

УДК 62(01)  
ББК 30  
Т38

Редакционная коллегия сборника:  
*М.Н. Ахметова, Ю.В. Иванова, К.С. Лактионов, М.Г. Комогорцев,  
В.В. Ахметова, В.С. Брезгин, А.В. Котляров, А.С. Яхина*

Ответственный редактор: *О.А. Шульга*

**Технические** науки в России и за рубежом (II): материалы междунар. заоч. науч. конф.  
Т38 (г. Москва, ноябрь 2012 г.). — Москва: Буки-Веди, 2012. — vi, 166 с.

ISBN 978-5-4465-0022-2

В сборнике представлены материалы международной заочной научной конференции «Технические науки в России и за рубежом (II)».

Предназначен для научных работников, преподавателей, аспирантов и студентов технических специальностей, а также для широкого круга читателей.

УДК 62(01)  
ББК 30

# 1. ИНФОРМАТИКА И КИБЕРНЕТИКА

## Прогнозирование урожайности на основе многолетних космических наблюдений за развитием вегетации

Бондур Валерий Григорьевич, доктор технических наук, академик РАН, профессор;

Гороховский Константин Юрьевич, старший научный сотрудник;

Игнатьев Владимир Юрьевич, аспирант;

Мурынин Александр Борисович, кандидат технических наук

Государственное учреждение «Научный центр аэрокосмического мониторинга «Аэрокосмос», Минобрнауки России и РАН (г. Москва)

*Предлагается метод прогнозирования урожайности почв с помощью индекса NDVI на основе многоспектральных космических изображений. Приводятся и анализируются результаты вычислительного эксперимента по прогнозированию урожайности для различных сельскохозяйственных культур для различных областей. Проводится сравнение двух моделей плодородия почв, полученных в ходе исследования.*

**Ключевые слова:** дистанционное зондирование, прогнозирование урожайности, многомерная линейная регрессия, индекс NDVI.

*The method of forecasting soils' productivity using NDVI and based on multispectral space images is suggested. The results of computational experiments are analyzed for different crops and different regions. It is carried out a comparison of two models of soil fertility, which were figure out during the work.*

**Keywords:** remote sensing, yield forecasting, multiple linear regression, NDVI.

### Введение

В настоящей работе приводятся две модели оценки плодородия, с помощью которых исследуется возможность успешного прогнозирования урожайности различных культур.

Подход к решению задачи прогнозирования урожайности заключается в построении модели плодородия на основе космических данных и применении различных методов оптимизации с использованием набора предикторов, полученных в результате компьютерного анализа мультиспектральных изображений из космоса. Предложенный подход предусматривает также дальнейшее расширение и уточнение модели по мере накопления статистических данных об урожайности прошлых лет для различных областей Российской Федерации [20].

Индекс NDVI является одним из наиболее распространенных индикаторов роста и плотности растительности, рассчитываемых по данным ДЗЗ [2–4, 6–8]. Для вычисления этого индекса используется следующая формула:

$$NDVI = \frac{R_2 - R_1}{R_2 + R_1},$$

где  $R_2$  – значение отражения в ближней инфракрасной области спектра,  $R_1$  – значение отражения в красной области спектра. Согласно этой формуле, плотность расти-

тельности (NDVI) в определенной точке изображения равна разнице интенсивностей отраженного света в красном и инфракрасном диапазоне, деленной на сумму их интенсивностей.

Расчет NDVI базируется на двух наиболее стабильных (не зависящих от прочих факторов) участках спектральной кривой отражения сосудистых растений (Рис. 1). В красной области спектра (0,6–0,7 мкм) лежит максимум поглощения солнечной радиации хлорофиллом высших сосудистых растений, а в инфракрасной области (0,7–1,0 мкм) находится область максимального отражения клеточных структур листа. То есть высокая фотосинтетическая активность (связанная, как правило, с густой растительностью) ведет к меньшему отражению в красной области спектра и большему в инфракрасной. Отношение этих показателей друг к другу позволяет четко отделять и анализировать растительные от прочих природных объектов [2,3]. Использование же не простого отношения, а нормализованной разности между минимумом и максимумом отражений увеличивает точность измерения, позволяет уменьшить влияние таких явлений как различия в освещенности снимка, облачности, дымки, поглощение радиации атмосферой и пр.

