



Math-Net.Ru

Общероссийский математический портал

А. С. Степанов, Е. А. Фомина, Л. В. Илларионова, К. Н. Дубровин, Д. В. Федосеев, Аппроксимация временных рядов индексов вегетации (NDVI и EVI) для мониторинга сельхозкультур (посевов) Хабаровского края, *Информатика и автоматизация*, 2023, выпуск 22, том 6, 1473–1498

DOI: 10.15622/ia.22.6.8

Использование Общероссийского математического портала Math-Net.Ru подразумевает, что вы прочитали и согласны с пользовательским соглашением

<http://www.mathnet.ru/rus/agreement>

Параметры загрузки:

IP: 3.147.43.65

17 сентября 2024 г., 20:15:57



А.С. СТЕПАНОВ, Е.А. ФОМИНА, Л.В. ИЛЛАРИОНОВА, К.Н. ДУБРОВИН,
Д.В. ФЕДОСЕЕВ

АППРОКСИМАЦИЯ ВРЕМЕННЫХ РЯДОВ ИНДЕКСОВ ВЕГЕТАЦИИ (NDVI И EVI) ДЛЯ МОНИТОРИНГА СЕЛЬХОЗКУЛЬТУР (ПОСЕВОВ) ХАБАРОВСКОГО КРАЯ

Степанов А.С., Фомина Е.А., Илларионова Л.В., Дубровин К.Н., Федосеев Д.В.
Аппроксимация временных рядов индексов вегетации (NDVI и EVI) для мониторинга сельхозкультур (посевов) Хабаровского края.

Аннотация. Аппроксимация рядов сезонного хода индексов вегетации является основой для эффективного мониторинга сельскохозяйственных культур, их идентификации и автоматизированной классификации пахотных земель. Для пахотных земель Хабаровского края в период с мая по октябрь 2021 года по мультиспектральным снимкам Sentinel-2A (20 м) с использованием маски облачности были построены временные ряды NDVI и EVI. Для приближения временных рядов были использованы пять видов аппроксимирующих функций: функция Гаусса; двойная гауссиана; двойная синусоида; ряд Фурье; двойная логистическая. Были построены и рассчитаны характеристики экстремумов аппроксимированных временных рядов для разных типов пахотных земель: гречихи, многолетних трав, сои, залежи и пара. Было показано, что для каждой сельхозкультуры аппроксимированные кривые сезонного хода имели характерный вид. Как было достоверно установлено ($p < 0,05$), наиболее высокую точность аппроксимации рядов NDVI и EVI показал ряд Фурье (средняя ошибка составила, соответственно, 8,5% и 16,0%). Аппроксимация рядов NDVI с использованием двойной синусоиды, двойной гауссианы и двойной логистической функции приводила к увеличению ошибки до 8,9-10,6%. Аппроксимация рядов EVI на основе двойной гауссианы и двойной синусоиды способствовала росту средней ошибки до 18,3-18,5%. Проведенный апостериорный анализ с использованием критерия Тьюки показал, что для полей с соей, парующих и залежных земель для приближения индексов вегетации достоверно лучше использовать ряд Фурье, двойную гауссиану или двойную синусоиду, для полей с гречихой целесообразно применять ряд Фурье или двойную гауссиану. В целом, средняя ошибка аппроксимации сезонных временных рядов NDVI в 1,5-4 раза меньше, чем ошибка аппроксимации рядов EVI.

Ключевые слова: индекс вегетации, Хабаровский край, аппроксимация, пахотные земли, сельскохозяйственная культура, временной ряд.

1. Введение. Эффективный мониторинг сельскохозяйственных культур с использованием данных дистанционного зондирования Земли является основой успешного перехода к точному и цифровому земледелию. Одно из важнейших направлений мониторинга связано с разработкой подходов по применению вегетационных индексов (VI) для автоматизированной классификации пахотных земель, выявления залежных земель, оценки состояния посевов, прогнозирования урожайности [1 – 4].

К наиболее популярным индексам вегетации, расчет которых базируется на анализе мультиспектральных изображений, относятся

Normalized Difference Vegetation Index (NDVI) и Enhanced Vegetation Index (EVI) [5 – 8].

Отмечено, что для решения задач автоматизированной классификации пахотных земель важнейшую роль имеет исследование графиков временных рядов VI разных культур, предварительное выявление их сходства между собой, а также выбор параметра – типа VI, который продемонстрирует самую высокую точность классификации [9 – 10]. Так, например, было показано, что для верной идентификации полей с рисом в провинции Хубей (Китай) целесообразно применять не NDVI и EVI, а специальную комбинацию красного, ближнего инфракрасного, и коротковолнового инфракрасного каналов [9]. В другой работе, посвященной картографированию пахотных земель северных провинций Китая, было установлено, что NDVI превзошел EVI в точности распознавания хлопка, пшеницы, кукурузы, поскольку NDVI обладал лучшей разделяемостью для различения культур на этапах созревания [10]. На примере сельскохозяйственных земель индийского штата Одиша были исследованы временные ряды NDVI и EVI для зернобобовых, хлопка, риса, а также парующих земель [11]. Были определены характерные параметры каждой кривой – значения максимумов, продолжительность фенологических циклов, а также рассчитана степень сходства между временными рядами разных видов сельхозкультур, а также вариации внутри одного класса. Моделирование фенологических циклов культур с помощью временных рядов VI также имеет большое значение при прогнозировании урожайности. В частности, рассматривались показатели, основанные на значениях NDVI и EVI для разных фаз роста сои в долине Миссисипи [12]. Модели прогнозирования урожайности сои, основанные как на NDVI, так и на EVI, имели точность около 95%. Утверждается, что разработанные подходы к моделированию урожайности полезны для мониторинга различных сельскохозяйственных регионов мира и могут быть применены к различным культурам для составления карт урожайности или моделирования жизненного цикла культур. В другом исследовании, при изучении полей кукурузы и сои в штате Небраска (США), была построена система, описывающая связь между фенологическими показателями, полученными с помощью дистанционного зондирования – на основе индекса NDVI, и ключевыми стадиями роста сельскохозяйственных культур [13]. В целом, это исследование продемонстрировало, что разработанная система обладает большим потенциалом для оценки сроков различных стадий роста кукурузы и

сои с приемлемой точностью (ошибка модели не превышала 8%). Прогнозирование дат наступления разных стадий роста сельскохозяйственных культур с опережением по времени весьма полезно для управления ростом культур и мониторинга в течение сезона [14]. Для создания общеевропейской базы фенологических данных по разным сельскохозяйственным культурам также тестировалась возможность использования NDVI и EVI как временных показателей, характеризующих наступление и продолжительность фенологических фаз. Было установлено, что EVI имел более тесную корреляционную связь с фенологическими циклами и меньшую ошибку, в сравнении с NDVI для большинства культур.

В качестве важнейшей задачи, связанной с обработкой значений VI, полученных в ходе одного сезона, часто рассматривается получение гладкой формы временного ряда индексов и подавление шума [15 – 16]. Это дает возможность рассчитывать параметры временного ряда и использовать их на практике (например, определить точку наступления и завершения, длительность вегетационного периода и прочие фенологические характеристики) [15, 17]. Также при отсутствии некоторых значений временного ряда применяется аппроксимация при помощи разных функций [16]. Применение аппроксимирующих функций для обработки временных рядов может быть использовано в дальнейшем для построения композитов и аналитического описания изменения VI в ходе одного сезона.

Наиболее простым способом является сглаживание скользящим средним; достаточно часто применяется фильтр Савицкого-Голлея, являющийся по сути развитием метода скользящего среднего с использованием полиномов разных степеней; также используется вейвлет-преобразование и фильтр Уиттакера [18]. При этом применение алгоритма скользящего окна с заполнением пропусков одновременно решает задачи интерполяции и фильтрации [19]. Другим подходом к сглаживанию является использование аппроксимирующих функций (асимметричная гауссова, двойная логистическая (DL) и полиномиальные функции) [20 – 23]. Например, сравнение точности классификации сельхозкультур при предварительном сглаживании временных рядов NDVI дискретным преобразованием Фурье, асимметричной функцией Гаусса, двойной логистической функцией (DL), а также фильтрами Савицкого-Голлея и Уиттакера показало, что наименьшая ошибка классификации была достигнута с помощью преобразования Фурье и фильтра Уиттакера (точность классификации выросла на 2-6%) [18]. Вместе с тем, использование асимметричной

