

Н.В. Гавриловская, Л.А. Хворова

Информационно-прогностическая система сбора, обработки, анализа и обобщения агрометеорологической информации*

Ключевые слова: информационно-прогностическая система, база данных, год-аналог, урожайность, вегетационный период, прогноз, кластерный анализ, модель.

Key words: information-prediction system, database, year-analogue, the productivity, the growing season, the forecast, cluster analysis, model.

Введение. Один из важных показателей, характеризующих деятельность сельскохозяйственных организаций, – урожайность сельскохозяйственных культур. Урожайность – комплексный показатель, являющийся, с одной стороны, исходной информацией для построения планов, прогнозов и принятия управленческих решений, с другой – один из основных результирующих показателей сельскохозяйственного производства. Поэтому особая роль в системе оперативного агрометеорологического обеспечения сельскохозяйственного производства отводится информационно-прогностическим системам сбора, обработки и анализа агрометеорологической информации, позволяющим в результате обобщения этой информации осуществлять прогноз урожайности и валового сбора сельскохозяйственных культур.

Разработанная авторами информационно-прогностическая система включает:

- 1) базу агрометеорологических данных;
- 2) комплекс эмпирико-статистических и динамических моделей;
- 3) технологию определения лет-аналогов;
- 4) методы оценки погрешности и оправданности прогнозов.

1. База агрометеорологических данных. Разработка и формирование базы данных проводились на основе многолетних материалов наблюдений Алтайской гидрометеорологической станции (с 1927 по 2008 г.) и средней урожайности яровой пшеницы (с 1961 по 2008 г.). Множество данных наблюдений формировалось по принципу: оно должно быть достаточно информативным для того, чтобы обеспечить выбор или построение хорошей прогностической модели урожайности и обеспечить качество определения лет-аналогов.

Таким образом, база данных содержит:

- 1) многолетние ежедневные метеорологические данные за 81 год (минимальная и максимальная температура воздуха (град С); минимальная относитель-

ная влажность воздуха (%); осадки (мм); скорость ветра (м/с); солнечная активность (число Вольфа);

- 2) агрометеорологическую информацию (данные наблюдений за посевом): влажность почвы по слоям (см); температура почвы (град С); даты наступления фенофаз;

- 3) гидрофизические характеристики по полям: тип почвы и мехсостав, удельная масса почвы (г/см³); объемная масса почвы (г/см); максимальная гигроскопичность (%); влажность устойчивого завядания (%); наименьшая влагоемкость (%); капиллярная влагоемкость (%); полная влагоемкость (%); нижняя и верхняя границы реальных значений влажности почвы;

- 4) информацию по технологиям: дата сева и дата уборки; урожайность культур.

Предварительный анализ исходных данных показал, что данные агрометеорологических наблюдений являются неоднородными и обладают широким диапазоном вариации, сильной колеблемостью и связностью. Это осложняет их обработку стандартными статистическими методами.

В процессе анализа временного ряда урожайности яровой пшеницы были изучены факторные признаки формирования урожая: сумма эффективных температур и ее влияние на формирование урожая в различные фазы вегетационного периода, сумма осадков в осенне-зимний, предпосевный и другие периоды вегетации, солнечная активность. Изучены динамики времени временного ряда, включая циклические свойства урожайности яровой пшеницы, показана нестационарность временного ряда урожайности.

Анализ качества исходных данных состоял в выявлении аномальных данных, их влиянии на результаты последующей обработки, определении причин их возникновения.

2. Комплекс прогностических моделей. К началу 90-х гг. XX в. в практике Росгидромета сформировался комплекс методов оценок состояния и прогноза продуктивности посевов. Он состоит из множества методов, которые отличаются друг от друга методической основой, набором исходной информации, заблаговременностью составления прогноза, масштабом территории посевов и технологией составления прогноза. Прогностические зависимости устанавливаются для урожаев сельскохозяйственных культур с учетом большого количества наблюдений инерционных факторов (запасов влаги в почве, фаз развития растений, их высоты, количества стеблей и т.д.).