Из результатов исследований, приведенных в [4], [5], [6], [7] и [8] годовые колебания урожайности наиболее точно предсказываются индексами NDVI в течение ве-

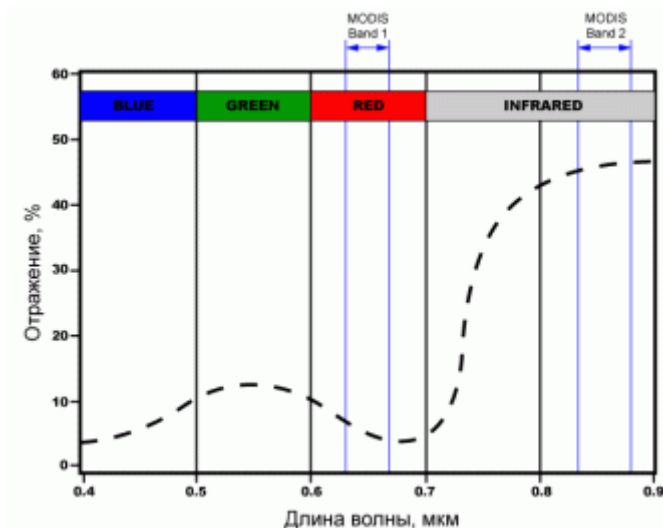


Рис. 1. Участки характеристической кривой отражения растительности (усредненной), используемые для расчета NDVI с помощью данных MODIS

гетацонного сезона — период года, в течение которого максимально проявляются все жизненные процессы растений.

В основу модели плодородия была положена гипотеза о том, что урожай зависит от показателей индекса NDVI в течение вегетативного сезона. Для любого участка с посевами эти значения NDVI имеют различную предикативную силу в зависимости от календарной даты, сельскохозяйственной культуры и особенностей исследуемого региона.

### Краткий обзор существующих методов

За последнее десятилетие в данной научной области сформировались несколько характерных методов и подходов прогнозирования урожайности.

Для прогноза урожайности в настоящее время используется несколько подходов:

- анализ тренда и цикличности в динамичности урожайности [9–13];
- выявление года-аналога [6–8];
- моделирование прироста биомассы растений [13,14];
- на основе анализа синоптических процессов [15–18];
- регрессионный метод [13,19].

Подход, основанный на регрессионном анализе, при наличии достаточно протяженного ряда качественных спутниковых данных позволяет получить достаточно хорошие результаты. В настоящее время работы по развитию данного подхода и анализу возможностей его применения для прогнозирования урожайности различных культур ведутся достаточно активно.

В данной работе сравниваются две модели предсказания урожайности

1. Глобальная линейная модель
2. Линейная модель для отдельных областей

Далее объясняются общие концепции, применяемые при построении этих моделей, а также объясняются особенности каждой модели в отдельности.

### Общие концепции, примененные для построения моделей плодородия

В основе моделей плодородия изложенных здесь заложена зависимость между состоянием растительности в вегетационный период и собранным урожаем. Другими словами если состояние растительности хорошее на протяжении всего периода роста, то вероятность получить хороший урожай увеличивается. В том же случае, если в какой либо промежуток произрастания состояние растительности отклонилось в сторону ухудшения, то вероятность получить хороший урожай уменьшается.

Разработанный метод можно описать следующим образом. Урожай определенной культуры на заданной территории должен достаточно достоверно предсказываться функцией, параметрами которой являются усредненные (по этой территории) значения индекса NDVI в течение роста и созревания сельскохозяйственной культуры. Чем полнее история изменения индекса, тем точнее можно предсказывать урожай. Соответственно, чем отдаленней период измерения индекса от момента созревания культуры, тем меньше будет предикативная сила модели. Исследование зависимости точности предсказания от отдаленности даты созревания является темой отдельной статьи. В данной работе дата предсказания фиксируется как конец Июля и в ней используются показания индекса NDVI в период Март-Июнь.