функции Гаусса и DL напротив, способствовало даже некоторому снижению точности классификации пахотных земель. В схожем исследовании, выполненном в Индии, проводилась оценка применимости четырех методов (преобразования Фурье, асимметричной гауссовой модели, DL-модели и фильтра Уиттакера) для моделирования сезонных колебаний VI [20]. Асимметричная функция Гаусса и DL показали хорошие результаты только для кривых с одним выраженным экстремумом (в том числе для регионов, где в год получают только один урожай). С другой стороны, некоторыми авторами отмечалось преимущество использования логистических функций для моделирования EVI и NDVI не только сельскохозяйственных культур, но и естественных лугов, а также залежных земель [21, 22]. Так, например, для моделирования временных рядов NDVI, полученных для полей кукурузы и сои в США применялись две логистические кривые, одна для начального периода, а другая последующего периода роста [22]. Для залежей предлагался итеративный метод логистической подгонки для моделирования EVI. Было установлено, что именно устойчивость этого метода к изменению уровня шума обеспечивала ему преимущество [21]. Для временных рядов значений NDVI, рассчитанных по данным MODIS для основных сельхозкультур Самарской области, проводилась оценка точности аппроксимации при помощи кусочно-линейных функций, асимметричных функций Гаусса и DL, рядов Фурье, многочленов и кубического сплайна [23]. Было показано, что наиболее высокая точность аппроксимации достигалась при использовании кубического сплайна. В другом исследовании, для подавления шума в дополнение к асимметричным функциям Гаусса и DL использовали фильтр Савицкого-Голлея, а также три модифицированных фильтра для аппроксимации временных рядов NDVI в канадской провинции Альберта (при этом учитывался тип поверхности – горы, леса, пашни, луга и т.д.) [24]. В среднем по региону использование асимметричных функций Гаусса и DL наилучшим образом снижало уровень шума при сохранении соответствующей целостности сигнала NDVI. Однако в некоторых особых случаях (например, в горных регионах) альтернативные фильтры обеспечивали более высокую точность аппроксимации. Вместе с тем, другими авторами было отмечено, что применение асимметричной функции Гаусса и DL, а также гармонического анализа и некоторых прочих методов восстановления данных временного ряда NDVI гарантировало хорошую точность для малооблачных территорий [25]. В то же время моделирование рядов NDVI с использованием разных аппроксимирующих функций в

районах, подверженных облачности, в том числе южной части российского Дальнего Востока, требует дополнительных исследований. В частности, ранее были проведены исследования, связанные с оценкой возможности использования функции Гаусса, DL и полиномов второй и третьей степени для решения задачи раннего прогнозирования урожайности сои для Амурской области, Хабаровского и Приморского краев, а также Еврейской Автономной области (ЕАО) [26, 27]. Было установлено, что точность прогнозирования непосредственно связана с ошибкой аппроксимации временных рядов, а полиномиальные функции продемонстрировали недостаточную точность. Вместе с тем, в связи с особенностями севооборота, такими как посев овса с подсевом многолетних трав, поздним севом гречихи, обеспечивающими более одного максимума на кривых сезонного хода VI, целесообразно рассмотреть аппроксимацию временных рядов в модельном регионе Дальнего Востока на основе периодических функций (ряд Фурье, двойная синусоида), а также двойной логистической функции и двойной функции Гаусса.

Таким образом, основной целью исследования явилась оценка возможности аппроксимации временных рядов индексов вегетации пахотных земель Хабаровского края с использованием нелинейных функций. Для достижения этой цели в рамках настоящей работы решались следующие задачи: построение временных рядов NDVI и EVI для разных сельскохозяйственных культур Хабаровского края; описание основных характеристик сезонного хода VI для каждой культуры; расчет аппроксимированных значений временных рядов NDVI и EVI с помощью нелинейных функций; оценка ошибки аппроксимации для разных сельскохозяйственных культур и функций, а также определение статистических различий способов аппроксимации; выбор оптимальной модели аппроксимации для пахотных земель Хабаровского края; оценка применимости модели в других сельскохозяйственных регионах южной части Дальнего Востока.

2. Материалы и методы. В качестве области исследования рассматривались пахотные земли Хабаровского района (2021 г): более 3500 га (155 отдельных полей), в том числе посеvy сои (1260 га), гречихи (238 га), многолетних трав (114 га), а также парующие (496 га) и залежные земли (1613 га). В период с 1 мая по 31 октября 2021 г для исследуемой области были получены 37 снимков (Sentinel-2A, размер пикселя 20 м). Выбранный временной интервал является достаточным для исследования изменений в фенологических циклах

сельскохозяйственных культур Хабаровского края – первые всходы появляются не раньше первой декады мая, уборка обычно заканчивается не позже середины октября. Для каждого пикселя формировались временные ряды NDVI и EVI, где значения VI определялись по формулам [28]:

$$NDVI = \frac{NIR-RED}{NIR+RED}, \quad (1)$$

$$EVI = 2,5 \frac{NIR-RED}{NIR+6RED-7,5BLUE+1}, \quad (2)$$

где NIR – отражение в ближней инфракрасной области спектра;

RED , $BLUE$ – отражения в красной и синей областях спектра.

С использованием маски облачности осуществлялась фильтрация полученных временных рядов NDVI и EVI. На следующем этапе рассчитывались средние значения VI для 155 отдельных полей.

Для аппроксимации временных рядов NDVI и EVI были использованы следующие функции: Гаусса (G); двойная гауссиана (DG); двойная синусоида (DS); ряд Фурье (DF); двойная логистическая (DL). Аппроксимация проводилась нелинейным методом наименьших квадратов, на основе алгоритма Левенберга-Марквардта [29].

Функция Гаусса представляет выражение следующего вида [30]:

$$G = a e^{-\frac{(x-b)^2}{2c^2}}, \quad (3)$$

где a , b , c – параметры.

DG представляет сумму двух гауссиан:

$$DG = a_1 e^{-\frac{\left(\frac{x-b_1}{c_1}\right)^2}{2}} + a_2 e^{-\frac{\left(\frac{x-b_2}{c_2}\right)^2}{2}}, \quad (4)$$

где a_1 , b_1 , c_1 , a_2 , b_2 , c_2 – параметры.

Значения функции DS рассчитываются по следующей формуле:

$$DS = a_1 \sin(b_1 x + c_1) + a_2 \sin(b_2 x + c_2), \quad (5)$$

где $a_1, b_1, c_1, a_2, b_2, c_2$ – параметры.

DF вычисляется по формуле:

$$DF = a_0 + a_1 \cos(xw) + b_1 \sin(xw) + a_2 \cos(2xw) + b_2 \sin(2xw), \quad (6)$$

где $a_0, a_1, w, b_1, a_2, b_2$ – параметры.

DL может быть задана следующей формулой:

$$DL = c_1 + c_2 \left(\frac{1}{1 + e^{-\frac{a_1 - x}{a_2}}} - \frac{1}{1 + e^{-\frac{a_3 - x}{a_4}}} \right), \quad (7)$$

где $c_1, c_2, a_1, a_2, a_3, a_4$ – параметры.

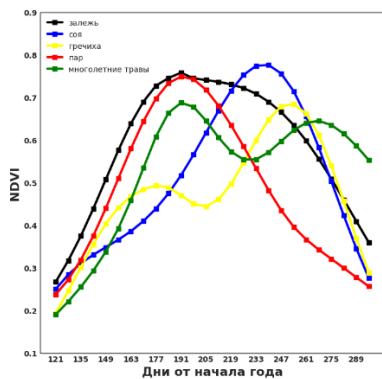
Для сравнительной оценки точности аппроксимации разными функциями рассчитывалась средняя абсолютная ошибка модели, MAPE, % (Mean absolute percentage error) по формуле:

$$MAPE = \frac{1}{n} \sum_{i=1}^n \frac{|Y^i - Y_m^i|}{Y^i} 100, \quad (8)$$

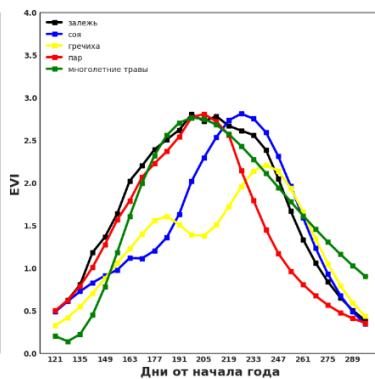
где n – количество членов временного ряда, Y^i – реальное значение VI, Y_m^i – аппроксимированное (смоделированное) значение VI.

В дальнейшем, с использованием двухфакторного дисперсионного анализа с апостериорным попарным сравнением на основе критерия Тьюки ($p=0,05$) по рассчитанным для 155 полей MAPE для NDVI и EVI, определялся оптимальный тип аппроксимирующей функции для каждого вида пахотных земель, а также тип функции, наилучшим образом моделирующий все типы сельскохозяйственных культур.

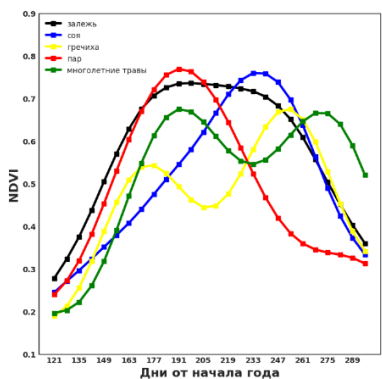
3. Результаты. Для каждого поля Хабаровского района в 2021 году были сформированы ряды средних значений NDVI и EVI. Далее для каждой культуры по усредненным временным рядам VI проводилась аппроксимация функциями DG, DS, DF, DL, G (рисунк 1).



