

МОДЕЛИРОВАНИЕ УРОЖАЙНОСТИ ЗЕРНОВЫХ С ИСПОЛЬЗОВАНИЕМ МЕТОДА СОВОКУПНОСТИ ДОКАЗАТЕЛЬСТВ В РАМКАХ КОНЦЕПЦИИ ТОЧНОГО ЗЕМЛЕДЕЛИЯ

Васильев Е.П.¹, Орешков В.И.²

¹ ФГБОУ ВПО «Рязанский государственный агротехнологический университет им. П.А. Костычева», Рязань, Россия (390044, г. Рязань, ул. Костычева, 1), e-mail: university@rgatu.ru.

² ФГБОУ ВПО «Рязанский государственный радиотехнический университет», Рязань, Россия (390005, г. Рязань, ул. Гагарина, 59/1), e-mail: rgrtu@rsreu.ru.

Рассмотрена возможность применения метода совокупности доказательств для построения бинарных классификационных моделей урожайности на основе данных агрохимического обследования почв. В качестве входных показателей модели использовались кислотность почв (рН), среднее содержание подвижного азота, калия и фосфора в мг/100 г, а также процент пашни с уклоном к югу и средний угол уклона. Исходный набор данных включал 56 земельных участков, для каждого из которых известна фактическая средняя урожайность (ц/га), на основе которой была сформирована бинарная переменная класса: высокая (15 участков) и низкая (41 участок). На основе ведомости агрохимического обследования для каждого измеренного показателя были вычислены коэффициенты WoE и информационные индексы, на основе которых определены характер зависимости бинарной переменной класса от агрохимических свойств участков, а также значимость (классифицирующая сила) каждого показателя в рамках классификационной модели. Предложен способ разбиения на интервалы для вычисления коэффициентов WoE как поиск компромисса между их равномерным заполнением, что обеспечивает наибольшую интерпретируемость модели, и максимизацией информационного индекса. Определено, что наиболее значимыми агрохимическими показателями в рамках построенной модели обладают кислотность почв, содержание в них азота и фосфора, а наименее – процент пашни с уклоном к югу и средний угол уклона.

Ключевые слова: точное земледелие, агрохимическое обследование, классификация, совокупность доказательств.

MODELING OF GRAIN YIELD USING WEIGHTS OF EVIDENCE METHOD IN THE CONCEPT OF PRECISION AGRICULTURE

Vasilyev E.P.¹, Oreshkov V.I.²

¹ Ryazan State Agrotechnological University n. a. P.A. Kostychev (390044, Ryazan, Russia, Kostychev Str., 1), e-mail: university@rgatu.ru.

² Ryazan State Radio Engineering University, Ryazan, Russia (390005 Gagarina Str. 59/1), e-mail: rgrtu@rsreu.ru.

The possibility of applying the method of weights of evidence to build a binary classification models of yield on the basis of agrochemical examination of soil are considered. As input parameters the model used by soil acidity (pH), the average content of available nitrogen, potassium and phosphorus in mg/100 g, and the percentage of arable land with a slope to the south and the average angle of slope. The original data set consisted of 56 fields, each of which is known for the actual average yield (cwt/ha), on which was formed by a binary class variable: high (15 fields) and low (41 fields). Based on the register of agrochemical survey for each of the measured parameters were calculated coefficients of WoE and information indices, which are defined on the basis of the dependence of the binary class variable from the agrochemical properties of fields, and significance (classifying the strength) of each indicator in the classification model. A method of partitioning intervals to calculate the coefficients WoE as finding a compromise between their uniform filling that provides the greatest interpretability of the model, and the maximization of the information index are proposed. It was determined that the most important indicators of agrochemical in the constructed models have soil acidity, content of nitrogen and phosphorus, and the least - the percentage of arable land with a slope to the south and the average angle of slope.

Key words: precision agriculture, agrochemical examination, classification, weights of evidence.

