(i+k).$$

На рис. 1 показана процедура выделения тренда для двух различных окон скользящего среднего размером s на примере синтезированного ряда случайных данных. Сплошной линией на рис. 1 показан график профиля, штриховая (красная) и штрихпунктирная (зеленая) линии отмечают тренд, полученный при использовании двух различных окон $s = 90$ и $s = 50$ соответственно.

Для оставшихся после удаления тренда флуктуационных составляющих профилей определяются дисперсии

$$F_s^2 = \frac{1}{N} \sum_{i=1}^N Y_s^2(i)$$

и попарные ковариации. Нормировка ковариационной матрицы позволяет получить матрицу коэффициентов взаимной корреляции:

$$R(s) = \begin{bmatrix} R_{1,1}(s) & R_{1,2}(s) & \dots & R_{1,m}(s) \\ R_{2,1}(s) & R_{2,2}(s) & \dots & R_{2,m}(s) \\ \dots & \dots & \dots & \dots \\ R_{m,1}(s) & R_{m,2}(s) & \dots & R_{m,m}(s) \end{bmatrix},$$

где $R_{j,j} = 1; j=1, 2, \dots, m$.

Переход к матрице частных коэффициентов корреляции осуществляется путем вычисления обратной матрицы коэффициентов взаимной корреляции [11]

$$C(s) = R^{-1}(s) = \begin{bmatrix} C_{1,1}(s) & C_{1,2}(s) & \dots & C_{1,m}(s) \\ C_{2,1}(s) & C_{2,2}(s) & \dots & C_{2,m}(s) \\ \dots & \dots & \dots & \dots \\ C_{m,1}(s) & C_{m,2}(s) & \dots & C_{m,m}(s) \end{bmatrix}$$

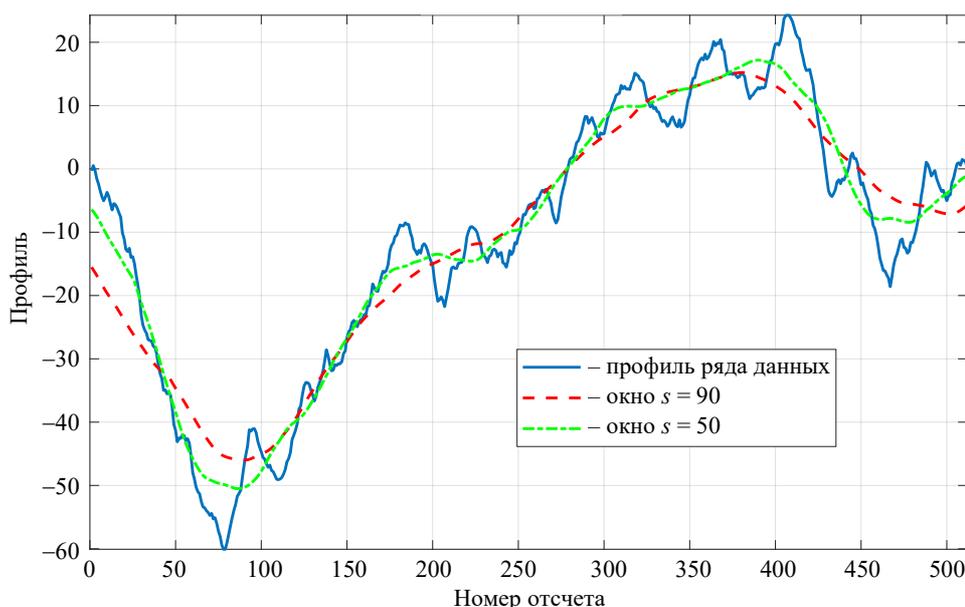


Рис. 1. Выделение тренда методом центрального скользящего среднего

Fig. 1. Trend identification by the method of the central moving average

с последующей нормировкой:

$$P_{j_1, j_2}(s) = \frac{-C_{j_1, j_2}(s)}{\sqrt{C_{j_1, j_1}(s) \cdot C_{j_2, j_2}(s)}}.$$

Мультимасштабность взаимного корреляционного анализа с удалением тренда достигается повторением указанных действий для различных размеров окна скользящего среднего s .

Мультиспектральное зондирование и определение дендрологических показателей.

В рамках данной статьи описанный ранее математический аппарат был применен к задаче экологического мониторинга, направленного на исследование выживаемости и роста сосны обыкновенной на участке сфагнового болота "Долгое" в Волжско-Камском государственном природном биосферном заповеднике. Изучаемый район имеет локальный градиент высот с запада на восток, приводящий к соответствующему градиенту местных гидрологических условий от относительно сухих до более влажных. В соответствии с этим в изучаемом районе были выделены три области – западная, центральная и восточная, обозначенные в связи с различными условиями увлажнения как Dry, Mid и Wet (сухая, средняя и влажная).

В зоне исследования было отобрано 101 дерево, для которых согласно методике, описанной в [16], была определена ширина годичных колец Tree Ring Width (TRW). Точное местоположение каждого изучаемого дерева определяли с помощью GPS-приемника Garmin GPSMAR 62S. Местоположение каждого дерева характеризовалось широтой (latitude – Lat) и долготой (longitude – Lon). Для выделенных деревьев определялись год рождения (Birth Year), среднее значение прироста (Mean growth), среднеквадратическое значение прироста (Std growth), а также тренд прироста годичных колец за последние 35 лет наблюдений (Growth trend), определяемый для каждого дерева с помощью линейного регрессионного анализа как разность ординат последней и первой точек линии регрессии, деленная на среднеквадратическое отклонение прироста годичных колец:

$$\Delta x = (x_{\max} - x_{\min}) / \sigma.$$

Мультиспектральное дистанционное зондирование выполнялось с помощью аэрофотосъемочного комплекса Геоскан 401 Геодезия, оснащенного мультиспектральной камерой MicaSense RedEdge-MX. Частотные диапазоны каналов камеры следующие: синий (Blue, 475 ± 32 нм), зелёный (Green, 560 ± 27 нм), красный (Red, 668 ± 16 нм), красный край (Red Edge, 717 ± 12 нм), ближний инфракрасный (NIR, 842 ± 57 нм).

В основе дистанционного зондирования природных объектов лежит возможность измерения электромагнитной энергии на различных длинах волн в процессе взаимодействия с изучаемыми объектами. В зависимости от физических свойств объекта при взаимодействии с ним электромагнитные волны разных частот в той или иной степени отражаются или поглощаются. Растительность взаимодействует с солнечным излучением иначе, чем другие природные объекты. Спектр растительности обычно отличается поглощением в красном и синем диапазонах, отражением в зеленом диапазоне и сильным отражением в ближнем инфракрасном диапазоне. Различные типы растительности, содержание воды, пигмента, углерода, азота и другие свойства объектов и окружающей их территории обуславливают различия, отражающиеся на спектральном составе электромагнитных колебаний. Измерение спектральных составляющих, изучение их взаимосвязи друг с другом, основанное на оценивании вегетационных индексов, могут предоставить важную информацию о здоровье растений, содержании воды, экологическом стрессе и других важных характеристиках.

Чтобы оценить местные гидрологические условия и связанную с ними активность растительности, согласно методике, приведенной в [17], были рассчитаны различные вегетационные индексы, из которых в результате были использованы три индекса, проявивших значимые отличия на анализируемых площадках:

$$NDVI = \frac{NIR - Red}{NIR + Red}$$

– нормализованный разностный вегетационный индекс (Normalized Difference Vegetation Index – NDVI);

$$NDRE = \frac{NIR - RedEdge}{NIR + RedEdge}$$

– нормализованный разностный индекс красного края (Normalized Difference Red Edge – NDRE);

$$NDWI = \frac{Green - NIR}{Green + NIR}$$

– нормализованный разностный водный индекс (Normalized Difference Water Index – NDWI).

