

Министерство сельского хозяйства Российской Федерации  
ФГБОУ ВО «КУБАНСКИЙ ГОСУДАРСТВЕННЫЙ АГРАРНЫЙ  
УНИВЕРСИТЕТ имени И.Т. ТРУБИЛИНА»

Кафедра компьютерных технологий и систем

## КОНТРОЛЬНАЯ РАБОТА

по дисциплине: Интеллектуальные информационные системы

на тему: АСК-анализ влияния условий ДТП на количество погибших и пострадавших

выполнил студент группы: ПИЗ1401 Канакин Егор Андреевич

Проверил: д.э.н., профессор ВАК Луценко Евгений Вениаминович

Защищена \_\_\_\_\_ Оценка \_\_\_\_\_  
(дата)

Краснодар, 2017

## СОДЕРЖАНИЕ

ПОДГОТОВКА ДАННЫХ ДЛЯ АНАЛИЗА .....	3
СИНТЕЗ И ВЕРИФИКАЦИЯ МОДЕЛЕЙ .....	6
РЕЗУЛЬТАТЫ ВЕРИФИКАЦИИ МОДЕЛЕЙ .....	8
ВЫВОД .....	11
СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ .....	12

## ПОДГОТОВКА ДАННЫХ ДЛЯ АНАЛИЗА

Для проведения АСК-анализа использовались статистические данные о количестве погибших и пострадавших при различных условиях ДТП. Информация о ДТП была взята с сайта <http://provodim24.ru/statistika-dtp.html>.

Информация была занесена в Excel-файл, а затем импортирована в систему «ЭЙДОС-X++» через универсальный программный интерфейс импорта данных режима 2.3.2.2 (рис. 1).

