

**“САНКТ-ПЕТЕРБУРГСКИЙ НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
УНИВЕРСИТЕТ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ,
МЕХАНИКИ И ОПТИКИ” (УНИВЕРСИТЕТ ИТМО)**

Факультет технологического менеджмента и инноваций

Направление подготовки Прикладная информатика

Магистерская программа Управление государственными информационными системами

Квалификация (степень) магистр

Кафедра УГИС

Группа 6901

МАГИСТЕРСКАЯ ДИССЕРТАЦИЯ
на тему

Поддержка принятия управленческих решений для развития урбанизированной территории на
примере противодействия распространению наркомании в Санкт-Петербурге

Автор магистерской диссертации Великий Д.П. _____ (подпись)
(Фамилия, И., О.)

Научный руководитель Захаров Ю.Н. _____ (подпись)
(Фамилия, И., О.)

Руководитель магистерской программы Чугунов А.В. _____ (подпись)
(Фамилия, И., О.)

:

К з а щ и т е д о п у с т и т ь

Зав. кафедрой Чугунов Андрей Владимирович _____
(Фамилия, И., О.)

“ ____ ” 2015г.

Санкт-Петербург, 2015 г.

Магистерская диссертация выполнена с оценкой _____

Дата защиты “___” 2015 г.

Секретарь ГЭК _____

Листов хранения _____

Чертежей хранения _____

Оглавление

Введение	3
1 Международный опыт использования информационных систем при реализации государственной антинаркотической политики	6
Введение	6
1.1 Международный опыт использования информационных систем при реализации государственной антинаркотической политики	7
1.1.1 Виды антинаркотических политик	7
1.1.2 Использование информационных систем при реализации антинаркотической политики	9
1.1.3 Зарубежный опыт прогнозирования распространения наркомании	12
Выводы по главе 1	16
2 Обзор методов прогнозирования временных рядов	17
Введение	17
2.1 Задачи прогнозирования и интеллектуальный анализ временных рядов	18
2.2 Сфера применения методов интеллектуального анализа временных рядов	20
2.3 Программные средства интеллектуального прогнозирования	23
2.4 Обзор литературы по нечеткому социальному прогнозированию	26
2.5 Теорема о нечеткой аппроксимации	28
Выводы по главе 2	29
3 Разработка модели прогнозирования численности наркозависимых в Санкт-Петербурге на основе нечеткой	

модели с многими переменными	31
Введение	31
3.1 Вычислительный эксперимент нечеткой авторегрессии .	32
3.1.1 Используемые программные средства	32
3.1.2 Методика проведения эксперимента	34
3.1.3 Результаты эксперимента	37
3.2 Разработка модели прогнозирования численности наркозависимых в Санкт-Петербурге на основе нечеткой модели с многими переменными	39
3.2.1 Теоретико-методологические основы анализа наркоситуации	40
3.2.2 Использование индикаторов мониторинга наркоситуации для прогнозирования доли наркозависимых в населении Санкт-Петербурга .	43
3.3 Формальное описание модели	47
Выводы по главе 3	52
4 Применение модели прогнозирования численности наркозависимых для информационно-аналитической поддержки принятия решений органами государственной власти	54
4.1 Текущая модель организации антинаркотического мониторинга и анализа наркоситуации на региональном уровне на примере Санкт-Петербурга	54
4.2 Сравнение эффективности применяемых на практике моделей прогнозирования наркоситуации с предлагаемой новой моделью на основе нечеткой логики	68
Заключение	71
Список литературы	73

Введение

В современном мире усиливается потребность в автоматизации процессов управления, в частности административных процессов государственного уровня. Для этого необходимо создавать информационные системы, отвечающие потребностям лиц, принимающих решения. Управление социально-экономическим развитием города — комплексная задача, включающая в себя управление различными сферами и подсистемами жизнедеятельности города с целью повышения его экономического, творческого, научного потенциала, повышения качества жизни горожан, привлечения инвестиций, создания комфортных условий для предпринимательства, обеспечения социальной защищённости и благоприятного эмоционально-психологического климата в городе.

