

Сравнительный анализ результатов оценки причинного эффекта использования интенсивной технологии в растениеводстве

Тарасов К.О., Понькина Е.В.

Алтайский государственный университет, г. Барнаул
kirilltarasov99@gmail.com, ponkinaelena77@mail.ru

Аннотация

Задача оценки причинных эффектов представляет собой сравнение состояния объекта с учетом и без учета вмешательства и оценки ожидаемой величины полученных различий целевого признака. В работе рассматривается сравнительный анализ методов оценки причинных эффектов на примере производства продукции растениеводства в Алтайском крае. Оценке подлежит причинный эффект от применения интенсивной технологии, предполагающей внесение удобрений и применение средств защиты растений. Тестировались следующие методы: парное сравнение средних, линейная регрессия и матчинг методы (Propensity Score Matching), позволяющие сбалансировать выборку по основным индикативным признакам. Результаты показали, что согласно всем рассматриваемым методам интенсификация производства, даже в засушливых условиях 2012 года привела в ожидаемому росту продуктивности пшеницы, средний ожидаемый эффект от применения интенсивной технологии варьируется от 3,12 ц/га до 3,71 ц/га. Полученные результаты демонстрируют возможность получения более корректных оценок причинных эффектов на основе сбалансированных выборок, использование простого сравнения средних приводит к недооценке или переоценке получаемого эффекта. Также в статье проанализированы ограничения и особенности применения метода Propensity Score Matching в социально-экономических исследованиях.

Ключевые слова: причинный эффект, парное сравнение средних, Propensity Score Matching, урожайность пшеницы, интенсивная технология, Алтайский край, Россия.

1. Введение

Одним из разновидностей анализа эффектов вмешательства (целенаправленного воздействия) в производственно-хозяйственную систему является причинный вывод (*Causal Inference*). Анализ причинности помогает получить заключение о каких-либо гипотезах путем нахождения различий между фактами и гипотетическими ситуациями (контрфактуалами). Фундаментальной основой исследования причинных эффектов в прикладных научных исследованиях является теория о потенциальных исходах (Potential Outcome Theory), основанная на исследованиях Дж. Неймана и Д. Рубина [1]. Центральная проблема теории потенциальных исходов – проблема количественной оценки причинного эффекта воздействия на состояние объекта.

Причинный эффект (causal effect) – это разность между двумя потенциальными исходами объекта $Y(1)$ и $Y(0)$, где при прочих равных условиях $Y(1)$ – оценка результирующего признака объекта в результате вмешательства, $Y(0)$ – оценка результирующего признака без учета вмешательства [2].

Причинный эффект обозначает эффект, изменение значения результирующего признака в результате какого-либо воздействия на объект выборки. Она заключается в том, что обозреваемый причинный эффект не обязательно равен истинному причинному эффекту. Оценка причинного эффекта, осуществляемая по набору эмпирических наблюдений, несбалансированных по значению контрольных признаков (например таких как возраст, образование респондента) как сравнение средних значений или применении регрессионных зависимостей на основе двух группе объектов (с вмешательством и без – *Treated Group* и *Control Group*) приводят к получению смещенных оценок причинных эффектов (Bias).

Для решения этой проблемы Дж. Нейман и Д. Рубин [3] предложили модель Рубина-Неймана, которая позволяет осуществлять псевдорандомизацию данных, с целью приблизить условия исследования на базе наблюдационных данных к условиям экспериментального исследования (т.е. сбалансировать контрольную и тестовую выборки по значениям контрольных переменных), тем самым снизить долю bias в оценке причинного эффекта и приблизить полученную оценку к истинному причинному эффекту.

Целью исследования является сравнительный анализ методов оценки причинных эффектов на примере оценки вклада элементов интенсификации технологии возделывания яровой пшеницы на территории Кулундинской степи в границах Алтайского края в засушливых условиях 2012 года.

