

УДК 634.13:631.523:004.8

UDC 634.13:631.523:004.8

DOI 10.30679/2219-5335-2020-1-61-16-32

DOI 10.30679/2219-5335-2020-1-61-16-32

**ПРИМЕНЕНИЕ МЕТОДА
КЛАСТЕРИЗАЦИИ
В СИСТЕМНО-КОГНИТИВНОМ
АНАЛИЗЕ И ИНТЕЛЛЕКТУАЛЬНОЙ
СИСТЕМЕ «ЭЙДОС»
В СЕЛЕКЦИИ ГРУШИ
НА ЗАДАННЫЕ ПРИЗНАКИ**

**APPLICATION
OF CLUSTERIZATION METHOD
IN THE SYSTEM-COGNITIVE
ANALYSIS AND INTELLECTUAL
SYSTEM «EIDOS»
IN PEAR BREEDING
ON THE SET SIGNS**

Киселёва Наталья Станиславовна
канд. биол. наук
старший научный сотрудник
лаборатории селекции плодовых культур
e-mail: nskiselyeva_05@mail.ru

Kiseleva Natalia Stanislavovna
Cand. Biol. Sci.
Senior Research Associate
of Garden Breeding crops Laboratory
e-mail: nskiselyeva_05@mail.ru

*Федеральное государственное
бюджетное научное учреждение
«Всероссийский научно-
исследовательский институт
цветоводства и субтропических
культур», Сочи, Россия*

*Federal State Budgetary
Scientific Institution
«All-Russian Research
Institute of Floriculture
and Subtropical Crops»,
Sochi, Russia*

В селекции плодовых культур, в частности груши, отбор гибридного потомства очень трудоёмок и обусловлен многолетним циклом, что не даёт возможность выявить ценные признаки растения в первые годы. Это влечёт необходимость дальнейшего изучения большего количества выборки, что в итоге не всегда даёт положительные результаты. Поэтому возможность применения системно-когнитивного анализа (АСК-анализа) позволяет выявлять закономерности в имеющейся информации, производить её анализ и на основании этого давать возможные прогнозы. В работе показаны результаты количественной оценки качества плодов сортов и гибридов груши по комплексу признаков с применением системно-когнитивного анализа (АСК-анализа) и его программного инструментария интеллектуальной системы «Эйдос». Созданы 3 статистические и 7 системно-когнитивных моделей, в которых сформированы обобщённые образы классов по количественной оценке комплекса

In the breeding of fruit crops, in particular pears, the selection of hybrid offspring is very time-consuming and limited by the long-term cycle and it causes the inability to identify the valuable traits of plant in the first years. This entails the need for further study of a larger number of samples, which ultimately does not lead to the positive results. Therefore, the possibility of applying the systemic cognitive analysis (ASK analysis) allows you to identify the patterns in the available information, analyze it and, based on this, give the possible forecasts. The quantitative estimation results of fruit quality of pear varieties and hybrids on signs complex with using an automated system-cognitive analysis (ASC-analysis) and its software toolbox of intellectual system «Eidos» have been shown in this work. 3 statistical and 7 system-cognitive models have been generated, in which generalized the images of classes by quantitative estimation of signs complex of fruit «quality» of pear's genotypes. Influence of values

признаков «качества» плода генотипов груши. Изучено влияние значений генотипических особенностей. Наиболее достоверной оказалась модель INF4, основанная на частном критерии ROI – Return On Investment, первый вариант расчёта относительных частот: N_j – суммарное количество признаков по j -му классу при интегральном критерии «Сумма знаний». С помощью созданной модели решены задачи прогнозирования (диагностика, классификация, распознавание, идентификация). Определена степень детерминированности классов и значимость градаций описательных шкал. Кластерным анализом выделено 2 группы генотипов по комплексу признаков «качество». Определены возможности использования новых алгоритмов кластерного анализа в методе когнитивной кластеризации, основанном на АСК-анализе и реализованном в интеллектуальной системе «Эйдос» для количественной оценки качества плодов в селекции груши. Показаны результаты агломеративной кластеризации, отличающейся от ранее известных классических методов.

Ключевые слова: ГРУША, СЕЛЕКЦИЯ, ПЛОД, АВТОМАТИЗИРОВАННЫЙ СИСТЕМНО-КОГНИТИВНЫЙ АНАЛИЗ, АСК-АНАЛИЗ, СИСТЕМА «ЭЙДОС»

of genotyping features has been studied. The most reliable turned out to be model INF4, based on private criterion of ROI – Return On Investment, the first variant of calculation of relative frequencies has appeared the most authentic: N_j - total quantity of signs on j class at integrated criterion of «Sum of knowledges». By means of created model have been the forecasting problems (diagnostics, categorizations, recognitions, identification). Degree of classes determinancy and gradation importance of descriptive scales has been defined. Clusterized analysis allocates 2 groups of genotypes on complex of «quality» signs. Possibilities of use of new algorithms clusterized analysis in a method of cognitive clusterization, based on the ASK-analysis and realised in the intellectual «Eidos» system for a quantitative estimation of fruits quality in pear breeding have been defined. The results of agglomerative clusterization, differing from earlier known classical methods has been shown.

Key words: PEAR, BREEDING, FRUIT, AUTOMATED SYSTEM-COGNITIVE ANALYSIS, ASC-ANALYSIS, «EIDOS» SYSTEM

Введение. При селекции методом отбора для следующих поколений отбираются лучшие по генотипу растения. Оптимальным считается генотип, детерминирующий лучшие по заданным критериям потребительские свойства растений. В селекции плодовых культур, в частности груши, отбор гибридного потомства очень трудоёмок и обусловлен многолетним циклом культур и невозможностью выявить ценные признаки в первые годы жизни выделенного растения [1]. Это влечёт необходимость дальнейшего изучения большего количества выборки, что в конечном итоге далеко не всегда даёт положительные результаты. Актуальность применения существующих ком-

пьютерных методов в селекции – в существенном снижении затрат как времени, так и финансовых средств на проведение селекционной работы методом отбора лучших по генотипу растений и использования их для формирования следующего селекционного поколения. Поэтому возможность применения системно-когнитивного анализа (АСК-анализа) и его программного инструментария интеллектуальной системы «Эйдос» позволяет выявлять закономерности в имеющейся информации, производить её анализ, обработку и на основании этого давать возможные прогнозы.

Целью современной селекции является получение сортов, обладающих комплексом хозяйственно ценных признаков. У плодовых растений в число таких показателей входят: продуктивность, устойчивость к болезням и вредителям, качество, срок созревания плодов и другие [2-10]. В списке требований, предъявляемых к современному сорту, наряду с урожайностью высокое качество плодов является, как правило, не менее важным [11, 12]. В задачу наших исследований входила сравнительная оценка генотипов и их идентификация по комплексу признаков качества плодов.

Объекты и методы исследований. В целях определения возможности использования метода кластеризации в АСК-анализе и системе «Эйдос» в селекции для количественной оценки качества плодов сортов и гибридов груши проводились исследования 18 сортов и 3 гибридов с высокими вкусовыми качествами и оптимальными биологическими характеристикам.