Предложенная модель урожайности в общей формулировке выглядит как:

$$y_{kr} = f_{kr}(x_{NDVI}(t), x_{NDVI}(t+1), x_{NDVI}(t+2), \dots)$$

где  $Y_k$  — прогнозируемое значение урожайности на окончание текущего сезона для территориального региона  $r$  и сельскохозяйственной культуры  $k$ ,

$f_{kr}$  — искомая функция прогноза урожайности для региона  $r$  и сельскохозяйственной культуры  $k$ ,

$x_{NDVI}(t)$  — значение вегетативного индекса NDVI для данного участка посевов. Для данной модели в качестве единицы регионального деления были выбраны административные области, так как информация об урожайности за предыдущие годы по отдельным культурам официально предоставляется службой государственной статистики Российской Федерации, и она находится в открытом доступе [20]. Наличие этой информации позволяет настроить свободные параметры модели на конкретный регион и сельскохозяйственную культуру.

Для упрощения выбрана линейная модель, то есть  $f_{kr}$  представляет собой линейную комбинацию от  $x_{NDVI}(t)$ :

$$\langle y_k \rangle_{region} = \sum_{t=1}^T \alpha_{kr}(t) \cdot \langle x_{NDVI}(t) \rangle_{region}, \quad (1)$$

где  $\langle \dots \rangle_{region}$  — усреднение по области,  $\alpha_{kr}(t)$  — параметры модели для территориального региона  $r$  и сельскохозяйственной культуры  $k$ , определяемые посредством многопараметрической линейной регрессией.

Информативными для прогнозирования урожайности считаются изображения, соответствующие месяцам, предшествующим началу сбора урожая. Имеется набор мультиспектральных космических изображений за годовой период. Также имеется статистические данные об урожайности. На основании имеющихся данных уточняются коэффициенты модели урожайности для различных сельскохозяйственных культур заданных областей.

### Глобальная линейная модель

Предположим, что почвенно-климатические характеристики мало изменяются для рассматриваемых областей Российской Федерации. При этом модель плодородия упрощается и становится одинаковой для рассматриваемых областей. Она будет выглядеть как:

$$y_{rk} = \sum_{t=1}^T \alpha_k(t) \cdot \langle x_{NDVI}(t) \rangle_r, \quad (2)$$

где  $k$  — индекс, указывающий на сельскохозяйственную культуру

$r$  — индекс, указывающий на область (region) Российской Федерации

$y_{rk}$  — оценка урожайности для данной области ( $r$ ) и культуры ( $k$ )

$\langle x_{NDVI}(t) \rangle_r$  — усредненное значение индекса NDVI для отдельно взятой области Российской Федерации.

$\alpha_k(t)$  — настраиваемые параметры модели для отдельных интервалов времени года.

Параметры модели настраиваются на основе исторических данных об урожайности и соответствующих исторических значениях индекса NDVI.

Подстройка параметров модели осуществляется с помощью линейной регрессии.

### Линейная модель для отдельных областей

В том случае если предположение об минимальных различиях между рассматриваемыми областями неверно, необходимо использовать модель, учитывающую различие между областями, а именно:

$$y_{rk} = \sum_{t=1}^T \alpha_{rk}(t) \cdot \langle x_{NDVI}(t) \rangle_r, \quad (3)$$

где  $\alpha_{rk}(t)$  — параметры модели, которые теперь меняются не только от времени года ( $t$ ), культуры к культуре ( $k$ ), но и от области к области ( $r$ ). Изменения параметров в зависимости от области должны учитывать не только изменения в составе почвы а и средние климатические условия в этом регионе. Остальные переменные в этой формулы уже были описаны ранее.

Достоинство этой модели заключается как раз в том, что выбранная формула позволяет тоньше настраивать параметры модели под особенности конкретной области.