Территория Кулундинской степи характеризуется засушливым климатом, при чем сумма осадков за вегетационный период является главным лимитирующим фактором урожайности яровой пшеницы. 2012 год характеризовался низкой величиной осадков и рекордно низкой продуктивностью яровой пшеницы. Оценка эффекта ожидаемого прироста урожайности пшеницы в случае применения средств защиты растений и/или внесения удобрений под посевы пшеницы основанной на данных реальной практики работы предприятия является важной прикладной задачей для развития растениеводства в Алтайском крае. Алтайский край, несмотря на крупнейшую площадь посева в Российской Федерации, находится на четвертом месте по общему валовому сбору зерновых и зернобобовых культур. По урожайности Алтайский край стабильно находится в конце рейтинга по регионам РФ, как и по объему внесения удобрений.

В связи с этим ключевые прикладные вопросы исследования заключаются в следующем:

- Каков ожидаемый прирост урожайности пшеницы в Кулундинской степи Алтайского края в случае внесения удобрений под посевы пшеницы в условиях засухи 2012 года?
- Каков ожидаемый прирост урожайности яровой пшеницы в случае применения средств защиты растений в условиях Кулундинской степи Алтайского края в условиях засухи 2012 года?

2. Методы анализа причинности

Помимо метода сопоставления оценок склонностей (*Propensity Score Matching - PSM*) в работе используются методы попарного сравнения средних – t-тест Стьюдента сравнение средних [4] и U-тест Манна-Уитни-Вилкоксона [5]. Эти методы предоставляют возможность быстро оценить причинный эффект, однако не позволяют утверждать приближенность оценки причинного эффекта к его истинному значению. Также эти методы позволяют оценить эффект воздействия только посредством сравнения пары выборок, следовательно, невозможно оценить эффект других факторов как на переменную результата, так и на фактор воздействия. Те же ограничения относятся к методу однофакторной регрессии. Метод многофакторной регрессии, однако, позволяет оценить причинный эффект других факторов, но все еще не позволяет утверждать приближенность оценки к истинному значению эффекта. Для всех этих методов используется предположение о стабильности эффектов воздействия на объект (*Stable Unit Treatment Value Assumption - SUTVA*) [6].

Это предположение подразумевает отсутствие интерференции между объектами, а также существование только одного варианта воздействия. При соблюдении этих условий, можно утверждать, что оценка причинного эффекта приближена к его истинному значению.

Метод сопоставления оценок склонностей (PSM) был разработан Дональдом Рубином совместно с Полом Розенбаумом для того, чтобы уменьшить предвзятость, связанную с переменными-конфаундерами, которая может быть обнаружена при расчете эффекта воздействия. Конфаундеры (Confounder) – это переменные, которые могут быть связаны как с изучаемым фактором воздействия, так и с выводом [7]. Их действие создает различия между рассчитанным и фактическим эффектом воздействия. Так как распределение воздействий на объекты выборки в обсервационных исследованиях не удовлетворяет требованиям рандомизированности, проблема появления предвзятости требует нахождения специального решения для ее уменьшения или устранения. Метод сопоставления оценок склонностей позволяет имитировать рандомизированную выборку с помощью оценок склонности (*Propensity Score*), а после этого использовать стандартные матчинг-техники для получения лучших результатов в сравнении с применением матчинг-техник на необработанную контролируруемую выборку. Метод сопоставления оценок склонностей также позволяет провести исследования, связанные с потенциально вредным воздействием на объекты, такие как исследования воздействия курения на здоровье человека, с более точными результатами и при этом не нарушая этичности, так как вредное воздействие, например курение, не назначается человеку в рамках эксперимента, а было выбрано самим объектом по своей воле.

Основные этапы реализации данного метода включают пять базовых шагов.

Пусть результат вмешательства/невмешательство будет обозначен как Y , а воздействие – X .