Одним из проблемных явлений в жизни общества является проблема наркомании, наркотической зависимости. Наркомания наносит ощутимый экономический ущерб обществу. Но не менее опасен урон общественной морали и принципам благородства потому, что наркотики разрушительно влияют на мозг человека. А ведь ещё Рене Декарт провозгласил классический принцип рационализма Нового времени: «*Cogito ergo sum*», или «Мыслю, следовательно существую». Таким образом, можно предположить, что наркотизм причиняет человеку поистине экзистенциальные страдания. Но несмотря на чисто гуманистические соображения о стремлении помочь ближнему нужно понимать, что наркотизм — сложный феномен, полностью победить который не смогли ещё ни в одной стране мира.

Характерными чертами наркоситуации являются скрытность, обусловленная её нелегальной природой, сетевой характер распространения, рискованность, но и прибыльность наркотико-торговли, наличие связей между наркоситуацией и другими негативными явлениями: безработицей, коррупцией, преступностью и др. Все эти особенности требуют подходов, которые способны были бы адресовать их.

Общепринятым подходом к организации антинаркотических программ является использование наукоемкой аналитики для поддержки принятия управленческих решений ответственными лицами. И здесь большую **актуальность** имеют модели, способные не только описать свойства процесса, но и объяснить внутренние механизмы его протекания, связь наркотизации с другими сферами города, генерировать т.н. «actionable insights» — идеи, которые позволяют действовать: применять конкретные рычаги управления, выявленные при помощи моделирования, для влияния на наркоситуацию. Более того, интересны модели, способные к самообучению, универсальной подстройке под исходные данные.

Однако, в практике повседневного информационно-аналитического обеспечения органов государственной власти в Санкт-Петербурге, в основном, применяются традиционные модели прогнозирования, известные по эконометрике, математической статистике: модель авторегрессии с интегрированным скользящим средним, модель Хольта-Винтерса, модель линейной регрессии и др. Данные модели для своего корректного применения требуют наложения определённых ограничений на исходные данные, таких как их линейность, отсутствие т.н. outliers — случайных всплесков, наличие сезонных циклов и т.п. В то же время, данные, наблюдаемые при мониторинге наркоситуации, зачастую являются нелинейными, неполными, полученными из экспертных или вовсе неформальных источников. Для учета этих особенностей могут подходить интеллектуальные методы прогнозирования, одним из которых является нечеткая логика. Нечеткую логику зачастую применяют вместе с алгоритмами машинного обучения, такими, как искусственные нейронные сети и др. Такие авторы, как Н. А. Абдуллавева [43], М. Г. Мамедова, З.Г. Джабраилова [46], П. С. Пак, Г. Ким [30], А. Сасу [33] демонстрируют, хоть местами и противоречивые, но обнадеживающие результаты в применении интеллектуальных методов и нечеткой логики для прогнозирования социальных процессов.

Практическая значимость работы. Подход к прогнозированию с использованием методов интеллектуального анализа временных рядов, гипотетически, может эффективно решить задачу государственного управления и оценки стратегического развития региона в области борьбы с наркоманией. Для этого соответствующее программное обеспечение должно быть внедрено в действующие информационно-аналитические системы. Данная работаа выполняется по заказу Санкт-Петербургского информационно-аналитического центра, деятельность которого связана с разработкой информационно-аналитических систем поддержки принятия решений в органах государственной власти и местного самоуправления.

Цель ВКР состоит в повышении эффективности поддержки принятия решений при прогнозировании развития наркоситуации с Санкт-Петербургом. Для достижения поставленной цели необходимо решить следующие задачи:

- Обзор существующих в мире политик и моделей организации антинаркотической деятельности на государственном уровне;
- Описание используемой в настоящее время системы мониторинга и анализа наркоситуации в Санкт-Петербурге;
- Обзор методов и моделей прогнозирования временных рядов;
- Разработка и опытная проверка метода прогнозирования на основе нечеткой логики, выработка рекомендаций к применению;
- Разработка компонента государственной информационной системы, реализующего новый метод прогнозирования.

