

УДК 631.524.84

Н.В. Гавриловская

Методы, алгоритмы и технология прогнозирования агрометеорологических факторов в моделях продуктивности зерновых культур*

N.V. Gavrilovskaya

Methods, Algorithms and Technology Forecasting Agro-Meteorological Factors in Models of the Crops Productivity

Рассматривается актуальная научная задача, состоящая в разработке методов, алгоритмов и технологии поэтапного прогнозирования агрометеорологических факторов и урожайности зерновых культур, основанных на принципе аналогичности.

Ключевые слова: технология, алгоритм, агрометеорологические факторы, погодный сценарий, продуктивность.

Введение. Политика государства в последние годы направлена на подъем и стабильное функционирование сельского хозяйства и агропромышленного производства. Об этом свидетельствуют осуществление приоритетного национального проекта «Развитие агропромышленного комплекса», реализация Государственной программы развития сельского хозяйства и регулирования рынков сельскохозяйственной продукции, сырья и продовольствия на 2008–2012 гг. Кроме того, сами производители и сотрудники научно-исследовательских институтов и лабораторий испытывают новые методы и технологии ведения сельскохозяйственного производства.

Для поддержки сельскохозяйственных производителей и управления ценовыми рисками с сентября 2006 г. в России реализуется проект по организации рынка биржевых торгов зерном с использованием механизма форвардных и фьючерсных контрактов, которые заключаются начиная с марта каждого года. В связи с этим до даты заключения контрактов субъектам аграрной сферы необходимо обладать аналитической информацией о складывающихся погодных условиях, их возможном воздействии на объекты сельскохозяйственного производства и ожидаемую продуктивность зерновых культур.

Сдерживающим моментом в решении этих задач является погодный фактор, компоненты которого (среднесуточная температура воздуха, суточная сумма осадков и т.д.) могут изменяться в широком диапазоне от года к году и в течение самого периода вегетации растений.

The article discusses the urgent scientific problem to develop methods, algorithms and technology for stage-by-stage forecasting agro-meteorological factors and grain crop yields based on the principle of similarity.

Key words: technology, algorithm, agro-meteorological factors, weather scenario, productivity

В настоящее время разработаны эффективные моделирующие комплексы производственного процесса, такие как AGROTOOL (Агрофизический институт, Санкт-Петербург, Россия), EPIC (Soil & Water Research Laboratory, USDA-ARS), AGROSIM (Centre for Agricultural Landscape Research, MÜNCHENBERG, Germany) и другие, которые еще до практической реализации того или иного агротехнологического мероприятия способны спрогнозировать его последствия, встраиваясь непосредственно в технологию принятия решений. Однако их использование сдерживается отсутствием необходимой агрометеорологической информации будущего периода.

Разработка алгоритма поэтапного прогнозирования урожайности зерновых культур. Создание алгоритма поэтапного прогнозирования включает несколько этапов: от теоретических исследований до построения прогноза урожайности с помощью моделей продуктивности зерновых культур (рис. 1) и предусматривает использование современных компьютерных технологий.

Разработанный автором алгоритм состоит из следующих основных этапов:

1) *формирование информационного обеспечения:* сбор и расчет агрометеорологических показателей, предварительный статистический анализ и оценка достоверности экспериментальных данных, необходимых для идентификации алгоритма;

2) *технология определения лет-аналогов:* кластеризация данных, выбор оптимального разбиения на кластеры, образование класса лет-аналогов для исследуемого года;

* Работа выполнена при финансовой поддержке ведомственно-аналитической программы «Развитие научного потенциала высшей школы 2009–2011» (проект №2.2.2.4/4278).