Одним из современных направлений повышения эффективности агробизнеса является точное земледелие (ТЗ). В основе концепции ТЗ лежит тот факт, что земельные площади, используемые под те или иные культуры, не являются однородными по своим

агрохимическим и агрофизическим свойствам, степени увлажненности и т.д. Следовательно, использовать одну и ту же норму внесения удобрений, средств защиты растений и орошения на всей площади неэффективно. Действительно, некоторые участки могут располагаться выше или ниже, что обуславливает разную близость грунтовых вод. Поэтому, применяя орошение для всех участков в одинаковом объеме, мы рискуем «залить» те из них, для которых степень естественной увлажненности выше, или «засушить» участки, где наблюдается недостаток естественного увлажнения. Очевидно, что дифференцированный подход в данном случае обеспечит более эффективное использование ресурсов и увеличение урожайности. Результатами внедрения ТЗ в практику агробизнеса являются оптимизация использования ресурсов и повышение урожайности за счет целевого применения удобрений, пестицидов, орошения и других мероприятий. Также является важным и экологический аспект: ограничение применения нитратов, ядохимикатов, забора воды из водоемов и т.д. [5].

Ключевым этапом ТЗ является аналитическая обработка информации о локализации участков, полученной с помощью навигационных систем (GPS, GLONASS), данных агрохимического и агрофизического обследования почв, истории урожайности. Анализ данных в ТЗ позволяет обнаруживать зависимости, которые «объясняют» характер влияния тех или иных агрохимических и агрофизических характеристик на урожайность. Знание таких зависимостей позволяет принимать решения о целесообразности проведения мероприятий, направленных на повышение урожайности, их объем и локализацию.

Важную роль в анализе сельскохозяйственных данных с целью реализации ТЗ является моделирование урожайности [2]. Построенные модели позволяют не только оценивать ее потенциальный уровень для площадей с определенными агрохимическими параметрами, но и исследовать степень и характер влияния того или иного параметра на урожайность. Например, если модель отражает зависимость урожайности от содержания в почве подвижного азота, то с ее помощью можно оценить количество удобрений, которое требуется внести для получения желаемой урожайности. Кроме этого, такая модель позволит оценить ожидаемый экономический эффект, т.е. соотнести затраты на внесение удобрений и дополнительную прибыль от увеличения урожайности, и сделать вывод о целесообразности проведения данного вида агрохимических мероприятий.

Построение моделей урожайности является достаточно сложной задачей, поскольку требует учета большого количества агрохимических и агрофизических факторов, которые характеризуются высокой степенью взаимных связей, обычно сложных и нелинейных. Например, известно, что высокая кислотность почв не только сама по себе негативно влияет на рост растений, но и ухудшает усвояемость ими содержащихся в почве макроэлементов

(азота, фосфора, калия). Поэтому внесение одного и того же количества удобрений на участки с различной кислотностью даст совершенно различный результат.

Сложность и нелинейность зависимостей между признаками часто делает применение классических методов прикладной статистики для построения моделей урожайности малоэффективным, а результаты моделирования трудно интерпретируемыми. Кроме этого, сельхозпредприятия не всегда содержат в своем штате аналитиков, имеющих соответствующую математическую подготовку. В этой связи практический интерес представляет разработка методик применения различных эвристических методов, которые, хотя и не являются полностью математически обоснованными, позволяют получить приемлемое решение в большинстве практически значимых случаев. В данной работе рассматривается метод моделирования урожайности на основе данных агрохимического обследования почв с помощью технологии, известной в зарубежной литературе как weight of evidence (WoE), что можно перевести как совокупность доказательств или вес доказательства [6]. Изначально метод использовался для поддержки принятия решений в кредитном скоринге, позже нашел применение в медицинской диагностике, поиске полезных ископаемых, выявлении мошенничеств и других областях. Рассмотрим ведомость агрохимического обследования почв (яровой ячмень [2]), фрагмент которой представлен на рис. 1.