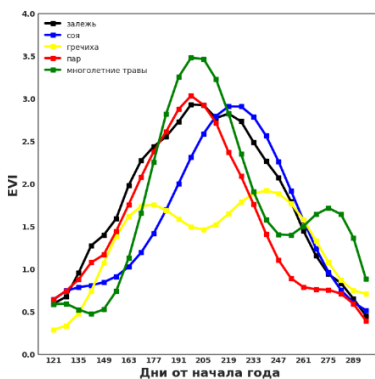
а)



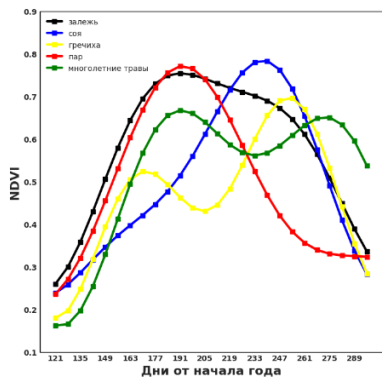
б)



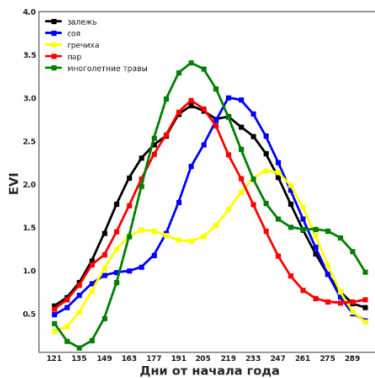
в)



г)



д)



е)

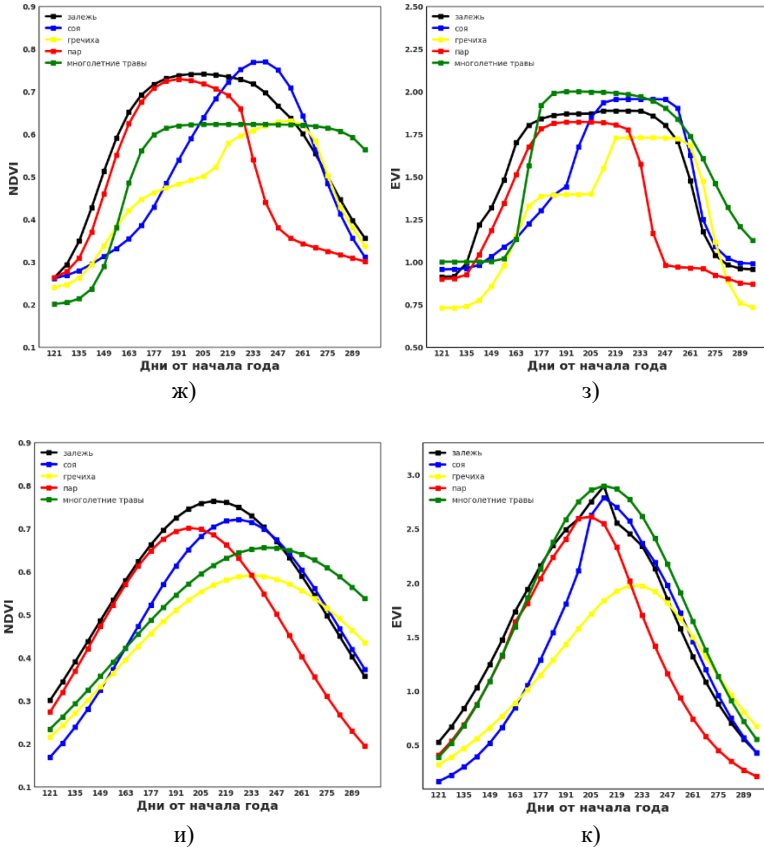


Рис. 1. Аппроксимация временных рядов VI гречихи, залежи, многолетних трав, пара, сои (2021 г, Хабаровский район) разными функциями: а) б) DG; в), г) DS; д), е) DF; ж), з) DL; и), к) G

Как видно из рисунков 1(а) и (б), для сои характерен один пик на аппроксимированной кривой NDVI и один пик на аппроксимированной кривой EVI (при использовании функции DG). День наступления максимума для сои на графике сезонного хода NDVI – 233 календарный день с начала года, EVI – 226, а численные значения $NDVI_{max}$ и EVI_{max} составили, соответственно, 0,78 и 2,81. Также наличие выраженного пика было характерно для парующих земель, при этом наступление максимума являлось достаточно ранним, в сравнение с соей: максимум NDVI приходился на 191 календарный день и достигал 0,75, а максимум EVI соответствовал 205

календарному дню и превышал 2,80. Сезонный ход EVI залежных земель, аппроксимированных DG, в целом схож с динамикой EVI растительного покрова парующих полей: максимум EVI пришелся на 198 день с начала года, но, при этом, в течение 5–7 календарных недель наблюдалось плато. Еще более выраженное плато в период со 184 по 247 календарный день соответствовало сезонному ходу NDVI. Сезонный ход NDVI для гречихи и многолетних трав (рисунок 1(а)) имел двухпиковый характер, что в первом случае отражало поздний сев гречихи (конец июня – начало июля), а во втором – укос многолетних трав, в частности, тимофеевки луговой, в конце июля.

Необходимо отметить, что график аппроксимированного ряда сезонного хода EVI для многолетних трав при помощи DG имел однопиковый вид (рисунок 1(б)), в то время как применение DS и DF (рисунок 1(г, е)) для аппроксимации приводило к двухпиковой форме, что более точно отражало особенности культивирования многолетних трав в Хабаровском крае. Вместе с тем, аппроксимация функциями DS и DF сезонного хода индексов вегетации soi показала, что вид кривых и характеристики экстремумов вполне соответствовали показателям для DG (рисунок 1(в, г, д, е)). Наступление максимума NDVI soi при аппроксимации DS и DF также пришлось на 233 календарный день (численные значения составили 0,76 и 0,78), а максимума EVI soi – на 226 день (EVI_{max} , соответственно, 2,90 и 2,98). Также схожие результаты наблюдались при аппроксимации функциями DG, DS, DF индексов вегетации гречихи и парующих полей. Второй максимум NDVI гречихи приходился на 254 календарный день, значения $NDVI_{max}$ находились в диапазоне 0,67-0,69 при использовании DG, DS, DF. Второй максимум EVI соответствовал 240 дню от начала года, при этом разброс значений EVI_{max} был более значителен – от 1,91 при аппроксимации DS до 2,20 при аппроксимации DG. Как видно из рисунка 1(в, д), для графика сезонного хода NDVI залежи, аппроксимированного функциями DS и DF, также характерно наличие плато в период со 177 по 247 календарный день. Значения NDVI на плато при аппроксимации DS находились в интервале 0,68-0,74, а при аппроксимации DF – в диапазоне 0,67-0,75. Как следует из рисунка 1(ж), аппроксимация с использованием DL для рядов сезонного хода NDVI парующих земель, многолетних трав и гречихи способствовала образованию плато на графике. Аналогично, применение DL для аппроксимации рядов сезонного хода EVI приводило к потере выраженного максимума для всех изучаемых типов пахотных земель Хабаровского района (рисунок 1(з)). В результате аппроксимации сезонного хода NDVI soi с помощью DL среднее значение максимума, пришедшегося на 233

календарный день, также составило 0,77, как и в случае DG, DS, DF. Использование для сглаживания сезонного хода NDVI и EVI функции Гаусса приводило к потере одного экстремума для гречихи и многолетних трав, а также к исчезновению плато для залежи (рисунок 1(и, к)). Аппроксимация сезонного хода NDVI сои на основе G привела к снижению значения максимума до 0,72 в сравнении с двойными функциями; аналогично, численное значение $NDVI_{max}$ для парующих полей составило 0,70, при этом максимум достигался к 205 дню с начала года.

Как видно из диаграммы размаха, ошибка аппроксимации сезонного хода NDVI выбранными функциями зависит от вида произрастающей культуры (рисунок 2). Так, при использовании DG, DS, DF и DL ошибка аппроксимации сезонного хода NDVI гречихи находилась в диапазоне от 10 до 25%, залежи и многолетних трав – от 4 до 13%, сои и пара – от 5 до 18%. MAPE при применении функции Гаусса находилась в пределах 10-30% для сезонного хода NDVI сои, 15-30% для гречихи, 12-28% для пара. Усреднение проводилось по полям с разными культурами, поэтому для многолетних трав, представленных двумя полями большого размера, на диаграмме размаха значения MAPE отображались точками. Максимальная вариативность MAPE наблюдалась при аппроксимации сезонного хода NDVI гречихи функциями DS, DF – размах средних квартилей превышал 5%.