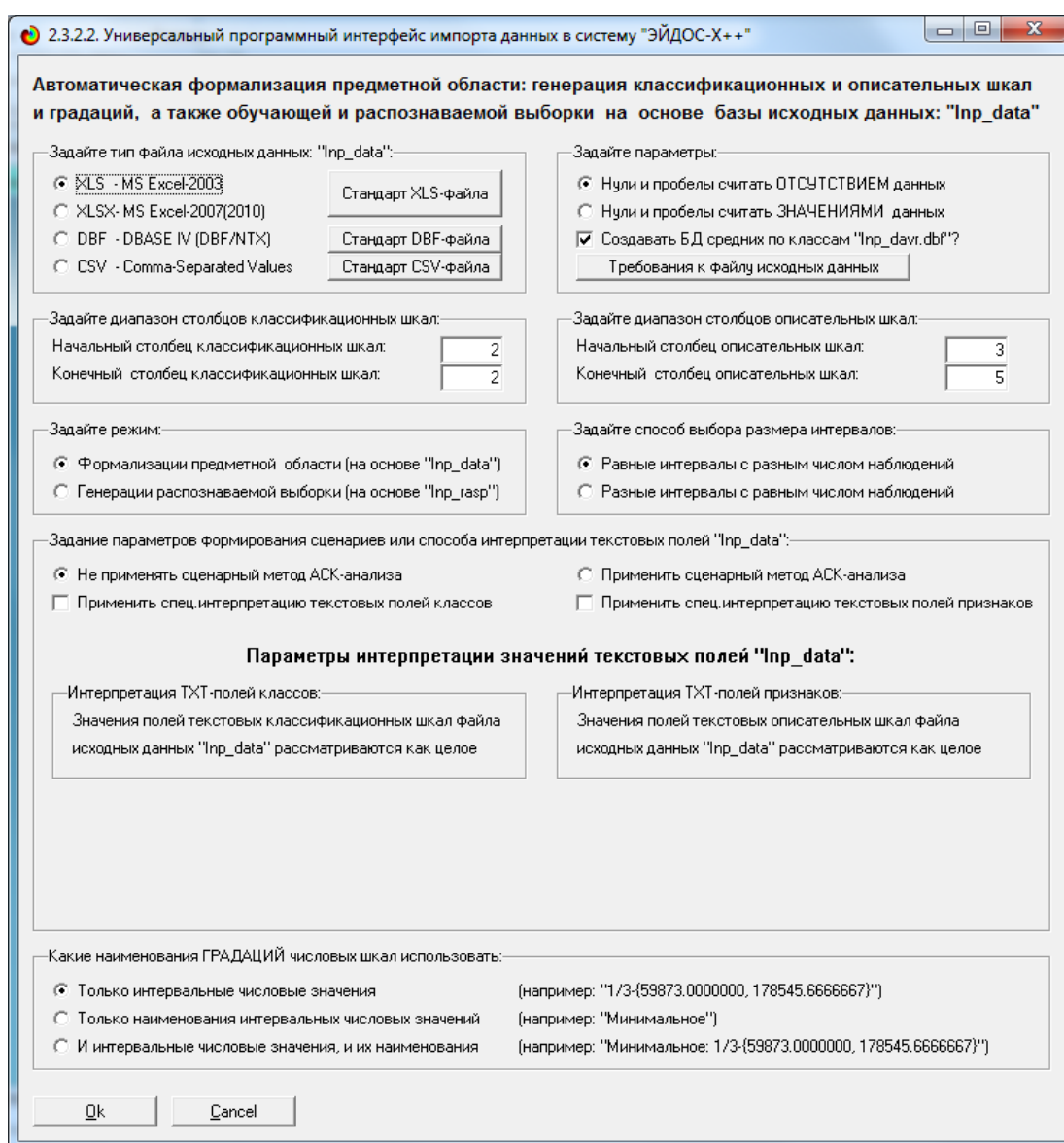


Рисунок 1 – Универсальный программный интерфейс импорта данных в систему «ЭЙДОС-X++»

В экранной форме режима 2.3.2.2 был выбран формат файла для импорта «XLS – MS Excel-2003», а также заданы номера столбцов, содержащие классификационные (2) и описательные (3-5) шкалы.

После импорта данных загруженные классификационные шкалы можно просмотреть в режиме 2.1 (рис. 2).

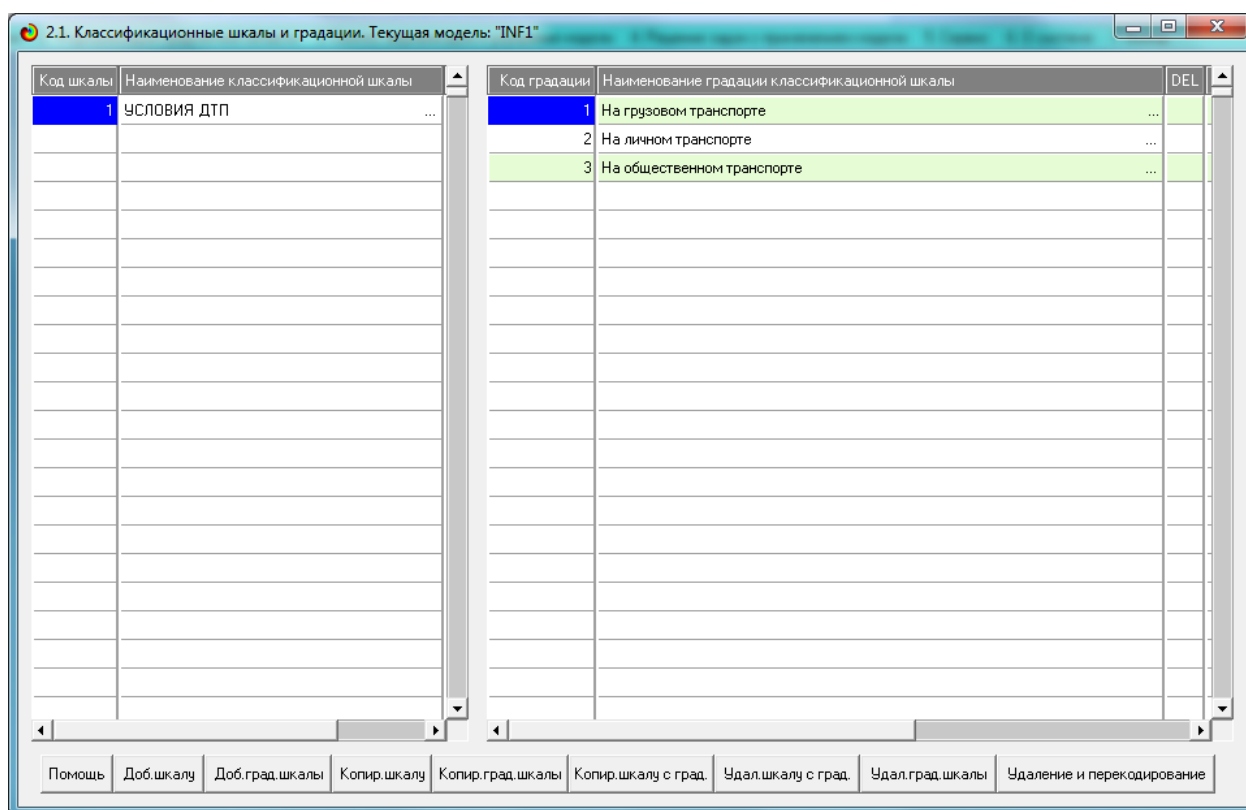


Рисунок 2 – Классификационные шкалы и градации В загруженных данных присутствует одна классификационная шкала текстового типа «Условия ДТП», которая включает 3 класса: «На грузовом транспорте», «На личном транспорте», «На общественном транспорте».