В течение 2002-2018 гг. проводилась количественная оценка генотипического разнообразия груши и условий среды (коллекционные насаждения ФГБНУ ВНИИЦиСК, г. Сочи). Материалом для исследования стали данные по комплексу хозяйственно-биологических признаков 21 изучаемого генотипа (*Pyrus communis* L.). Изучение биологических и хозяйственных свойств и оценка качества плодов проводились по методическим изданиям: «Программа и методика сортоизучения плодовых, ягодных и оре-

хоплодных культур» [3], «Программа Северо-Кавказского центра по селекции плодовых, ягодных, цветочно-декоративных культур и винограда на период до 2030 года» [4]. В качестве метода решения задачи предлагается применить новую инновационную интеллектуальную технологию: автоматизированный системно-когнитивный анализ (АСК-анализ) и его программный инструментарий – систему «Эйдос».

Обсуждение результатов. Груша – вторая по широте распространения плодовая культура, нуждающаяся в детальном изучении для отбора сортов, наиболее пригодных к почвенно-климатическим условиям влажных субтропиков Краснодарского края. Плоды изучаемых сортов и гибридов отличаются размерами – от среднего (80-110 г) до крупного (250-450 г). Биохимический состав плодов характеризуется следующими параметрами: содержание сахаров (%); общих кислот (%); витамина С (мг/100г); сухого вещества (%); сахарокислотный коэффициент.

Предлагаемый для применения АСК-анализ представляет собой метод искусственного интеллекта, разработанный проф. Е.В. Луценко [13] для решения широкого спектра таких задач, как идентификация, классификация, диагностика, прогнозирование, и имеет свой программный инструментарий – универсальную когнитивную аналитическую систему «Эйдос-Х++» (система «Эйдос») [13, 14, 15].

Система Эйдос находится в полном открытом бесплатном доступе (http://lc.kubagro.ru/aidos/_Aidos-X.htm) с актуальными исходными текстами (http://lc.kubagro.ru/_AIDOS-X.txt). Особенностью АСК-анализа и системы «Эйдос» является возможность сопоставимости различных количественных и качественных оценок, так как они рассматриваются по количеству содержащихся в них информации, а исходные данные, приведённые в различных единицах измерения, могут быть обработаны в одной модели путём метризации шкал и представления в одних единицах измерения [14].

Для построения модели, способствующей ускорению отбора в селекции груши на качество плодов и тем самым снижению трудоемкости и затрат, было решено использовать данные по комплексу признаков «качество» [3, 16, 17].

Для того, чтобы применить выбранный метод и инструмент для решения поставленной задачи, необходимо было выполнить следующие этапы:

- когнитивная структуризация предметной области;
- формализация предметной области;
- синтез и верификация модели;
- повышение качества модели и выбор наиболее достоверной;
- решение в наиболее достоверной модели задач диагностики (классификации, распознавания, идентификации), поддержки принятия решений и исследования моделируемой предметной области путём исследования её модели.

На этапе когнитивно-целевой структуризации предметной области определяется, что будет учитываться в качестве факторов, действующих на моделируемый объект (причин), а что – в качестве результатов действия этих факторов (последствий). В данной работе мы собираемся на основе комплекса признаков качества плодов груши определять степенную оценку этого качества с распределением генотипов по группам. Соответственно, для формализации задачи, выполняемой на следующем этапе АСК-анализа, мы будем использовать 2 классификационные шкалы – «генотип», «срок созревания» и 8 описательных – «масса плодов», «дегустационная оценка», «привлекательность внешнего вида», «содержание сахаров», «общая кислотность», «аскорбиновая кислота», «количество сухих веществ», «сахарокислотный коэффициент».

Исходные данные для статьи получены в результате сортоизучения и селекционной работы с генотипами груши, хорошо адаптированными по своим свойствам для условий влажных субтропиков России. В соответ-

ствии с методологией АСК-анализа решение поставленной задачи выполнено в три этапа:

- преобразование исходных данных из файла MS Excel в базы данных системы «Эйдос»;
- синтез и верификация моделей предметной области;
- применение моделей для решения задач идентификации, прогнозирования и исследования предметной области.

Для решения поставленной задачи проведено преобразование исходных данных из файла MS Excel в базы данных системы «Эйдос» (рис.1).

	В	С	Д	Е	Ф	Г	Н	И	Ж	К	Л	М
	ГЕНОТИП	Срок созревания	Средняя урожайность, ц/га	Периодичность	Сумма сахаров, %	Общая кислотность, %	Сухое вещество, %	Витамин «С», мг/%	Сахаро-кислотный индекс	Средняя масса плода, г	Привлекательность внешнего вида, балл	Дегустационная оценка, балл
2	Вильямс	летний	104,9	54,99	9,62	0,39	15	8,75	24,66	130	4,3	4,1
3	Кюре	осенне-зимний	101,5	37,25	11,2	0,33	18,9	5,63	33,93	120	4	3,5
4	Рассвет	осенне-зимний	115,8	65,95	11,31	0,26	14,8	5,81	43,5	160	4,5	4,7
5	гибрид 85	осенне-зимний	118,3	23,21	11,41	0,77	16,9	10,74	14,82	120	4,2	4,4
6	Южанка	летний	85	52,98	11,73	0,26	14,8	5,81	45,11	140	4	4,3
7	Сочинская	ранне-летний	140	75,72	11,95	0,67	18,8	10,67	17,83	140	4,6	4,7
8	Бере Бос	осенне-зимний	120	118,72	12,59	0,38	20,6	6,34	33,13	120	4,2	4,4
9	Хостинская	осенне-зимний	92,5	40,2	12,59	0,38	16,1	5,61	33,13	200	4,4	4,6
10	Нектарная	летний	83,3	71,2	13,01	0,38	15,5	8,92	34,24	200	4	4,5
11	Черномор	летний	85,5	29,1	13,39	0,42	19,2	7,42	31,88	120	4,4	4,5
12	Яснотка	летний	129,9	53,68	13,57	0,32	23,1	8,98	42,4	110	4,2	4,6
13	Лучистая	осенне-зимний	85,8	77,42	17,07	0,22	18,5	4,05	77,59	130	4,4	4,4
14	Славянка	осенне-зимний	61,6	40,28	11,95	0,26	14,8	5,81	45,96	130	4	4
15	Вербена	осенне-зимний	96,6	37,1	13,33	0,58	15,5	8,92	22,98	150	4,4	4,7
16	Вега	ранне-летний	180	25,8	24,93	0,35	18,5	7,39	71,23	120	4	4,4
17	гибрид 22	летний	100	84,2	22,68	0,29	20	4,75	78,21	230	4,2	4,4
18	гибрид 11	летний	115	75	21,7	0,31	16	8,84	70	250	4,2	4,2
19	Душица	осенне-зимний	113,3	89	12,27	0,32	22	0,35	38,34	120	4,2	4,5
20	Сочинская	летний	109,1	131,16	12,27	0,32	22	3,17	38,34	160	4,2	4,7
21	Медуница	летний	86,6	91,38	12,59	0,32	22,5	9,15	39,35	120	4,4	4,3
22	Хохлатка	летний	83,5	89,62	11,2	0,38	18,7	10,56	29,47	200	4,3	4,2

Рис. 1. Исходные данные для ввода в систему «Эйдос»

В данном случае мы хотели бы идентифицировать генотипы груши по определенному признаку. Вставляем xls-файл с таблицей в папку Inp_data. Запускаем режим 2.3.2.2 системы «Эйдос», представляющий собой автоматизированный программный интерфейс (API) с внешними данными табличного типа.