ГЛАВА 1

Международный опыт использования информационных систем при реализации государственной антинаркотической политики

Введение

Наркоситуация как разновидность социальной ситуации представляет собой ограниченную временными и пространственными рамками совокупность социальных процессов, складывающихся в результате взаимодействия различных сторон общественных отношений. Рассматривая наркоситуацию как разновидность социального взаимодействия, важно в первую очередь использовать системный подход, выделяя элементы, составляющие наркоситуацию как систему, а также интерпретируя характер взаимодействия между ними.

Полноценное моделирование наркоситуации, по нашему мнению, должно учитывать взаимовлияние социально-экономических индикаторов, описывающих наркоситуацию, тем самым оно может представлять интуитивно понятную логику протекания процессов, и выявлять необходимые для изменения ситуации рычаги управления. Немалую роль здесь играет и сам характер доминирующей в стране системы управления в области противодействия наркомании. В данной главе рассматривается процесс антинаркотического регулирования «сверху-вниз» с тем, чтобы согласовать прикладное моделирование с целями, выдвигаемыми государством при реализации стратегий развития наркоситуации.

Цель главы: разработка модели прогнозирования доли наркозависимых в населении Санкт-Петербурга с помощью индикаторов наркотизации.

Поставленные задачи для достижения цели данной главы:

- обзор существующих антинаркотических политик и способов их информационной поддержки.

1.1 Международный опыт использования информационных систем при реализации государственной антинаркотической политики

Обеспечение национальной безопасности — одна из важнейших функций государства. В Стратегии национальной безопасности Российской Федерации [48] наркомания определяется как социально значимое заболевание, требующее разработки «единых общероссийских подходов к диагностике, лечению и реабилитации пациентов». Снижение уровня наркомании должно способствовать «повышению качества жизни российских граждан».

Однако, существующие модели организации антинаркотической деятельности далеко не всегда достигают цели, а бюджетные средства, которые инвестируются в их реализацию, не окупаются. Поэтому для исследования актуально обратиться к накопленному международному опыту. Предмет данной главы — соотношение между антинаркотической политикой государства и информационными системами, призванными обеспечить её выполнение. В данной главе приводится обзор основных видов наркополитик, обзор и классификация некоторых из существующих в мире информационных систем, обеспечивающих противодействие распространению наркомании в обществе, а также применяемых в аналитике наркоситуации методов прогнозирования.

1.1.1 Виды антинаркотических политик

Рассмотрим виды наркополитик. Политика **снижения потребления** (use reduction) нацелена на устранение, или по меньшей мере снижение употребления наркотиков в обществе. Уменьшение потребления — основная цель контроля над наркотиками в США, что отражено в Национальной стратегии по борьбе с наркотиками (National Drug

Control Strategy) [28]. Она также занимает видное место в антинаркотических политиках европейских стран, таких как Швеция и Франция, а также странах Ближнего Востока. Политика уменьшения потребления основывается на мысли, что проблемы, связанные с наркотиками, возникающие в обществе, семье, социальных группах, могут быть решены, только если прекратить или минимизировать употребление наркотиков.

Для людей, которые не употребляют наркотики, парадигма уменьшения потребления предполагает профилактические программы, разработанные с намерением предотвратить употребление наркотиков.

Политика снижения вреда на микроуровне (micro harm reduction) нацелена на снижение среднего вреда отдельным наркозависимым лицам, не употребляющим наркотики. Эта политика опирается на мысль, что само по себе употребление наркотиков лишь умеренно опасно, и можно предпринять шаги для уменьшения рисков, сопутствующих употреблению наркотиков. Термин «снижение вреда» может трактоваться широко, однако эта гибкость может стать источником путаницы, давая возможность ввода в действие двух диаметрально противоположных политик, направленных при этом одну цель — снижение вреда (например, программы обмена шприцев для наркозависимых с целью уменьшения распространения инфекционных заболеваний и, следовательно, снижения вреда для наркозависимых, и обязательное лишение свободы для лиц, употребляющих наркотики, с целью снижения вреда для тех, кто не употребляет). Поэтому при анализе нужно четко определять задачи, поскольку есть различные негативные эффекты, связанные с употреблением наркотиков, каждый со своим контекстом употребления и противодействия наркотикам.