Как показывает опыт использования этих прогностических зависимостей для оценки урожайности сельскохозяйственных культур, они, как прави-

* Работа выполнена при поддержке ведомственно-аналитической программы «Развитие научного потенциала Высшей школы 2009–2010» (проект №2.2.2.4/4278).

ло, не применимы в других регионах. Кроме того, используемые в практике методы прогнозов с течением времени устаревают и перестают удовлетворять современным требованиям к качеству, возможностям и заблаговременности прогнозов.

Поэтому для прогноза урожайности авторами были разработаны и исследованы эмпирические [1] и статистические модели [2]. Выбор эмпирико-статистических моделей объясняется тем, что потенциальными пользователями (подразделениями Росгидромета, управлениями сельского хозяйства, агрономами) в основном используются статистические, синоптико-статистические, физико-статистические, динамико-статистические модели. Их достоинством служат простота и общедоступность использования. Разработанные авторами прогностические модели отличаются от уже разработанных и используемых в оперативной практике Росгидромета методической основой, набором исходной информации, заблаговременностью составления прогноза и технологией составления прогноза.

Особое место в технологии прогноза урожайности зерновых культур должны занимать прогнозы, позволяющие давать заблаговременную оценку ожидаемой продуктивности посевов поэтапно: до сева яровых культур, спустя две декады после сева и в другие фазы вегетационного периода [3]. Это значительно увеличит надежность прогнозов урожайности и эффективность использования прогностической информации для решения вопросов формирования резервных фондов продовольствия, наличия необходимых мощностей для хранения полученного урожая, строить адекватную и эффективную политику внешней торговли.

Методика заблаговременных прогнозов создавалась в рамках технологии определения года-аналога и подробно описана в работах авторов [1, 2, 4]. Согласно потребностям в прогностической информации о будущем урожае, были определены следующие временные интервалы прогноза: по прошествии осенне-зимнего периода (конец марта – начало апреля), предпосевной период (начало мая), спустя две декады после сева. В дальнейшем величина урожая может быть уточнена и в другие периоды вегетационного цикла.

Учитывая влияние агрометеорологических факторов на будущий урожай, в качестве основных параметров моделей были взяты количество осадков за осенний (x_1) и зимний периоды (x_2), сумма осадков за весенний (предпосевной) период (x_3); за первые две декады вегетационного периода: сумма эффективных температур $>10^\circ\text{C}$ (x_4), сумма осадков (x_5), количество дней с осадками (x_6); урожайность (y).

Выбор данных параметров объясняется тем, что они отражают климатические условия возделывания культуры на конкретной территории и, следовательно, параметры моделей прогноза будут отражать особенности этой конкретной территории.

Для построения прогностических моделей использовали аппарат классического регрессионного

анализа (метод наименьших квадратов – МНК) и компонентный анализ [5, 6] или метод главных компонент (МГК). Подробный алгоритм построения моделей с использованием МГК описан в работе [2]. В таблице 1 приведены прогностические модели по этапам осуществления прогноза, расчетная величина урожая для 2007 и 2008 гг., абсолютная ошибка прогноза. Для построения моделей использованы данные с 1961 по 1997 г. Необходимая метеорологическая информация считалась известной только на момент составления прогноза. Фактическая урожайность в 2007 г. составила 17,6 ц/га, в 2008 г. – 18,9 ц/га.

Из таблицы видно, что ошибки прогноза по уравнениям, построенным с помощью МГК, вполне приемлемы (не выходят за пределы 12%), поэтому можно считать, что результаты численных расчетов величины ожидаемой урожайности хорошо согласуются с фактическими данными. Анализ результатов прогнозирования приведен в четвертом разделе.