На рисунке 2 приведены реально использованные параметры таблицы Inp_data.xls. В результате работы режима сформированы 2 классификационные шкалы с суммарным количеством градаций (классов) 24 и 8 описа-

тельных шкал с суммарным числом градаций 80. С использованием классификационных и описательных шкал и градаций исходные данные были закодированы и в результате получена обучающая выборка, которая представляет собой нормализованные исходные данные, то есть таблицу исходных данных, закодированную с помощью классификационных и описательных шкал и градаций.

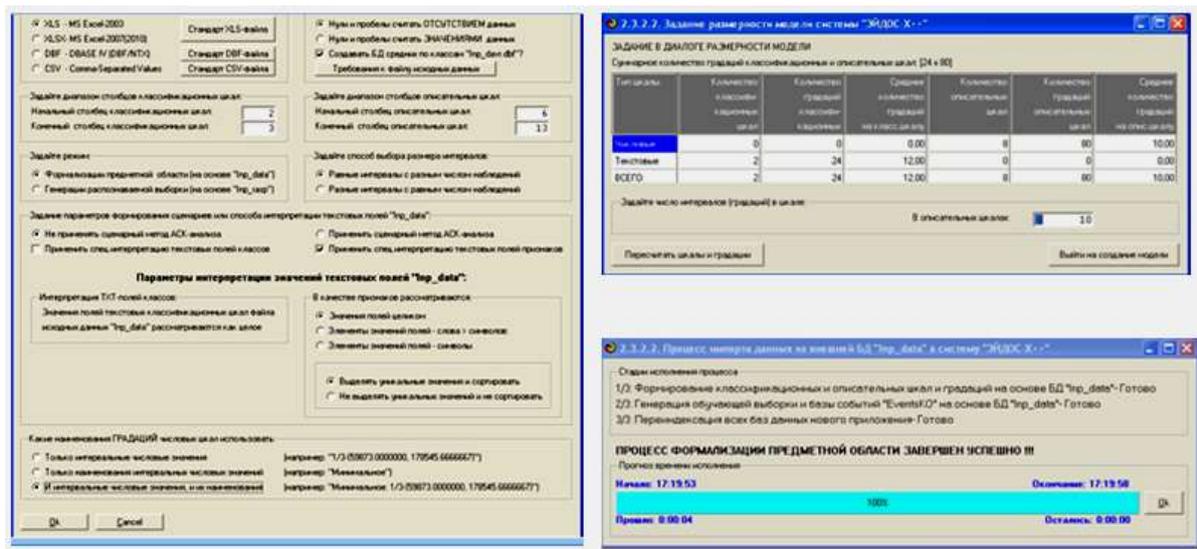


Рис. 2. Экранные формы универсального программного интерфейса (API) 2.3.2.2 системы «Эйдос» с внешними данными табличного типа

Синтез и верификация статистических и системно-когнитивных моделей (СК-моделей) осуществляется в режиме 3.5 системы «Эйдос» (рис. 3).

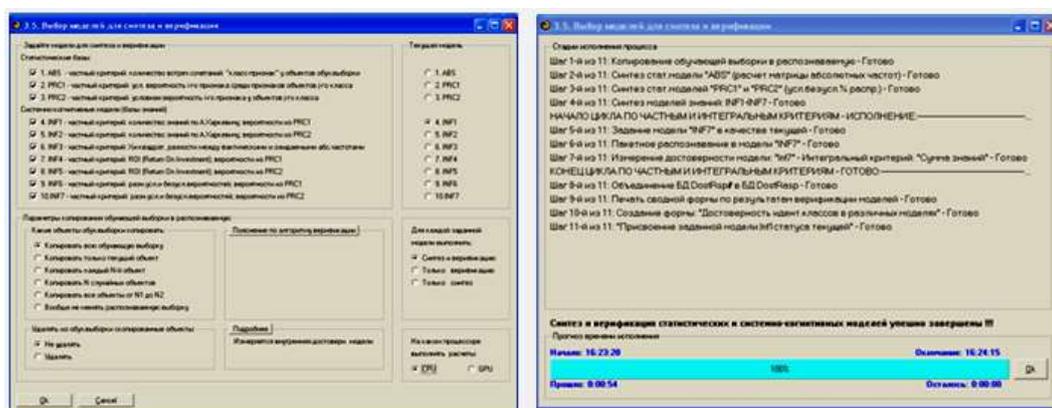


Рис. 3. Экранная форма режима синтеза и верификации статистических и системно-когнитивных моделей системы «Эйдос»

Из рисунка 3 видно, что весь процесс синтеза и верификации моделей занял 54 секунды, так как операция выполнена на центральном процессоре (CPU), что занимает больше времени, чем на графическом процессоре (GPU) видеокарты. Фрагменты самих созданных статистических и системно-когнитивных моделей (СК-модели) приведены на рисунке 4.

The image shows two side-by-side screenshots of spreadsheets. The left spreadsheet (INF3) has columns for 'Модель', 'Параметры модели', 'Критерий Л1', 'Критерий Л2', 'Критерий Л3', 'Критерий Л4', 'Критерий Л5', 'Критерий Л6', 'Критерий Л7', 'Критерий Л8', 'Критерий Л9', 'Критерий Л10'. The right spreadsheet (INF4) has columns for 'Модель', 'Параметры модели', 'Критерий Л1', 'Критерий Л2', 'Критерий Л3', 'Критерий Л4', 'Критерий Л5', 'Критерий Л6', 'Критерий Л7', 'Критерий Л8', 'Критерий Л9', 'Критерий Л10', 'Сумма', 'Среднеарифметическое значение'. Both spreadsheets contain numerical data for various models and their corresponding criteria.

Рис. 4. Модели INF3 и INF4 (фрагменты)

Оценка достоверности моделей в системе «Эйдос» осуществляется путем решения задачи классификации объектов обучающей выборки по обобщенным образам классов и подсчета количества истинных положительных и отрицательных, а также ложных положительных и отрицательных решений по F-мере Ван Ризбергена, а также по критериям L1 и L2-мерам проф. Е.В. Луценко, смягчающим и преодолевающим недостатки F-меры [13]. В режиме 3.4 системы «Эйдос» изучается достоверность каждой частной модели в соответствии с этими мерами достоверности (рис. 5).

Из рисунка 5 видно, что в данном интеллектуальном приложении по критерию L2 проф. Е.В. Луценко [13] наиболее достоверными являются СК-модели INF4 и INF5 с интегральным критерием «Сумма знаний» ($L2 = 0,908$ при максимуме 1,000 и 97,619 % достоверности идентификации), что является достаточно хорошим результатом. На основании этого можно говорить о наличии достоверной зависимости параметров качества плодов

груши с учётом периода созревания в разных погодных условиях от генотипических особенностей изучаемых сортов и гибридов.