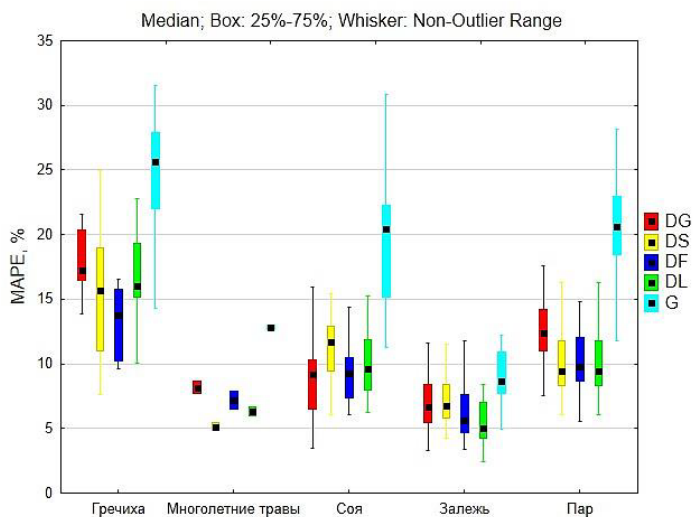


Рис. 2. Диаграмма размаха ошибки аппроксимации временных рядов сезонного хода NDVI для пахотных полей Хабаровского края в 2021 г.

Как видно из рисунка 3, в среднем значения MAPE аппроксимации сезонного хода EVI были существенно выше, в сравнении с NDVI. Ошибка аппроксимации временных рядов EVI полей с залежью составила от 8 до 18%, а для многолетних трав – от 15 до 24% при использовании DG, DS, DF и G. Применение DL для обработки исходных рядов в среднем способствовало увеличению средней ошибки, при этом можно отметить достаточно большую вариативность MAPE по отдельным полям – от 22 до 54% для залежи, от 28 до 48% для гречихи, от 33 до 58% для сои, и от 28 до 82% для парующих земель. Аппроксимация сезонного хода EVI с использованием функции Гаусса также приводила к достаточно большой средней ошибке и вариативности этого показателя: для гречихи в интервале от 18 до 52%, для сои – от 21 до 43%, а для пара – от 20 до 48%. При этом MAPE аппроксимации временных рядов EVI находилась в диапазоне 11-33% при применении DG, DS и DF (поля с соей), 11-40% (пар), 13-48% (гречиха).

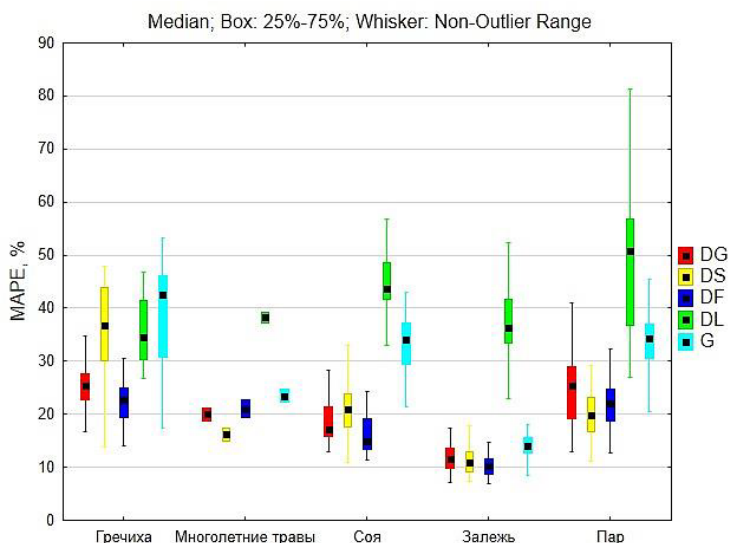


Рис. 3. Диаграмма размаха ошибки аппроксимации временных рядов сезонного хода EVI для пахотных полей Хабаровского края в 2021 г.

На следующем этапе для каждого вида произрастающей культуры проводилась оценка влияния типа функции в сочетании с NDVI или EVI на точность аппроксимации. Методами двухфакторного дисперсионного анализа было установлено, что значимо на ошибку

аппроксимации для каждой культуры влияют и тип функции, и показатель VI (таблица 1). При этом, как видно из таблицы, вклад фактора VI в точность аппроксимации в 2-4 раза выше, чем фактор типа функции для гречихи, многолетних трав и пара, и в 1,1-1,4 раза выше для залежи и сои.

Таблица 1. Оценка значимости факторов (VI и тип функции) на точность аппроксимации сезонного хода VI для разных культур

Вид культуры	Фактор	Характеристики двухфакторного анализа			
		SS	MS	F	p
Гречиха	VI	6402,0	6402,0	135,1	<0,05
	Функция	3012,4	753,1	15,9	<0,05
Многолетние травы	VI	1254,6	1254,6	50,6	<0,05
	Функция	319,6	79,9	3,2	<0,05
Соя	VI	23660,9	23660,9	444,3	<0,05
	Функция	17027,3	4256,8	79,9	<0,05
Залежь	VI	17658,5	17658,5	377,4	<0,05
	Функция	15665,8	3916,4	83,7	<0,05
Пар	VI	33669,0	33669,0	501,0	<0,05
	Функция	15621,6	3905,4	58,1	<0,05

Апостериорный анализ на основе критерия Тьюки для полей с гречихой показал, что использование DF давало существенно меньшую ошибку аппроксимации в сравнении с DS, DL и G (таблица 2). В целом, для аппроксимации сезонного хода VI гречихи целесообразно было использовать DG (ошибка 21,5%) или DF (ошибка 17,5%).

Таблица 2. Апостериорный анализ (критерий Тьюки) по аппроксимирующим функциям для полей с гречихой

	DG	DS	DF	DL	G
DG		$p > 0,05$	$p > 0,05$	$p > 0,05$	$p < 0,05$
DS	$p > 0,05$		$p < 0,05$	$p > 0,05$	$p > 0,05$
DF	$p > 0,05$	$p < 0,05$		$p < 0,05$	$p < 0,05$
DL	$p > 0,05$	$p > 0,05$	$p < 0,05$		$p > 0,05$
G	$p < 0,05$	$p > 0,05$	$p < 0,05$	$p > 0,05$	
MAPE, %	21,5	25,5	17,5	26,3	31,9

Как следует из таблицы 3, значимых различий между аппроксимацией сезонного хода VI многолетних трав пятью различными функциями выявлено не было. Очевидно, что в связи с тем, что в среднем ошибка аппроксимации NDVI многолетних трав в 2-5 раза меньше, чем EVI, необходимо оценить значимость различий отдельно для NDVI и EVI.

Таблица 3. Апостериорный анализ (критерий Тьюки) по аппроксимирующим функциям для полей с многолетними травами

	DG	DS	DF	DL	G
DG		p>0,05	p>0,05	p>0,05	p>0,05
DS	p>0,05		p>0,05	p>0,05	p>0,05
DF	p>0,05	p>0,05		p>0,05	p>0,05
DL	p>0,05	p>0,05	p>0,05		p>0,05
G	p>0,05	p>0,05	p>0,05	p>0,05	
MAPE,%	14,1	10,7	14,1	22,3	18,1

С помощью критерия Тьюки было установлено, что для полей с соей статистически значимы различия точности аппроксимации VI с применением функций G и DL (MAPE составила 26,3% и 28,2%), с одной стороны, и DG, DS, DF (ошибка от 12,7% до 16,4%) (таблица 4).

Таблица 4. Апостериорный анализ (критерий Тьюки) по аппроксимирующим функциям для полей с соей

	DG	DS	DF	DL	G
DG		p>0,05	p>0,05	<i>p<0,05</i>	<i>p<0,05</i>
DS	p>0,05		p>0,05	<i>p<0,05</i>	<i>p<0,05</i>
DF	p>0,05	p>0,05		<i>p<0,05</i>	<i>p<0,05</i>
DL	<i>p<0,05</i>	<i>p<0,05</i>	<i>p<0,05</i>		p>0,05
G	<i>p<0,05</i>	<i>p<0,05</i>	<i>p<0,05</i>	p>0,05	
MAPE,%	14,7	16,4	12,7	28,2	26,3

В тоже время для сезонного хода VI залежных земель ошибка аппроксимации функцией DF была равной 8,4%, функцией DS – 9,5%, функцией DG – 10,0%, а гауссианой – 12,5%, что достоверно ниже, чем при использовании DL (таблица 5).