В загруженных данных присутствует одна классификационная шкала текстового типа «Условия ДТП», которая включает 3 класса: «На грузовом транспорте», «На личном транспорте», «На общественном транспорте».

Список описательных шкал импортированных данных доступен в режиме 2.2 (рис. 3).

2.2. Описательные шкалы и градации. Текущая модель: "INF1"

Код шкалы	Наименование описательной шкалы	Код градации	Наименование градации описательной шкалы
1	ПРОЦЕНТ ПОГИБШИХ	1	1/5-{2.7759418, 8.1782002}
2	ПРОЦЕНТ РАНЕННЫХ	2	2/5-{8.1782002, 13.5804587}
3	РЕГИОН РФ	3	3/5-{13.5804587, 18.9827171}
		4	4/5-{18.9827171, 24.3849756}
		5	5/5-{24.3849756, 29.7872340}

Помощь   Доб. шкалу   Доб. град. шкалы   Копир. шкалу   Копир. град. шкалы   Копир. шкалу с град.   Удал. шкалу с град.   Удал. град. шкалы   Перекодировать   Очистить

Рисунок 3 – Описательные шкалы и градации

Импортированные данные имеют 2 описательных шкалы числового типа: «Процент погибших» и «Процент раненых». Эти шкалы были автоматически разбиты системой «ЭЙДОС-Х++» на 5 равных интервалов. 3-я описательная шкала «Регион РФ» является текстовой и представляет собой список регионов, по которым были собраны статистические данные.

## СИНТЕЗ И ВЕРИФИКАЦИЯ МОДЕЛЕЙ

Выбор списка моделей для синтеза и верификации на основании импортированных данных в системе «ЭЙДОС-X++» производится в режиме 3.5 (рис. 4).