Наименование модели и частота критерия	Интегральный критерий	С	С	С	А-Генность модели A-Резонанс -АТР/АТР+	А-Полнота модели A-Искл -АТР/АТР+	L2-мера проф. Е.В. Луценко	Процент правых решений	Процент правых не решений
1. ABS - частный критерий: количество веток соцветий "х/кас."	Корреляция абс частот с обр...				0.827	1.000	0.905	100.000	15.17
1. ABS - частный критерий: количество веток соцветий "х/кас."	Средна абс частот по призна...				0.837	1.000	0.911	100.000	10.72
2. PFC1 - частный критерий: усл. вероятность н-го признака сред.	Корреляция усл.отн частот с о...				0.827	1.000	0.905	100.000	15.17
2. PFC1 - частный критерий: усл. вероятность н-го признака сред.	Средна усл.отн частот по призна...				0.741	1.000	0.851	100.000	10.72
3. PFC2 - частный критерий: условная вероятность н-го признака.	Корреляция усл.отн частот с о...				0.827	1.000	0.905	100.000	15.17
3. PFC2 - частный критерий: условная вероятность н-го признака.	Средна усл.отн частот по призна...				0.741	1.000	0.851	100.000	10.72
4. INF1 - частный критерий: количество значаев по А/Харкевичу, в.	Семантический резонанс знач...				0.835	0.842	0.838	90.476	68.58
4. INF1 - частный критерий: количество значаев по А/Харкевичу, в.	Средна значаев				0.795	0.903	0.846	95.238	42.70
5. INF2 - частный критерий: количество значаев по А/Харкевичу, в.	Семантический резонанс знач...				0.835	0.842	0.838	90.476	68.58
5. INF2 - частный критерий: количество значаев по А/Харкевичу, в.	Средна значаев				0.795	0.903	0.846	95.238	42.70
6. INF3 - частный критерий: Унклардот, равенство между фактис.	Семантический резонанс знач...				0.818	0.854	0.836	92.857	73.15
6. INF3 - частный критерий: Унклардот, равенство между фактис.	Средна значаев				0.853	0.756	0.802	92.857	72.88
7. INF4 - частный критерий: ROI (Return On Investment), вероятно.	Семантический резонанс знач...				0.849	0.870	0.859	92.857	78.46
7. INF4 - частный критерий: ROI (Return On Investment), вероятно.	Средна значаев				0.849	0.976	0.908	97.619	39.11
8. INF5 - частный критерий: ROI (Return On Investment), вероятно.	Семантический резонанс знач...				0.849	0.870	0.859	92.857	78.46
8. INF5 - частный критерий: ROI (Return On Investment), вероятно.	Средна значаев				0.849	0.976	0.908	97.619	39.11
9. INF6 - частный критерий: разн.усл.и безуслов.вероятностей, вер.	Семантический резонанс знач...				0.830	0.793	0.811	92.857	53.17
9. INF6 - частный критерий: разн.усл.и безуслов.вероятностей, вер.	Средна значаев				0.757	0.954	0.844	92.857	37.39
10. INF7 - частный критерий: разн.усл.и безуслов.вероятностей, ве.	Семантический резонанс знач...				0.830	0.793	0.811	92.857	53.17
10. INF7 - частный критерий: разн.усл.и безуслов.вероятностей, ве.	Средна значаев				0.757	0.954	0.844	92.857	37.39

Рис. 5. Экранная форма с информацией о достоверности моделей по F-критерию Ван Ризбергена и L1- и L2-критериям проф. Е.В. Луценко

Частотные распределения числа истинных и ложных положительных и отрицательных решений по результатам прогнозирования оценки качества плодов груши с учётом их генотипических особенностей и сроков созревания в СК-модели INF4 по данным обучающей выборки показаны на рисунке 6 и также подтверждают, что созданная модель не даёт ложных решений. Рисунок 6 содержит изображения двух частотных распределений, похожих на нормальные, сдвинутых относительно друг друга по фазе. Левое распределение, меньшее по амплитуде, включает истинноотрицательные и ложноположительные решения, а правое, большее по амплитуде, включает ложные отрицательные и истинноположительные решения.

Сдвиг этих распределений относительно друг друга и другие различия между ними и позволяют решать задачу прогнозирования и другие задачи. В нашем случае видно, что наиболее достоверная модель INF4 лучше определяет непринадлежность объекта к классу, чем принадлежность (от-

рицательные решения всегда истинные). В положительных решениях до 0 идут только ложные решения, при уровнях сходства от 0 до 20 % – ложные и истинные, после 20 % встречаются только истинные решения.

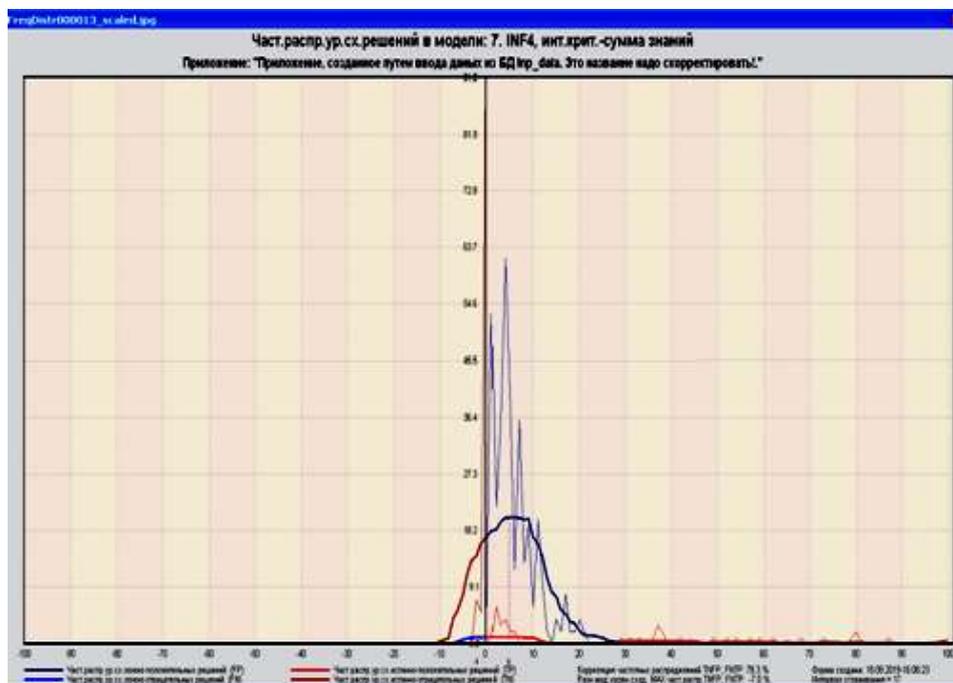


Рис. 6. Частотные распределения числа истинных и ложных положительных и отрицательных решений и их разности в СК-модели INF4 при интегральном критерии «Сумма знаний»

Если модель объекта достаточно адекватна, то её исследование корректно считать исследованием самого моделируемого объекта [13]. В нашем случае это именно так. Степень детерминированности класса тем выше, чем больше среднее количество информации в различных значениях факторов о переходе объекта моделирования в состояние, соответствующее классу.