Нормализованный разностный вегетационный индекс NDVI является наиболее известным и часто используемым индексом, характеризующим рассеяние зеленой растительной массы в ближнем инфракрасном диапазоне и поглощение хлорофилла в красном диапазоне. Высокая фотосинтетическая активность приводит к более низким значениям коэффициентов отражения в красной зоне спектра и большим значениям в ближней инфракрасной. Нормализованный разностный индекс красного края NDRE – это индекс вегетации для оценки состояния растительности с использованием крайнего красного канала, отвечающий за содержание хлорофилла и азота в листе. Нормализованный разностный водный индекс NDWI отражает изменения содержания влаги в растительном покрове, поскольку коэффициент отражения в зеленом и ближнем инфракрасном диапазонах имеет отличающиеся свойства по-

глощения воды. Значения рассмотренных вегетационных индексов варьируются от –1 до 1.

Общий вид исследованной области приведен на рис. 2. На рисунке местоположение каждого исследованного дерева выделено маркером, при этом разные цвета маркеров соответствуют деревьям, расположенным на различных площадках – западной, центральной и восточной. Отдельные изображения на рис. 2 соответствуют пяти каналам камеры и используемым вегетационным индексам.

Для оценивания мультиспектральных вегетационных индексов вокруг места произрастания каждого дерева была выделена область, в пять раз превышающая диаметр его наблюдаемой кроны. Для нежизнестойких деревьев и деревьев с истощенной кроной соответствующий диаметр кроны оценивался по диаметру ствола по приближенной модели [18]. В соответствующей области для каждого из анализируемых показателей были рассчитаны основные описательные статистики.

Для выявления статистически значимых различий вегетационных индексов был использован непараметрический критерий Крускала–Уоллиса, пороговый уровень значимости составлял 0.05. Статистически значимые расхождения между тремя характерными локальными участками изучаемой территории были выявлены как для средних значений, так и для среднеквадратических отклонений индексов NDVI, NDRE и NDWI.

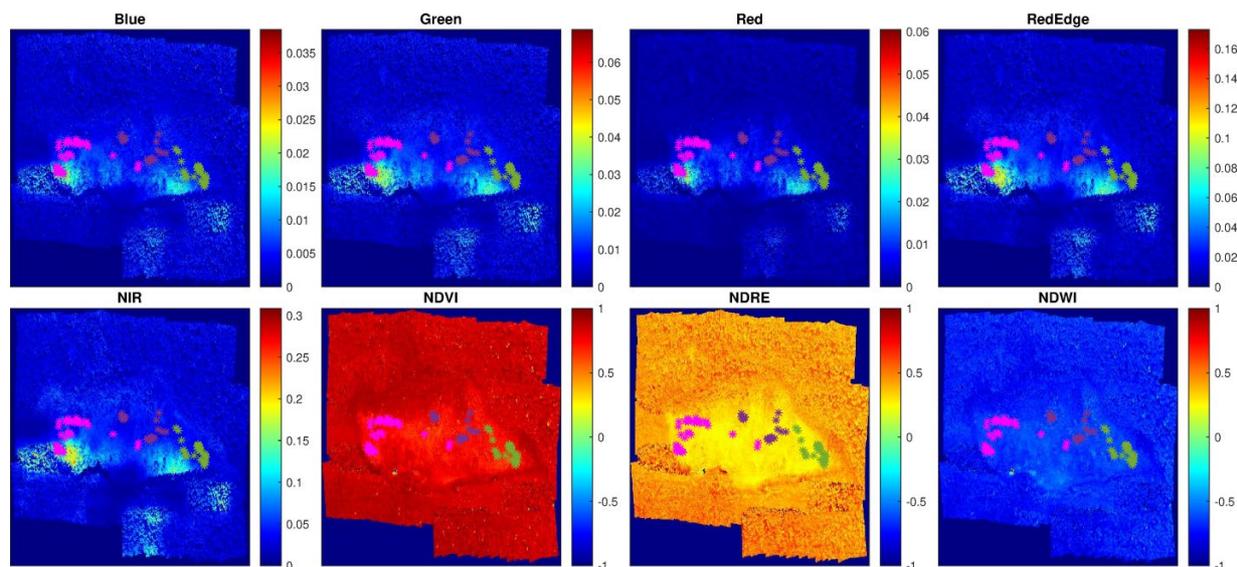


Рис. 2. Общий вид исследованной области при наблюдении с беспилотного летательного аппарата по данным мультиспектральной съемки

Fig. 2. General view of the studied area during unmanned aerial vehicle observation according to multispectral survey data

Взаимный корреляционный анализ прироста годичных колец и гидроклиматических показателей. Долгосрочные климатические записи и связанные с ними данные годичных колец демонстрируют сложное взаимодействие множества колебательных моделей и медленных трендов, что ограничивает применимость традиционных инструментов корреляционного и/или спектрального анализа.

С целью установления влияния климатических условий на прирост годичных колец деревьев был проведен мультимасштабный корреляционный анализ с удалением тренда методом СМА для временных рядов данных, соответствующих ширине годичных колец TRW, значениям индекса интенсивности засухи Палмера (Palmer Drought Severity Index) и температуры. Индекс интенсивности засухи Палмера относится к обобщенным индексам, определяющим баланс между потреблением и расходом влаги и учитывающим количество осадков, влажность почвы и воздуха, объем руслового стока, уровень подземных вод, высоту снежного покрова и иные показатели.

Размер окна скользящего среднего s , определяемый в годах, при выполнении процедуры удаления тренда в анализируемых рядах данных менялся от 2 до 128 лет. Анализировалась взаимосвязь прироста годичных колец с сезонными и годовыми рядами значений индекса засухи Палмера и температуры. Получаемые коэффициенты корреляции были обозначены, соответственно, PDSI и Temp. Результаты сравнения соответствующих коэффициентов корреляции по критерию Крускала–Уоллиса приведены на рис. 3 и 4. Порог уровня значимости был равен 0.05. Красными метками на рис. 3 и 4 отмечены уровни значимости, соответствующие статистически значимым отличиям между тремя участками изучаемой территории. Зелеными линиями на рис. 3 и 4 выделены совокупности боксовых диаграмм, для которых выявлены значимые отличия.

Как можно видеть из представленных на рис. 3 результатов, значимые отличия коэффициентов корреляции PDSI для трех площадок были определены для масштабов $s = 32, 64$ и 128 при использовании индекса засухи Палмера за каждый из сезонов (Winter, Spring, Sum-

mer, Autumn), так же, как и для годичных записей (Annual). Для масштаба $s = 4$ значимые отличия определены для записей, соответствующих осеннему и летнему сезонам.

Коэффициенты корреляции прироста годичных колец с температурой в зимний период для трех площадок значимо отличаются практически для всех масштабов s ; с температурой в весенний период – для масштабов $s = 64, s = 128$; с температурой осенью – для масштабов $s = 8, s = 16$.

Построение БС. Для создания БС все рассмотренные параметры были сначала объединены в три группы, включая глобальные гидроклиматические вариации (А), характеризующиеся коэффициентами корреляции PDSI и температуры, локальные гидрологические условия (Б), характеризующиеся мультиспектральными индексами, и условия и параметры роста деревьев (С), характеризующиеся данными годичных колец.

Чтобы уменьшить сложность модели, из нее были исключены все избыточные связи между параметрами с помощью трехэтапной процедуры. Сперва были отобраны только те параметры, для которых были выявлены значимые отличия в пределах трех характерных частей изучаемой территории. При построении графа на основании используемой корреляционной матрицы полных или частных корреляций выбирался порог, ниже которого корреляционные связи признавались несущественными. Затем некоторые связи были запрещены экспертной оценкой, чтобы учесть, что глобальные индексы (А) могут влиять на локальные условия (Б) и на рост и самочувствие деревьев (С), но не наоборот. Напротив, внутригрупповые связи допускались. Например, годовой прирост напрямую приводит к увеличению диаметра ствола, в то время как более старые деревья с большим диаметром ствола могут в определенные периоды жизни снижать скорость своего роста. На третьем этапе для создания графа, представляющего структуру БС, которая по определению подразумевает ориентированный ациклический граф, из модели были исключены оставшиеся петли. Для этого повторялась процедура, состоящая в нахождении любого существующего цикла и исключении в нем самого "слабого" звена до тех пор, пока результирующий граф не оказывался полностью ациклическим.