Среди недостатков этой модели следует отметить тот факт, что настройка параметров требует гораздо больше исходных данных, доступных для обучения.

### Описание процесса настройки моделей

Подстройка параметров выбранных моделей производилась с использованием композитных индексов NDVI за период с 2000 по 2009 годы. При подготовке исходных данных использовались шестнадцатидневные композитные изображения с разрешением 250 метров, полученные со сканера MODIS (спутник TERRA), за период в 10 лет [21].

Данные об урожайности за указанный промежуток времени были получены с интернет-сайта Федеральной службы государственной статистики. Настройка производилась для трех групп сельхоз. культур: пшеница, картофель, овощи.

Всего в эксперименте использовались данные по 13-ти региональным образованиям Российской Федерации:

- Московской области;
- Воронежской области;
- Владимирской области;
- Ивановской области;
- Нижегородской области;
- Чувашской республики;
- Мордовской республики;
- Рязанской области;
- Тульской области;
- Орловской области;
- Курской области;
- Тамбовской области;
- Липецкой области.

В предлагаемых моделях в качестве компонентов вектора признаков  $x_{NDVI}$  использовались усредненные по области значения вегетационного индекса NDVI. В каче-

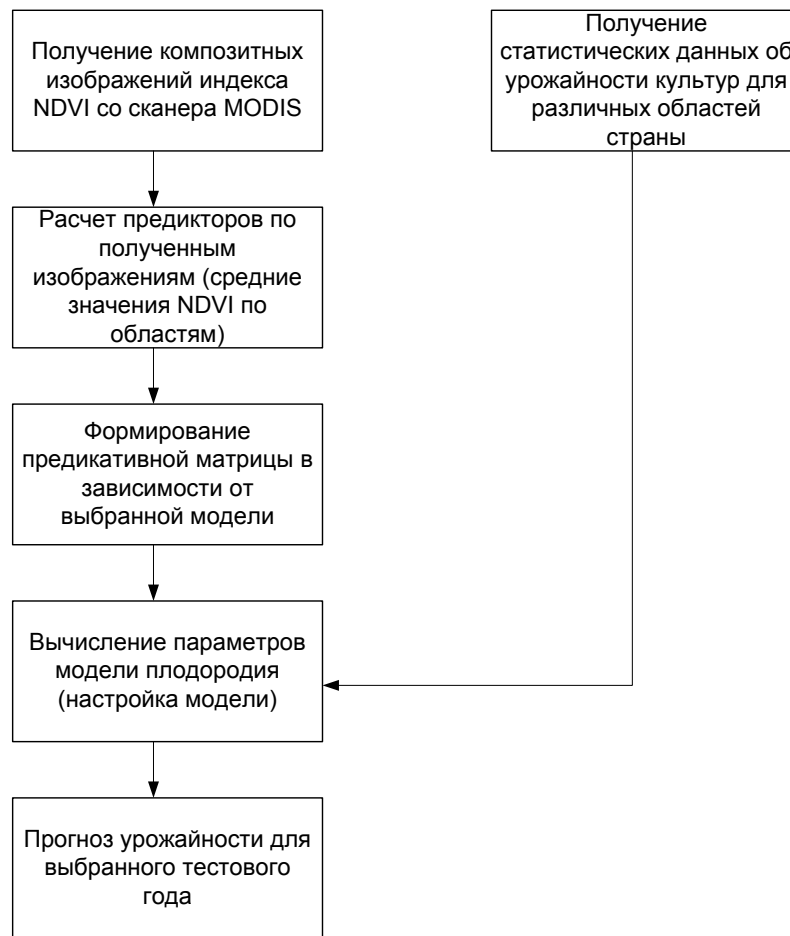


Рис. 2. Метод прогнозирования урожайности по многоспектральным космическим изображениям с помощью многомерной линейной регрессии

стве информативных использованы изображения, полученные за период времени шесть месяцев, которые предшествуют началу сбора урожая (февраль, март, апрель, май, июнь, июль).