3. Технологія определения года-аналога. Модели любого уровня, осуществляющие прогноз, должны опираться на предполагаемые погодные сценарии прогнозного года. В настоящее время важнейшие метеорологические факторы – осадки и температура воздуха, имеющие решающее значение в формировании урожая сельскохозяйственных культур, не поддаются надежному прогнозу даже на месяц вперед. Отсутствие фундаментальной теории, объясняющей наличие связи между процессами в отдельные сезоны и урожайностью, обуславливает применение в методиках в основном аппарата математической статистики, в частности кластерного и дискриминантного анализов. В результате использования данных процедур каждый год на основе различных сочетаний значений метеорологических характеристик может быть отнесен к одному из нескольких классов. Данные процедуры используются в технологии определения года-аналога.

Главную задачу в технологии определения года-аналога составляет отнесение исходных объектов к определенному классу на основании исследования системы признаков или показателей, характеризующих эти объекты. Иными словами, для оценки влияния погодных условий на формирование урожая требуется на основе исследования совокупности агрометеорологических параметров классифицировать ситуацию в определенный период вегетационного цикла, учитывая ее влияние на состояние растений, точнее на урожайность. Данная задача является типичной задачей распознавания образов.

Под объектом классификации будем понимать годы, а в качестве признаков или показателей, характеризующих эти объекты, выступают агрометеорологические параметры: сумма осадков, число дней с осадками, сумма эффективных температур за определенный период вегетации и урожайность культуры.

Продемонстрируем технологию определения года-аналога для конкретного объекта – 1997 г. (считаем, что погодный сценарий данного года известен на момент определения года-аналога).

Прогностические модели по этапам осуществления прогноза

№ периода	Период	Модель		Y расчетное, ц/га		Относительная ошибка прогноза, %	
				2007 г.	2008 г.	2007 г.	2008 г.
1	Осень–зима	МГК(1)	$y = 0,02x_1 - 0,02x_2 + 18,49$	17,69	17,90	0,52	5,28
		МНК	$y = 0,03x_1 - 0,01x_2 + 17,81$	18,89	18,76	7,34	0,72
2	Осень–зима–весна	МГК(2)	$y = 0,04x_1 - 0,004x_2 + 0,04x_3 + 14,91$	17,59	19,28	0,03	2,01
		МНК	$y = 0,02x_1 - 0,02x_2 + 0,07x_6 + 16,95$	16,75	20,16	4,82	6,65
3	Осень–зима–весна–две декады	МГК(4)	$y = 0,03x_1 - 0,01x_2 + 0,03x_3 - 0,01x_4 + 0,08x_5 + 0,36x_6 + 14,92$	17,92	16,69	1,81	11,72
		МНК	$y = 0,01x_1 - 0,03x_2 + 0,03x_3 + 0,001x_4 + 0,19x_5 + 0,01x_6 + 14,61$	17,57	14,63	0,19	22,58

Примечания: МГК (1) – прогностическое уравнение, построенное по методу главных компонент, с учетом одной компоненты; МГК (2) – прогностическое уравнение, построенное по методу главных компонент, с учетом двух первых компонент и т.д.; МНК – прогностическое уравнение, построенное по методу наименьших квадратов.

Классификацию осуществляли по фактическим данным за первые две декады вегетационных периодов 1971–1997 гг. Процедуру кластеризации проводили несколько раз с помощью программы статистической обработки данных Statistica 6.0 при различных значениях числа кластеров. После чего выбиралась лучшая группировка в смысле критерия минимума отношений средних межкластерных и внутрикластерных расстояний [7]. Лучший вариант – разбиение на шесть кластеров.

Анализ проведенной классификации позволяет предположить, что для 1997 г. годами-аналогами будут являться 1973, 1976, 1981, 1988 гг. (табл. 2).