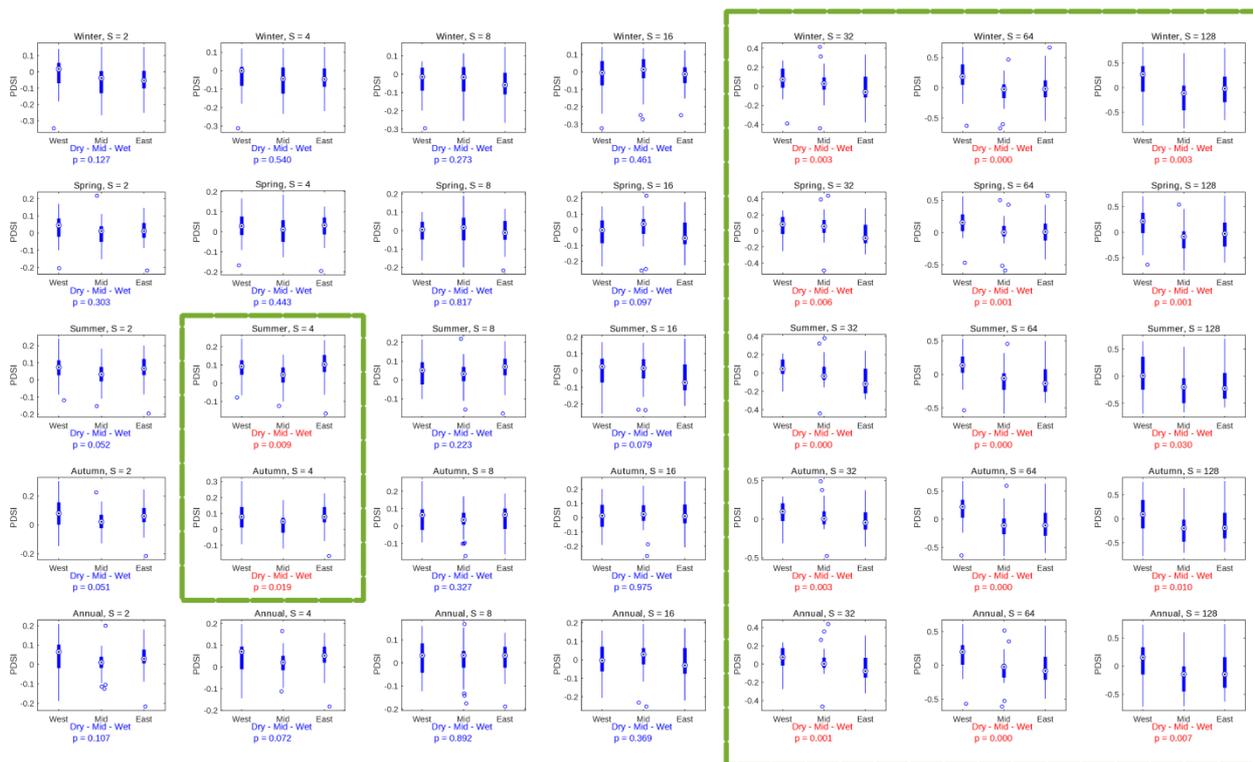


Рис. 3. Боксовые диаграммы коэффициентов корреляции прироста годовых колец с сезонными и годовыми данными индекса засухи Палмера

Fig. 3. Boxplots of the correlation coefficients of the tree rings growth and the seasonal and annual data of the Palmer drought index

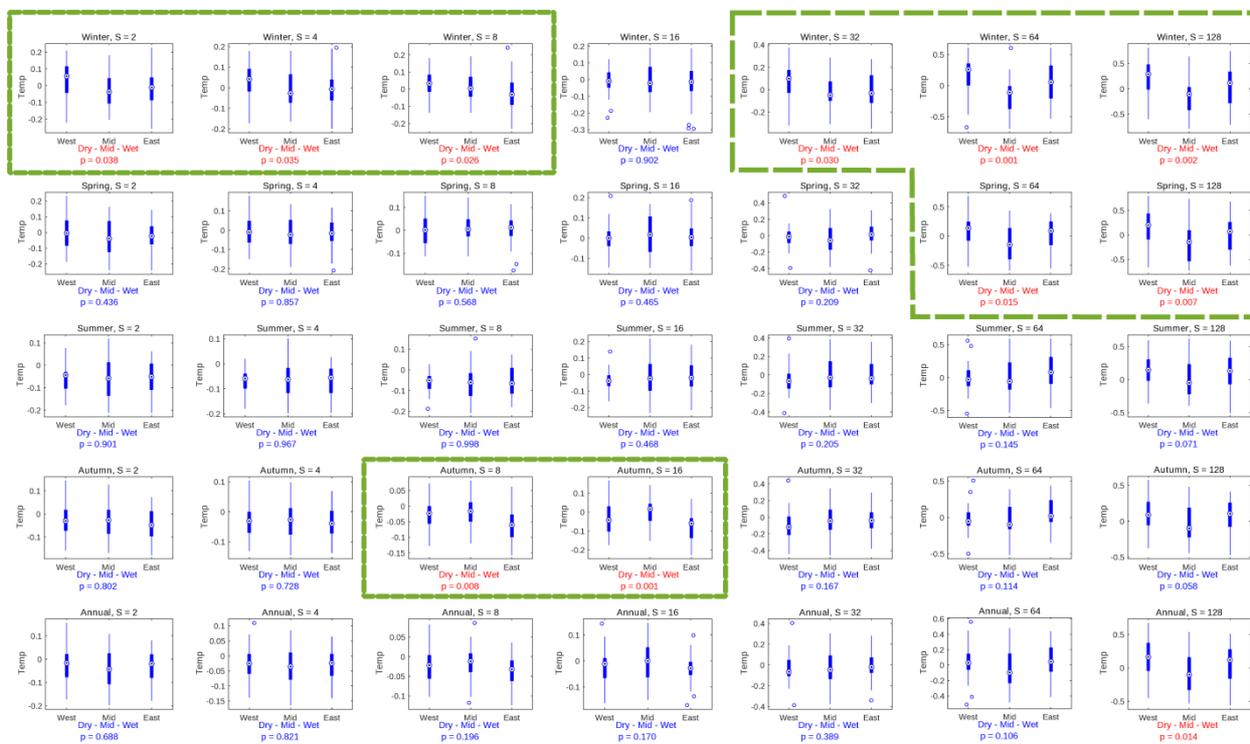


Рис. 4. Боксовые диаграммы коэффициентов корреляции прироста годовых колец с сезонными и годовыми данными температуры

Fig. 4. Boxplots of the correlation coefficients of the tree rings growth and the seasonal and annual data of the temperature

Хотя байесовские сетевые модели представляют собой распространенный инструмент для описания сложных взаимодействий в гетерогенных природных системах (см., например, [2, 7]), большинство доступных решений сосредоточено на полностью экспертно управляемом определении структуры сети. Напротив, быстро расширяющиеся и все более эффективные (хотя обычно требующие большого количества данных наблюдений для их начального обучения) подходы, основанные на машинном обучении, обычно фокусируются на полностью управляемых данными структурах, которые довольно часто оказываются трудно интерпретируемыми. Описанная трехэтапная процедура построения модели остается в промежуточной нише, поскольку она работает в терминах

обычных величин, допуская при этом ограниченное вмешательство экспертов в основном для устранения бессмысленных причинно-следственных связей.

На рис. 5 представлена БС, отражающая взаимосвязь рассмотренных параметров. Сеть параметризована коэффициентами корреляции Спирмена. Предполагаемые причинно-следственные связи направлены сверху вниз. Матрица коэффициентов корреляции Спирмена приведена на рис. 6.