Паретто-кривая степени детерминированности классов (градации классификационных шкал) в режиме 3.7.3 (рис. 7), наглядно показывает, что в модели INF4 около 50 % наиболее значимых классов обеспечивают 69 % суммарной значимости. В соответствии со схемой обработки данных присвоим СК-модели INF4 статус текущей модели. Для это запустим режим 5.6 с параметрами, приведёнными на экранной форме (рис. 8).

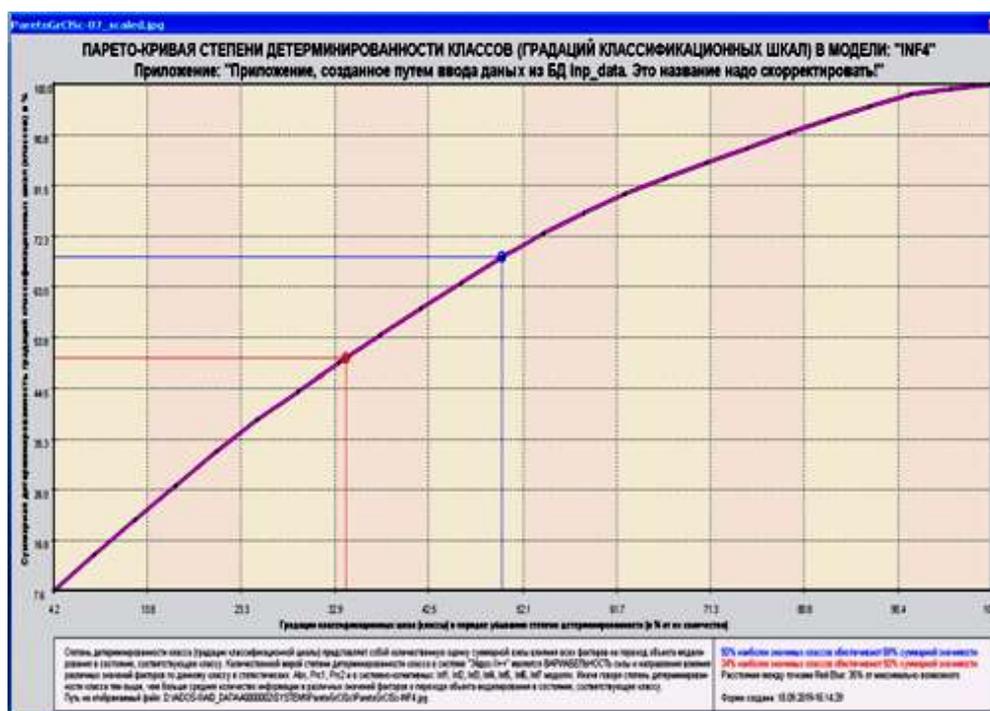


Рис. 7. Парето-кривая значимости влияния особенностей генотипа

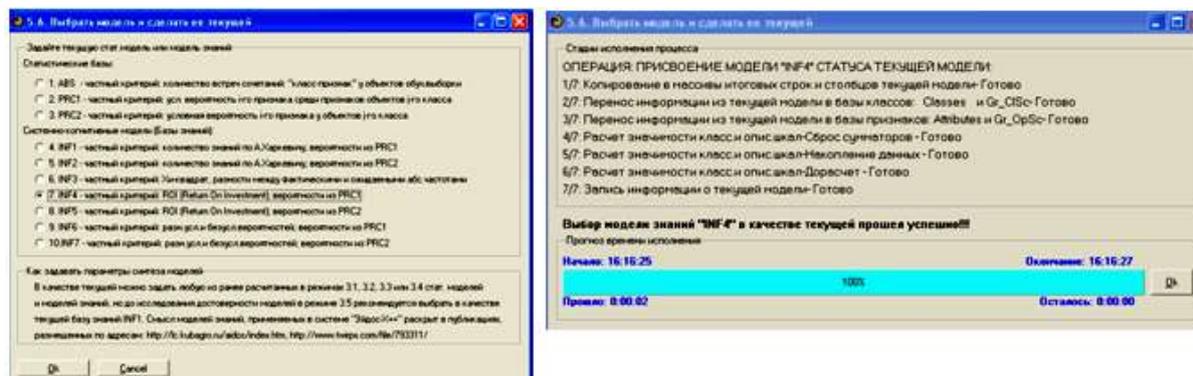


Рис. 8. Экранные формы придания наиболее достоверной по L2-критерию СК-модели INF4 статуса текущей модели

Для выполнения диагностики, то есть прогнозирования степени сходства конкретных генотипов с обобщёнными образами классов по комплексу показателей качества плодов на основе обучающей выборки, в наиболее достоверной СК-модели INF4 запускается режим 4.1.2. Вывод результатов распознавания проводится в режиме 4.1.3 с созданием 10 выходных форм на основе результатов этого прогнозирования. Эти формы отра-

жают результаты прогнозирования в различных разрезах и обобщениях. На рисунке 9 приведены две из этих 10 форм: 4.1.3.1 и 4.1.3.7. Символ «✓» стоит против тех результатов прогнозирования, которые подтвердились на опыте, то есть соответствуют факту.

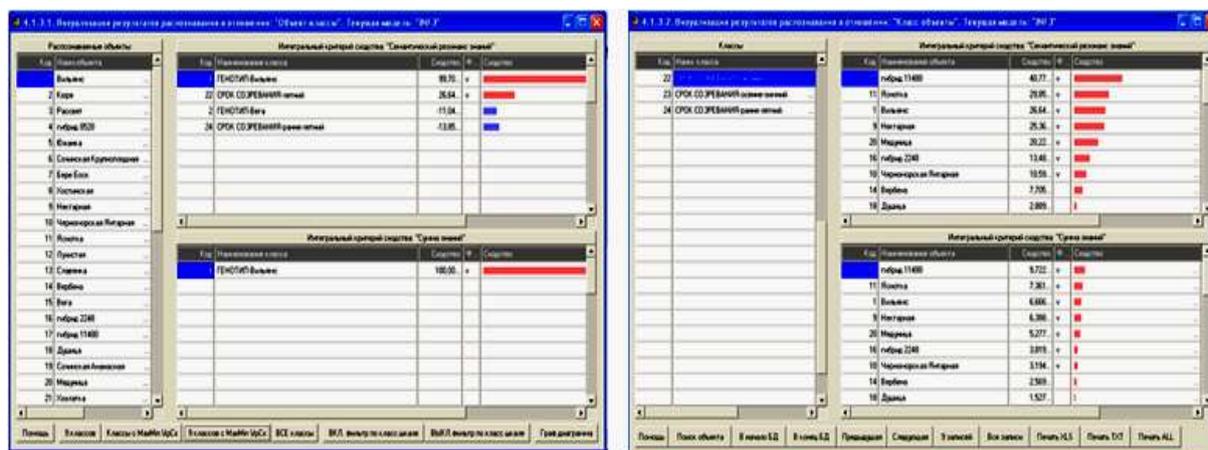


Рис. 9. Выходные формы по результатам прогнозирования количественных оценок параметров «качества» плодов с учётом сроков созревания на основе особенностей генотипов

Из рисунка 9 видно, что результаты хорошие, естественно при учёте информации из рисунка 6 о том, что достоверные прогнозы в данной модели имеют уровень сходства выше 20 %, то есть по сути прогнозы с более низким уровнем сходства можно просто игнорировать.