Первым шагом является сбор и анализ данных. На данном этапе, после сбора данных, необходимо понять, нужно ли применять данный метод. Для этого необходимо проанализировать степень сбалансированности исходных контрольной и тестовой групп данных, наличие выбросов в данных. Также анализируются пропуски в данных.

Второй шаг – это отбор ковариатов (вектор X) для использования в модели соответствия. Ковариаты – это факторы, связанные с изучаемой переменной и переменной воздействия. Необходимо обязательно включить все переменные, связанные с переменной воздействия, так как они могут оказывать какое-либо влияние на нее. После балансировки выборки по набору отобранных ковариатов должна улучшиться согласованность данных, если это произошло, то ковариаты были отобраны верно.

Третий шаг – это выбор модели для расчета оценок склонности объектов (*Propensity Score*), и расчет этих оценок. Существует несколько моделей для их расчета, однако самой популярной является логистическая регрессия. Эта модель описывает вероятность дискретного события $Y = 1$ при условии X :

$$\Pr(Y = 1|X) = \frac{1}{1 + e^{-aX}},$$

где $aX = a_1x_1 + a_2x_2 + \dots + a_nx_n$.

Оценка склонности объекта i осуществляется по формуле:

$$Score_i = \Pr(Y_i = 1|X_i).$$

Четвертый шаг – выбор метода анализа соответствий (*Matching*) и поиск соответствий. Выбор метода оценки соответствия во многом определяет состав рандомизированных тестовой и контрольной групп, поэтому на практике часто используют различные

методы. Выбор метода оценки близости основывается на сравнении результатов балансировки данных согласно разным методам. Существует несколько методов, часто используемые – точное соответствие (*Exact Matching*), субклассификация (*Subclassification*), метод поиска ближайшего соседа (*Nearest Neighbor*), оптимальное соответствие (*Optimal Matching*), полное соответствие (*Full Matching*), и генетическое соответствие (*Genetic Matching*) [8–10]. Все эти методы доступны в различных программных пакетах для работы со статистическими данными. Каждый из этих методов при псевдорандомизации ищет пару объектов, по одному из экспериментальной и контрольной группы, с помощью метрики дистанции. Дистанция измеряется как модуль разности оценок склонностей двух объектов:

$$Distance(X_i, X_j) = \left| Score_i - Score_j \right|, \quad i, j = 1, \dots, N.$$

Рассмотрим подробнее использованные методы поиска соответствий.

Метод поиска ближайшего соседа – самый популярный метод анализа соответствий. При его использовании рассчитывается приближенное значение среднего эффекта воздействия на объекты экспериментальной группы (АТТ) [11]:

$$E(Y_i(1) - Y_i(0) | T_i = 1).$$

Данный метод находит соответствия каждому объекту в экспериментальной группе, после чего отбрасывает объекты контрольной группы без пары. Этот метод называют жадным, так как каждый объект контрольной группы используется лишь раз.

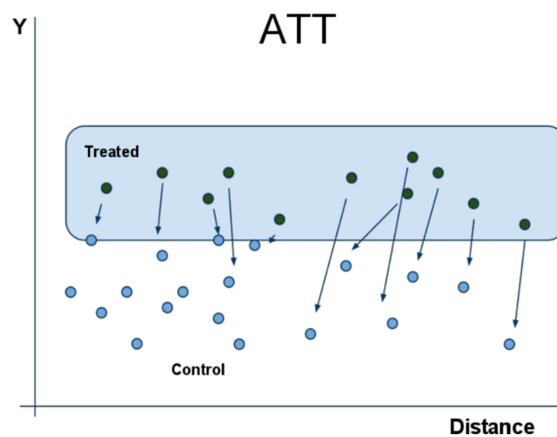


Рисунок 1. Визуализация метода поиска ближайшего соседа. *Источник:* [9]

Метод полного соответствия также находит пару каждому объекту, основываясь на метрике дистанции, но при его использовании рассчитывается АТТ или истинный эффект воздействия в каждой группе отдельно, после чего группы взвешиваются по числу объектов в них [9].