Иной вариант построения графа можно получить, используя матрицу частных коэффициентов корреляции (рис. 7, 8). Предполагаемые причинно-следственные связи для графа, изображенного на рис. 7, направлены сверху вниз. Отличие графа на рис. 7 от графа, представлен-

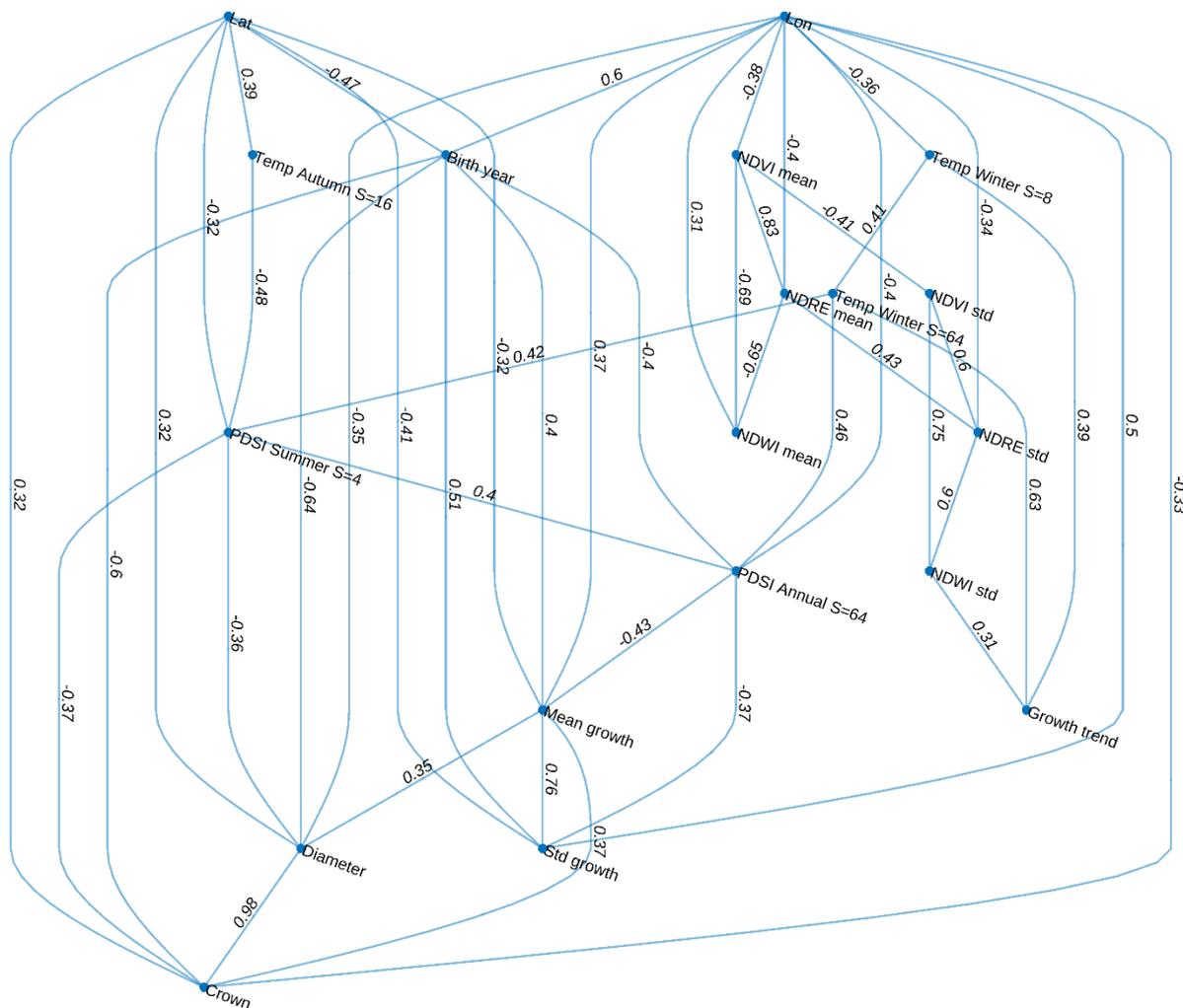


Рис. 5. Байесовская сеть, параметризованная корреляциями Спирмена между индексами гидроклиматической изменчивости, мультиспектральными индексами местной растительности и данными годичных колец

Fig. 5. Bayesian network parameterized by Spearman correlations between hydroclimatic variability indices, multispectral local vegetation indices and tree ring data

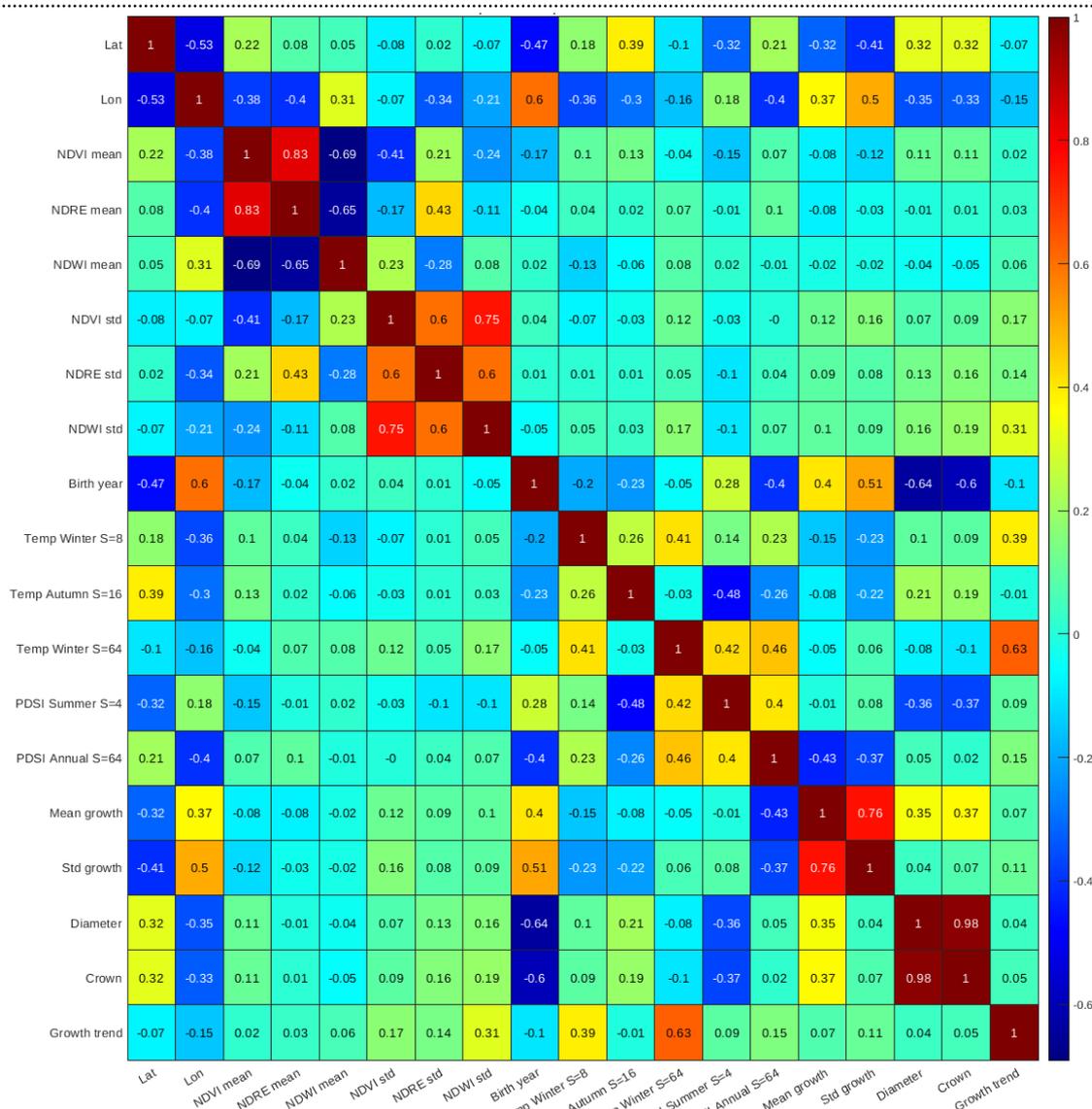


Рис. 6. Матрица коэффициентов корреляции Спирмена между рассмотренными параметрами

Fig. 6. Matrix of Spearman correlation coefficients between the considered parameters

ного на рис. 5, вызвано в первую очередь тем, что коэффициенты частной корреляции, как правило, имеют существенно меньшие значения, чем коэффициенты полной корреляции, за счет устранения общих влияний на анализируемые переменные. Например, когда на несколько параметров роста дерева влияют глобальный индекс гидроклиматических колебаний, такой как индекс засухи Палмера, и (или) вариации температуры, их влияние вносит дополнительный вклад в суммарные корреляции между всеми этими параметрами. При переходе к частным корреляциям этот дополнительный вклад исключается. Оставшиеся "внутренние" корреляции после устранения множественных перекрестных влияний зачастую уменьшаются,

и ими можно пренебречь без существенной потери точности модели. При этом структура графа изменяется.