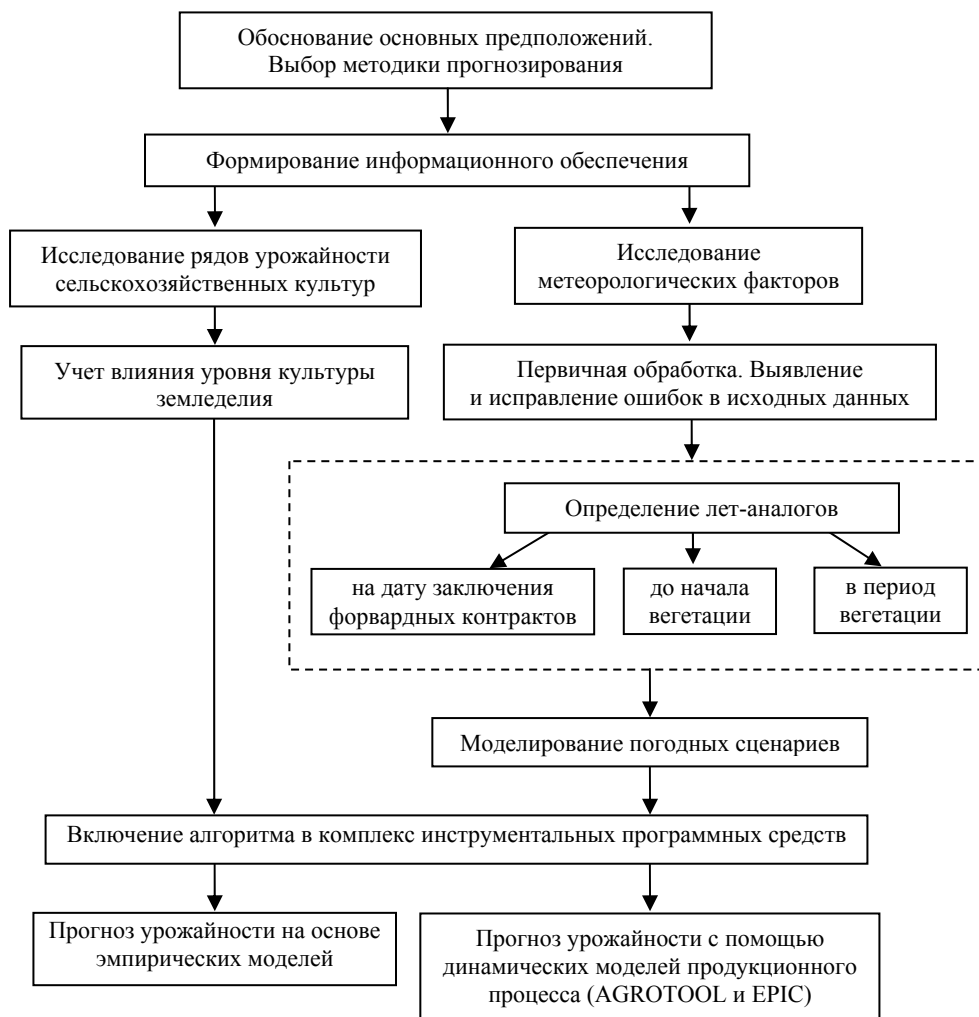


Рис. 1. Блок-схема алгоритма поэтапного прогнозирования урожайности зерновых культур

3) *моделирование погодных сценариев*, основанное на использовании двух подходов: принципа аналогичности и стохастических методов;

4) *разработка комплекса инструментальных программных средств*;

5) *прогноз урожайности на основе прогностических моделей*;

6) *оценка урожайности с помощью имитационно-моделирующего комплекса AGROTOOL и имитационной модели EPIC*.

Рассмотрим содержание и реализацию некоторых этапов.

1. Формирование информационного обеспечения. В рамках данного этапа разработанного алгоритма сформирована методика предварительного статистического анализа и оценки достоверности экспериментальных агрометеорологических данных. Центральное место в рассматриваемой методике занимают вопросы обработки и анализа накапливаемой информации.

Как показывает практика, данные агрометеорологических наблюдений являются неоднородными

и обладают широким диапазоном вариации, сильной колеблемостью и связностью. Это осложняет их обработку стандартными статистическими методами. При формировании информационного обеспечения исследуется многомерный ряд урожайности зерновых культур, что позволяет изучить показатели динамики временного ряда, включая циклические свойства и нестационарность.

2. Технология определения лет-аналогов. В основе данной технологии лежит принцип аналогичности и классификации агрометеорологических факторов [1]. Первый этап в технологии состоит в отнесении всей совокупности исходных объектов к определенному классу на основании исследования системы признаков или показателей, характеризующих эти объекты.

Пусть из множества Ω всех исследуемых объектов:

$$\Omega = \left\{ (X^r, Y) : X^r = \{x_{ij}^r\}, Y = \{y_r\}, \right. \\ \left. r = \overline{1, n}; i = \overline{1, t}; j = \overline{1, m}; x_{ij} \in R, y_r \in R \right\}$$

необходимо сформировать по решающему правилу S непересекающиеся подмножества $A_k \in \Omega$ – классы сходных между собой объектов – лет-аналогов вида $A_k = \{(X^N, Y_N) : X^N \in X^r, Y_N \in Y, N = \overline{1, a_k}, a_k \leq n\}$.