Метод генетического соответствия находит соответствия с помощью обобщенной версии дистанции Махаланобиса (GMD), у которой присутствует дополнительный весовой параметр W [8]:

$$GMD(X_i, X_j, W) = \sqrt{(X_i - X_j)^T (S^{-\frac{1}{2}})^T W S^{-\frac{1}{2}} (X_i - X_j)},$$

где W – матрица положительного определенного веса (баланса) размера $k \times k$ и $S^{-\frac{1}{2}}$ – разложение Холецкого параметра S . Все элементы W ограничены нулем, за исключением тех, что расположены на главной диагонали, которая состоит из k параметров, которые

должны быть выбраны. При применении формулы для оценок склонностей, X в формуле заменяется на Z , где Z – матрица, состоящая из оценок склонностей и основных ковариатов X [8].

Пятый шаг – анализ балансировки данных. Для этого используются разные способы, такие как построение QQ графиков распределения объектов выборки, гистограмм и точечных графиков распределения объектов по оценкам склонности в контрольной и экспериментальной группе до и после балансировки данных. QQ график – это график рассеяния, создаваемый размещением двух наборов квантилей напротив друг друга [12]. Программные пакеты зачастую также предоставляют сравнение с помощью таблиц параметров со средними значениями по группам с вычисленным улучшением значений в процентах.

Важно отметить, что после реализации всех пяти шагов данного метода, необходимо осуществить проверку остаточного дисбаланса. Главной проблемой метода сопоставления оценок склонностей является возможность при неправильной реализации ухудшить предвзятость данных после балансировки. Это было названо парадоксом сопоставления оценок склонностей (*Propensity Score Matching Paradox*) [13]. В связи с этим данный метод надо использовать с максимальной аккуратностью и обязательно сравнивать остаточную предвзятость с исходной, что позволит показать, что применение псевдорандомизации не ухудшает предвзятость.

После успешного выполнения этих пяти шагов можно переходить к расчету эффекта воздействия, что можно осуществлять с помощью различных статистических методов, например, с помощью логистической регрессии, теста t -критерия Стьюдента или U -критерия Манна-Уитни.

3. Данные и территория исследования

Кулундинская степная зона (далее – Кулунда) расположена в юго-восточной части Западной Сибири и простирается от центра до юга Алтайского края. Климат в Кулунде континентальный, многолетняя средняя температура – минус 18°C в самый холодный месяц (январь) и $+19^{\circ}\text{C}$ в самый теплый месяц (июль) [14].

Частые засухи в Алтайском крае значительно снижают урожайность сельскохозяйственных культур и существенно повышают климатические риски для сельскохозяйственных производителей [15, 16]. В целом в Кулундинской степной зоне можно выделить три агроэкологические подзоны: 1) лесостепь, в которой средняя урожайность пшеницы с 1996 по 2011 г. составила 15 ц/га, 2) типичная степь – с урожайностью 11 ц/га и 3) сухая степь с самой низкой средней урожайностью пшеницы – 8 ц/га [17].

В 2011 г. в Кулундинской степной зоне функционировало 576 сельскохозяйственных организаций, 1325 крестьянско-фермерских хозяйств (КФХ) и индивидуальных предпринимателей (ИП), основным товарным производством которых являлось производство пшеницы. Примерно 34% сельскохозяйственных организаций и крестьянско-фермерских хозяйств выращивали пшеницу на менее чем 2000 га, 28% – на 2000–6000 га площади пашни и 38% – на более чем 6000 га, в среднем на одно предприятие [18].

Данные используемые в работе представляют собой результаты опроса фермеров в комбинации со статистическими данными официальной статистики [18]. Общее количество наблюдений – 196, из них 111 предприятий относятся к категории обществ с ограниченной ответственностью. Средняя площадь земель сельскохозяйственного назначения составила 9333 га, а посевная площадь 7150 га. В дескриптивной статистике (таблица 1), помимо общей статистики по всем данным, был проведен анализ по двум группам объектов. Первая группа представляет предприятия, которые применяли в 2012 г. интенсивные технологии при возделывании пшеницы, в частности, средства защиты растений (СЗР) и удобрения, а вторая группа – предприятия, не применявшие интенсивных технологий.