Существенным достоинством БС является возможность их использования для прогнозирования показателей, представляющих интерес для исследователей. На рис. 9 приведены результаты прогнозирования диаметра ствола (Trunk diameter), диаметра кроны (Crown diameter) и тренда прироста годичных колец (Growth trend) на основе данных дистанционного зондирования и коэффициентов корреляции прироста годичных колец с индексом засухи Палмера (PDSI) и температурой (Temp). Для прогнозирования использована БС (рис. 5), однако данные среднего прироста (Mean growth)

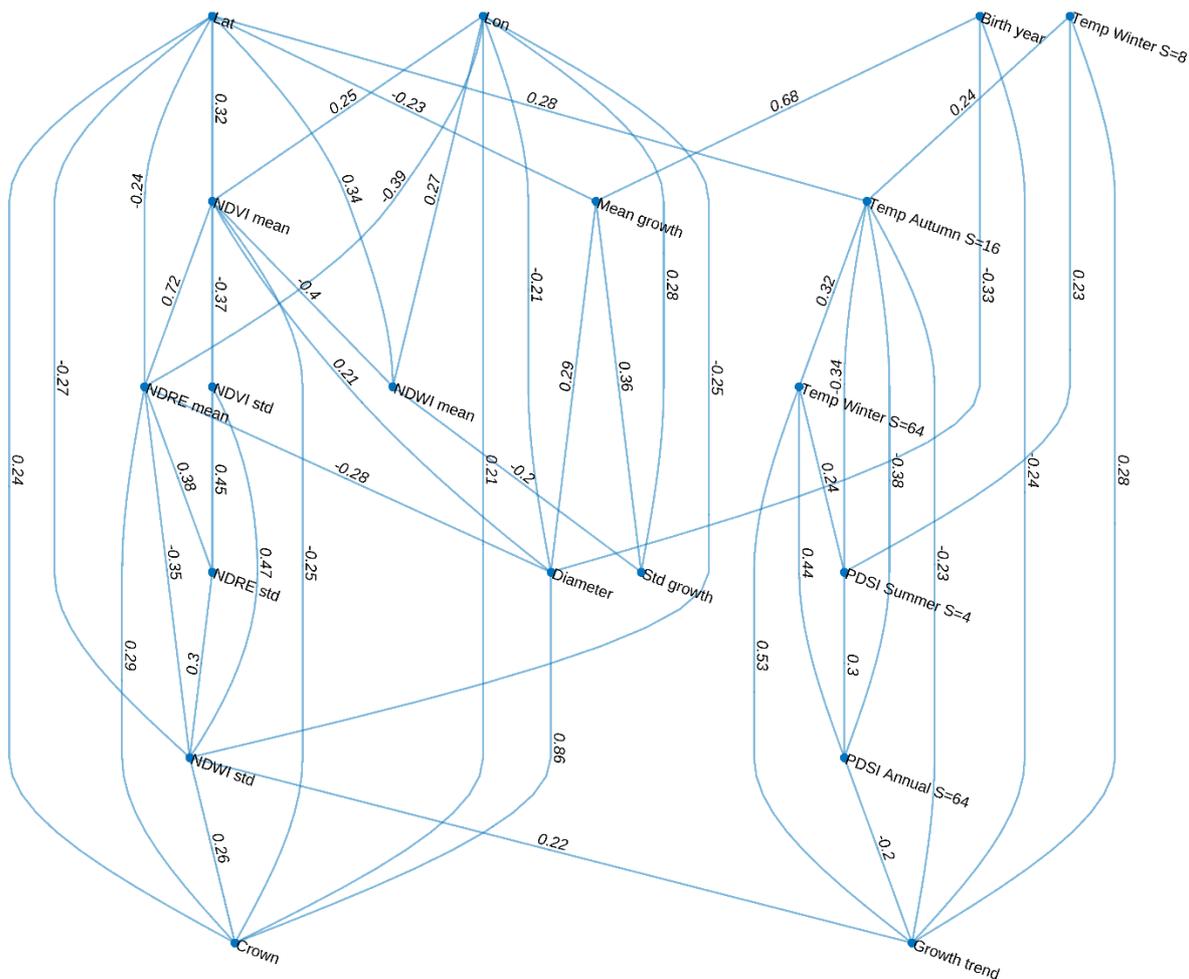


Рис. 7. Байесовская сеть, параметризованная частными корреляциями Спирмена между индексами гидроклиматической изменчивости, мультиспектральными индексами местной растительности и данными годовичных колец

Fig. 7. Bayesian network parameterized by partial Spearman correlations between hydroclimatic variability indices, multispectral local vegetation indices and tree ring data

и среднеквадратического отклонения прироста годовичных колец (Std growth) исключены из числа предикторов, поскольку их влияние на рассматриваемые целевые показатели очевидно. Для улучшения регрессионной модели удалены выбросы, к которым отнесены значения, превышающие три медианных абсолютных отклонения. На рис. 9 синими маркерами показаны значения, абсциссы которых соответствуют наблюдаемым показателям, а ординаты – их предсказываемым значениям. Красными маркерами обозначены выбросы. Синие штриховые линии соответствуют линиям регрессии до удаления выбросов, красные линии – линиям регрессии после удаления выбросов. В верхней части рисунка приведены полученные значения коэффициентов корреляции реальных и прогнозируемых значений до и после удаления выбросов. Приведенные значения коэффициен-

тов корреляции представляют собой квадратный корень из коэффициентов детерминации, характеризующих качество регрессии.

Как можно видеть, полученные коэффициенты корреляции превышают значения 0.6, а в случае анализа тренда прироста годовичных колец стремятся к значению 0.8, что позволяет считать качество прогнозирования вполне приемлемым.

Заклучение. Изменение климата является одной из глобальных проблем современности. Лесные экосистемы вносят определяющий вклад в локальный микроклимат, снижая как абсолютный нагрев, так и размах суточных колебаний температуры, и, таким образом, в значительной мере определяют сбалансированность окружающей среды и благополучие экотона в целом. Представленная работа направлена на частичную автоматизацию мониторинга