Среди стран, придерживающихся данного подхода, можно выделить Швейцарию, Нидерланды, Канаду.

Снижение вреда на микроуровне уходит корнями как в общественное здравоохранение, так и в более широкое движение за

нормализацию употребления наркотиков. Внедрение современных методов снижения вреда в населенных пунктах США является следствием более ранних европейских опытов с программами обмена шприцев и контролируемого распространения наркотиков.

Мероприятия по реализации антинаркотической политики могут оказывать влияние на распространенность, интенсивность и вред от употребления наркотиков. Понятие тотального вреда (total harm reduction) охватывает все эти явления. Политика **тотального снижения вреда** объединяет черты снижения потребления и снижения вреда на микроуровне [25]. Макровред определяется как произведение распространенности, интенсивности и среднего вреда от наркотиков. Тотальный вред определяется как сумма макровредов. Данный подход тесно связан с анализом рисков, экономического ущерба и т. д. На данный момент он существует, прежде всего, как теоретическая модель.

С практической точки зрения, можно выстроить антинаркотические мероприятия от более строгих (снижение употребления) к менее строгим (снижение вреда): полицейский надзор за уличной торговлей наркотиками, полицейские «облавы», ограничение оборота прекурсоров, арест за малые нарушения, в т.ч. выращивание марихуаны, тестирование на наркотики, вмешательство в частную жизнь наркозависимых, их обучение, наркосуды, лечение и реабилитация.

1.1.2 Использование информационных систем при реализации антинаркотической политики

Рассмотрим применение информационных систем при реализации антинаркотической политики на примере США, придерживающейся преимущественно политики снижения потребления, Европы, некоторые из стран которой придерживаются политики снижения

вреда, и ряда других стран.

Кратко опишем структуру государственного управления в области контроля за оборотом наркотиков и формирования наркополитики в США. Главным лицом, формирующим наркополитику в США (после Президента), является т.н. «Drug czar», глава ONDCP (Управления национальной политики по контролю за наркотиками), который отчитывается перед Конгрессом. К функциям ONDCP относится: создание координированной национальной стратегии противодействия наркотикам, создание центра управления антинаркотическими проектами и контроль антинаркотических бюджетов. Из правоохранительных органов главную роль в борьбе с контрабандой и употреблением наркотиков играет DEA (Управление по контролю за распространением наркотиков). В области исследований следует выделить как научные учреждения (Комитет по данным и исследованиям для политики в области незаконных наркотиков в составе Национальной академии наук), так и частные аналитические агентства (RAND и др.) [9; 32].

К категории мониторинговых систем можно отнести PDMP, созданную для идентификации лиц, злоупотребляющих наркосодержащими лекарствами, ограничения выписывания и продажи данной категории лекарств как государственными, так и частными клиниками и аптеками. В эту же категорию входят опросы, такие как ADAM (тестирование арестантов на наличие следов приема наркотиков), NSDUH, SAMHDA (общенациональные опросы), National Roadside Survey (тестирование водителей на алкоголь и наркотики).

NADDIS — это система сбора и индексирования данных, содержащая миллионы личных дел граждан, для доступа полиции и наркоаналитиков. ADNET — ИС Департамента обороны, функции которой — мониторинг, обеспечение мероприятий по снижению спроса на наркотики, обеспечение правоприменения как внутри, так и вне страны. Проект HDTA нацелен на координацию деятельности различных государственных агентств в т.н. «высокоинтенсивных

зонах», в частности на границе с Мексикой, с целью пресечения наркотрафика. STRIDE — система извлечения информации о результатах лабораторного анализа образцов наркотиков из материалов уголовных дел по наркопреступлениям.

