**Прогнозирование запасов продуктивной почвенной влаги на основе мультиномиальной логистической регрессии**

**Кизимова Т.А.**

Младший научный сотрудник

Сибирский Федеральный Научный Центр Агробиотехнологий РАН,

Новосибирск

E-mail: [tanya.luzhnykh@mail.ru](mailto:tanya.luzhnykh@mail.ru)

При возделывании сельскохозяйственных культур, влагообеспеченность является одним из основных факторов лимитирующим продуктивность растений. Получение высоких урожаев зависит от влагообеспеченности культуры во время вегетации. В данный период осадки не всегда обеспечивают оптимальные условия увлажнения для активного развития растений, вследствие они вегетируют за счет запасов продуктивной влаги (ЗПВ) накопленной в почве. Своевременная оценка ЗПВ в метровом слое почвы, даёт возможность планировать урожайность сельскохозяйственных культур на полевой сезон, принимать эффективные управленческие решения в выборе агротехнических мероприятий (обработка почвы, применение и корректировка доз удобрений, средства защиты растений и др.), тем самым рационально использовать ресурсы хозяйства. Поэтому становится актуальным заблаговременно определять ЗПВ в почве перед посевом культур [1-5]. В нашей работе для такой цели был применён метод прогноза на основе модели мультиномиальной логистической регрессии.

**Цель данной работы** – построить модель мультиномиальной логистической регрессии (МЛР), способную прогнозировать запас продуктивной влаги в метровом слое почвы перед посевом.

Моделирование проводили на данных многофакторного полевого опыта СиНИ-ИЗиХ СФНЦА РАН расположенном на территории ОС «Элитная» - филиала СФНЦА РАН Новосибирской области (центрально-лесостепная подзона). Данные включают результаты исследований четырёхпольного зернопарового севооборота (пар-пшеница-пшеница-пшеница) с разными вариантами основной обработки почвы с 2009 по 2018 год. Почва стационара – чернозём выщелоченный среднесуглинистого гранулометрического состава. Для моделирования использовали данные по запасам продуктивной влаги в метровом слое почвы.

Построение мультиномиальной логистической регрессии осуществляли с помощью программного пакета модулей SPSS версия «26». При моделировании 80% данных исходной выборки использовались для обучения моделей, 20% в качестве тестирования.

**Результаты исследований**

Для прогноза целевого показателя – запас продуктивной влаги в метровом слое почвы (мм), исследовали связь этого показателя (зависимая переменная) от набора переменных (факторов). При этом учитывались как качественные факторы: предшественник, способ обработки почвы, так и количественные, которые характеризуют погодные условия.

Качественные (категориальные) факторы имеют следующие градации:

- Способ обработки почвы с 3 градациями: 0-нулевая обработка, 1- безотвальная обработка,2-вспашка;

- Предшественник имеет 4 градации: 0-пар, 1- первая пшеница по пару, 2- вторая пшеница по пару, 3- третья пшеница по пару.

К количественным факторам относятся погодные условия, а именно сумма активных температур воздуха >0° (больше нуля градусов) и осадков за периоды: сентябрь-ноябрь, декабрь-февраль, март-апрель. Зависимую переменную можно классифицировать в качестве категориальной, при этом значения запаса продуктивной влаги в слое 0-100 см описывается тремя градациями: 1 – хорошие запасы от 130 мм и выше, данная градация объединена (включает в себя хорошие от 130-160 мм и очень хорошие> 160 мм ЗПВ, так как при построении модели был недостаток исходных данных этих градаций); 2 –удовлетворительные запасы от 90 до 130 мм; 3 – плохие запасы от 60 до 90 мм; 4 – очень плохие запасы <60 мм.

В модели МЛР одна из категорий зависимой переменной объявляется опорной (эталонной), а все другие категории сравниваются с ней. Независимые переменные могут быть категориальными или количественными. Уравнение МЛР прогнозирует вероятность принадлежности зависимой переменной к категориям по значениям независимых переменных. Окончательный выбор предсказываемой категории для зависимой переменной производится по правилу наибольшей вероятности принадлежности. Результаты прогноза модели представлены в таблице 2. Подробное описание структуры модели МЛР представлено в материалах [5].

Для получения коэффициентов МЛР (использующихся для прогнозирования), была использована выборка данных, сформированных в виде таблицы из 96 наблюдений (строк) и 8 факторов (столбцов), включая зависимую переменную. Число наблюдений распределено по показателю ЗПВ в 0-100 см слое почвы следующим образом: хорошие запасы от 130 мм и выше -12, удовлетворительные запасы от 90-130 мм - 36, плохие запасы от 60 до 90 мм – 35, очень плохие запасы менее 60 мм – 13.

Для оценки качества модели МЛР, использовали показатель показатель Псевдо R-квадрат, который описывает ту часть дисперсии, которую можно объяснить с помощью регрессии. Наиболее распространенными являются меры, предложенные Нэйджелкерком, Коксом и Снеллом. В нашем случае объясненная часть дисперсии составляет по мере Нэйджелкерка (как правило, наиболее применяемой) 88 %, а по мере Коксаи Снелла 81%, что говорит о высоких прогностических возможностях метода.