В системе «Эйдос» есть много возможностей для исследования достоверной модели, отнесённой к самому объекту моделирования, «перенесённой на него», но мы рассмотрим лишь результаты кластерно-конструктивного анализа классов и признаков (когнитивные диаграммы и дендрограммы). Эти диаграммы отражают сходство/различие классов и признаков. Мы получаем их в режимах 4.2.2.1 и 4.2.2.2 (рис. 10).

Отметим, что на когнитивной диаграмме, приведённой на рисунке 10, показаны количественные оценки сходства/различия классов груши по обуславливающим их свойствам, полученные с применением системно-когнитивной модели, созданной на основе имеющихся данных.

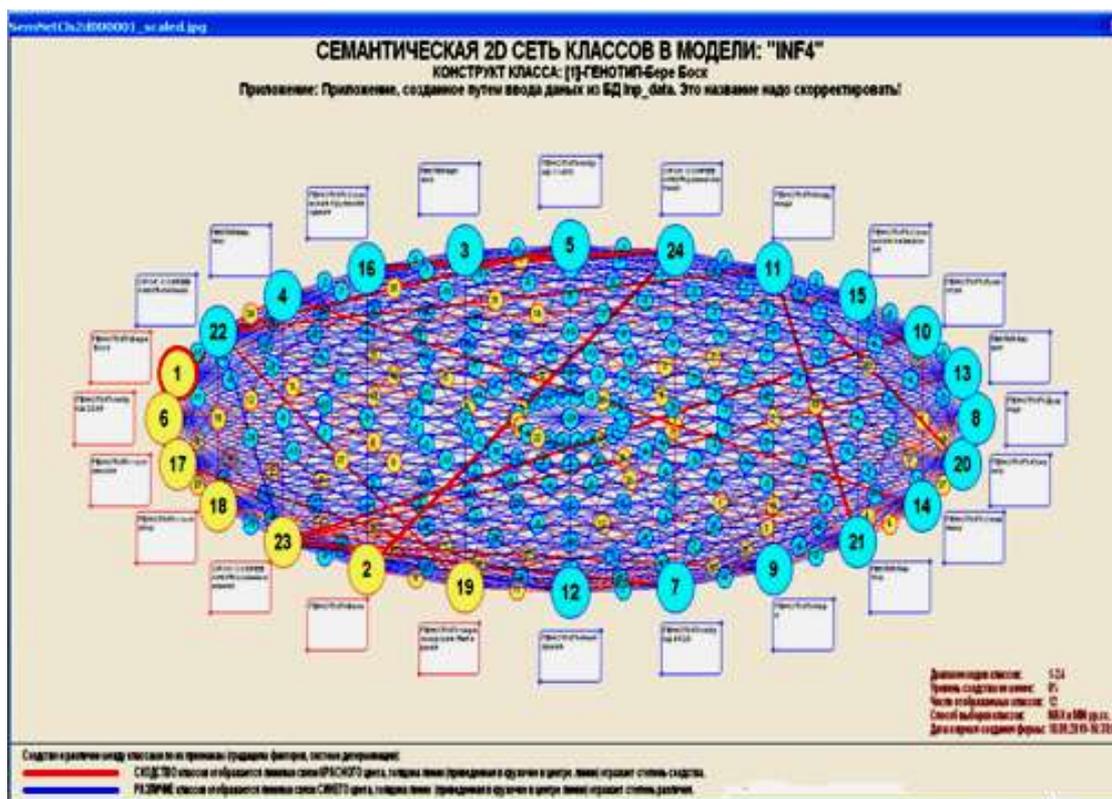
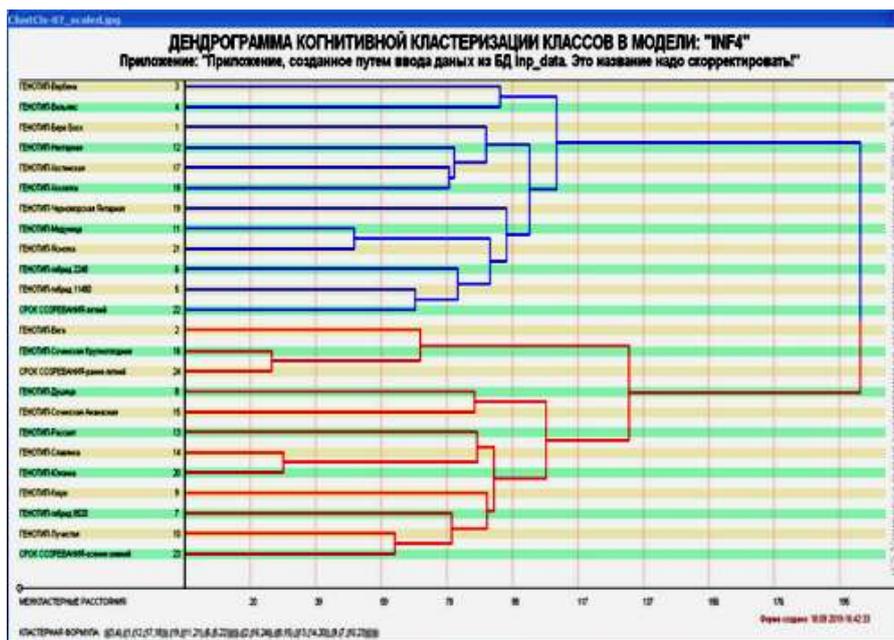
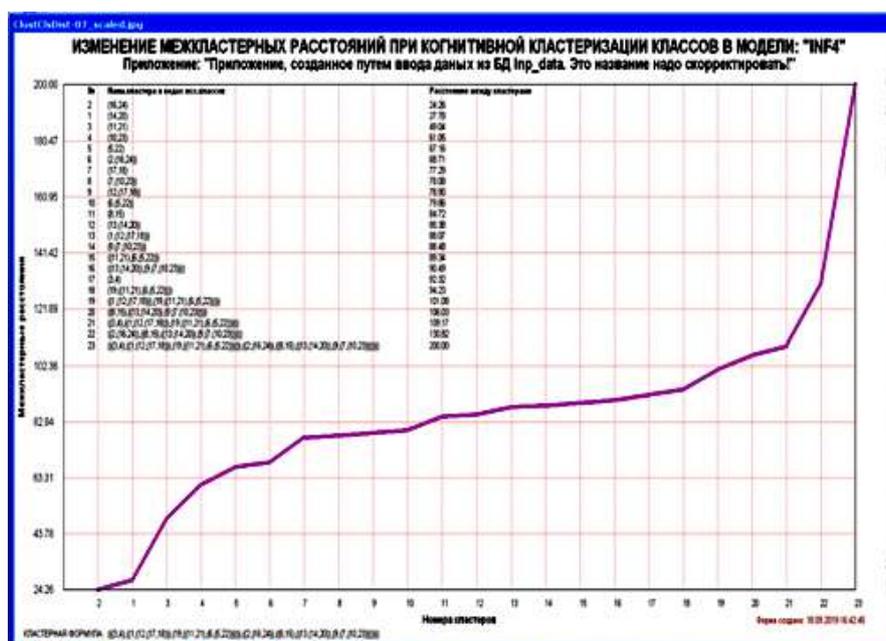