В Европейском союзе функционирует децентрализованная организация European Monitoring Centre for Drugs and Drug Addiction (EMCDDA), которая ежегодно публикует отчёты о состоянии наркотических проблем в государствах-членах Евросоюза, собирает и предоставляет актуальные эмпирические данные для как для учёных, так и для политиков. Деятельность EMCDDA во многом основана на информационной сети Reitox, составленной из назначенных национальный институтов, ответственных за сбор данных и формирование отчетов по проблемам наркотиков и наркомании. Эти институты носят название «национальных точек фокуса» или «национальных пунктов наблюдения за наркоситуацией». Эта мониторинговая система охватывает 30 стран, причем сбор и обмен данными стандартизирован. В Миссии организации указано, что она призвана «помогать определять направление наркополитик Евросоюза, и разрабатывать подходящие рекомендации странам для организации лечения, превентивных мер и деятельности по уменьшению вреда».

Среди других примеров информационных систем можно привести The Exchange on Drug Demand Reduction Action (EDDRA) — база данных с доступом через Интернет, предоставляющая официальным ответственным лицам данные по программам снижения спроса на наркотики в Евросоюзе.

В Канаде функционирует Drug Treatment Court Information System, обеспечивающая деятельность т.н. «наркосудов», которые являются типичным примером реализации политики снижения вреда.

Таким образом, по результатам краткого обзора, обнаруживается, что в странах, взявших на вооружение политику снижения употребления, среди ИС превалируют мониторинговые системы, отчасти — аналитические, и на последнем месте — медицинские и

ресоциализирующие программы. С другой стороны, в странах, придерживающихся политики снижения вреда, есть как мониторинговые системы, так и системы, обеспечивающие различные социальные программы по снижению вреда.

1.1.3 Зарубежный опыт прогнозирования распространения наркомании

Рассмотрим подходы к прогнозированию наркоситуации с использованием эмпирических данных (социологических опросов и др.) в странах Евросоюза и США.

В монографии центра EMCDDA дается обзор методов моделирования т.н. «потребления наркотиков с высокой степенью риска» (high-risk drug use) или «проблемного употребления наркотиков» (problem drug use) в Европе [11]. Описан метод прогнозирования распространения наркомании с использованием ГИС [11; 40], разработанный в рамках программы DIPEP — Drug Incidence & Prevalence Estimation Program. Распространение наркомании моделируется как процесс, схожий с заражением в эпидемиологии. В этой парадигме ключевыми понятиями являются пути передачи инфекции, группы риска, распространенность инфекции, её географическое распределение, эпидемический цикл. Зоны распространения наркомании моделируются с помощью метода обратных взвешенных расстояний. Таким образом моделируется изменение частоты и распространенности потребления наркотиков в Великобритании. На примере города Глазго показана первоначальная концентрация «наркоэпидемии» внутри крупного города с последующим распространением в окрестные города в течение пятилетних циклов.

Для оценки численности наркозависимых предлагается применение множества динамических моделей, использующих доступные

статистические данные. В данном подходе основополагающим тезисом является необходимость уменьшить неопределенность, связанную со скрытым характером процессов, характеризующих наркоситуацию. Для этого применяются методы стратифицированных калибровочных выборок, вероятностное моделирование, процессы Пуассона, метод множественного захвата и перезахвата, оценка по связи распространенности тяжелых наркотиков и частоте аквизитивных преступлений, марковские модели. Применение множества методов и моделей для оценки одной целевой популяции позволяет более точно судить о трендах и о порядке численности наркозависимых.

В качестве альтернативной модели прогнозирования наркоситуации предлагается модель с множеством индикаторов. В данной модели строится гипотеза о существовании связей типа «причина-следствие» между индикаторами. Выделяется три группы индикаторов:

- социальные;
- правовые;
- медицинские.

Авторами отмечается универсальность данной модели, возможность её использования для исследования различных социально-экономических сценариев, адаптация модели к исходным данным. Практическое применение данной модели позволило получить неожиданные и континтуитивные результаты.