Классификация агрометеорологических ситуаций была проведена повторно с привлечением дополнительных четырех лет (1967–1970 гг.), не участвующих в первой классификации. Результаты повторной классификации по первым двум декадам вегетационного периода остались прежними – в один кластер попали 1973, 1976, 1981, 1988 гг. Это говорит об устойчивости результатов классификации в пространстве погодных факторов от уровня их зашумленности.

Процедура определения года-аналога так же, как и оценка урожайности, осуществляется поэтапно. До сева яровых (в осенне-зимний и предпосевной периоды) по уже известным погодным данным осуществляется прогноз урожайности по разрабо-

танным эмпирико-статистическим моделям и определяются годы-аналоги (или год-аналог) для данного года, по погодному сценарию которого корректируется прогнозная оценка. Уточнение оценки урожайности производится через две декады после сева и в другие периоды вегетационного периода с корректировкой года-аналога. Рамки данной статьи не позволяют полностью привести все результаты поэтапной классификации и оценки урожайности.

4. Методы оценки погрешности и оправдываемости прогнозов. Анализ результатов прогноза.

Оценка достоверности и качества полученных прогностических значений урожайности (табл. 1) нельзя оценить по статистическим критериям, так как выполняются не все необходимые условия их применимости. Это, в принципе, объясняется специфичностью агрометеорологической информации, которая не всегда укладывается в рамки статистических законов. Поэтому оценка качества прогнозов и, соответственно, успешность разработанной технологии долгосрочного прогноза урожайности яровой пшеницы производилась по их оправдываемости относительно допустимой погрешности прогноза ($0,67\sigma$, σ – среднее квадратическое отклонение), согласно Инструкции по оценке оправдываемости агрометеорологических прогнозов¹ и Методическим указаниям².

Таблица 2

Годы-аналоги агроклиматических условий первых двух декад вегетационного периода для 1997 г.

Характеристики	1997 г.	Годы-аналоги			
		1973 г.	1976 г.	1981 г.	1988 г.
Сумма температур выше 5 °С	264,1	280,8	247,1	268,1	271,8
Сумма осадков, мм	15,9	22,4	4,2	8,6	9,3
Количество дней с осадками	5	6	4	6	7
Урожайность фактическая, ц/га	12,8	17,5	13,7	13,4	15,9

¹ Инструкция по оценке оправдываемости агрометеорологических прогнозов. М., 1983.

² Методические указания по проведению производственных (оперативных) испытаний новых и усовершенствованных методов гидрометеорологических и гелиогеофизических прогнозов. Л., 1991.

Из таблицы 1 видно, что средняя ожидаемая урожайность по трем этапам прогноза с использованием моделей МГК для 2007 г. – 17,7 ц/га; средняя относительная ошибка – 0,78%. Оправдываемость прогноза по всем трем этапам составляет 100%. По моделям МНК средняя ожидаемая урожайность по трем этапам прогноза – 17,7 ц/га; средняя относительная ошибка – 4,12%. Оправдываемость прогноза по всем трем этапам составляет 96%. Для 2008 г. средняя ожидаемая урожайность по трем этапам прогноза с использованием моделей МГК – 18 ц/га; средняя относительная ошибка – 6,34%. Оправдываемость прогноза по всем трем этапам составляет 94%. По моделям МНК средняя ожидаемая урожайность по трем этапам прогноза – 17,9 ц/га; средняя относительная ошибка – 9,98%. Оправдываемость прогноза по всем трем этапам составляет 90%.

Проанализируем результаты, полученные в третьем разделе. Годами-аналогами для 1997 г. будут являться 1973, 1976, 1981, 1988 гг. Урожайность 1997 г. по фактическим данным этих лет будет находиться в пределах 13,4–17,5 ц/га. По погодному сценарию наиболее близкие годы – 1976, 1981, 1988 гг. (табл. 2). Следовательно, интервал для ожидаемой урожайности можно сузить: 13,4–15,9 ц/га.