Здесь $X^r = \{x_{ij}^r\}$ – матрица значений агрометеорологических факторов размера $t \times m$, определенная для каждой совокупности наблюдаемых признаков, влияющих на вектор y_r ; y_r – вектор значений фактической урожайности; n – число исследуемых лет; m – число имеющихся агрометеорологических характеристик; t – дискретный момент времени; k – количество сформированных классов, обладающих совокупностью факторов, близких между собой по влиянию на результирующий признак объекта y_N ; a_k – количество лет-аналогов в соответствующем классе.

Под решающим правилом S будем понимать обеспечение экстремума функционала $F(A_k) \rightarrow \min(\max)_{A_k}$ – мера однородности объектов, где $F : X \times X \rightarrow R$. В случае зависимых признаков за меру однородности объектов принимают расстояние Махаланобиса

$$D_M = \sqrt{(x'_i - \bar{x})^T \Sigma^{-1} (x'_i - \bar{x})}, \quad i = \overline{1, t},$$

где x'_i – многомерный вектор признаков; Σ – корреляционная матрица; \bar{x} – центр класса (кластера).

Под объектами классификации понимаются годы, а в качестве признаков или показателей, характеризующих эти объекты, выступают агрометеорологические факторы.

Для сравнения нескольких типизаций и выбора наиболее оптимальной из них необходим критерий – численная мера качества классификации. Оценкой качества классификации может служить показатель

$$K = K_w / K_b,$$

где $K_w = \frac{2}{k \cdot (k-1)} \sum_{i=1}^k \overline{d_{ii}}$,

$$K_b = \frac{1}{k} \sum_{i=1}^k \sum_{j=i+1}^k \overline{d_{ij}}.$$

Здесь K_w – внутрикластерные и K_b – межкластерные расстояния; k – число кластеров; $\overline{d_{ii}}$ – среднее расстояние между точками внутри i -го кластера; $\overline{d_{ij}}$ – среднее расстояние между парами точек i -го и j -го кластеров.

В результате данной процедуры, во-первых, формируется обучающая выборка $X_{l_1}, X_{l_2}, \dots, X_{l_n}$, $l = \overline{1, k}$, где X_i – вектор многомерных наблюдений; k – общее число классов, выявленных в процессе предварительной типологизации; а о наблюдениях X_{li} известно, что все они характеризуют объекты, принадлежащие l -му классу. Во-вторых, строится классификатор (дискриминантная функция) каждого

классифицируемого объекта, заданного значениями своих описательных признаков.

3. Моделирование погодных сценариев. На третьем этапе алгоритма поэтапного прогнозирования урожайности зерновых культур рассмотрена система моделирования погодных сценариев, основанная на использовании двух подходов: принципа аналогичности и стохастических методов.

Моделирование погодных сценариев по принципу аналогичности – это задача второго этапа технологии определения лет-аналогов, которая заключается в том, чтобы из всех подмножеств $A_k \in \Omega$ выбрать класс объектов A_{k_0} , который лучше всего соответствует (согласно определенным критериям) новому элементу

$$X^{n+1} = \{x_{i_0 j}^{n+1}\}, \quad i_0 = \overline{1, l_0}; \quad j = \overline{1, m}, \quad l_0 < l;$$

l обычно принимают равным 365 дням; l_0 – номер дня, с которого производится моделирование погодного сценария. Для оценки влияния погодных условий на формирование урожая требуется на основе исследования совокупности агрометеорологических параметров классифицировать ситуацию в определенный период времени, учитывая ее влияние на состояние растений, точнее, на урожайность.

Вследствие этого сформированный класс объектов A_{k_0} образует ансамбль возможных реализаций погодных условий, который можно описать вектором $G = \{G^0, G^1, \dots, G^{k_0}\}$, где G^0 – исследуемый год; G^1, \dots, G^{k_0} – годы-аналоги.

Тогда прогнозный сценарий $\hat{G}^0(t+l')$, $l' = \overline{l_0, l}$ для исследуемого года можно построить с помощью оптимизационной процедуры:

$$O = \sum_{t=1}^{l_0-1} [\hat{G}^0(t) - G^0(t)]^2 \rightarrow \min,$$

где $\hat{G}^0(t) = \sum_{i=1}^{k_0} \alpha_i G^i$, $\sum_{i=1}^{k_0} \alpha_i = 1$, $\alpha_i \geq 0$ – параметры аналогичности.