Отдельно европейскими исследователями выделяется класс динамических моделей. Термин «динамическая модель» охватывает методы системного анализа и моделирования временных рядов, применяемые для оценки распространенности наркопотребления. Данные методы должны обеспечивать не только дескриптивный анализ, но и моделировать процессы, лежащие в основе наркоситуации. Динамические модели в отличие от статических

описывают процессы во времени. Например, исследуемым процессом может быть изменение состояния наркотребителей в модели «поимка-мечание-повторная поимка». Модели поимки-повторной поимки делятся на два класса: с открытой популяцией и с закрытой популяцией. В моделях с закрытой популяцией предполагается, что популяция не изменяется на протяжении исследуемого периода, а в моделях с открытой популяцией учитывается прибыль и убыль особей популяции. Модели с закрытой популяцией проще и предназначены скорее для анализа коротких временных периодов, но при достаточно большом временном отрезке они, как правило, необъективны, и в таких случаях применяются более сложные модели с открытой популяцией.

В отчете Национального Института правосудия США [3; 38] приводится анализ наркоситуации в 23 крупнейших городах США. Датасет состоит из данных уриноанализа (на предмет позитивного тестирования на кокаин, марихуану, опиаты) арестантов в разрезе по типам преступлений. Эта информация, в частности, была использована для установления и объяснения связи между наркотиками и преступностью. На основании снижения частоты позитивных тестов на кокаин среди мужчин 15-20 лет строится гипотеза о продолжительном снижении употребления кокaina по мере того, как представители данной возрастной группы стареют.

В работе Дж. Колкинса [2] производится оценка эластичности спроса на кокаин и героин на основании датасетов DUF и STRIDE, предоставленных Национальным институтом правосудия США и Управлением по борьбе с наркотиками. Автор пытается ответить на вопрос: насколько изменится потребление при увеличении цен? В качестве метода математического анализа использованы дифференциальные уравнения. Анализ показывает высокую эластичность спроса. В качестве выводов автор связывает подъем употребления кокaina и героина в 1980-х с существенным снижением цен на них. Как следствие, для проведения антинаркотической политики могут быть полезными мероприятия, приводящие к

повышению цен на наркотики.

Датасеты ADAM и ADAM II — результат опросов, проводимых с 1997 по 2013 годы в 35 (ADAM) и 10 (ADAM II) округах США по методу интервью с вопросами об употреблении и участии в рынке наркотиков и уриноанализа арестантов в течение 48 часов с момента ареста на предмет обнаружения наркотических веществ[4; 15]. Датасеты ADAM являются одними из важнейших источников информации для лиц, определяющих наркотическую политику в США, ввиду того, что они охватывают популяцию, мало представленную в иных опросах. С точки зрения аналитических инструментов, для оценки трендов и выявления годовых циклов применяются статистические методы, в частности, модель логистической регрессии. С помощью данной модели оценивается вероятность положительного тестирования на содержание в моче определенного наркотика в зависимости от типа преступления, сезона и года. В качестве вспомогательных методов используются преобразования Фурье. Существенное внимание уделяется заполнению пропусков во временных рядах, определению зависимых и независимых переменных.

Выводы по главе 1

В главе был проведен обзор видов наркополитик, в общем виде описывающих концептуальные основы и практические меры по урегулированию наркоситуации. Исследована связь между постулируемыми в развитых странах стратегиями антинаркотической политики и средствами их информационной поддержки. Обнаружено, что в целом имеют тенденции к стандартизации в антинаркотической аналитике: от проведения аналогичных опросов с определённой периодичностью до создания филиалов крупных исследовательских центров во всех регионах страны, отчитывающихся в центральный хаб с помощью стандартизованных отчетов.

В построении прогнозов имеются нерешенные проблемы с неоднородностью исходных данных, требующие для своего решения продвинутых методов статистического контроля данных и заполнения пробелов. Однако, сравнительная молодость дисциплины способствует большому разнообразию применяемых методов, большая часть которых заимствованы из других дисциплин, таких как эпидемиология, зоология, эконометрика и др.

ГЛАВА 2

Обзор методов прогнозирования временных рядов

Введение

Цель данной главы – анализ применимости новых моделей прогнозирования к задачам государственного управления.