Сравнительный анализ результатов классификации по метеорологическим характеристикам первых двух декад вегетационного периода (табл. 2) показал, что годом-аналогом для 1997 г. является 1981 г. Величина урожая по погодному сценарию 1981 г. составит 13,4 ц/га (фактическая – 12,8 ц/га). Абсолютная ошибка прогноза по фактической урожайности 1981 г. – 0,6 ц/га, что не превышает допустимую погрешность ($0,67\sigma$, $\sigma = 4,3$ вычислено по фактическому ряду урожайности).

Заключение. Серьезность влияния, оказываемого информацией на планирование и принятие решений организациями, привела к росту понимания того, что информация – это ресурс, обладающий определенной ценностью, и, следовательно, нуждающийся в упорядочении и управлении. В статье рассмотрена одна из практических задач использо-

вания агрометеорологической информации – поэтапная оценка урожайности зерновых культур на примере яровой пшеницы с использованием технологии года-аналога.

Разработанная авторами информационно-прогностическая система, использующая уникальную базу данных, может стать основополагающим средством адресного снабжения специалистов и руководителей точной и своевременной информацией.

Как показал анализ и всестороннее изучение проблемы осуществления заблаговременного прогноза урожайности сельскохозяйственных культур, разработанная авторами технология поэтапной оценки урожайности зерновых культур на основе определения года-аналога и расчетами по прогностическим моделям является новой и оригинальной. Кроме того, прогностические модели, построенные с использованием МГК, можно использовать для оценки урожайности года-аналога при ее отсутствии с достаточной степенью точности.

Анализ результатов применения данной технологии показал хорошее совпадение фактической урожайности с расчетной по данным различных этапов прогноза. Это говорит о том, что разработанные авторами прогностические модели могут успешно применяться для текущего планирования и перспективного прогнозирования урожайности яровой пшеницы на территории Алтайского края.

Метод поэтапной оценки ожидаемой урожайности, заключающийся в последовательном уточнении оценок, полученных на предыдущих этапах, может быть эффективно использован совместно с технологией определения года-аналога. Такой метод исследования позволяет резко сократить число сложных и дорогостоящих опытов и время на проведение научно-исследовательских работ.

Дальнейшее развитие описанной методики предполагает осуществление аналогичных расчетов для других почвенно-климатических условий и сельскохозяйственных культур. Меняя объект моделирования и географический район, можно исследовать принципиальную применимость и устойчивость данной технологии прогноза для различных культур и сельскохозяйственных регионов.

Библиографический список

1. Хворова, Л.А. Применение информационных технологий, математических методов и моделей для обработки и анализа многомерных данных / Л.А. Хворова, Н.В. Гавриловская // Известия АлтГУ. – 2006. – №1 (49).
2. Хворова, Л.А. Прогнозирование урожайности зерновых культур: методы и расчеты / Л.А. Хворова, Н.В. Гавриловская // Известия АлтГУ. – 2008. – №1 (57).
3. Русакова, Т.И. Современная технология поэтапного прогнозирования урожайности и валового сбора зерновых культур / Т.И. Русакова, В.М. Лебедева, И.Г. Грингоф, Н.М. Шкляева // Метеорология и гидрология. – 2006. – №7.
4. Хворова, Л.А. Разработка алгоритма определения года-аналога для оценки урожайности зерновых культур в условиях Алтайского края / Л.А. Хворова, Н.В. Гавриловская // Известия АлтГУ. – 2007. – №1 (53).
5. Дронов, С.В. Многомерный статистический анализ: учеб. пособие. / С.В. Дронов. – Барнаул, 2006.
6. Сиротенко, О.Д. Математическое моделирование водно-теплового режима и продуктивности агроэкосистем / О.Д. Сиротенко. – Л., 1981.
7. Груза, Г.В. О принципах автоматической классификации метеорологических объектов / Г.В. Груза, Э.Я. Ранькова // Метеорология и гидрология. – 1970. – №32.