Таблица 5. Апостериорный анализ (критерий Тьюки) по аппроксимирующим функциям для залежи

	DG	DS	DF	DL	G
DG		$p>0,05$	$p>0,05$	$p<0,05$	$p>0,05$
DS	$p>0,05$		$p>0,05$	$p<0,05$	$p>0,05$
DF	$p>0,05$	$p>0,05$		$p<0,05$	$p<0,05$
DL	$p<0,05$	$p<0,05$	$p<0,05$		$p<0,05$
G	$p>0,05$	$p>0,05$	$p<0,05$	$p<0,05$	
MAPE,%	10,0	9,5	8,4	21,7	12,1

Схожая ситуация была отмечена и для парующих земель (таблица 6). Точность аппроксимации VI была достоверно выше при применении функций DS и DF (ошибка, соответственно 15,4% и 15,9%) в сравнении с DL и G (ошибка 29,5% и 27,0%).

Таблица 6. Апостериорный анализ (критерий Тьюки) по аппроксимирующим функциям для пара

	DG	DS	DF	DL	G
DG		$p>0,05$	$p>0,05$	$p<0,05$	$p<0,05$
DS	$p>0,05$		$p>0,05$	$p<0,05$	$p<0,05$
DF	$p>0,05$	$p>0,05$		$p<0,05$	$p<0,05$
DL	$p<0,05$	$p<0,05$	$p<0,05$		$p>0,05$
G	$p<0,05$	$p<0,05$	$p<0,05$	$p>0,05$	
MAPE,%	19,1	15,4	15,9	29,5	27,0

В связи с тем, что аппроксимация сезонного хода индексов вегетации может применяться в задачах автоматизированной классификации пахотных земель, зачастую априори неизвестно, какой тип культуры произрастает на поле. Таким образом, с практической точки зрения, целесообразно найти метод, наилучшим образом аппроксимирующий сезонный ход каждого индекса для всех культур. С помощью двухфакторного дисперсионного анализа было установлено, что на точность аппроксимации NDVI значимо влияет и вид культуры, и аппроксимирующая функция (таблица 7). Как видно из таблицы, при аппроксимации временных рядов NDVI гречихи и сои наименьшая ошибка была характерна при использовании DF, для многолетних трав оптимальным методом явился DS, а для залежи и пара – DL. При этом различия в точности при применении этих трех функций для многолетних трав, залежи, пара и сои не являлось

достоверным ($p > 0,05$). Аналогичным образом было выявлено значимое влияние обоих факторов на сезонный ход EVI (таблица 8). В ходе попарного сравнения было установлено, что функция DF являлась оптимальным выбором для приближения временного ряда EVI гречихи и сои ($p < 0,05$), а для залежи, пара и многолетних трав возможно использование как DF, так и DS и DG ($p > 0,05$).

Таблица 7. Средние значения MAPE (%) для разных типов функций при аппроксимации сезонного хода NDVI гречихи, многолетних трав, сои, залежи и пара (Хабаровский край, 2021)

Вид культуры	DG	DS	DF	DL	G	p
Гречиха	17,5± 3,4	15,9 ±6,1	12,8 ±2,7	16,7± 3,3	24,7± 4,9	p<0,05
Многолетние травы	8,2 ± 0,6	5,2 ± 0,4	7,2 ± 1,0	6,3 ± 0,5	12,8± 0,2	
Соя	8,7 ± 2,6	11,8 ±3,4	9,3 ± 2,1	10,3± 3,4	19,3± 4,7	
Залежь	7,5 ± 3,5	7,2 ± 1,9	6,1 ± 1,8	5,7 ± 2,4	9,5 ± 3,2	
Пар	13,3± 4,0	10,1 ± 2,8	10,1± 2,3	10,1± 2,8	20,4± 3,3	
p	p<0,05					

Таблица 8. Средние значения MAPE (%) для разных типов функций при аппроксимации сезонного хода EVI гречихи, многолетних трав, сои, залежи и пара (Хабаровский край, 2021)

Вид культуры	DG	DS	DF	DL	G	p
Гречиха	25,4± 5,8	35,4 ±10,5	21,5 ±5,0	36,0± 6,6	39,7± 11,3	p<0,05
Многолетние травы	20,0± 1,8	16,2 ±1,6	20,9 ±2,3	38,3± 1,4	23,5± 1,7	
Соя	18,8± 4,9	21,1 ±5,9	16,2 ±3,6	46,5± 8,9	33,3± 5,5	
Залежь	12,5± 4,7	11,7 ±3,5	10,7 ±3,2	37,7± 8,2	14,8± 5,0	
Пар	25,0± 7,4	20,7± 6,0	21,7± 4,0	48,8± 13,8	33,6± 5,7	
p	p<0,05					

Как следует из таблицы 9, была установлена значимость различий средней ошибки аппроксимации NDVI для пахотных земель Хабаровского края для разных типов функций. В частности, наилучший результат – с ошибкой в 8,5% и 8,9% продемонстрировали DF и DL. Достоверно наиболее низкая точность аппроксимации –

16,1% – была получена при использовании гауссианы. Приближение кривой NDVI функциями DG и DS в принципе показало достаточно хорошие результаты: ошибка, соответственно 10,6% и 9,8%.

Таблица 9. Апостериорный анализ (критерий Тьюки) по аппроксимирующим функциям для сезонного хода NDVI пахотных земель Хабаровского края

	DG	DS	DF	DL	G
DG		$p>0,05$	$p<0,05$	$p<0,05$	$p<0,05$
DS	$p>0,05$		$p<0,05$	$p>0,05$	$p<0,05$
DF	$p<0,05$	$p<0,05$		$p>0,05$	$p<0,05$
DL	$p<0,05$	$p>0,05$	$p>0,05$		$p<0,05$
G	$p<0,05$	$p<0,05$	$p<0,05$	$p<0,05$	
MAPE, %	10,6	9,8	8,5	8,9	16,1

В таблице 10 представлены результаты апостериорного анализа – попарного сравнения точности аппроксимации сезонного хода EVI пахотных земель Хабаровского края в зависимости от типа функции.

Таблица 10. Апостериорный анализ (критерий Тьюки) по аппроксимирующим функциям для сезонного хода EVI пахотных земель Хабаровского края

	DG	DS	DF	DL	G
DG		$p>0,05$	$p<0,05$	$p<0,05$	$p<0,05$
DS	$p>0,05$		$p<0,05$	$p<0,05$	$p<0,05$
DF	$p<0,05$	$p<0,05$		$p<0,05$	$p<0,05$
DL	$p<0,05$	$p<0,05$	$p<0,05$		$p<0,05$
G	$p<0,05$	$p<0,05$	$p<0,05$	$p<0,05$	
MAPE, %	18,5	18,3	16,0	42,8	26,4

Как видно, наиболее низкая ошибка также наблюдалась при использовании DF (16,0%), при этом она в два раза превышала соответствующее значение MAPE для кривой NDVI. Наименее удачной для аппроксимации кривой EVI представляется функция DL, со средней ошибкой 42,8%. MAPE при применении гауссианы составила 26,4%, а функции DG и DS продемонстрировали примерно одинаковые результаты – 18,5% и 18,3%.

Представляет интерес оценка потенциальной возможности применения полученных результатов для аппроксимации временных рядов VI сельскохозяйственных земель других территорий, в частности районов со схожими климатическими характеристиками. На рисунке 4 представлены аппроксимированные с помощью DF

усредненные значения сезонного хода NDVI для полей с соей Амурской области (Архаринский, Октябрьский, Михайловский, Тамбовский р-ны), Еврейской Автономной области (Ленинский, Октябрьский р-ны), Хабаровского края (Хабаровский, Вяземский, им. Лазо р-ны).

Как видно из рисунка, кривые сезонного хода NDVI сои схожи для трех регионов южной части Дальнего Востока, при этом как общая форма кривой, так и период наступления максимума в 2022 году (219-233 календарные дни) также соответствовали 2021 году. Для пахотных земель провинции Хэйлундзян (северо-восточный Китай) основные фазы развития сои совпадали на временной шкале с фазами развития культуры в субъектах ДФО, а графики сезонного хода NDVI вполне схожи с представленными выше [31, 32]. Таким образом, предложенный подход к аппроксимации временных рядов может быть применен в регионах с климатическими условиями, соответствующими условиям южной части Хабаровского края.

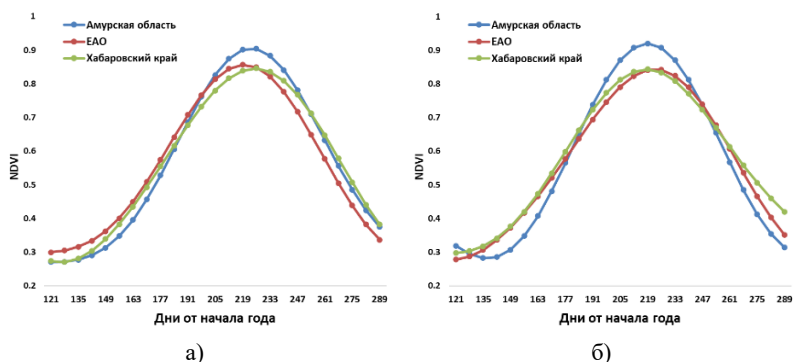


Рис. 4. Аппроксимация функцией DF временных рядов NDVI сои в Амурской области, ЕАО, Хабаровском крае: а) 2021 г., б) 2022 г.