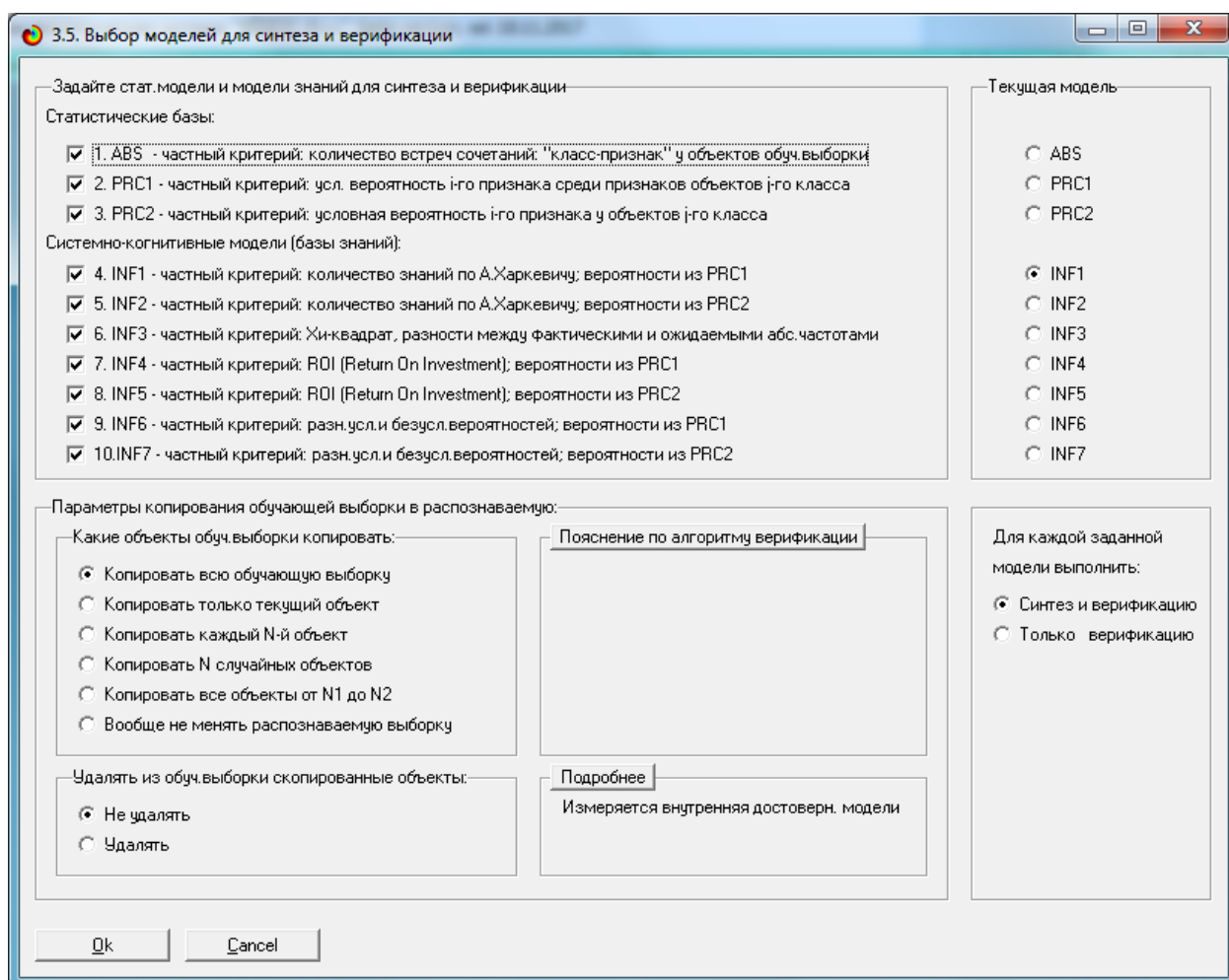


Рисунок 4 – Выбор моделей для синтеза и верификации. В результате выполнения режима 3.5 были синтезированы все выбранные 10 моделей, причем на выполнение данной процедуры системе «ЭЙДОС-X++» потребовалось всего 33 секунды (рис. 5).

В результате выполнения режима 3.5 были синтезированы все выбранные 10 моделей, причем на выполнение данной процедуры системе «ЭЙДОС-X++» потребовалось всего 33 секунды (рис. 5).