№ поля	Площадь, га	Кислотность, рН	Азот, мг/100г	Калий, мг/100г	Фосфор, мг/100г	Юг, %	Ср. угол	Урожайность (ц/га)	Урожайность (класс)
1	183,01	4,8	3	19,3	18,52	91,36	1,28	2,2	Низкая
2	185,99	5,5	5	10,83	19,15	67,9	0,83	4,8	Низкая
3	47,52	5,6	6	12,76	16	44,24	0,92	6,3	Высокая
4	63,04	5,6	6	12,76	16	92,8	0,69	6,4	Высокая
5	141,08	5,2	4	24,42	19,55	26,26	1,05	3,5	Низкая
6	100,06	5,2	4	14,89	10,8	20,05	1,51	3,4	Низкая
7	18,15	5	3	7,4	6	1,33	1,78	2,2	Низкая
8	99,64	5,4	5	15,97	25,2	0,12	1,38	4,7	Низкая
9	91,36	5,1	4	9,31	6,94	12,98	1,36	3,5	Низкая
10	118,06	5,6	6	11,17	13,73	0,14	1,44	6,3	Высокая
11	73,24	4,9	3	13,15	12,32	8,81	1,14	2,2	Низкая
12	128,73	6,1	9	6,91	16,11	50,15	1,28	5,6	Высокая
13	40,87	6,1	9	6,43	12,93	99,97	1,8	5	Низкая
14	47,29	0	0	0	0	74,68	1,13	0,5	Низкая
15	191	6	9	17,78	19,74	14,34	1,21	11,5	Высокая
16	35,48	4,9	3	9,3	7,25	0,02	1,41	2,2	Низкая
17	17,7	4,4	1	6,8	10	46,39	2,32	0,4	Низкая

Рис. 1. Фрагмент ведомости агрохимического обследования почв.

Она содержит следующие показатели:

- № поля – уникальный идентификатор записи;
- площадь – площадь поля в гектарах;
- кислотность – средняя кислотность почв, рН;

- азот – среднее содержание подвижного азота в почве, мг/100 г;
- калий – среднее содержание подвижного калия в почве, мг/100 г;
- фосфор – среднее содержание фосфора в почве, (мг/100 г);
- юг – процент пашни с уклоном к югу;
- средний угол – средний угол уклона, градусов;
- урожайность – наблюдаемая средняя урожайность, ц/га;
- урожайность (класс) – бинарная переменная, которая принимает значение «Высокая» для урожайности больше 6 ц/га и «Низкая» – в противном случае.

На основе этих данных можно построить два типа моделей – численного предсказания и классификации. Для обеих моделей в качестве независимых переменных используются площадь, кислотность, калий, фосфор, юг и средний угол. Для модели численного предсказания выходной переменной будет урожайность числом, а для классификационной – класс урожайности. Во многих практических приложениях нет необходимости знания точных значений целевой переменной. Для принятия решения достаточно знать общий характер ее зависимости от входных переменных. Кроме этого, классификационные модели более удобны для интерпретации, чем численные, поскольку могут генерировать правила, формируемые на естественном языке [3].

Рассмотрим каждое наблюдение набора данных как эксперимент, исходом которого будет класс урожайности (высокая/низкая). Ведомость агрохимического обследования содержит 41 поле с низкой урожайностью и 15 – с высокой. Предположим, что целью классификации является обнаружение полей с высокой урожайностью для поддержки принятия решений о целесообразности их засева либо планирования соответствующих агротехнологических мероприятий. Тогда появление в наблюдении значения переменной класса «Высокая» будем называть событием, а значения «Низкая» – не-событием.

Произведем квантование диапазона значений каждой независимой переменной на заданное число интервалов. Рассмотрим переменную «Азот». Ее минимальное наблюдаемое значение составляет 2, а максимальное – 9. Разобьем данный диапазон на 4 интервала, каждый из которых соответствует изменению переменной на 2, т.е. 2–3, 4–5, 6–7 и 8–9. Для каждого интервала подсчитаем количество событий P_i и не-событий N_i для наблюдений, в которых значение переменной «Азот» попало в соответствующий интервал. Затем рассчитаем коэффициенты, называемые весами доказательств, по формуле [4]:

$$WoE_i = \ln \frac{(N_i/N)}{(P_i/P)}, \quad (1)$$

где i – номер интервала, $P = 15$ общее число положительных исходов в наборе данных, $N = 41$ – общее число отрицательных исходов. Результаты расчетов представлены в таблице 1.

Таблица 1 – Результаты формирования интервалов для вычисления коэффициентов WoE для содержания азота

№ интервала	Интервал	P_i	N_i	Вес интервала	WoE
1	2–3	0	18	0,32	3,89
2	4–5	0	16	0,29	3,78
3	6–7	6	3	0,16	-1,71
4	8–9	9	4	0,23	-2,19

Вес интервала – отношение числа наблюдений, попавших в интервал, к общему числу наблюдений. Чтобы избавиться от 0 в знаменателе выражения под логарифмом, определим $P_i/P = 1/(2n) = 0,089$, где $n = 56$ – общее число наблюдений. Такая подстройка удобна тем, что ее значение всегда больше нуля и гарантированно меньше любого возможного P_i/P (или N_i/N , если аналогичную подстройку проводить для не-событий).