Для модели, где учитываются региональные поправки, вектор признаков расширяется бинарными признаками  $b_1, b_2, \dots, b_R$  (где  $R$  — количество областей), определяющими принадлежность объекта к определенной области. Например, если изображение принадлежит московской области, то  $b_1 = 1$ , в противном случае —  $b_1 = 0$ .

#### **Метод прогнозирования урожайности по многоспектральным космическим изображениям с помощью многомерной линейной регрессии**

Метод прогнозирования урожайности по многоспектральным изображениям с помощью многомерной линейной регрессии представлен на рисунке 2.

#### **Результаты вычислительных экспериментов и их анализ**

В результате проведенных вычислительных экспериментов для обеих рассматриваемых моделей получены

графики зависимости урожайности от года для каждой из трех сельскохозяйственных культур и каждой из 13-ти административных областей, рассмотренных в работе. В качестве примера прогнозирования урожайности приведены графики для Московской и Воронежской областей по трем с/х культурам.

На рисунках 3 и 4 представлены графики зависимости урожайности (в центнерах с гектара) для определенной сельскохозяйственной культуры от года за одиннадцатилетний период. Пунктирная линия соответствует фактическим значениям урожайности, а сплошная — спрогнозированным значениям урожайности, вычисленным с помощью полученной линейной модели.

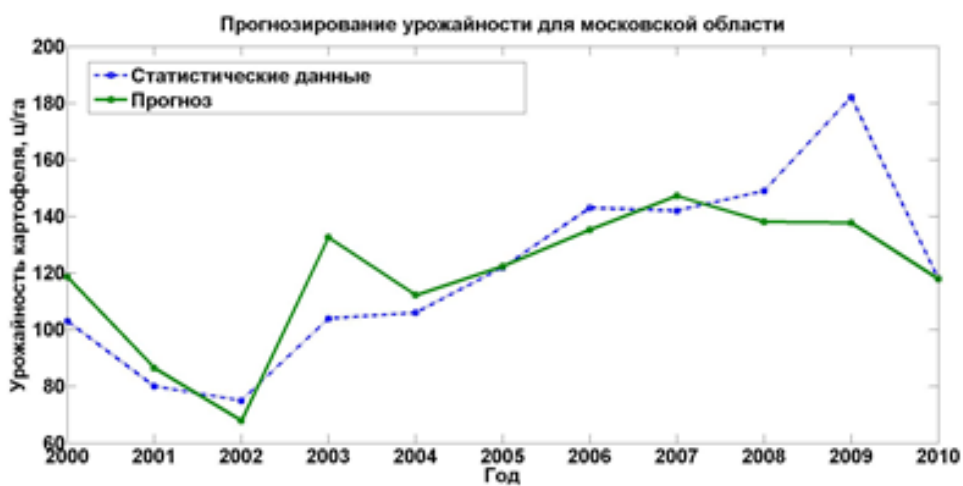
Анализируя результаты прогнозов урожайности, представленные на графиках рисунках 3 и 4, можно сделать следующие выводы:

- предложенный метод прогнозирования урожайности дает возможность заблаговременно предсказывать показатели урожайности для будущего сезона при условии наличия статистических данных об урожайности за прошлые года и мультиспектральных космических изображений для месяцев, предшествующих началу сбора урожая;