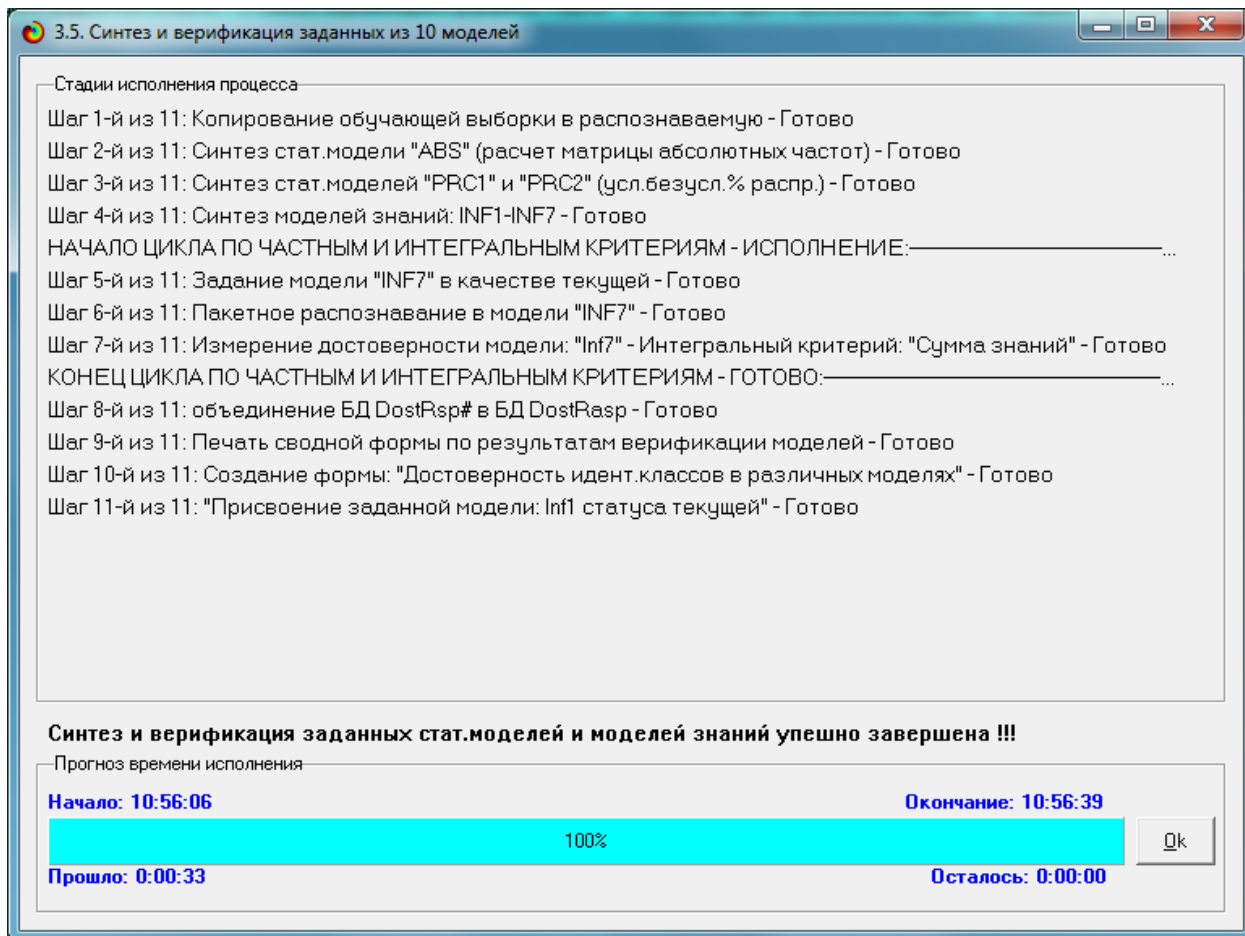


Рисунок 5 – Синтез и верификация заданных моделей

## РЕЗУЛЬТАТЫ ВЕРИФИКАЦИИ МОДЕЛЕЙ

Для просмотра обобщенной таблицы результатов верификации (оценки достоверности) моделей в системе «ЭЙДОС-X++» используется режим 4.1.3.7 (рис. 6).

Наименование модели и частного критерия	Наименование интегрального критерия	Дифференциальная достоверность модели (+1, -1)	Колво логических объектов выборки	Число истинно-положительных решений (TP)	Число истинно-отрицательных решений (TN)	Число ложноположительных решений (FP)	Число ложноотрицательных решений (FN)	Точность модели	Полнота модели	F-мера Ван Ринбергена	Сумма моделей: усреднен. истинно-полож. решений (STP)	Сумма моделей: усреднен. истинно-отриц. решений (STN)	Сумма моделей: усреднен. ложнополож. решений (SFP)
1. ABS - частный критерий: количество встреч сооптаний "ч.лас...	Корреляция абс частот с обр...	-0.235	54	52	10	98	2	0.347	0.963	0.510	34.330	2.356	54.278
1. ABS - частный критерий: количество встреч сооптаний "ч.лас...	Сумма абс частот по призна...	-0.333	54	54	108	108	0	0.333	1.000	0.500	36.565	60.087	60.087
2. PRC1 - частный критерий: усл. вероятность того признака сред...	Корреляция усложн частот с о...	-0.235	54	52	10	98	2	0.347	0.963	0.510	34.330	2.356	54.278
2. PRC1 - частный критерий: усл. вероятность того признака сред...	Сумма усложн частот по при...	-0.333	54	54	108	108	0	0.333	1.000	0.500	36.710	60.400	60.400
3. PRC2 - частный критерий: условная вероятность того признака...	Корреляция усложн частот с о...	-0.235	54	52	10	98	2	0.347	0.963	0.510	34.328	2.356	54.274
3. PRC2 - частный критерий: условная вероятность того признака...	Сумма усложн частот по при...	-0.333	54	54	108	108	0	0.333	1.000	0.500	36.565	60.087	60.087
4. INF1 - частный критерий: количество знаний по А.Харкевичу в...	Семантический резонанс зна...	0.272	54	41	62	46	13	0.471	0.759	0.582	21.829	29.419	18.981
4. INF1 - частный критерий: количество знаний по А.Харкевичу в...	Сумма знаний	0.272	54	43	60	48	11	0.473	0.796	0.593	10.990	15.907	7.564
5. INF2 - частный критерий: количество знаний по А.Харкевичу в...	Семантический резонанс зна...	0.272	54	41	62	46	13	0.471	0.759	0.582	21.883	29.374	18.971
5. INF2 - частный критерий: количество знаний по А.Харкевичу в...	Сумма знаний	0.272	54	43	60	48	11	0.473	0.796	0.593	11.240	16.479	7.904
6. INF3 - частный критерий: Хинвадват, разности между фактич...	Семантический резонанс зна...	0.284	54	41	63	45	13	0.477	0.759	0.586	22.238	29.653	21.185
6. INF3 - частный критерий: Хинвадват, разности между фактич...	Сумма знаний	0.284	54	41	63	45	13	0.477	0.759	0.586	16.051	27.231	15.034
7. INF4 - частный критерий: ROI (Return On Investment); вероятно...	Семантический резонанс зна...	0.272	54	39	64	44	15	0.470	0.722	0.569	20.133	19.253	16.435
7. INF4 - частный критерий: ROI (Return On Investment); вероятно...	Сумма знаний	0.210	54	43	55	53	11	0.448	0.796	0.573	8.888	6.156	5.051
8. INF5 - частный критерий: ROI (Return On Investment); вероятно...	Семантический резонанс зна...	0.272	54	39	64	44	15	0.470	0.722	0.569	20.179	19.241	16.431
8. INF5 - частный критерий: ROI (Return On Investment); вероятно...	Сумма знаний	0.210	54	43	55	53	11	0.448	0.796	0.573	9.085	6.419	5.320
9. INF6 - частный критерий: разн усил и безуслов вероятностей; вер...	Семантический резонанс зна...	0.173	54	40	55	53	14	0.430	0.741	0.544	24.137	25.801	24.753
9. INF6 - частный критерий: разн усил и безуслов вероятностей; вер...	Сумма знаний	0.160	54	41	53	55	13	0.427	0.759	0.547	15.800	22.665	16.648
10. INF7 - частный критерий: разн усил и безуслов вероятностей; вер...	Семантический резонанс зна...	0.160	54	41	53	55	13	0.427	0.759	0.547	24.163	25.660	24.816
10. INF7 - частный критерий: разн усил и безуслов вероятностей; вер...	Сумма знаний	0.160	54	41	53	55	13	0.427	0.759	0.547	15.920	22.680	16.920