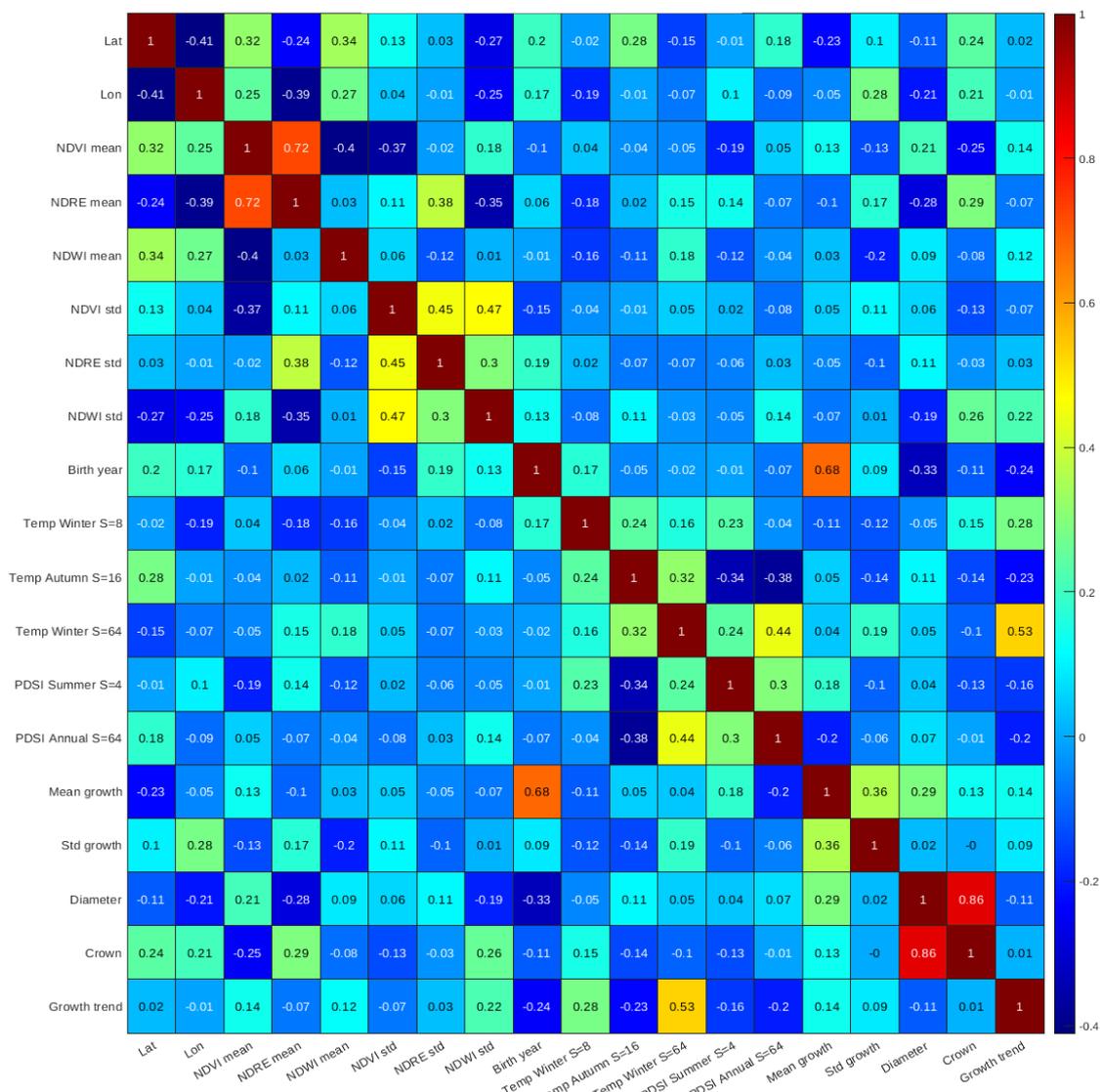


Рис. 8. Матрица частных коэффициентов корреляции Спирмена между рассмотренными параметрами
 Fig. 8. Matrix of partial Spearman correlation coefficients between the considered parameters

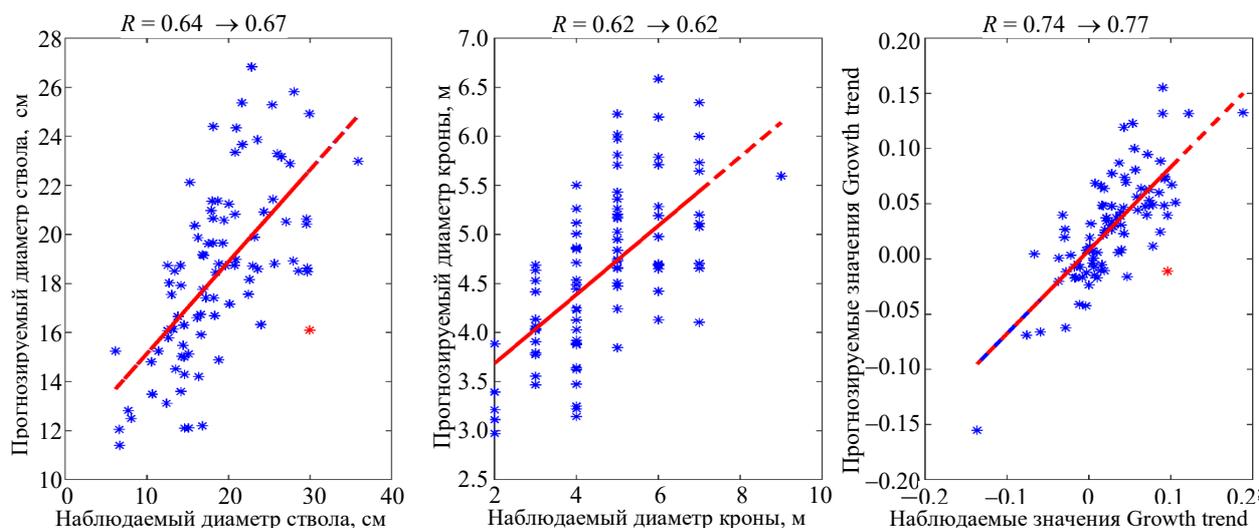


Рис. 9. Линейные регрессии, отражающие связь наблюдаемых и прогнозируемых показателей

Fig. 9. Linear regressions showing the relationship between observed and predicted indicators

благополучия лесного фонда на примере исследования территории сфагнового болота и анализа роста сосны обыкновенной.

Параметры, характеризующие рост и текущее состояние деревьев в области анализа, были получены с помощью дендрологических измерений. Для выбранного региона исследований характерны риски усыхания деревьев вследствие продолжительных засух. Поэтому многолетние данные прироста годичных колец деревьев были рассмотрены в комплексе с данными изменения температуры и индекса засухи Палмера. Для выявления взаимосвязи указанных показателей был проведен мультимасштабный взаимный корреляционный анализ с удалением тренда.

На основании мультиспектрального дистанционного зондирования исследуемой территории были определены вегетационные индексы, характеризующие локальные гидрологические условия.

Проведенный статистический анализ позволил выделить показатели, значимо отличающиеся для трех площадок в зоне исследования,

отличающихся градиентом высот и степенью увлажнения. Для выявленного набора показателей были построены НПБС, параметризованные как полными, так и частными коэффициентами корреляции Спирмена. Построенные НПБС могут быть использованы как для анализа и визуализации, так и для прогнозирования состояния лесных насаждений.

Полученные результаты показывают, что слабый контраст увлажненности в зонах исследований тем не менее явно отразился на динамике роста деревьев, включая изменение их связи с жарой и уровнем засухи. Это подчеркивает важность местных гидрологических условий для сохранения и благополучия лесных экосистем при повышенной нагрузке на окружающую среду и должно учитываться при планировании экологического управления. Предложенная байесовская сетевая модель отражает взаимосвязи между различными факторами, влияющими на лесную экосистему, и может использоваться для оценки рисков и улучшения планирования экологического управления.

Авторский вклад

Пыко Никита Сергеевич – выполнение статистического анализа данных, разработка математической модели, визуализация и статистическая оценка значимости полученных результатов.

Тишин Денис Владимирович – планирование и выполнение экспериментальных работ, дендрологическая и гидроклиматическая характеристика участка исследований, анализ годичного прироста деревьев.

Искандиров Павел Юрьевич – выполнение наземных дендрологических исследований.

Гафуров Артур Маратович – планирование и первичная обработка данных дистанционного зондирования.

Усманов Булат Мансурович – выполнение дистанционного зондирования с использованием БПЛА.

Богачев Михаил Игоревич – разработка концепции исследований и интерпретация результатов.

Author's contribution

Nikita S. Pyko, implementation of the statistical data analysis, development of the mathematical model, visualization and statistical evaluation of the significance of the obtained results.

Denis V. Tishin, planning and implementation of experimental work, dendrological and hydroclimatic description of the study area, analysis of the annual growth of trees.

Pavel Yu. Iskandirov, implementation of the ground-based dendrological studies.

Artur M. Gafurov, planning and primary processing of remote sensing data.

Bulat M. Usmanov, remote sensing using UAVs.

Mikhail I. Bogachev, development of the research concept and interpretation of the results.

Список литературы

1. Monitoring cliff erosion with LiDAR surveys and Bayesian network-based data analysis / P. Terefenko, D. Paprotny, A. Giza, O. Morales-Nápoles, A. Kubicki, S. Walczakiewicz // Remote Sens. 2019. № 11. P. 843. doi: 10.3390/rs11070843

2. Paprotny D., Morales-Nápoles O. Estimating extreme river discharges in Europe through a Bayesian network // Hydrology and Earth System Sciences.