Проинтерпретируем полученные результаты. Интервалы № 1 и 2, которые не содержат положительных исходов (т.е. наблюдений с высокой урожайностью), соответствуют низкому уровню содержания азота в почве. Интервалы № 3 и 4, соответствующие высокому содержанию азота, содержат большую часть наблюдений с высокой урожайностью. При этом коэффициенты WoE для интервалов № 1 и 2 являются положительными, а для интервалов № 3 и 4 – отрицательными. Из формулы (1) видно, что если относительная доля отрицательных исходов (не-событий) больше, чем положительных (событий), то значение под знаком логарифма больше 1 и коэффициент WoE оказывается положительным, а в противном случае – отрицательным. Следовательно, $WoE < 0$ указывает на большую вероятность появления положительного исхода, а $WoE > 0$ – отрицательного.

Таким образом, если новое наблюдение, т.е. земельный участок с известными агрохимическими параметрами, но с неизвестной урожайностью, имеет содержание азота в почве 2–5 мг/100 г, то модель классифицирует урожайность как низкую, а если 6–9, то как высокую. Модель показала, что с увеличением содержания азота урожайность растет, что согласуется с теорией растениеводства.

Следует, однако, учитывать, что урожайность зависит не только от содержания азота, но и от других агрохимических показателей, представленных в наборе данных на рис. 1. Для каждого из них можно провести аналогичные вычисления и определить характер его влияния

на урожайность. Открытым остается вопрос о силе этого влияния, т.е. значимости каждого признака с точки зрения классификации. Для оценки такой значимости вводится показатель, называемый информационным индексом [4]:

$$IV = \sum_{i=1}^k \left\{ \left(\frac{N_i}{N} - \frac{P_i}{P} \right) \cdot WoE_i \right\}. \quad (2)$$

Информационный индекс отражает степень связи признака и бинарной выходной переменной. Для данных из таблицы 1 его значение составит:

$$IV = \sum_{i=1}^4 \left\{ \left(\frac{N_i}{41} - \frac{P_i}{15} \right) \cdot WoE_i \right\} = 4,83.$$

С помощью значения IV можно оценить значимость (предсказательную силу) переменной для бинарной классификации: $IV < 0,02$ – отсутствует; $0,02 \leq IV < 0,1$ – низкая; $0,1 \leq IV < 0,3$ – средняя; $IV > 0,3$ – высокая, выше 3 – очень высокая. Таким образом, можно сделать вывод о высокой значимости влияния содержания азота в почве на урожайность.

Средством визуализации результатов WoE-анализа являются диаграммы, где для каждого интервала значение коэффициента WoE отражается в виде столбца соответствующей высоты. Например, для азота WoE-диаграмма будет иметь вид, представленный на рис. 2. Интерпретация WoE-диаграмм проста: для интервалов значений входного признака с коэффициентами $WoE < 0$ появление положительных исходов классификации (событий) более вероятно, а вероятность пропорциональна значению коэффициента.

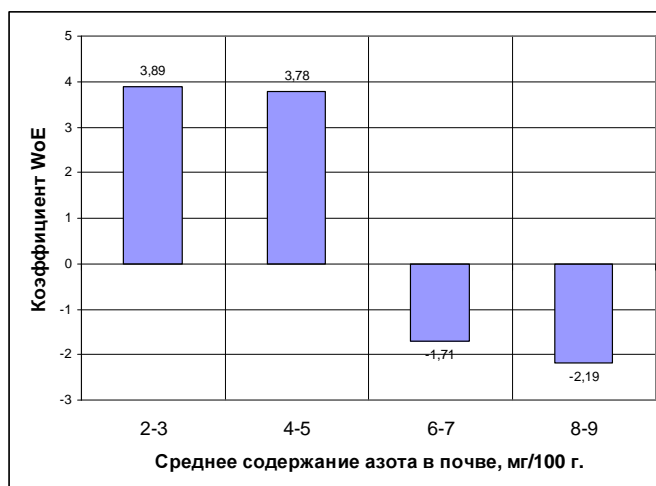


Рис. 2. Диаграмма WoE для содержания азота в почве.