Далее проводилась сравнительная оценка точности классификации на основе временных рядов NDVI с использованием рассмотренных аппроксимирующих функций, а также полиномиальной функции третьего порядка (куб) на примере пахотных земель Хабаровского района (2021-2022 гг.). Для классификации был использован метод RF (Random Forest), в обучающую выборку случайным образом включались 70% от общего числа пикселей, в тестовую 30%. В таблице 11 представлены значения общей точности (ОА, %), полученные по результатам трехкратной кросс-валидации.

Таблица 11. Точность классификации пахотных земель Хабаровского района на основе аппроксимированных рядов NDVI (RF, 2021-2022 гг.)

Общая точность	Полином 3-й степени	DG	DS	DF	DL	G
$OA_{cp} \pm \Delta OA_{cp}, \%$	$85,7 \pm 0,7$	$90,6 \pm 0,6$	$91,7 \pm 0,8$	$92,3 \pm 0,6$	$90,2 \pm 0,8$	$89,2 \pm 1,0$
p	$p < 0,05; p_{DS-DF} > 0,05; p_{DG-DF} > 0,05; p_{DL-DF} < 0,05; p_{G-DF} < 0,05; p_{куб-DF} < 0,05;$					

Методами однофакторного дисперсионного анализа с использованием апостериорного попарного сравнения было установлено, что ошибка классификации при использовании DF, DG и DS была существенно ниже, чем при применении для подготовки временных рядов DL, G, а также кубического полинома. OA классификации на основе DF превысила 92%.

4. Заключение. В ходе исследования по спутниковым снимкам Sentinel 2A за 2021 год были построены временные ряды NDVI и EVI для пахотных земель Хабаровского края. Было установлено, что аппроксимированные графики сезонного хода NDVI и EVI для основных сельскохозяйственных культур Хабаровского края имели характерный вид, позволяющий в дальнейшем использовать ряды сезонного хода в задачах идентификации культур и классификации пахотных земель, в том числе в регионах со схожими климатическими условиями. Методами двухфакторного дисперсионного анализа с апостериорным попарным сравнением на основе критерия Тьюки было установлено, что наиболее высокая точность аппроксимации временного ряда NDVI для всех культур достигалась при применении DF и DL (ошибка не превышала 9%), а временного ряда EVI – при применении DF (ошибка составила 16%). Таким образом, при необходимости подготовки временных рядов VI пахотных земель Хабаровского края для классификации (при отсутствии части данных, необходимости синхронизации данных нескольких лет наблюдения и т.д.) целесообразно использование DF. Вместе с тем, для большинства сельхозкультур функции DG и DS также показали хорошие результаты для приближения рядов сезонного хода NDVI и EVI. Использование разработанного подхода, в отличие от методов сглаживания временного ряда, позволит, во-первых, не только повысить точность автоматизированной классификации, но и использовать аппроксимированные ряды для выявления отклонений в развитии, а также прогнозирования урожайности. Изучение временных рядов VI, характерных для южной части Дальнего Востока, обеспечит возможность экстраполяции исследований на прочие субъекты ДФО, а также сопредельные провинции КНР.

Литература

1. Ferchichi A., Abbes A.B., Barra V., Farah I.R. Forecasting vegetation indices from spatio-temporal remotely sensed data using deep learning-based approaches: A systematic literature review // *Ecological Informatics*. 2022. vol. 68. no. 101552.
2. Liu J., Song W. Mapping abandoned cropland using Within-Year Sentinel-2 time series // *Catena*. 2023. vol. 223. no. 106924.
3. Терехин Э.А. Оценка сезонных значений вегетационного индекса (NDVI) для детектирования и анализа состояния посевов сельскохозяйственных культур // *Исследование земли из космоса*. 2015. № 1. DOI: 10.7868/S0205961415010108.
4. Савин И.Ю., Лупян Е.А., Бартаев С.А. Оперативный спутниковый мониторинг состояния посевов сельскохозяйственных культур в России // *Геоматика*. 2011. № 2. С. 69–76.
5. Черепанов А.С., Дружинин Е.Г. Спектральные свойства растительности и вегетационные индексы // *Геоматика*. 2009. № 3. С. 28–32.
6. Воронина П.В., Мамаш Е.А. Классификация тематических задач мониторинга сельского хозяйства с использованием данных дистанционного зондирования MODIS // *Вычислительные технологии*. 2014. Т. 19. № 3. С. 76–102.
7. Zhang C., Marzougui A., Sankaran S. High-resolution satellite imagery applications in crop phenotyping: An overview // *Computers and Electronics in Agriculture*. 2020. vol. 175. no. 105584.
8. Kong D., Zhang Y., Gu X., Wang D. A robust method for reconstructing global MODIS EVI time series on the Google Earth Engine // *Journal of Photogrammetry and Remote Sensing*. 2019. vol. 155. pp. 13–24.
9. Chen Y., Hu J., Cai Z., Yang J., Zhou W., Hu Q., Wang C., You L., Xu B. A phenology-based vegetation index for improving ratoon rice mapping using harmonized Landsat and Sentinel-2 data // *Journal of Integrative Agriculture*. 2023. 20 p. DOI: 10.1016/j.jia.2023.05.035.
10. Hao P., Tang H., Chen Z., Meng Q., Kang Y. Early-season crop type mapping using 30-m reference time series // *Journal of Integrative Agriculture*. 2020. vol. 19. no. 7. pp. 1897–1911.
11. Singh R.K., Rizvi J., Behera M.D., Biradar C. Automated crop type mapping using time-weighted dynamic time warping-A basis to derive inputs for enhanced food and Nutritional Security // *Current Research in Environmental Sustainability*. 2021. vol. 3. no. 100032.
12. Shammi S.A., Meng Q. Use time series NDVI and EVI to develop dynamic crop growth metrics for yield modeling // *Ecological Indicators*. 2021. vol. 121. no. 107124.
13. Bandaru V., Yaramasu R., Koutilya P.N.V.R., He J., Fernando S., Sahajpal R., Wardlow B.D., Suyker A., Justice C. PhenoCrop: An integrated satellite-based framework to estimate physiological growth stages of corn and soybeans // *International Journal of Applied Earth Observation and Geoinformation*. 2020. vol. 92. no. 102188.
14. Tian F., Cai Z., Jin H., Hufkens K., Scheifinger H., Tagesson T., Smets B., Van Hoolst R., Bonte K., Ivits E., Tong X., Ardo J., Eklundh L. Calibrating vegetation phenology from Sentinel-2 using eddy covariance, PhenoCam, and PEP725 networks across Europe // *Remote Sensing of Environment*. 2021. vol. 260. no. 112456.
15. Катаев М.Ю., Бекеров А.А., Медвецкий Д.В. Методика сглаживания временного ряда вегетационного индекса NDVI // *Доклады ТУСУРа*. 2017. Т. 20. № 1. С. 85–88.
16. Li S., Xu L., Jing Y., Yin H., Li X., Guan X. High-quality vegetation index product generation: A review of NDVI time series reconstruction techniques // *International*