Рисунок 6 – Обобщенная таблица результатов верификации моделей

Наиболее достоверная пара модели и интегрального критерия выделена в таблице зеленым цветом. Для импортированных данных эта оказалась модель с частным критерием ROI и вероятностью PRC1 (INF4) по интегральному критерию «Сумма знаний».

Для просмотра и анализа синтезированных моделей в системе «ЭЙДОС-X++» используется режим 4.1.3.11 (рис. 7).

В режиме 4.1.3.11 выделяем наиболее достоверную модель из сводной таблицы и нажимаем кнопку нужного типа графической визуализации результатов верификации выбранной модели. В нашем случае посмотрим график «TP, TN, FP, FN, сумма» для модели INF4 и интегрального критерия «Сумма знаний» (рис. 8).



4.1.3.11. Част.распр.уровн.сход. TP, TN, FP, FN решений при разных моделях и инт.критериях. Текущая модель: "INF1"

Наименование частного и интегрального критерия	-100	-99	-98	-97	-96	-95	-94	-93	-92	-91	-90	-89	-88
Част.распр.Уровней Сходства ложно-положительных решений (FP)													
Част.распр.Уровней Сходства ложно-отрицательных решений (FN)													
Интегральный критерий: СУММА ЗНАНИЙ													
Уровни сходства (Ур.Сх.) [%]:	-100	-99	-98	-97	-96	-95	-94	-93	-92	-91	-90	-89	-88
Част.распр.Уровней Сходства истинных решений (TP+TN)	15												
Част.распр.Уровней Сходства ложных решений (FP+FN)													
Част.распр.Уровней Сходства истинно-положительных решений (TP)													
Част.распр.Уровней Сходства истинно-отрицательных решений (TN)	15												
Част.распр.Уровней Сходства ложно-положительных решений (FP)													
Част.распр.Уровней Сходства ложно-отрицательных решений (FN)													
7. INF4 - ЧАСТНЫЙ КРИТЕРИЙ: ROI (RETURN ON INVESTMENT); ВЕРОЯТНОСТИ ИЗ PRC1 ...													
Интегральный критерий: СЕМАНТИЧЕСКИЙ РЕЗОНАНС ЗНАНИЙ													
Уровни сходства (Ур.Сх.) [%]:	-100	-99	-98	-97	-96	-95	-94	-93	-92	-91	-90	-89	-88
Част.распр.Уровней Сходства истинных решений (TP+TN)													
Част.распр.Уровней Сходства ложных решений (FP+FN)													
Част.распр.Уровней Сходства истинно-положительных решений (TP)													
Част.распр.Уровней Сходства истинно-отрицательных решений (TN)													
Част.распр.Уровней Сходства ложно-положительных решений (FP)													
Част.распр.Уровней Сходства ложно-отрицательных решений (FN)													
Интегральный критерий: СУММА ЗНАНИЙ													
Уровни сходства (Ур.Сх.) [%]:	-100	-99	-98	-97	-96	-95	-94	-93	-92	-91	-90	-89	-88
Част.распр.Уровней Сходства истинных решений (TP+TN)													
Част.распр.Уровней Сходства ложных решений (FP+FN)													

Помощь по графикам | TP, TN, FP, FN, резонанс | TP, TN, FP, FN, сумма | (TP-FP), (TN-FN), резонанс | (TP-FP), (TN-FN), сумма | (T-F)/(T+F)\*100, резонанс | (T-F)/(T+F)\*100, сумма