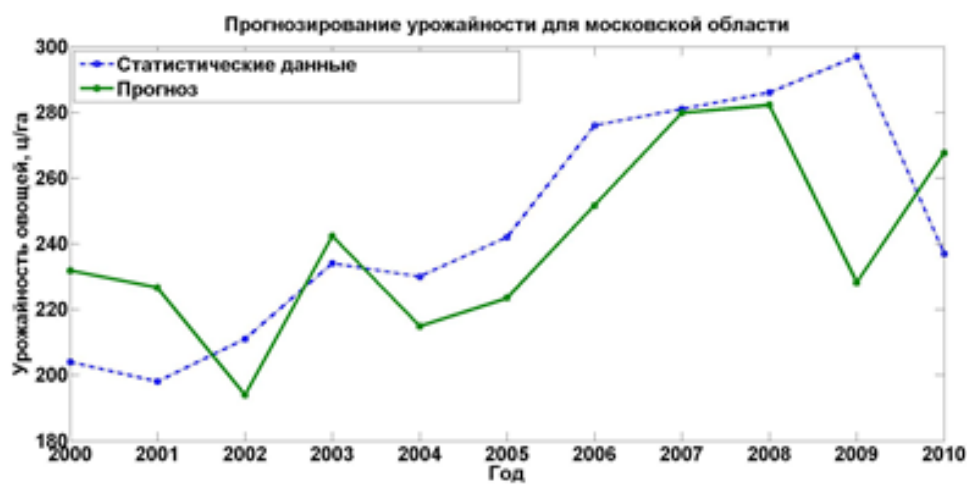
- невысокая точность прогноза урожайности обусловлена отсутствием достаточного количества статистиче-



(а)

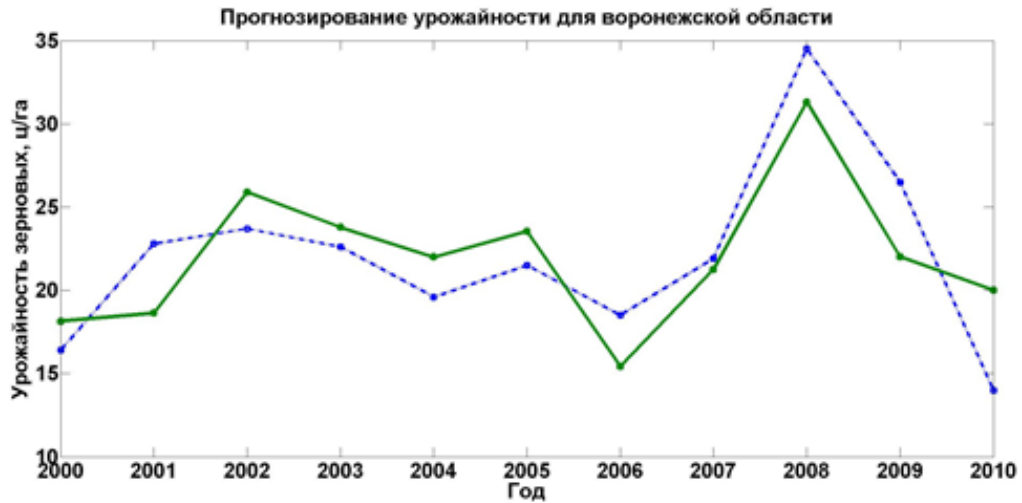


(б)

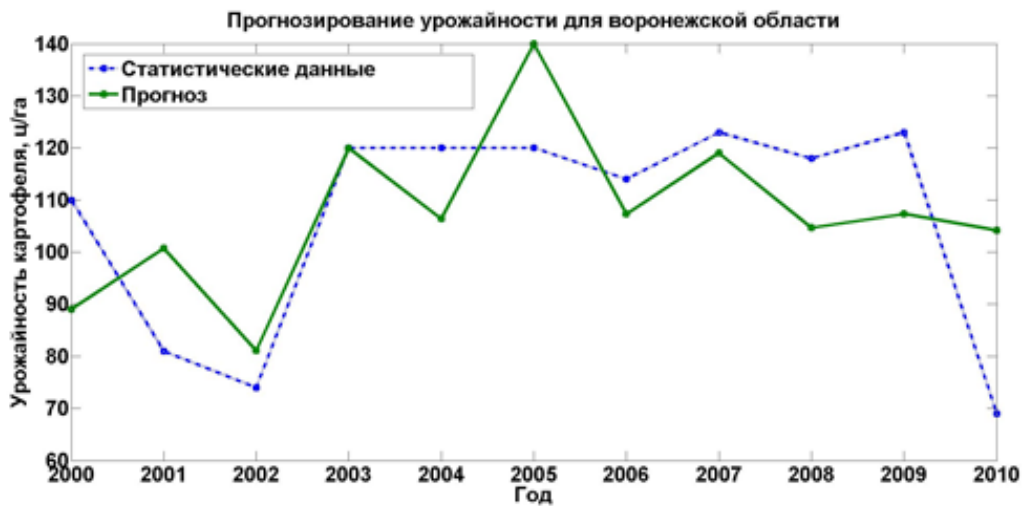


(в)