Использование этого метода предполагает, что при работе модели на ее вход подаются фактические погодные данные до того момента, с которого начинается прогнозирование. Для состыковки (сглаживания) фактических данных и данных лет-аналогов фиксируются отклонения фактических данных на дату прогноза и отклонения данных года-аналога. Эти отклонения сглаживаются с помощью динамического звена первого порядка, отфильтровывающего флуктуации.

Второй метод имеет *стохастический (вероятностный) характер*. Источником новых погодных реализаций служит так называемый *генератор погоды*, в котором ежесуточные метеорологические данные, такие как максимальная и минимальная температуры воздуха, минимальная влажность воздуха, скорость ветра, осадки и коэффициент ослаб-

ления солнечной радиации, моделируются как многомерный случайный нестационарный процесс [2]. Такой подход базируется на авторегрессионной модели, введенной Ричардсоном [3] и модифицированной в лаборатории математического моделирования агроэкосистем Агрофизического института Санкт-Петербурга [4].

Стохастический генератор погоды моделирует синтетические ежесуточные ряды метеорологических элементов со статистическими характеристиками, близкими к тем, что и в исторических данных фактических реализаций погоды за 20–30 прошлых лет.

Итогом моделирования погодных сценариев является «веер» возможных траекторий формирования урожая и соответствующие ему множества возможных значений потенциального вероятностного прогноза ресурсов продуктивности. Понятие выработки сценария погоды не означает, что в результате будет получена реализация метеоусловий, когда-либо встречавшихся в данной местности.

Основная цель моделирования погодных сценариев заключается в том, что эта процедура, будучи использована в качестве входных данных математических моделей продукционного процесса, даст результат, решающий задачу прогноза того или иного параметра продукционного процесса. Таким образом, полученные в результате моделирования метеорологические ситуации присоединяются к имеющимся фактическим метеоусловиям, образуя полный набор ежедневных входных данных для математических моделей продуктивности зерновых культур.

4. Разработка комплекса инструментальных программных средств. Особая роль в системе оперативного агрометеорологического обеспечения

сельскохозяйственного производства отводится информационно-прогностическим системам обработки и анализа агрометеорологической информации, позволяющим в результате обобщения этой информации осуществлять прогноз агрометеорологических факторов и урожайности зерновых культур.

В связи с этим возникла необходимость разработки программного комплекса для обработки экспериментальных агрометеорологических данных (Program Complex for Processing and Forecasting Agrometeorological Factors) и информационного обеспечения моделей продуктивности зерновых культур [5]. В состав данного программного комплекса входят:

- 1) база экспериментальных данных;
- 2) блок формирования и первичной обработки агрометеорологических факторов;
- 3) блок реализации технологии определения лет-аналогов (применение принципа аналогичности для формирования погодных сценариев с помощью кластерного и дискриминантного анализов).

Разработка и формирование базы экспериментальных данных проводились на основе многолетних материалов наблюдений Агрометеорологической станции Барнаул и АНИИСХОЗ им. В.В. Докучаева.

Для компьютерной реализации комплекса и реляционной модели базы данных выбрана вычислительная платформа Java (NetBeans). В системе разработан интерфейс, позволяющий осуществлять экспорт и импорт во внешние источники, редактировать и формировать данные в базе.

База экспериментальных данных AMDB (AgroMeteorological Data Base) представлена в виде иерархически организованной совокупности управляющих и подчиненных таблиц данных (рис. 2).