Рис. 10. Когнитивные диаграммы классов и конструктор класса «Генотип Бере Боск»

Информация о сходстве/различии классов, содержащаяся в матрице сходства, может быть визуализирована не только в форме когнитивных диаграмм, пример которой приведён на рисунке 10, но и в форме агломеративных дендрограмм, полученных в результате когнитивной кластеризации [13, 15] (рис. 11). В результате работы данного режима формируются дендрограммы результатов агломеративной когнитивной древовидной кластеризации классов и графики пошагового изменения межкластерного расстояния (рис 11 а, б). Из рисунка 11а видно, что когнитивная кластеризация может начинаться как с монообъектных, так и с полиобъектных классов. Во втором случае классы создаются путём объединения объектов на основе имеющейся информации, необходимой для формирования исходной модели. На рисунке 10б приведён график пошагового изменения межкластерного расстояния при когнитивной кластеризации в текущей модели INF4.



а)



б)

Рис. 11. а) дендрограмма когнитивной кластеризации, полученная в режиме 4.2.2.3. системы «Эйдос»; б) график пошагового изменения межкластерного расстояния при когнитивной кластеризации, полученный в режиме 4.2.2.3 системы «Эйдос»

Выводы. В результате проделанной работы, с помощью системы «Эйдос» были созданы 3 статистические и 7 системно-когнитивных моделей, в которых непосредственно на основе эмпирических данных сформированы

обобщённые образы классов по количественной оценке комплекса признаков «качества» плода генотипов груши, изучено влияние значений генотипических особенностей на эти результаты, и на основе этого решены задачи идентификации, принятия решений и исследования моделируемой предметной области путём исследования её модели. Определена степень детерминированности классов и значимость градаций описательных шкал. Кластерным анализом выделено 2 группы генотипов по комплексу признаков «качество».

Таким образом, в данной статье определены возможности использования новых алгоритмов кластерного анализа в методе когнитивной кластеризации, основанном на АСК-анализе и реализованном в интеллектуальной системе «Эйдос» для количественной оценки качества плодов в селекции груши и показаны результаты агломеративной кластеризации, отличающиеся от ранее известных подробной интерпретацией, не всегда возможной для классических методов.

Литература

1. Bell R.L., Zimmerman R.H. Combining Ability Analysis of Juvenile Period in Pear// Hort. Science, 1990. Vol. 25(11). P. 1425-1427.
2. Quamme, H. A. Resistance to Environmental Stress. / H. A. Quamme, C. Stushnoff // In J.N. Moore and J. Janick (eds.) Methods in fruit breeding. W. Lafayette, Ind.: Purdue Univ. Press, 1983. P.242-266.
3. Программа и методика сортоизучения плодовых, ягодных и орехоплодных культур / под общ. ред. Е.Н. Седова, Т.П. Огольцовой. Орел: ВНИИСПК, 1999. 608 с.
4. Программа Северо-Кавказского центра по селекции плодовых, ягодных, цветочно-декоративных культур и винограда на период до 2030 года / под ред. Е.А. Егорова. Краснодар: ГНУ СКЗНИИСиВ, 2013. С. 57-59. ISBN: 972-5-98272-096-2
5. Современные методологические аспекты организации селекционного процесса в садоводстве и виноградарстве. Краснодар: СКЗНИИСиВ, 2012. 569 с.
6. Комплексная программа по селекции семечковых культур в России на 2001-2020 гг.: Постановление межд. науч.-метод. конф. «Основные направления и методы селекции семечковых культур» (Орел, 31 июля-3 авг. 2001 г.). Орел: ВНИИСПК, 2003. 30 с.
7. Программа селекционных работ по плодовым, ягодным, цветочно-декоративным культурам и винограду Союза селекционеров Северного Кавказа на период до 2010 г. / отв. ред.: Еремин Г.В., Луговской А.П. Т.1. Краснодар: СКЗНИИСиВ, 2005. 342 с.
8. Nilsson, F. Practical results from fruit tree breeding // Svensk. Jordbr. Forsk, 1974. P. 109-119.
9. Alston F.H. Strategy for apple and pear breeding // Improving vegetatively propagated crops, 1978. P. 113-123.
10. Hanter D.M. Pear breeding for the 21 st century-program and progress at Harrow// Acta Horticulturae 338, 1993. P. 377-383.

11. Ульяновская Е.В., Богданович Т.В., Щеглов С.Н. Оценка качества плодов сортов и кребов яблони [Электронный ресурс] // Плодоводство и виноградарство Юга России. 2017. № 48(6). С. 56-63. URL: <http://journalkubansad.ru/pdf/17/06/07.pdf>. (дата обращения: 02.12.2019).

12. Можар, Н.В. Совершенствование сортимента груши на основе конкурентоспособного генофонда // Научные труды СКЗНИИСиВ. Т. 9. Краснодар: ФГБНУ СКЗНИИСиВ, 2016. С 71-77.

13. Луценко Е.В. Автоматизированный системно-когнитивный анализ в управлении активными объектами (системная теория информации и её применение в исследовании экономических, социально-психологических, технологических и организационно-технических систем): монография (научное издание). Краснодар: КубГАУ. 2002. 605 с. <http://elibrary.ru/item.asp?id=18632909>

14. Луценко Е.В. Метризация измерительных шкал различных типов и совместная сопоставимая количественная обработка разнородных факторов в системнокогнитивном анализе и системе «Эйдос» // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. Краснодар: КубГАУ, 2013. №08(092). С. 859-883. IDA [article ID]: 0921308058. Режим доступа: <http://ej.kubagro.ru/2013/08/pdf/58.pdf>.

15. Луценко Е.В., Коржаков В.Е. Метод когнитивной кластеризации или кластеризация на основе знаний (кластеризация в системно-когнитивном анализе и интеллектуальной системе «Эйдос») // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. Краснодар: КубГАУ, 2011. №07(071). С. 528-576. Шифр Информрегистра: 0421100012\0253, IDA [article ID]: 0711107040. Режим доступа: <http://ej.kubagro.ru/2011/07/pdf/40.pdf>.

16. Киселёва Н.С. Выделение ценных генотипов по степени близости к модели в сравнительной оценке коллекции сортов и гибридов груши // Селекция и сорторазведение садовых культур. Конкурентоспособные сорта и технологии для высокоэффективного садоводства: материалы Международной научно-практической конференции, посвящённой 170-летию ВНИИСПК (02-05 июня 2015 г.). Орел: ВНИИСПК, 2015. С. 97-100.