Для достижения указанной цели необходимо решить следующие задачи:

1. обзор литературы по теме прогнозирования временных рядов, и в особенности, интеллектуального прогнозирования
2. выявление сфер применения новых методов прогнозирования, степени их теоретической проработки и практических результатов их использования.
3. обзор программных пакетов, позволяющих реализовывать методы нечеткого прогнозирования.
4. обзор литературы по теме нечеткого прогнозирования социально-экономических процессов.

Работа выполняется по заказу Санкт-Петербургского информационно-аналитического центра.

2.1 Задачи прогнозирования и интеллектуальный анализ временных рядов

Количество информации, порождаемой человечеством, непрерывно растет [12]. В связи с этим все важнее становится умение анализировать данные, извлекать из них знания, чтобы превратить в движущую силу социально-экономического развития.

Одним из методов анализа данных является прогнозирование. Прогнозная аналитика использует методы математического моделирования, искусственного интеллекта и теории игр, анализирует текущие и исторические факты для составления предсказаний о будущих событиях. Фиксируются связи между разными факторами, на основании их выявления идентифицируются риски и возможности. Задача прогнозирования лежит в основе финансового планирования в экономике, управления объемами производства, принятия решений в сфере управления социальными системами.

На сегодняшний день существует множество моделей прогнозирования временных рядов [51]: регрессионные и авторегрессионные, нейросетевые, экспоненциального сглаживания, на базе цепей Маркова, классификационные и др. (Рис. 2.1). Прогнозирование в государственном управлении осуществляется посредством разных методов: экстраполяции, факторного прогнозирования, модельного прогнозирования, экспертного оценивания. Органы государственной власти в последние годы все чаще обращаются к методу экспертного прогнозирования [44]. В этом случае эксперт дает прогноз, опираясь на опыт, аналогии, интуицию. В то же время интуитивная природа данного метода заставляет при его использовании полагаться лишь на квалификацию и репутацию эксперта, тем самым внося дополнительную неопределенность в процесс прогнозирования. Существуют исследования, показывающие, что прогнозы экспертов по точности уступают прогнозам на основе

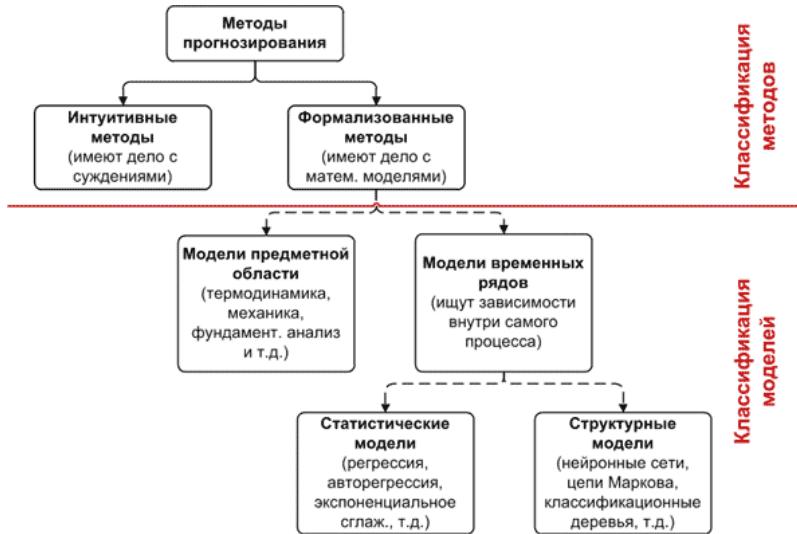


Рис. 2.1: классификация моделей и методов прогнозирования [50]

моделирования. П. Майл на примере 20 прогнозов из области клинической диагностики [26] и Дж. Сойер на примере 45 прогнозов в социальной сфере [34] демонстрируют тот факт, что экспертные оценки ни в одном из перечисленных случаев не были существенно точнее статистических моделей прогнозирования. Кроме того, наблюдается нехватка специалистов по работе с данными [10]. Все это побуждает задуматься о других подходах к анализу данных.