В частности, диаграмма на рис. 2 однозначно указывает на то, что повышенное содержание азота в почве способствует увеличению урожайности.

Очевидной проблемой метода, обуславливающей его эвристическую сущность, является отсутствие формально обоснованных рекомендаций по выбору числа интервалов для вычисления коэффициентов WoE и информационного индекса. В рассмотренном выше примере разделение диапазона изменения содержания азота на 4 равных интервала, очевидно, не является оптимальным. Действительно, до значения 6 мг/100 г встречаются только не-события, поэтому интервалы № 1 и 2 можно объединить в один с $WoE = 2,26$. Тогда информационный индекс для признака составит $IV = 3,53$. Он несколько снизился, но по-прежнему указывает на высокую значимость признака, хотя модель существенно упростилась.

На практике можно использовать два способа разбиения признаков на WoE-интервалы: 1) так, чтобы в каждый интервал попадало примерно одинаковое число наблюдений (событий и не-событий); 2) так, чтобы максимизировать информационный индекс. Первый способ является логически более прозрачным для пользователя, интервалы проще интерпретируются. При выборе второго способа, поскольку значение информационного индекса возрастает с увеличением числа интервалов, алгоритм будет стремиться максимизировать его, и если не ввести ограничений, то каждый интервал будет сформирован для отдельного значения. Ограничением является вес интервала – процент попавших в него наблюдений относительно их общего числа. Тогда разбиение на интервалы закончится, как только будет получен интервал с весом ниже заданного (скажем, 10%). Таким образом, для второго способа, варьируя минимальный вес интервала, аналитик может добиваться построения наиболее «адекватной», с его точки зрения, WoE-модели.

Еще одним важным вопросом является выбор числа интервалов. Можно использовать два основных подхода.

1. Максимизация информационного индекса. Как правило, в результате получается очень большое число интервалов (несколько десятков), что влечет три недостатка:

а) ухудшается интерпретируемость модели: WoE-диаграмма будет содержать большое количество столбцов для коэффициентов с различными знаками;

б) модель окажется «переобученной», т.е. слишком точно «подогнанной» под исходные данные. Это приведет к ухудшению ее обобщающей способности при работе с новыми наблюдениями: даже небольшие вариации исходных данных могут привести к изменению класса наблюдения и ошибочным выводам;

в) большое число интервалов в соответствии с формулой (2) приведет к росту информационного индекса и завышенной оценке значимости признака. Данное явление в зарубежной литературе известно под синтетическим термином *overprediction*, который пока не имеет устоявшегося перевода на русский язык.

2. Субъективные представления аналитика. Аналитик может выбрать число интервалов, исходя из представлений о том, насколько значительным должно быть изменение признака, чтобы повлиять на результаты классификации объекта. Например, в кредитном скоринге одним из признаков для оценки кредитоспособности клиента является возраст. Очевидно, что за 1–2 года социальный статус и доход клиента в большинстве случаев значительно не меняется, а вот за 10–15 лет это вполне вероятно. Следовательно, весь наблюдаемый диапазон возраста, скажем, с 18 до 75, следует разбить на интервалы по 10 лет, что позволит получить 5–6 интервалов. Аналогичным образом содержание макроэлементов в почве также классифицируется на очень низкую, низкую, среднюю, повышенную, высокую и очень высокую, а кислотность почв на очень сильнокислую, среднекислую, слабокислую, нейтральную и щелочную. Поэтому для соответствующих признаков агрохимической ведомости на рис. 1 при проведении WoE-анализа можно использовать 6 интервалов и ассоциировать с перечисленными классами. Это сделает модель более интерпретируемой, а информационный индекс реально отражающим вклад признака в классифицирующую способность модели.

Таким образом, построение WoE-модели может потребовать многочисленных экспериментов с целью подбора числа интервалов и выбора способа их формирования, минимального веса интервалов. Поэтому «ручная» реализация WoE-анализа, например в электронных таблицах, может оказаться весьма трудоемкой, особенно на больших массивах данных. В настоящее время WoE-анализ включается как в предметно-ориентированные программные средства анализа данных, так и в универсальные аналитические платформы (например, SAS Enterprise Miner). Единственной отечественной разработкой, в которой доступен алгоритм реализации WoE-анализа и эффективная визуализация его результатов, является аналитическая платформа Deductor Enterprise компании Basegroup Labs (www.basegroup.ru) [4].