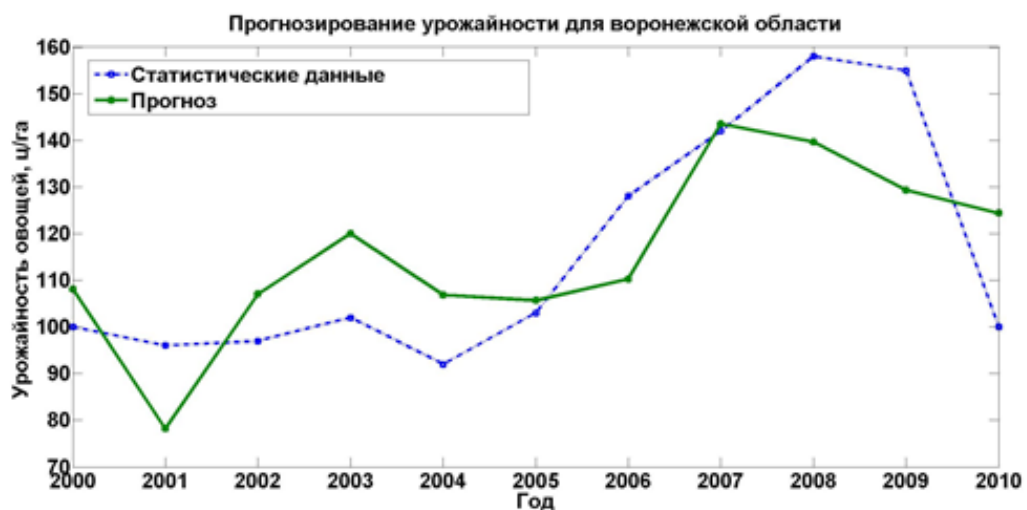
Рис. 3. Результаты прогнозирования урожайности для московской области на 11 лет: (а) – урожайность зерновых, (б) – урожайность картофеля, (в) – урожайность овощей.



(а)



(б)



(в)

Рис. 4. Результаты прогнозирования урожайности для воронежской области на 11 лет: (а) – урожайность зерновых, (б) – урожайность картофеля, (в) – урожайность овощей.

ских данных, используемых при обучении. При этом метод оказывается достаточно устойчивым к использованию его для прогнозирования урожайности основных сельскохозяйственных культур для нескольких областей.

Как видно из графиков, динамика изменения урожайности за 11 лет разная для московской и воронежской областей. Этот факт свидетельствует о том, что для более точного прогнозирования следует использовать в обучении данные по областям близким по территориальным и климатическим характеристикам.

### Заключение

На основании результатов проведенных вычислительных экспериментов продемонстрирована возможность

прогнозирования урожайности по многоспектральным космическим изображениям с использованием реальных статистических данных по урожайности для различных сельскохозяйственных культур и разных областей страны. При этом показано, что описанный подход позволяет прогнозировать урожайность как в условиях минимального различия для областей в глобальной линейной модели, так и его отсутствие в линейной модели для отдельных областей.

Проведен анализ точности прогнозирования урожайности, который показал достоинства и недостатки предложенного метода.

В будущих исследованиях предполагается модификация построенных моделей с целью улучшения точности и общности при прогнозировании урожайности для большего количества областей.