2017. Vol. 21, № 6. P. 2615–2636. doi: 10.5194/hess-21-2615-2017

3. A continuous Bayesian network for earth dams' risk assessment: An application / D.-J. Delgado-Hernández, O. Morales-Nápoles, D. De-León-Escobedo, J.-C. Arteaga-Arcos // Structure and Infrastructure Engineering. 2014. Vol. 10, № 2. P. 225–238. doi: 10.1080/15732479.2012.731416

4. Morales Nápoles O., Steenbergen R. Analysis of axle and vehicle load properties through Bayesian networks based on weigh-in-motion data // Reliability Engineering & System Safety. 2014. Vol. 125. P. 153–164. doi: 10.1016/j.res.2014.01.018

5. Cooke R. M., Wielicki B. Probabilistic reasoning about measurements of equilibrium climate sensitivity: Combining disparate lines of evidence // Climatic Change. 2018. № 151. P. 541–154. doi: 10.1007/s10584-018-2315-y

6. Overview on Bayesian Network applications for dependability, risk analysis and maintenance areas / P. Weber, G. Medina-Oliva, C. Simon, B. Iung // Engineering Applications of Artificial Intelligence. 2012. Vol. 25, № 4. P. 671–682. doi: 10.1016/j.engappai.2010.06.002

7. Hanea A., Morales Nápoles O., Ababei D. Non-parametric Bayesian networks: Improving theory and reviewing applications // Reliability Engineering & System Safety. 2015. Vol. 144. P. 265–284. doi: 10.1016/j.res.2015.07.027

8. Morales O., Kurowicka D., Roelen A. Eliciting conditional and unconditional rank correlations from conditional probabilities // Reliability Engineering & System Safety. 2008. Vol. 93, № 5. P. 699–710. doi: 10.1016/j.res.2007.03.020

9. Baba K., Shibata R., Sibuya M. Partial correlation and conditional correlation as measures of conditional independence // Australian & New Zealand J. of Statistics. 2004. Vol. 46, № 4. P. 657–664. doi: 10.1111/j.1467-842X.2004.00360.x

10. Baba K., Sibuya M. Equivalence of Partial and Conditional Correlation Coefficients // J. of the Japan Statistical Society. 2005. Vol. 35, № 1. P. 1–19. doi: 10.14490/JJSS.35.1

11. Detrended partial-cross-correlation analysis: a new method for analyzing correlations in complex

system / N. Yuan, Z. Fu, H. Zhang, L. Piao, J. Luterbacher // Scientific reports. 2015. Vol. 5, № 1. P. 8143. doi: 10.1038/srep08143

12. Detrended Partial Cross-Correlation Analysis of Two Nonstationary Time Series Influenced by Common External Forces / X.-Yu. Qian, Y.-M. Liu, Zh.-Q. Jiang, B. Podobnik, W.-X. Zhou, H. E. Stanley // Physical Review. 2015. Vol. 91. P. 06281. doi: 10.1103/PhysRevE.91.062816

13. Zhou W. X. Multifractal detrended cross-correlation analysis for two nonstationary signals // Physical Review E. 2008. Vol. 77. P. 066211. doi: 10.1103/PhysRevE.77.066211

14. Horvatic D., Stanley H. E., Podobnik B. Detrended cross-correlation analysis for non-stationary time series with periodic trends // Europhysics Lett. 2011. Vol. 94, № 1. P. 18007. doi: 10.1209/0295-5075/94/18007

15. Alvarez-Ramirez J., Rodriguez E., Echeverría J. C. Detrending fluctuation analysis based on moving average filtering // Physica A: statistical mechanics and its applications. 2005. Vol. 354. P. 199–219. doi: 10.1016/j.physa.2005.02.020

16. Rinn F. TSAP-Win time series analysis and presentation for dendrochronology and related applications // User Reference Version 0.53. Heidelberg: Rinntech, 2005. P. 1–88.

17. Article Evaluating Multispectral Images and Vegetation Indices for Precision Farming Applications from UAV Images / S. Candiago, F. Remondino, M. De Giglio, M. Dubbini, M. Gettelli // Remote Sens. 2015. Vol. 7, № 4. P. 4026–4047. doi: 10.3390/rs70404026

18. Pyataev A. S., Vais A. A. Pine Crown and Trunk Diameter Dependence Research // CEUR Workshop Proc. 2019. P. 160–165.

Информация об авторах

Пыко Никита Сергеевич – магистр по направлению "Инфокоммуникационные технологии и системы связи" (СПбГЭТУ "ЛЭТИ", 2019), аспирант кафедры радиотехнических систем, младший научный сотрудник научно-образовательного центра "Цифровые телекоммуникационные системы" Санкт-Петербургского государственного электротехнического университета "ЛЭТИ" им. В. И. Ульянова (Ленина). Автор 36 научных работ. Сфера научных интересов – статистический анализ данных; математическое моделирование.

Адрес: Санкт-Петербургский государственный электротехнический университет "ЛЭТИ" им. В. И. Ульянова (Ленина), ул. Профессора Попова, д. 5 Ф, Санкт-Петербург, 197022, Россия

E-mail: goststalker13@gmail.com

<https://orcid.org/0000-0002-6668-9512>

Тишин Денис Владимирович – кандидат биологических наук (2006), доцент кафедры общей экологии Института экологии и природопользования Казанского (Приволжского) федерального университета, старший научный сотрудник Информационно-методического центра факультета компьютерных технологий и информатики Санкт-Петербургского государственного электротехнического университета "ЛЭТИ" им. В. И. Ульянова (Ленина). Автор 62 научных работ. Сфера научных интересов – сегнетоэлектрики; электрокалорический эффект; пирозлектрический эффект; мультиферроики.

Адрес: Казанский (Приволжский) федеральный университет, ул. Кремлевская, д. 18/1, Казань, 420008, Россия

E-mail: kpfuecology@gmail.com

<https://orcid.org/0000-0002-4790-2840>

Искандиров Павел Юрьевич – эколог по специальности "Экология" (КФУ, 2013), аспирант кафедры общей экологии Института экологии и природопользования Казанского (Приволжского) федерального университета. Автор 12 научных работ. Сфера научных интересов – дендрохронология; фенология.

Непараметрические байесовские сети как инструмент комплексирования данных

Мультимасштабного анализа временных рядов и дистанционного зондирования

Nonparametric Bayesian Networks as a Tool of Multiscale Time Series Analysis

and Remote Sensing Data Integration

Адрес: Казанский (Приволжский) федеральный университет, ул. Кремлевская, д. 18/1, Казань, 420008, Россия
E-mail: iskandirovpj@stud.kpfu.ru
<https://orcid.org/0009-0004-3824-6708>

Гафуров Артур Маратович – старший преподаватель кафедры ландшафтной экологии Института экологии и природопользования, старший научный сотрудник Научно-исследовательского центра превосходства киберфизических систем Института физики Казанского (Приволжского) федерального университета. Автор 50 научных работ. Сфера научных интересов – геоморфология и эволюционная география; геоэкология; аэрокосмические исследования Земли; фотограмметрия.

Адрес: Казанский (Приволжский) федеральный университет, ул. Кремлевская, д. 18/1, Казань, 420008, Россия
E-mail: AMGafurov@kpfu.ru
<https://orcid.org/0000-0002-0812-1750>

Усманов Булат Мансурович – старший преподаватель кафедры ландшафтной экологии Института экологии и природопользования Казанского (Приволжского) федерального университета. Автор 72 научных работ. Сфера научных интересов – геоморфология и эволюционная география; геоэкология; аэрокосмические исследования Земли; фотограмметрия.

Адрес: Казанский (Приволжский) федеральный университет, ул. Кремлевская, д. 18/1, Казань, 420008, Россия
E-mail: BUsmanof@kpfu.ru
<https://orcid.org/0000-0001-9354-1348>

Богачев Михаил Игоревич – доктор технических наук (2018), доцент (2011) кафедры радиотехнических систем Санкт-Петербургского государственного электротехнического университета "ЛЭТИ" им. В. И. Ульянова (Ленина). Автор 200 научных работ. Сфера научных интересов – статистический анализ данных; математическое моделирование.