В качестве критерия оценки прогнозной модели определена величина отклонения (погрешность) фактического ЗПВ от прогнозируемого. В таблице 1 представлены сравнительные прогностические способности данного метода, проверенные на исходной выборке.

Таблица 1. –Классификационная таблица МЛР.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Количество наблюдений | Предсказанные | | | | % правильных прогнозов |
| Хорошие запасы от 130 мм и выше | Удовлетворите-  льные запасы от 90 до 130 мм | Плохие запасы от 60 до 90 мм | Очень плохие запасы менее 60 мм |
| Хорошие запасы от 130 мм и выше (12) | 9 | 3 | 0 | 0 | 75,0 |
| Удовлетворите-  льные запасы от 90 до 130 мм (36) | 1 | 29 | 6 | 0 | 80,6 |
| Плохие запасы от 60 до 90 мм (35) | 0 | 6 | 29 | 0 | 82,9 |
| Очень плохие запасы менее 60 мм (13) | 0 | 0 | 0 | 13 | 100 |
| Общая доля |  |  |  |  | 83,3 |

Для выборки наблюдений с хорошим ЗПВ (130 мм и выше) доля правильных прогнозов составила 75%, из них верно спрогнозировано 9 наблюдение, неправильными были 3 прогноза, которые попали в категорию с удовлетворительными ЗПВ (от 90 до 130 мм). Для категории удовлетворительные ЗПВ правильно было спрогнозировано 29 наблюдений из 36, при этом процент правильных прогнозов составил 80,6%. Самый большой процент (соответствует 100%) правильных прогнозов у градации с ЗПВ менее 60 мм. Общая доля правильных прогнозов по всем категориям составляет 83,6%. Верификацию построенной модели провели на данных 2018 года (табл. 2).

Таблица 2. Прогноз ЗПВ в слое почвы 0-100 см в 2018 г

|  |  |  |  |
| --- | --- | --- | --- |
| Предшественник | Обработка почвы | Фактический  ЗПВ, мм | ЗПВ (прогноз МЛР), мм |
| Пар | Вспашка | 133,7 | Хорошие запасы от 130 мм и выше |
| Пар | Безотвальная | 166,9 | Хорошие запасы от 130 мм и выше |
| Пар | Безотвальная | 166,9 | Хорошие запасы от 130 мм и выше |
| Пар | Нулевая | 135,0 | Хорошие запасы от 130 мм и выше |
| 1-ая пшеница по пару | Вспашка | 133,0 | Хорошие запасы от 130 мм и выше |
| 1-ая пшеница по пару | Безотвальная | 138,8 | Хорошие запасы от 130 мм и выше |
| 1-ая пшеница по пару | Нулевая | 106,9 | Плохие запасы от 60 до 90 мм**\*** |
| 2-ая пшеница по пару | Вспашка | 125,0 | Хорошие запасы от 130 мм и выше**\*** |
| 2-ая пшеница по пару | Безотвальная | 147,7 | Хорошие запасы от 130 мм и выше |
| 2-ая пшеница по пару | Нулевая | 138,1 | Хорошие запасы от 130 мм и выше |
| 3-я пшеница по пару | Вспашка | 115 | Хорошие запасы\* от 130 мм и выше |
| 3-я пшеница по пару | Безотвальная | 132,4 | Хорошие запасы от 130 мм и выше |
| 3-я пшеница по пару | Нулевая | 124,5 | Хорошие запасы \*от 130 мм и выше |

Примечание: \*- неверно спрогнозированные градации

Учитывая небольшой объем статистической выборки и малое количество предикторов, прогностические свойства модели можно признать удовлетворительными.

**Заключение**

В ходе исследования с помощью МЛР построена и обучена модель, позволяющая с допустимой достоверностью прогнозировать запас продуктивной влаги перед посевом в метровом слое почвы, используя данные о погодных условиях, обработке почвы, предшественнике. На основе прогноза, можно разрабатывать рекомендации по ориентировочным дозам внесения удобрений в зависимости от спрогнозированной градации, планировать урожайность и возможную прибавку урожайности от удобрений.

1. Абрамов Н. В. Формирование водного режима в севооборотах интенсивного типа // Агропродовольственная политика России. – 2020. – № 1-2. – С. 2-8.
2. Rodrigo T, Diego F. S., Glauber C, Hitoshi N., Gilliard C. Soil moisture forecast for smart irrigation: The primetime for machine learning, Expert Systems with Applications. 2022. Vol 207. Р. – 117653. https://doi.org/10.1016/j.eswa.2022.117653.
3. Тарасова Л. Л. Новый метод прогноза запасов влаги // Материалы научной конференции «Вклад Агрофизики в решение фундоментальных задач сельскохозяйственной науки, 01-02 октября. Санкт-Петербург. – 2020 г. – С. 262.
4. Неверов А. А. (2020). Прогнозирование почвенных влагозапасов на основе статистического моделирования природных процессов // Известия Оренбургского государственного аграрного университета. –2020. – №2 (82). – С.14-18.
5. Long J. S. Regression models for categorical and limited dependent variables, Thousand Oaks: Sage Publ, 1997. – 328p.