- Journal of Applied Earth Observation and Geoinformation. 2021. vol. 105. no. 102640.
17. Миклашевич Т.С., Барталев С.А. Метод определения фенологических характеристик растительного покрова на основе временных рядов спутниковых данных // Современные проблемы дистанционного зондирования Земли из космоса. 2016. Т. 13. № 1. С. 9–24.
 18. Shao Y., Lunetta R.S., Wheeler B., Liame J.S., Campbella J.B. An evaluation of time-series smoothing algorithms for land-cover classifications using MODIS-NDVI multi-temporal data // Remote Sensing of Environment. 2016. vol. 174. pp. 258–265.
 19. Плотников Д.Е., Миклашевич Т.С., Барталев С.А. Восстановление временных рядов данных дистанционных измерений методом полиномиальной аппроксимации в скользящем окне переменного размера // Современные проблемы дистанционного зондирования Земли из космоса. 2014. Т. 11. № 2. С. 103–110.
 20. Atkinson P.M., Jeganathan C., Dash J., Atzberger C. Inter-comparison of four models for smoothing satellite sensor time-series data to estimate vegetation phenology // Remote Sensing of Environment. 2012. vol. 123. pp. 400–417.
 21. Cao R., Chen J., Shen M., Tang Y. An improved logistic method for detecting spring vegetation phenology in grasslands from MODIS EVI time-series data // Agricultural and Forest Meteorology. 2015. vol. 200. pp. 9–20.
 22. Seo B., Lee J., Lee K., Hong S., Kang S. Improving remotely-sensed crop monitoring by NDVI-based crop phenology estimators for corn and soybeans in Iowa and Illinois, USA // Field Crops Research. 2019. vol. 238. pp. 113–128.
 23. Воробьёва Н.С., Чернов А.В. Аппроксимация временных рядов NDVI в задаче раннего распознавания видов сельскохозяйственных культур по космическим снимкам // Сборник трудов III международной конференции и молодежной школы «Информационные технологии и нанотехнологии» (ИТНТ-2017). Самара: Новая техника, 2017. С. 390–399.
 24. Hird J.N., McDermid G.J. Noise reduction of NDVI time series: An empirical comparison of selected techniques // Remote Sensing of Environment. 2009. vol. 113. no. 1. pp. 248–258.
 25. Julien Y., Sobrino J.A. Optimizing and comparing gap-filling techniques using simulated NDVI time series from remotely sensed global data // International Journal of Applied Earth Observation and Geoinformation. 2019. vol. 76. pp. 93–111.
 26. Stepanov A., Dubrovin K., Sorokin A., Aseeva T. Predicting soybean yield at the regional scale using remote sensing and climatic data // Remote Sensing. 2020. vol. 12(12). no. 1936.
 27. Stepanov A., Dubrovin K., Sorokin A. Function fitting for modeling seasonal normalized difference vegetation index time series and early forecasting of soybean yield // The Crop Journal. 2022. vol. 10. no. 5. pp. 1452–1459.
 28. Lin S., Hao D, Zheng Y., Zhang H., Wang C., Yuan W. Multi-site assessment of the potential of fine resolution red-edge vegetation indices for estimating gross primary production // International Journal of Applied Earth Observation and Geoinformation. 2022. vol. 113. no. 102978.
 29. Gavin H.P. The Levenberg-Marquardt method for nonlinear least squares curve-fitting problems. Department of civil and environmental engineering, Duke University, 2019. vol. 19.
 30. Zhao J., Zhang X. An Adaptive Noise Reduction Method for NDVI Time Series Data Based on S-G Filtering and Wavelet Analysis // Journal Indian Society Remote Sensing. 2018. vol. 46. no. 12. pp. 1975–1982.

31. Han Y., Meng J., Xu J. Soybean growth assessment method based on NDVI and phenological calibration // Transactions of the Chinese Society of Agricultural Engineering. 2017. vol. 33. no. 2. pp. 177–182.
32. Zhang H., Kang J., Xu X., Zhang L. Accessing the temporal and spectral features in crop type mapping using multi-temporal Sentinel-2 imagery: A case study of Yi'an County, Heilongjiang province, China // Computers and Electronics in Agriculture. 2020. vol. 176. no. 105618.

Степанов Алексей Сергеевич — д-р фармацевт. наук, ведущий научный сотрудник, лаборатория селекции зерновых и колосовых культур, Дальневосточный научно-исследовательский институт сельского хозяйства (ДВ НИИСХ). Область научных интересов: интересов сельскохозяйственная экономика прогнозирование урожайности сельскохозяйственных культур, математическое моделирование в сельском хозяйстве. Число научных публикаций — 80. stepanfx@mail.ru; улица Клубная, 13, 680521, Восточное, Россия; р.т.: +7(924)210-9102.

Фомина Елизавета Андреевна — инженер, лаборатория численных методов математической физики, Вычислительный центр Дальневосточного отделения Российской академии наук (ВЦ ДВО РАН). Область научных интересов: математическое моделирование, аппроксимация. Число научных публикаций — 3. eliz37@mail.ru; улица Ким Ю Чена, 65, 680000, Хабаровск, Россия; р.т.: +7(914)316-0703.

Илларионова Любовь Викторовна — канд. физ.-мат. наук, старший научный сотрудник, лаборатория численных методов математической физики, Вычислительный центр Дальневосточного отделения Российской академии наук (ВЦ ДВО РАН). Область научных интересов: машинное обучение, программирование, численные методы. Число научных публикаций — 40. illarionova_l@list.ru; улица Ким Ю Чена, 65, 680000, Хабаровск, Россия; р.т.: +7(924)200-4940.

Дубровин Константин Николаевич — младший научный сотрудник, лаборатория численных методов математической физики, Вычислительный центр Дальневосточного отделения Российской академии наук (ВЦ ДВО РАН). Область научных интересов: математическое моделирование, машинное обучение, применение методов дистанционного зондирования в сельском хозяйстве. Число научных публикаций — 14. pobforward@gmail.com; улица Ким Ю Чена, 65, 680000, Хабаровск, Россия; р.т.: +7(909)859-0881.

Федосеев Денис Владимирович — аспирант, Вычислительный центр ДВО РАН. Область научных интересов: система управления базами данных, программирование, численные методы. d.fedoseev@mail.ru; Ким Ю Чена, 65, 680000, Хабаровск, Россия; р.т.: +7(962)151-34-35.

Поддержка исследований. Исследования выполнены при финансовой поддержке РФФ, проект № 23-76-00007 на тему «Разработка научных методов и подходов устойчивого управления почвенными ресурсами на основе технологий дистанционного зондирования Земли (на примере юга Дальнего Востока)».

A. STEPANOV, E. FOMINA, L. ILLARIONOVA, K. DUBROVIN, D. FEDOSEEV
**VEGETATION INDICES (NDVI AND EVI) TIME SERIES
APPROXIMATION FOR MONITORING CROPS OF
Khabarovsk Territory**

Stepanov A., Fomina E., Illarionova L., Dubrovin K., Fedoseev D. Vegetation Indices (NDVI and EVI) Time Series Approximation for Monitoring Crops of Khabarovsk Territory.

Abstract. Approximation of the series of the seasonal vegetation index time series is the basis for monitoring agricultural crops, their identification and cropland classification. For cropland of the Khabarovsk Territory in the period from May to October 2021, NDVI and EVI time series were constructed using Sentinel-2A (20 m) multispectral images using a cloud mask. Five functions were used to approximate time series: Gaussian function; double Gaussian; double sine wave; Fourier series; double logistic. Characteristics of extremums for approximated time series for different types of arable land were built and calculated: buckwheat, perennial grasses, soybeans, fallow and ley. It was shown that each type requires a characteristic species. It was found ($p < 0.05$) that Fourier approximation showed the highest accuracy for NDVI and EVI series (average error, respectively, 8.5% and 16.0%). Approximation of the NDVI series using a double sine, double Gaussian and double logistic function resulted in an error increase of 8.9-10.6%. Approximation of EVI series based on double Gaussian and double sine wave causes an increase in average errors up to 18.3-18.5%. The conducted a posteriori analysis using the Tukey criterion showed that for soybean, fallow and ley lands, it is better to use the Fourier series, double Gaussian or double sine wave to approximate vegetation indices, for buckwheat it is advisable to use the Fourier series or double Gaussian. In general, the average approximation error of the NDVI seasonal time series is 1.5-4 times less than the approximation error of the EVI series.

Keywords: vegetation index, Khabarovsk Territory, approximation, arable land, crop, time series.

References

1. Ferchichi A., Abbes A.B., Barra V., Farah I.R. Forecasting vegetation indices from spatio-temporal remotely sensed data using deep learning-based approaches: A systematic literature review. *Ecological Informatics*. 2022. vol. 68. no. 101552.
2. Liu J., Song W. Mapping abandoned cropland using Within-Year Sentinel-2 time series. *Catena*. 2023. vol. 223. no. 106924.
3. Terekhin E.A. [Assess the seasonal NDVI values for detection and analysis of agricultural crops]. *Issledovanie Zemli iz kosmosa – Earth exploration from space*. 2015. no. 1. DOI: 10.7868/S0205961415010108. (In Russ.).
4. Savin I.Y., Bartalev S.A., Loupian E.A. [Operational satellite monitoring of the state of agricultural crops in Russia]. *Geomatika – Geomatics*. 2011. no. 2. pp. 69–76. (In Russ.).
5. Cherepanov A.S., Druzhinina E.G. [Spectral properties of vegetation and vegetation indices]. *Geomatika – Geomatics*. 2009. no. 3. pp. 28–32. (In Russ.).
6. Voronina P.V., Mamash E.A. [Classification of thematic monitoring for agriculture problems using remote sensing MODIS data]. *Vychislitel'nye tekhnologii – Computing technologies*. 2014. vol. 19. no. 3. pp. 76–102. (In Russ.).
7. Zhang C., Marzougui A., Sankaran S. High-resolution satellite imagery applications in crop phenotyping: An overview. *Computers and Electronics in Agriculture*. 2020. vol. 175. no. 105584.