Рисунок 7 – Выбор графика результатов верификации моделей

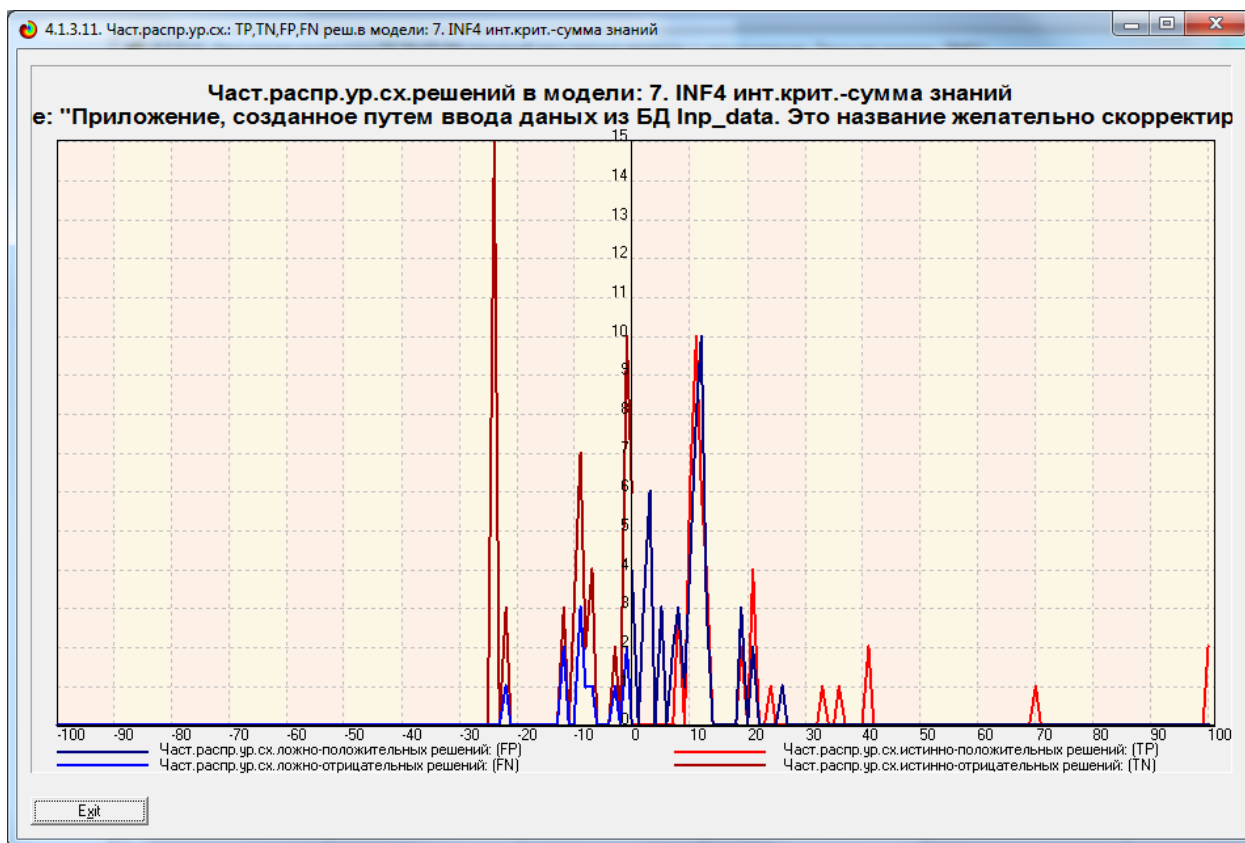


Рисунок 8 – Визуализация результатов верификации моделей

После разделения классов на типичную и нетипичную части (режим 3.7.6) достоверность модели возросла (рисунок 9):