Дескриптивная статистика показывает, что только 24% фермеров вносило удобрения в технологии возделывания пшеницы, а 69% применяли средства защиты растений. Результаты работы предприятий показали низкую среднюю урожайность пшеницы – 6,01 ц/га в 2012 году, при средней урожайности за 2008-2012 гг. – 9,36 ц/га. Значительно более низкая продуктивность пшеницы в 2012 году объясняется малым количеством осадков – 86 мм в среднем по территории в период с апреля по август.

В таблице 1 приведены характеристики используемых данных в анализе.

Таблица 1

Дескриптивная статистика объектов выборки, Алтайский край, 2012 г.

Показатель	Статистика	Итого по выборке	Группа 1 – применяющие интенсивные технологии	Группа 2 – не применяющие интенсивные технологии
Общие данные о предприятии:				
Организационно-правовая форма: 1 – ООО 0 – прочие	Количество (%)	111 (57%) 85 (44%)	30 (64%) 17 (36%)	81 (54%) 68 (46%)
Площадь земель сельхозназначения, га	Среднее (ст. откл.)	9333 (7072)	10419 (6801)	8990 (7143)
Площадь пашни, га	Среднее (ст. откл.)	7150 (5407)	8340 (5400)	6775 (5373)
Общая численность работников, чел.	Среднее (ст. откл.)	135 (402)	136 (125)	135 (456)
Структура производства:				
Доля пшеницы в структуре посевных площадей, %	Среднее (ст. откл.)	45 (21)	42 (19)	46 (22)
Доля паров в структуре посевных площадей, %	Среднее (ст. откл.)	15 (16)	18 (15)	14 (17)
Доля растениеводства в общей выручке предприятия, %	Среднее (ст. откл.)	54 (31)	55 (29)	53 (31)
Технология возделывания пшеницы:				
Внесение удобрений под посевы пшеницы, (Да / Нет)	Количество (%)	47 (24%) 149 (76%)	47 (100%) 0 (0%)	0 (0%) 60 (100%)
Применение средств защиты растений под посевы пшеницы, (Да / Нет)	Количество (%)	136 (69%) 60 (31%)	136 (100%) 0 (0%)	0 (0%) 60 (100%)
Продуктивность пшеницы:				
Урожайность пшеницы в 2012 г., ц/га	Среднее (ст. откл.)	6,01 (3,74)	8,03 (4,30)	5,38 (3,31)
Урожайность пшеницы за 2008-2012 гг., ц/га	Среднее (ст. откл.)	9,36 (3,87)	12,27 (4,24)	8,45 (3,26)
Характеристики менеджера предприятия:				
Возраст руководителя, лет	Среднее (ст. откл.)	53 (8,95)	53 (9,86)	52 (8,66)
Наличие высшего образования (Да / Нет)	Количество (%)	131 (67%) 65 (33%)	28 (60%) 19 (40%)	103 (69%) 46 (31%)

Источник: Подготовлено авторами.

Тестирование гипотез о нормальности распределения по непрерывным переменным на основе теста Шапиро-Уилка показало, что гипотеза о нормальности распределения отвергается для всех исследуемых величин (таблица 2).

Таблица 2
Результаты теста Шапиро-Уилка для непрерывных переменных выборки

Переменная	p-значение
Сумма осадков за апрель-август, 2012 г.	2.033e-12
Урожайность пшеницы в 2012 г., ц/га	1.022e-9
Средняя урожайность пшеницы в 2008-2012 гг., ц/га	1.526e-5
Площадь земель сельхозназначения в 2012 г., га	4.477e-9
Площадь посева в 2012 г., га	1.616e-9
Доля паров в структуре посевных площадей в 2012 г., %	3.48e-16
Нагрузка на 1 рабочего, га	< 2.2e-16
Возраст менеджера, лет	3.643e-4

Источник: Подготовлено авторами.