8. Kong D., Zhang Y., Gu X., Wang D. A robust method for reconstructing global MODIS EVI time series on the Google Earth Engine. *Journal of Photogrammetry and Remote Sensing*. 2019. vol. 155. pp. 13–24.
9. Chen Y., Hu J., Cai Z., Yang J., Zhou W., Hu Q., Wang C., You L., Xu B. A phenology-based vegetation index for improving ratoon rice mapping using harmonized Landsat and Sentinel-2 data. *Journal of Integrative Agriculture*. 2023. 20 p. DOI: 10.1016/j.jia.2023.05.035.
10. Hao P., Tang H., Chen Z., Meng Q., Kang Y. Early-season crop type mapping using 30-m reference time series. *Journal of Integrative Agriculture*. 2020. vol. 19. no. 7. pp. 1897–1911.
11. Singh R.K., Rizvi J., Behera M.D., Biradar C. Automated crop type mapping using time-weighted dynamic time warping-A basis to derive inputs for enhanced food and Nutritional Security. *Current Research in Environmental Sustainability*. 2021. vol. 3. no. 100032.
12. Shammi S.A., Meng Q. Use time series NDVI and EVI to develop dynamic crop growth metrics for yield modeling. *Ecological Indicators*. 2021. vol. 121. no. 107124.
13. Bandaru V., Yaramasu R., Koutilya P.N.V.R., He J., Fernando S., Sahajpal R., Wardlow B.D., Suyker A., Justice C. PhenoCrop: An integrated satellite-based framework to estimate physiological growth stages of corn and soybeans. *International Journal of Applied Earth Observation and Geoinformation*. 2020. vol. 92. no. 102188.
14. Tian F., Cai Z., Jin H., Hufkens K., Scheifinger H., Tagesson T., Smets B., Van Hoolst R., Bonte K., Ivits E., Tong X., Ardo J., Eklundh L. Calibrating vegetation phenology from Sentinel-2 using eddy covariance, PhenoCam, and PEP725 networks across Europe. *Remote Sensing of Environment*. 2021. vol. 260. no. 112456.
15. Kataev M.Yu., Bekerov A.A., Medveckiy D.V. [Method to smooth time series of NDVI]. *Doklady TUSURA – TUSUR reports*. 2017. vol. 20. no. 1. pp. 85–88. (In Russ.).
16. Li S., Xu L., Jing Y., Yin H., Li X., Guan X. High-quality vegetation index product generation: A review of NDVI time series reconstruction techniques. *International Journal of Applied Earth Observation and Geoinformation*. 2021. vol. 105. no. 102640.
17. Miklashevich T.S., Bartalyov S.A. [Method for estimating vegetation cover phenological characteristics from satellite data time series]. *Sovremennyye problemy distantsionnogo zondirovaniya Zemli iz kosmosa – Modern problems of remote sensing of the Earth from space*. 2016. vol. 13. no. 1. pp. 9–24. (In Russ.).
18. Shao Y., Lunetta R.S., Wheeler B., Liame J.S., Campbella J.B. An evaluation of time-series smoothing algorithms for land-cover classifications using MODIS-NDVI multi-temporal data. *Remote Sensing of Environment*. 2016. vol. 174. pp. 258–265.
19. Plotnikov D.E., Miklashevich T.S., Bartalev S.A. [Using local polynomial approximation within moving window for remote sensing data time-series smoothing and data gaps recovery]. *Sovremennyye problemy distantsionnogo zondirovaniya Zemli iz kosmosa – Modern problems of remote sensing of the Earth from space*. 2014. vol. 11. no. 2. pp. 103–110. (In Russ.).
20. Atkinson P.M., Jeganathan C., Dash J., Atzberger C. Inter-comparison of four models for smoothing satellite sensor time-series data to estimate vegetation phenology. *Remote Sensing of Environment*. 2012. vol. 123. pp. 400–417.
21. Cao R., Chen J., Shen M., Tang Y. An improved logistic method for detecting spring vegetation phenology in grasslands from MODIS EVI time-series data. *Agricultural and Forest Meteorology*. 2015. vol. 200. pp. 9–20.

22. Seo B., Lee J., Lee K., Hong S., Kang S. Improving remotely-sensed crop monitoring by NDVI-based crop phenology estimators for corn and soybeans in Iowa and Illinois, USA. *Field Crops Research*. 2019. vol. 238. pp. 113–128.
23. Vorobyova N.S., Chernov A.V. [Approximation of NDVI time series in the problem of early recognition of crop species from satellite images] *Sbornik trudov III mezhdunarodnoj konferencii i molodezhnoj shkoly «Informacionnye tekhnologii i nanotekhnologii»* [Proceedings of the III International Conference and Youth School «Information Technology and Nanotechnology»]. Samara: Novaya tekhnika, 2017. pp. 390–399. (In Russ.).
24. Hird J.N., McDermid G.J. Noise reduction of NDVI time series: An empirical comparison of selected techniques. *Remote Sensing of Environment*. 2009. vol. 113. no. 1. pp. 248–258.
25. Julien Y., Sobrino J.A. Optimizing and comparing gap-filling techniques using simulated NDVI time series from remotely sensed global data. *International Journal of Applied Earth Observation and Geoinformation*. 2019. vol. 76. pp. 93–111.
26. Stepanov A., Dubrovin K., Sorokin A., Aseeva T. Predicting soybean yield at the regional scale using remote sensing and climatic data. *Remote Sensing*. 2020. vol. 12(12). no. 1936.
27. Stepanov A., Dubrovin K., Sorokin A. Function fitting for modeling seasonal normalized difference vegetation index time series and early forecasting of soybean yield. *The Crop Journal*. 2022. vol. 10. no. 5. pp. 1452–1459.
28. Lin S., Hao D., Zheng Y., Zhang H., Wang C., Yuan W. Multi-site assessment of the potential of fine resolution red-edge vegetation indices for estimating gross primary production. *International Journal of Applied Earth Observation and Geoinformation*. 2022. vol. 113. no. 102978.
29. Gavin H.P. The Levenberg-Marquardt method for nonlinear least squares curve-fitting problems. Department of civil and environmental engineering, Duke University, 2019. vol. 19.
30. Zhao J., Zhang X. An Adaptive Noise Reduction Method for NDVI Time Series Data Based on S–G Filtering and Wavelet Analysis. *Journal Indian Society Remote Sensing*. 2018. vol. 46. no. 12. pp. 1975–1982.
31. Han Y., Meng J., Xu J. Soybean growth assessment method based on NDVI and phenological calibration. *Transactions of the Chinese Society of Agricultural Engineering*. 2017. vol. 33. no. 2. pp. 177–182.
32. Zhang H., Kang J., Xu X., Zhang L. Accessing the temporal and spectral features in crop type mapping using multi-temporal Sentinel-2 imagery: A case study of Yi'an County, Heilongjiang province, China. *Computers and Electronics in Agriculture*. 2020. vol. 176. no. 105618.

Stepanov Alexey — Ph.D., Dr.Sci., Leading researcher, Laboratory of breeding cereals and ear crops, Far Eastern Research Institute of Agriculture (FEARI). Research interests: agricultural economics, forecasting the yield of crops, mathematical modeling in agriculture. The number of publications — 80. stepanfx@mail.ru; 13, Clubnaya St., 680521, Vostochnoye, Russia; office phone: +7(924)210-9102.

Fomina Elizaveta — Engineer, Laboratory of numerical methods of mathematical physics, Computing Center of the Far Eastern Branch of the Russian Academy of Sciences (CC FEB RAS). Research interests: mathematical modeling, approximation. The number of publications — 3. eliz37@mail.ru; 65, Kim Yu Chen St., 680000, Khabarovsk, Russia; office phone: +7(914)316-0703.

Illarionova Lyubov — Ph.D., Senior researcher, Laboratory of numerical methods of mathematical physics, Computing Center of the Far Eastern Branch of the Russian Academy of Sciences (CC FEB RAS). Research interests: machine learning, numerical methods. The number of publications — 40. illarionova_l@list.ru; 65, Kim Yu Chen St., 680000, Khabarovsk, Russia; office phone: +7(924)200-4940.

Dubrovin Konstantin — Junior researcher, Laboratory of numerical methods of mathematical physics, Computing Center of the Far Eastern Branch of the Russian Academy of Sciences (CC FEB RAS). Research interests: mathematical modeling, machine learning, application of remote sensing methods in agriculture. The number of publications — 14. nobforward@gmail.com; 65, Kim Yu Chen St., 680000, Khabarovsk, Russia; office phone: +7(909)859-0881.

Fedoseev Denis — Post-graduate student, Computing Center of the Far Eastern Branch of the Russian Academy of Sciences (CC FEB RAS). Research interests: database management system, programming, numerical methods. d.fedoseev@mail.ru; 65, Kim Yu Chen St., 680000, Khabarovsk, Russia; office phone: +7(962)151-34-35.

Acknowledgements. The studies were financially supported by the Russian Science Foundation project № 23-76-00007 «Development of scientific methods and approaches for sustainable management of soil resources based on remote sensing technologies (in the south of the Far East)».