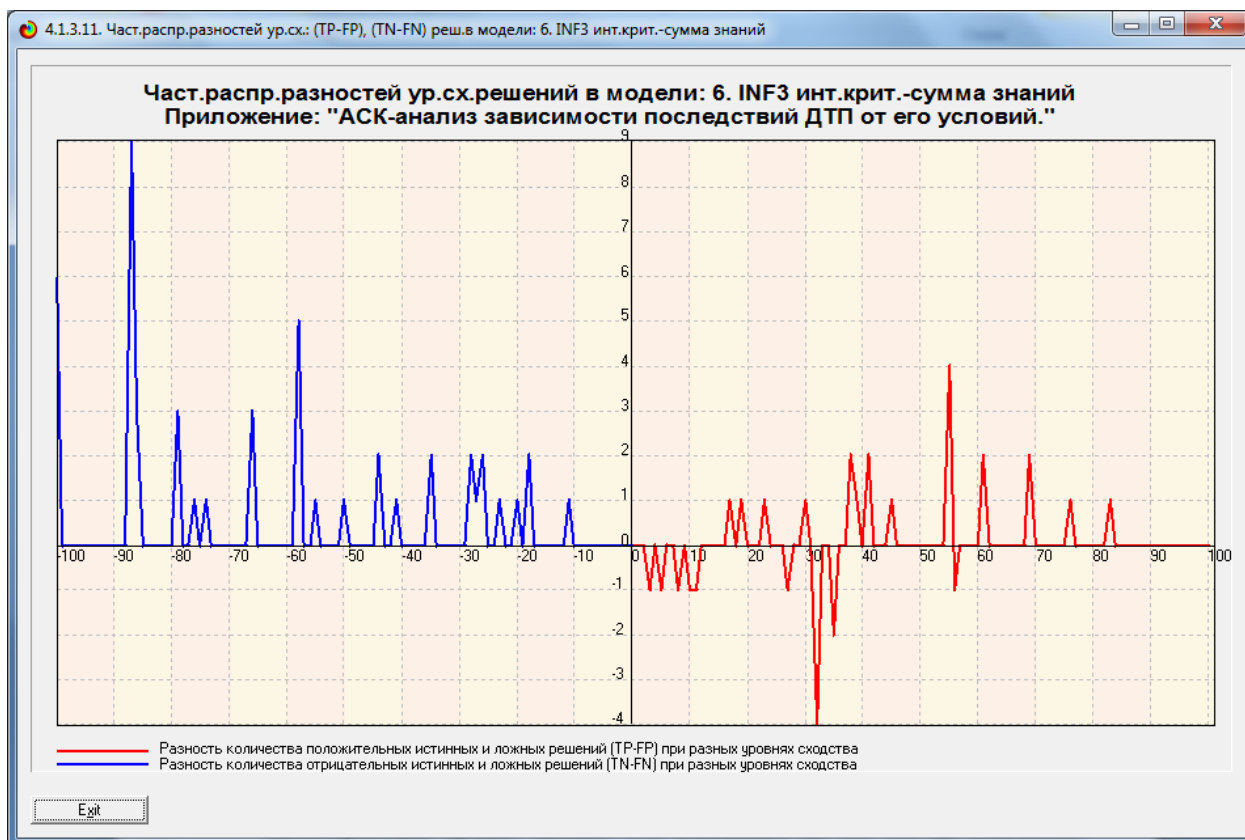


Рисунок 9. Разность между частотными распределениями уровней сходства истинных и ложных решений.

Из рисунка 9 видно, что при уровнях сходства выше 35% практически все положительные решения являются истинными, отрицательные решения все являются истинными, а ложно-положительные решения имеют в основном очень низкий уровень сходства до 10%.

На графике, изображенном на рисунке 8, видно, что частное распределение сходства-различия верно и ошибочно идентифицированных и неидентифицированных состояний в выбранной модели не имеет существенных различий и при повышении уровня сходства-различия степень идентификации крайне низкая. Это указывает на то, что на основании импортированных данных не удалось построить модель, позволяющую по

описанию имеющихся признаков достоверно отнести объект к какому-либо из изучаемых классов.

## ВЫВОД

Система АСК-анализа «ЭЙДОС-Х++» предоставляет автоматизированный инструментарий для синтеза различных математических моделей на основании импортируемых данных о классификационных и описательных шкалах объектов. После синтеза моделей система в автоматическом режиме производит их верификацию и сравнение, выявляя наиболее достоверную модель с соответствующим интегральным критерием.

Система «ЭЙДОС-Х++» также предоставляет возможности визуализированного представления результатов анализа моделей в виде различных графиков и диаграмм, что существенно облегчает их понимание и интерпретацию.

Проведение АСК-анализа данных о количестве погибших и пострадавших в ходе ДТП в различных регионах РФ при различных условиях в системе «ЭЙДОС-Х++» показало, что имеющихся данных недостаточно для построения достоверной математической модели. Для получения более достоверной модели необходимо провести дополнительный сбор информации об изучаемых объектах, дополнить списки классификационных и описательных шкал, а также отбросить случайно попавшие в выборку недостоверные данные.

## СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. Луценко Е.В. Универсальная когнитивная аналитическая система «Эйдос». – Краснодар: КубГАУ, 2014. – 600 с.
2. Луценко Е.В., Коржаков В.Е., Лаптев В.Н. Теоретические основы и технология применения системно-когнитивного анализа в автоматизированных системах обработки информации и управления. – Майкоп: АГУ, 2009. – 536 с.
3. Луценко Е.В., Лойко В.И., Лаптев В.Н. Современные информационно-коммуникационные технологии в научно-исследовательской деятельности и образовании: учеб. пособие. – Краснодар: КубГАУ, 2017. – 450 с.
4. Орлов А.И., Луценко Е.В. Системная нечеткая интервальная математика. – Краснодар: КубГАУ. 2014. – 600 с.
5. <http://provodim24.ru>