### Литература:

1. Кендалл М., Стюарт А. Многомерный статистический анализ и временные ряды (том 3) // М.: Наука, 1976
2. Sellers, P.J. Canopy reflectance, photosynthesis, and transpiration // International Journal of Remote Sensing, 6, 1335–1372, 1985.
3. Myneni, R.B., F.G. Hall, P.J. Sellers, and A.L. Marshak The interpretation of spectral vegetation indexes // IEEE Transactions on Geoscience and Remote Sensing, 33, 481–486, 1995.
4. Белоруцева Е.В. Мониторинг состояния сельскохозяйственных угодий Нечерноземной зоны Российской Федерации // Современные проблемы дистанционного зондирования Земли из космоса. 2012. Т.9, №1, с. 57–64.
5. Бондур В.Г., Мурнин А.Б., Рихтер А.А., Шманьян, М.А. Разработка алгоритма оценки степени деградации почвы по мультиспектральным изображениям // Известия ЮФУ. 2012. №6, с. 130–134.
6. Savin I., Negre T. (ed.) Agro-meteorological monitoring in Russia and Central Asian Countries, OPOCE EUR 22210EN, Ispra (Italy), 2006. – 214pp.
7. Savin I. Crop yield prediction with SPOT VGT in Mediterranean and Central Asian countries. -ISPRS Archives XXXVI-8/W48 Workshop proceedings: Remote sensing support to crop yield forecast and area estimates. Commission VIII, WG VIII/10. 2007. Stresa, Italy. pp.130–134.
8. Rembold F., Savin I., Negre T., 2005. Developing a simple operational multistep procedure for quantitative yield/production estimation. Proceedings of the AfricaGIS2005 Conference, 31 October to 4 November 2005, The Geo-Information Society of South Africa Tshwane (Pretoria), South Africa ISBN 1–920–01710–0, pp. 257–269.
9. Батырева О.В. Расчет значимости коэффициента множественной корреляции и выбор оптимального числа предсказателей. // Метеорология и гидрология. – 1969. – № 3. – С. 49–57.
10. Пасов В.М. Изменчивость урожаев и оценка ожидаемой продуктивности зерновых культур. -Л.: Гидрометеоздат, 1980. – 152 с.
11. Пасов В.М. Изменчивость урожая яровых зерновых культур в различных климатических зонах СССР. // Метеорология и гидрология. – 1973. – №.7. – с. 82–86.
12. Пасов В.М. Климатическая изменчивость урожаев озимой пшеницы. // Метеорология и гидрология. – 1973. – №2. – с. 94–103.
13. Methodology of the MARS Crop Yield Forecasting System (2008). Eur Rep 21291 EN/1–4.
14. Сиротенко О.Д., Абашина Е.А. Об оценке использования динамических моделей для оценки агрометеорологических условий формирования урожая. Метеорология и гидрология, № 8, 1982, с. 95–101.
15. Пасов В.М., Аксарина Е.А. Синоптико-статистический метод прогноза урожая зерновых культур до сева. // Труды ИЭМ. – 1979. – Вып. 13 (91). – с. 16–41.
16. Воронина В.Ф., Худякова А.И. Особенности атмосферной циркуляции в годы с высокой и низкой урожайностью картофеля в Приморье. // Труды ДВНИГМИ. – 1979 – Вып. 85. – С. 62–69.
17. Steyaert L.T., Le Duc S.K., McQuigg J.D. Atmospheric pressure and wheat yield modeling. Agric. Meteorology, 1978, v.19, N.1, p.23–24.
18. Wang Duo, Zhang Tan. Analysis of the long-term fluctuation of the large-scale circulation and corn yield in the North and the South-West of China. Acta Meteorologica Sinica, 1983, v.41, N.4, p.460–471.
19. Руководство по агрометеорологическим прогнозам. Под ред. Улановой Е.С., Моисейчик В.А., Полевого А.Н. – Л.: Гидрометеоздат, т. 1–2, 1984.



- 
20. Федеральная служба государственной статистики. Регионы России. Социально-экономические показатели. // [http://www.gks.ru/bgd/regl/b11\\_14p/Main.htm](http://www.gks.ru/bgd/regl/b11_14p/Main.htm).
21. U.S. Geological Survey // [ftp://e4ftl01.cr.usgs.gov/MODIS\\_Composites/](ftp://e4ftl01.cr.usgs.gov/MODIS_Composites/).