4. Результаты

В ходе анализа данных были использованы t-тест сравнения средних, тест Манна-Уитни-Вилкоксона, модели регрессии при одном и множестве факторов, а также метод псевдорандомизации PSM. Результаты оценки причинных эффектов интенсификации технологии возделывания пшеницы приведены в таблице 3.

Результаты показывают, что применение средств защиты растений и удобрений помогло бы поднять урожай с каждого гектара как минимум на 3,12 центнера, что принесло бы дополнительно 2652 рублей прибыли с каждого гектара земли в ценах 2012 г., при средней цене реализации за тонну зерна в 8500 рублей [19]. Затраты на СЗР составили бы около 280 рублей на гектар земли, затраты на удобрения – 328 рублей на гектар земли в ценах 2012 года. Прирост чистой прибыли при применении СЗР составил бы 2044 рублей с гектара земли. Результаты также показали, что эффект интенсификации технологии возделывания яровой пшеницы характеризуется нелинейностью, в частности сумма эффектов от применения удобрений и СЗР отдельно не равна эффекту применения обоих элементов интенсивной технологии, что означает, что эти эффекты не являются аддитивными. В целом оценки причинных эффектов варьируются от 0,39 до 3,23 ц/га в случае внесения удобрений, применение СЗР дает эффект варьируемый от 0,94 до 3,52 ц/га. Интенсификация технологии и применении обоих компонент дает эффект от 1,34 до 5,68 ц/га. Таким образом, можно констатировать позитивный эффект интенсификации, обнаруженный по данным практики возделывания пшеницы в Алтайском крае.

Таблица 3
Результаты оценки причинных эффектов интенсификации технологии возделывания пшеницы, Алтайский край

Статистики	Уровни интенсификации технологии возделывания пшеницы:		
	Внесение удобрений	Применение СЗР	Применение СЗР и удобрений
Парное сравнение средних:			

Продолжение таблицы 3				
t-тест с различными дисперсиями	Ожидаемое значение; 95% доверительный интервал; р-значение	1,85 (0,68; 3,02) 0,002	2,11 (1,09; 3,13) 0,001	3,47 (2,02; 4,51) 0,001
Тест Манна-Уитни-Вилкоксона	Ожидаемое значение; 95% доверительный интервал; р-значение	1,60 (0,60; 2,6) 0,002	2,00 (1,10; 2,90) 0,001	3,20 (2,00; 4,50) 0,001
Регрессионный анализ:				
Однофакторная регрессия	Ожидаемое значение; 95% доверительный интервал; р-значение	1,85 (0,79; 2,91) 0,001	2,11 (1,05; 3,18) 0,001	3,47 (2,00; 4,94) 0,001
Многофакторная регрессия	Ожидаемое значение; 95% доверительный интервал; р-значение	1,50 (0,48; 2,58) 0,007	2,06 (0,94; 3,18) 0,001	3,12 (1,34; 4,91) 0,001
Propensity Score Matching:				
Метод полного соответствия	Ожидаемое значение; 95% доверительный интервал; р-значение	1,50 (0,41; 2,58) 0,007	2,06 (0,94; 3,18) 0,001	3,12 (1,34; 4,91) 0,001
Метод генетического соответствия	Ожидаемое значение; 95% доверительный интервал; р-значение	1,81 (0,39; 3,23) 0,013	2,30 (1,07; 3,52) 0,001	3,71 (1,73; 5,68) 0,001

Источник: Подготовлено авторами.