Адрес: Санкт-Петербургский государственный электротехнический университет "ЛЭТИ" им. В. И. Ульянова (Ленина), ул. Профессора Попова, д. 5 Ф, Санкт-Петербург, 197022, Россия
E-mail: rogex@yandex.com
<https://orcid.org/0000-0002-0356-5651>

References

1. Terefenko P., Paprotny D., Giza A., Morales-Nápoles O., Kubicki A., Walczakiewicz S. Monitoring Cliff Erosion with LiDAR Surveys and Bayesian Network-Based Data Analysis. *Remote Sens.* 2019, no. 11, p. 843. doi: 10.3390/rs11070843
2. Paprotny D., Morales-Nápoles O. Estimating Extreme River Discharges in Europe Through a Bayesian Network. *Hydrology and Earth System Sciences.* 2017, vol. 21, no. 6, pp. 2615–2636. doi: 10.5194/hess-21-2615-2017
3. Delgado-Hernández D.-J., Morales-Nápoles O., De-León-Escobedo D., Arteaga-Arcos J.-C. A Continuous Bayesian Network for Earth Dams' Risk Assessment: An Application. *Structure and Infrastructure Engineering.* 2014, vol. 10, no. 2, pp. 225–238. doi: 10.1080/15732479.2012.731416
4. Morales Nápoles O., Steenbergen R. Analysis of Axle and Vehicle Load Properties through Bayesian Networks Based on Weigh-in-Motion Data. *Reliability Engineering & System Safety.* 2014, vol. 125, pp. 153–164. doi: 10.1016/j.res.2014.01.018
5. Cooke R. M., Wielicki B. Probabilistic Reasoning about Measurements of Equilibrium Climate Sensitivity: Combining Disparate Lines of Evidence. *Climatic Change.* 2018, no. 151, pp. 541–154. doi: 10.1007/s10584-018-2315-y
6. Weber P., Medina-Oliva G., Simon C., Iung B. Overview on Bayesian Network Applications for Dependability, Risk Analysis and Maintenance Areas. *Engineering Applications of Artificial Intelligence.* 2012, vol. 25, no. 4, pp. 671–682. doi: 10.1016/j.engappai.2010.06.002
7. Hanea A., Morales Nápoles O., Ababei D. Non-parametric Bayesian Networks: Improving Theory and Reviewing Applications. *Reliability Engineering & System Safety.* 2015, vol. 144, pp. 265–284. doi: 10.1016/j.res.2015.07.027
8. Morales O., Kurowicka D., Roelen A. Eliciting Conditional and Unconditional Rank Correlations from Conditional Probabilities. *Reliability Engineering & System Safety.* 2008, vol. 93, no. 5, pp. 699–710. doi: 10.1016/j.res.2007.03.020
9. Baba K., Shibata R., Sibuya M. Partial Correlation and Conditional Correlation as Measures of Conditional Independence. *Australian & New Zealand J. of Statistics.* 2004, vol. 46, no. 4, pp. 657–664. doi: 10.1111/j.1467-842X.2004.00360.x
10. Baba K., Sibuya M. Equivalence of Partial and Conditional Correlation Coefficients. *J. of the Japan Statistical Society.* 2005, vol. 35, no. 1, pp. 1–19. doi: 10.14490/JJSS.35.1
11. Yuan N., Fu Z., Zhang H., Piao L., Luterbacher J. Detrended Partial-Cross-Correlation Analysis: A New Method for Analyzing Correlations in Complex System. *Scientific Reports.* 2015, vol. 5, no. 1, p. 8143. doi: 10.1038/srep08143
12. Qian X.-Yu., Liu Y.-M., Jiang Zh.-Q., Podobnik B., Zhou W.-X., Stanley H. E. Detrended Partial Cross-Correlation Analysis of Two Nonstationary Time Series Influenced by Common External Forces. *Physi-*

cal Review. 2015, vol. 91, p. 06281. doi: 10.1103/PhysRevE.91.062816

13. Zhou W. X. Multifractal Detrended Cross-Correlation Analysis for Two Nonstationary Signals. *Physical Review E*. 2008, vol. 77, p. 066211. doi: 10.1103/PhysRevE.77.066211

14. Horvatic D., Stanley H. E., Podobnik B. Detrended Cross-Correlation Analysis for Non-Stationary Time Series with Periodic Trends. *Europhysics Let.* 2011, vol. 94, no. 1, p. 18007. doi: 10.1209/0295-5075/94/18007

15. Alvarez-Ramirez J., Rodriguez E., Echeverría J. C. Detrending Fluctuation Analysis Based on Moving Average Filtering. *Physica A: Statistical Me-*

chanics and Its Applications. 2005, vol. 354, pp. 199–219. doi: 10.1016/j.physa.2005.02.020

16. Rinn F. TSAP-Win Time Series Analysis and Presentation for Dendrochronology and Related Applications. User Reference Version 0.53. Heidelberg, Rinntech, 2005, pp. 1–88.

17. Candiago S., Remondino F., De Giglio M., Dubbini M., Gettelli M. Article Evaluating Multispectral Images and Vegetation Indices for Precision Farming Applications from UAV Images. *Remote Sens*. 2015, vol. 7, no. 4, pp. 4026–4047. doi: 10.3390/rs70404026

18. Pyataev A. S., Vais A. A. Pine Crown and Trunk Diameter Dependence Research. *CEUR Workshop Proc*. 2019, pp. 160–165.

Information about the authors

Nikita S. Pyko, Master in information and communication technology (2019), Postgraduate Student of the Department of Radio Engineering Systems, Junior Researcher at the Scientific and Educational Center "Digital Telecommunication Technologies" of Saint Petersburg Electrotechnical University. The author of 36 scientific publications. Area of expertise: statistical data analysis, mathematical modeling.

Address: Saint Petersburg Electrotechnical University, 5 F, Professor Popov St., St Petersburg 197022, Russia

E-mail: goststalker13@gmail.com

<https://orcid.org/0000-0002-6668-9512>

Denis V. Tishin, Can. Sci. (Biol.) (2006), Associate Professor of the Department of General Ecology of the Institute of Environmental Sciences of Kazan Federal University, Senior Researcher at the IMC FKTI of Saint Petersburg Electrotechnical University. Author of 62 scientific publications. Area of expertise: dendrochronology, phenology, dendroclimatology, paleoecology, carbon balance.

Address: Kazan Federal University, 18/1, Kremlyovskaya St., Kazan 420008, Russia

E-mail: kpfuecology@gmail.com

<https://orcid.org/0000-0002-4790-2840>

Pavel Yu. Iskandirov, Ecologist (KFU, 2013), Postgraduate Student of the Department of General Ecology of the Institute of Environmental Sciences of Kazan Federal University. The author of 12 scientific publications. Area of expertise: dendrochronology, phenology.

Address: Kazan Federal University, 18/1, Kremlyovskaya St., Kazan 420008, Russia

E-mail: iskandirovpj@stud.kpfu.ru

<https://orcid.org/0009-0004-3824-6708>

Artur M. Gafurov, Senior Lecturer at the Department of Landscape Ecology of the Institute of Environmental Sciences, Senior Researcher of the Research Center for Superiority of Cyber-Physical Systems of the Institute of Physics of Kazan Federal University. The author of 50 scientific publications. Area of expertise: geomorphology and evolutionary geography, geoecology, aerospace research of the Earth, photogrammetry.

Address: Kazan Federal University, 18/1, Kremlyovskaya St., Kazan 420008, Russia

E-mail: AMGafurov@kpfu.ru

<https://orcid.org/0000-0002-0812-1750>

Bulat M. Usmanov, Senior Lecturer at the Department of Landscape Ecology of the Institute of Environmental Sciences of Kazan Federal University. The author of 72 scientific publications. Area of expertise: geomorphology and evolutionary geography, geoecology, aerospace research of the Earth, photogrammetry.

Address: Kazan Federal University, 18/1, Kremlyovskaya St., Kazan 420008, Russia

E-mail: BUusmanof@kpfu.ru

<https://orcid.org/0000-0001-9354-1348>

Mikhail I. Bogachev, Dr Sci. (Eng.) (2018), Associate Professor (2011) at the Department of Radio Engineering Systems, Chief Researcher of the Scientific and Educational Center "Digital Telecommunication Technologies" of Saint Petersburg Electrotechnical University. The author of 200 scientific publications. Area of expertise: statistical data analysis, mathematical modeling.

Address: Saint Petersburg Electrotechnical University, 5 F, Professor Popov St., St Petersburg 197022, Russia

E-mail: rogex@yandex.com

<https://orcid.org/0000-0002-0356-5651>