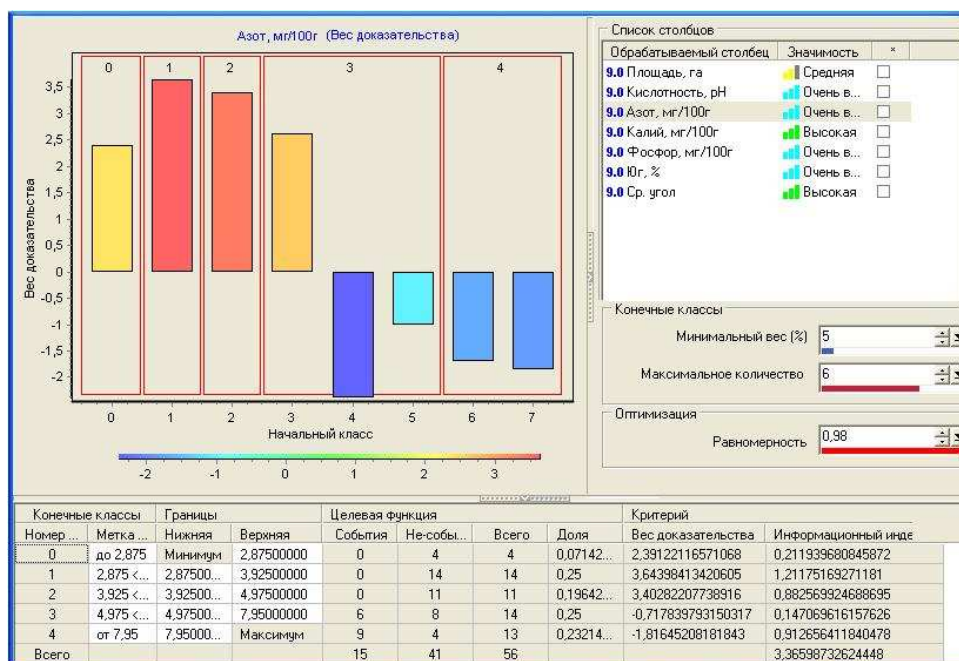


Рис. 2. Окно реализации WoE-анализа аналитической платформы Deductor.

Общий вид окна обработчика, реализующего WoE-анализ в аналитической платформе Deductor, представлен на рис. 2. В нем содержится список переменных, на которых строилась модель; диаграмма, где отображаются коэффициенты WoE для выделенной переменной; таблица, содержащая границы интервалов, количество наблюдений (событий и не-событий), попавших в каждый интервал, вес и WoE-коэффициент интервала, а также значение информационного индекса. Обработчик является интерактивным: изменяя минимальный вес и выбирая количество интервалов, аналитик подбирает наиболее подходящий вид модели. Критерием при этом может служить субъективное представление исследователя, максимальная значимость всех признаков и интерпретируемость WoE-диаграммы в смысле отражения зависимости класса урожайности от агрохимических показателей.

Приведем обобщенные результаты анализа ведомости агрохимического обследования почв в таблице 2 для следующих параметров: минимальный вес класса – 5%, число интервалов – 6 (для всех переменных), метод разбиения на интервалы – равномерный.

Таблица 2 – Результаты WoE-анализа для ведомости агрохимического обследования

Показатель	WoE средний	Коэф. корреляции	Информационный индекс	Значимость
Площадь участка, га	-0,082	0,009	0,297	Средняя
Кислотность, pH	3,000	0,638	4,637	Оч. высокая

Азот, мг/100 г	5,700	0,671	4,963	Оч. высокая
Калий, мг/100 г	0,152	0,207	0,664	Высокая
Фосфор, мг/100 г	3,800	0,481	3,928	Оч. высокая
Уклон к югу, %	0,132	0,097	0,066	Низкая
Средний угол, °	0,092	0,188	0,293	Средняя

По таблице можно сделать следующие выводы. Высокой и очень высокой значимостью для урожайности обладают кислотность, содержание азота, калия и фосфора, т.е. агрохимические параметры почвы, непосредственно влияющие на рост растений. Агрофизические факторы, такие как доля пашни с уклоном к югу и средний угол уклона, показывают среднюю и низкую значимость, что позволяет исключить их из модели.