17. Киселёва, Н.С. Количественная оценка качества плода для селекции груши // Субтропическое и декоративное садоводство. 2015. № 52. С. 41-48. ISSN: 2225-3068

18. Donald F. Morrison. Multivariate Statistical Methods, Duxbury Press; 4th Edition. N.Y.: McGraw-Hill, 2004: 496 p. ISBN-10: 9780534387785, ISBN-13: 978-0534387785.

19. Blashfield R.K. The growth of cluster analysis: Tryon, Ward, and Johnson. Multivariate Behavioral Research :15, 1980: PP.439-458.

20. Mojena R. and D. Wishart "Stopping rules for Ward's clustering method," in Proceedings of COMPSTAT, 1980. W.Urzburg, West Germany, Physika-Verlag. PP. 426-432.

21. Jae-On Kim, Charles W. Mueller. Factor Analysis: Statistical Methods and Practical Issues (Eleventh Printing, 1986). PP.78-87:139-201.

References

1. Bell R.L., Zimmerman R.H. Combining Ability Analysis of Juvenile Period in Pear// Hort. Science, 1990. Vol. 25(11). P. 1425-1427.

2. Quamme, H. A. Resistance to Environmental Stress. / H. A. Quamme, C. Stushnoff // In J.N. Moore and J. Janick (eds.) Methods in fruit breeding. W. Lafayette, Ind.: Purdue Univ. Press, 1983. P.242-266.

3. Programma i metodika sortoizuchenie plodovyh, yagodnyh i orekhoplodnyh kul'tur / pod obshch. red. E.N. Sedova, T.P. Ogol'covej. Orel: VNIISPК, 1999. 608 s.

4. Programma Severo-Kavkazskogo centra po selekcii plodovyh, yagodnyh, cvetochno-dekorativnyh kul'tur i vinograda na period do 2030 goda / pod red. E.A. Egorova. Krasnodar: GNU SKZNIISiV, 2013. S. 57-59. ISBN: 972-5-98272-096-2

5. Sovremennye metodologicheskie aspekty organizacii selekcionnogo processa v sadovodstve i vinogradarstve. Krasnodar: SKZNIISiV, 2012. 569 s.

6. Kompleksnaya programma po selekcii semechkovykh kul'tur v Rossii na 2001-2020 gg.: Postanovlenie mezhd. nauch.-metod. konf. «Osnovnye napravleniya i metody selekcii semechkovykh kul'tur» (Orel, 31 iyulya-3 avg. 2001 g.). Orel: VNIISPK, 2003. 30 s.
7. Programma selekcionnykh rabot po plodovym, yagodnym, cvetochno-dekorativnym kul'turam i vinogradu Soyuza selekcionerov Severnogo Kavkaza na period do 2010 g. / otv. red.: Eremin G.V., Lugovskoj A.P. T.1. Krasnodar: SKZNIISiV, 2005. 342 s.
8. Nilsson, F. Practical results from fruit tree breeding // Svensk. Jordbr. Forsk, 1974. P. 109-119.
9. Alston F.H. Strategy for apple and pear breeding // Improving vegetatively propagated crops, 1978. P. 113-123.
10. Hanter D.M. Pear breeding for the 21 st century-program and progress at Harrow// Acta Horticulturae 338, 1993. P. 377-383.
11. Ul'yanovskaya E.V., Bogdanovich T.V., Shcheglov S.N. Ocenka kachestva plodov sortov i krebov yabloni [Elektronnyj resurs] // Plodovodstvo i vinogradarstvo Yuga Rossii. 2017. № 48(6). S. 56-63. URL: <http://journalkubansad.ru/pdf/17/06/07.pdf>. (data obrashcheniya: 02.12.2019).
12. Mozhar, N.V. Sovershenstvovanie sortimenta grushi na osnove konkurentosposobnogo genofonda // Nauchnye trudy SKZNIISiV. T. 9. Krasnodar: FGBNU SKZNIISiV, 2016. S 71-77.
13. Lucenko E.V. Avtomatizirovannyj sistemno-kognitivnyj analiz v upravlenii aktivnyimi ob"ektami (sistemnaya teoriya informacii i eyo primenenie v issledovanii ekonomicheskikh, social'no-psihologicheskikh, tekhnologicheskikh i organizacionno-tekhnicheskikh sistem): monografiya (nauchnoe izdanie). Krasnodar: KubGAU. 2002. 605 s. <http://elibrary.ru/item.asp?id=18632909>
14. Lucenko E.V. Metrizaciya izmeritel'nykh shkal razlichnykh tipov i sovместnaya сопоставимая количественная обработка разнородных факторов в системнокогнитивном анализе i системе «Ejdos» // Politematicheskij setevoj elektronnyj nauchnyj zhurnal Kubanskogo gosudarstvennogo agrarnogo universiteta (Nauchnyj zhurnal KubGAU) [Elektronnyj resurs]. Krasnodar: KubGAU, 2013. №08(092). S. 859-883. IDA [article ID]: 0921308058. Rezhim dostupa: <http://ej.kubagro.ru/2013/08/pdf/58.pdf>.
15. Lucenko E.V., Korzhakov V.E. Metod kognitivnoj klasterizacii ili klasterizaciya na osnove znaniy (klasterizaciya v sistemno-kognitivnom analize i intellektual'noj sisteme «Ejdos») // Politematicheskij setevoj elektronnyj nauchnyj zhurnal Kubanskogo gosudarstvennogo agrarnogo universiteta (Nauchnyj zhurnal KubGAU) [Elektronnyj resurs]. Krasnodar: KubGAU, 2011. №07(071). S. 528-576. Shifr Informregistra: 0421100012\0253, IDA [article ID]: 0711107040. Rezhim dostupa: <http://ej.kubagro.ru/2011/07/pdf/40.pdf>.
16. Kiselyova N.S. Vydelenie cennykh genotipov po stepeni blizosti k modeli v sravnitel'noj ocenke kollekcii sortov i gibridov grushi // Selekcija i sortorazvedenie sadovykh kul'tur. Konkurentosposobnye sorta i tekhnologii dlya vysokoeffektivnogo sadovodstva: materialy Mezhdunarodnoj nauchno-prakticheskoy konferencii, posvyashchyonnoj 170-letiyu VNIISPK (02-05 iyunya 2015 g.). Orel: VNIISPK, 2015. S. 97-100.
17. Kiselyova, N.S. Kolichestvennaya ocenka kachestva ploda dlya selekcii grushi // Subtropicheskoe i dekorativnoe sadovodstvo. 2015. № 52. S. 41-48. ISSN: 2225-3068
18. Donald F. Morrison. Multivariate Statistical Methods, Duxbury Press; 4th Edition. N.Y.: McGraw-Hill, 2004: 496 p. ISBN-10: 9780534387785, ISBN-13: 978-0534387785.
19. Blashfield R.K. The growth of cluster analysis: Tryon, Ward, and Johnson. Multivariate Behavioral Research :15, 1980: PP.439-458.
20. Mojena R. and D. Wishart "\"Stopping rules for Ward's clustering method\"" in Proceedings of COMPSTAT, 1980. WUrzburg, West Germany, Physika-Verlag. PP. 426-432.
21. Jae-On Kim, Charles W. Mueller. Factor Analysis: Statistical Methods and Practical Issues (Eleventh Printing, 1986). PP.78-87:139-201.