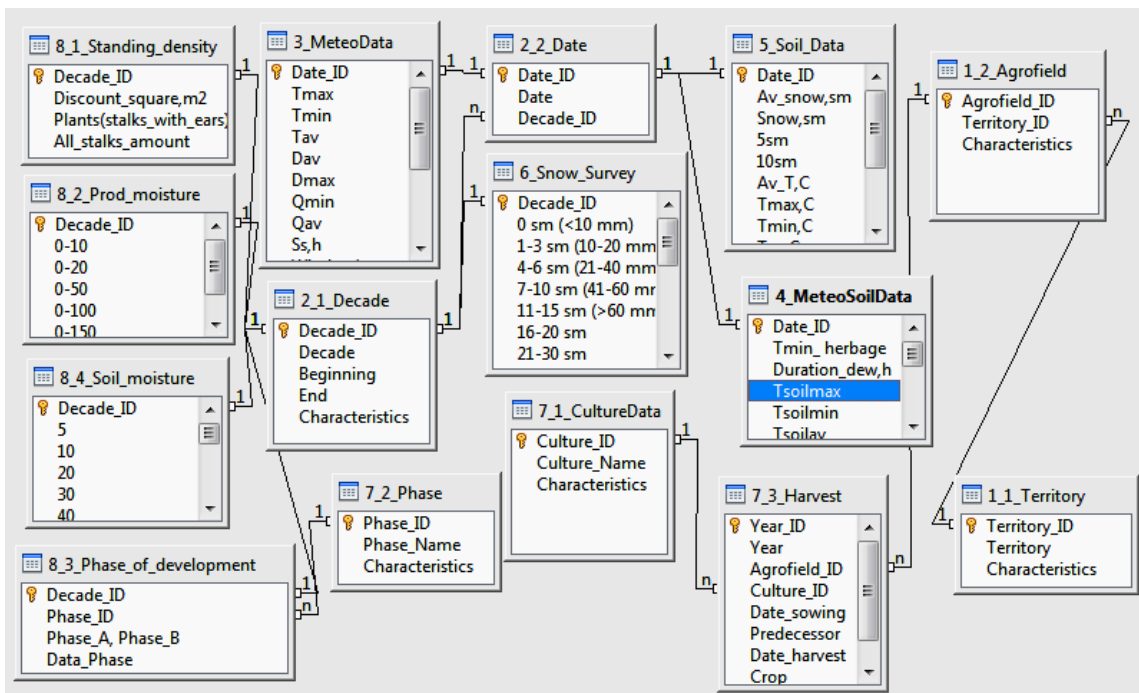


Рис. 2. Схема концептуальной модели базы данных AMDB

Информационная модель AMDB спроектирована с помощью специализированной библиотеки SQLite, позволяющей в полной мере использовать современный инструментальный средств работы с базами данных по SQL-стандарту.

В программном комплексе (рис. 3) реализован блок формирования и первичной обработки агро-

метеорологических факторов, позволяющий проводить оценку достоверности экспериментальных данных: анализ, выбраковку и восстановление аномальных (сбитых) или пропущенных наблюдений, а также блок реализации технологии определения лет-аналогов по алгоритму поэтапного прогнозирования урожайности, описанного в п. 2.



Рис. 3. Интерфейс программного комплекса обработки экспериментальных агрометеорологических данных

Разработанный программный комплекс может быть непосредственно использован при создании программно-алгоритмического обеспечения системы поддержки принятия управленческих решений и в существующих моделях продуктивности и прогнозирования урожайности зерновых культур.

Заключение. Разработанный автором алгоритм включает методы, алгоритмы и технологию прогно-

зирования агрометеорологических факторов в моделях продуктивности зерновых культур.

В статье подробно описаны математическая формализация принципа аналогичности и алгоритм формирования сценариев из погодных реализаций лет-аналогов и генераторов погодных данных. Предложенная технология определения лет-аналогов может быть использована для оценки урожайности зерновых культур.

Библиографический список

1. Гавриловская Н.В., Топаж А.Г., Хворова Л.А. Моделирование погодных сценариев для оценки урожайности зерновых культур в условиях Западной Сибири // Известия АлтГУ. – Барнаул, 2011. – №1 (69).
2. Хворова Л.А., Брыксин В.М., Гавриловская Н.В. Использование информационных технологий при прогнозировании урожайности зерновых культур // Научно-технические ведомости СПбГПУ: Информатика. Телекоммуникации. Управление. – СПб., 2009. – №5 (86).
3. Richardson C.W., Wright D.A. WGEN: A model for generating daily weather variables // US Department of

Agriculture, Agricultural Research Service, ARS-8. USDA. – Washington, 1984.

4. Топаж А.Г. Моделирование суточных метеоданных как входного сигнала модели продукционного процесса // Почва и растение – процессы и модели: сб. науч. тр. – СПб., 1992.

5. Гавриловская Н.В. Программный комплекс обработки и прогнозирования агрометеорологических факторов // Студент и научно-технический прогресс. Информационные технологии: материалы XLIX Междунар. науч. студ. конф. – Новосибирск, 2011.