Для сравнения в таблице 2 представлен коэффициент корреляции Пирсона для каждой входной переменной и урожайности. Он отражает линейные зависимости между величинами, поэтому при наличии нелинейных зависимостей не позволяет делать корректные выводы о значимости связи между ними. Кроме этого, коэффициент корреляции ограничен диапазоном от -1 до 1, что не всегда позволяет четко разделить значимые и не значимые факторы. Например, коэффициент корреляции с урожайностью для переменных калий (0,207) и средний угол (0,188) очень близки. В то же время информационные индексы, равные соответственно 0.664 и 0.293, различаются более чем в два раза. Благодаря нелинейной обработке при вычислении коэффициентов WoE с помощью логарифмической функции различия в значимости переменных являются более выраженными и легко обнаруживаются.

И, наконец, средние значения коэффициента WoE, вычисленные по всем интервалам признака, позволяют увидеть, на изменение каких агрохимических параметров урожайность будет реагировать наиболее сильно. В нашем случае это содержание азота, кислотность и содержание фосфора. Поэтому агрохимические мероприятия, направленные на увеличение содержания данных элементов в почве, являются наиболее перспективными с точки зрения повышения урожайности.

Заключение. Таким образом, в статье рассмотрен метод моделирования урожайности зерновых (на примере ярового ячменя) по данным агрохимического обследования почв на множестве земельных участков с различной площадью на основе WoE-анализа. WoE-анализ представляет собой простой в использовании, понимании и интерпретации метод моделирования многомерных данных с целью обнаружения связей между бинарной выходной переменной класса (урожайность) и набором входных переменных (агрохимических показателей), на основе распределения относительных частот появления

положительных и отрицательных исходов по диапазону изменения входной величины. Он позволяет обнаруживать значимость влияния признака на переменную класса и исключать наименее значимые признаки, сокращая размерность исходных данных, упрощая их интерпретируемость и дальнейшую обработку.

В процессе моделирования урожайности на основе ведомости агрохимического обследования почв было обнаружено, что наибольшее влияние на урожайность оказывают содержание азота, фосфора и кислотность почв, и в несколько меньшей степени – содержание калия. Именно регулирование содержания данных веществ обеспечивает основной потенциал повышения эффективности агробизнеса в рамках концепции точного земледелия.

Список литературы

1. Васильев Е.П., Орешков В.И. Современные аналитические платформы для задач АПК // Вестник Рязанского гос. агротехнологического университета имени П.А. Костычева. – Рязань : РГАТУ. – 2011. – № 1. – С. 68-75.
2. Евстропов А.С. Системы управления производством сельскохозяйственной продукции на основе информационно-инновационных технологий / А.С. Евстропов, В.А. Артамонов. – Рязань : ГНУ ВНИИМС, 2009. – 196 с.
3. Паклин Н.Б. Бизнес-аналитика: от данных к знаниям (+CD) / Паклин Н.Б., Орешков В.И. – Изд. 2-е, перераб. и доп. – СПб. : Питер, 2010. – 700 с.
4. Anderson R. The Credit Scoring Toolkit. – Oxford University Press Inc., New York, 2007. – P. 731.
5. Rub G. Data Mining in Precision Agriculture: Management of Spatial Information / Rub G., Brenning A. // IPMU 2010, LNAI 6178. – Springer-Verlag Berlin Heidelberg 2010. – P. 350–359.
6. Weed D.L. Weight of Evidence: A Review of Concept and Methods // Risk Analysis. – 2005. – Vol. 25. – №. 6. – P. 1545-1557.

Рецензенты

Солдак Юрий Максимович, доктор экономических наук, профессор кафедры экономики, менеджмента и организации производства ФГБОУ «Рязанский государственный радиотехнический университет», г. Рязань.

Текучев Владимир Васильевич, доктор экономических наук, профессор, заведующий кафедрой экономической кибернетики ФГБОУ «Рязанский государственный агротехнологический университет им. П.А. Костычева», г. Рязань.