5. Заключение

Преимущества метода PSM относительно других методов заключаются в первую очередь в применении псевдорандомизации данных, что позволяет считать полученную оценку причинного эффекта приближенной к истинному эффекту без использования предположения о стабильности воздействия на объект. Также этот метод позволяет проводить приближенные к экспериментальным условиям исследования на больших массивах данных, а также с данными из тех сфер, в которых провести экспериментальное исследование невозможно или крайне сложно, например в области сельского хозяйства или медицины. Полученные оценки отличаются друг от друга не значительно, однако можно предположить, что оценка методом сопоставления оценок склонностей будет более точной в сравнении с другими методами за счет балансировки данных.

Список литературы

1. Rubin D.B. Basic Concepts of Statistical Inference for Causal Effects in Experiments and Observational Studies.
2. Yamamoto T. Statistical Models for Causal Analysis.

3. Imbens G.W., Rubin D.B. Rubin Causal Model // The New Palgrave Dictionary of Economics / Под ред. Palgrave Macmillan. — London : Palgrave Macmillan UK, 2008. — С. 1–10.
4. Kalpić D., Hlupić N., Lovrić M. Student's t-Tests // International Encyclopedia of Statistical Science / Ed. by M. Lovrić. — Berlin, Heidelberg : Springer Berlin Heidelberg, 2011. — P. 1559–1563.
5. Neuhäuser M. Wilcoxon–Mann–Whitney Test / Ed. by M. Lovrić. — Berlin, Heidelberg : Springer Berlin Heidelberg, 2011. — P. 1656–1658.
6. Rubin D.B. Randomization Analysis of Experimental Data: The Fisher Randomization Test Comment // Journal of the American Statistical Association. — 1980. — Vol. 75. — Randomization Analysis of Experimental Data, no. 371. — P. 591.
7. Clinical Pharmacy Education, Practice and Research. — 2019.
8. Diamond A., Sekhon J.S. Genetic Matching for Estimating Causal Effects: A General Multivariate Matching Method for Achieving Balance in Observational Studies // Review of Economics and Statistics. — 2013. — Vol. 95. — Genetic Matching for Estimating Causal Effects, no. 3. — P. 932–945.
9. Pettigrew S. The sample size-imbalance frontier.
10. Exact Matching [Электронный ресурс]. — URL: <https://r.iq.harvard.edu/docs/matchit/2.4-15/Exact{ }Matching.html>.
11. Imai K. Statistics and Causal Inference.
12. Understanding Q-Q Plots | University of Virginia Library Research Data Services + Sciences [Электронный ресурс]. — URL: <https://data.library.virginia.edu/understanding-q-q-plots/>.
13. King G., Nielsen R. Why Propensity Scores Should Not Be Used for Matching // Political Analysis. — 2019. — Vol. 27, no. 4. — P. 435–454.
14. AgroAtlas - Главная [Электронный ресурс]. — URL: <http://www.agroatlas.ru/ru/index.html>.
15. Dronin N.M., Bellinger E.G. Climate dependence and food problems in Russia, 1900-1990: the interaction of climate and agricultural policy and their effect on food problems. Climate dependence and food problems in Russia, 1900-1990. — Budapest; New York : Central European University Press, 2005. — 366 p.
16. Kogan F., Guo W. Early twenty-first-century droughts during the warmest climate // Geomatics, Natural Hazards and Risk. — 2016. — Vol. 7, no. 1. — P. 127–137.
17. Понькина Е. и др. Экономические, экологические, технологические факторы и результаты деятельности сельскохозяйственных предприятий в условиях Кулундинской степи. — Барнаул : Изд-во Алтайского государственного университета, 2014. — 138 с.
18. Prishchepov A.V. and etc. Revealing the determinants of wheat yields in the Siberian breadbasket of Russia with Bayesian networks // Land Use Policy. — 2019. — Vol. 80. — P. 21–31.
19. Мониторинг цен на пшеницу организаций 01.10.2012: АгроНовости Ассет [Электронный ресурс]. — URL: <https://agro-bursa.ru/prices/wheat/01-10-2012/>.