В последние два десятилетия активно развивается направление прогнозирования, связанное с интеллектуальным анализом временных рядов [52]. Основными целями этого направления являются, во-первых, анализ и моделирование процессов, характеризующихся высокой степенью неопределенности, в том числе в областях, слабо подверженных формализации, во-вторых, повышение уровня интеллектуальной поддержки современных специалистов, и, в-третьих, выявление скрытых закономерностей и извлечение новых знаний из временных рядов.

2.2 Сфера применения методов интеллектуального анализа временных рядов

В основе новых методов анализа временных рядов лежит нечеткая модель временного ряда. Теория нечетких множеств была впервые изложена Л. Заде в 1965 г. [42]. В 1973 г. предложена теория нечеткой логики [41]. Вскоре после этого теория стала популярна. В конце 1980-х гг. это направление стало бурно развиваться в Японии. Нечеткое управление стало применяться в промышленности, железнодорожном транспорте и разработке потребительской техники.

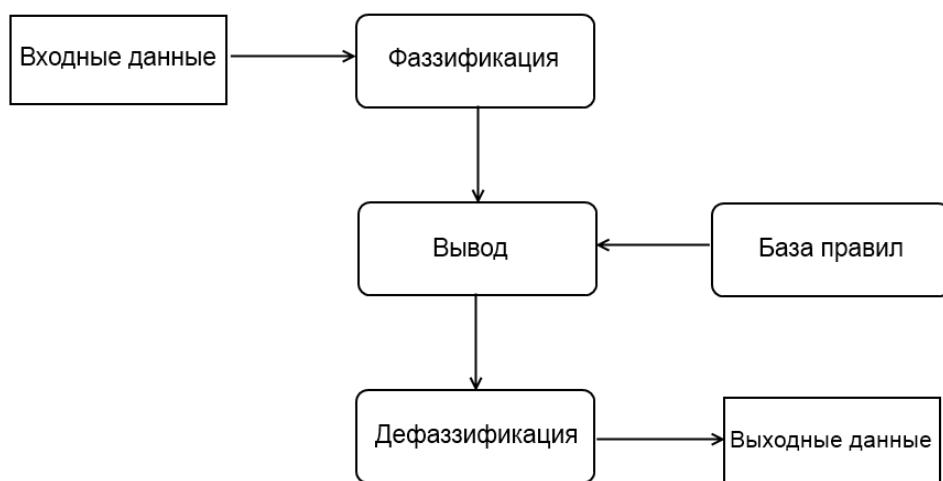


Рис. 2.2: Система нечеткого вывода

Основной идеей нечеткой логики является многозначность. Высказывание может иметь любое истинностное значение в промежутке от 0 до 1. Таким образом, воспроизводится неточность человеческого мышления.

Модели статических и динамических систем, построение, использование и анализ которых базируется на положениях теории нечетких множеств и нечеткой логики, называют нечеткими моделями или нечеткими системами. Целью нечеткого моделирования сложных явлений является приближенное описание зависимости.

В основе нечетких производственных моделей лежит совокупность нечетких правил «ЕСЛИ, ТО», описывающих зависимости между нечеткими переменными предметной области, композиционное правило вывода и способ вычисления значений нечетких переменных (способ нечеткого вывода).

Модель описания поведения систем на естественном (или близком к естественному) языке, в виде приближенных рассуждений в теории нечетких множеств и нечеткой логики, основанная на композиционном правиле вывода, называется системой нечеткого логического вывода.

В систему нечеткого логического вывода входят следующие объекты (рис. 2.2):

1. совокупность нечетких производственных правил (база правил);
2. блок фазификации;
3. блок дефазификации;
4. блок вывода.

На основании общей модели, приведенной выше, создаются контроллеры для управления преимущественно техническими объектами.

В начале 1990-х гг. была предложена теория нечетких временных рядов [36]. Нечетким временным рядом называют упорядоченную в равноотстоящие моменты времени последовательность наблюдений над некоторым процессом, состояния которого изменяются во времени, если значение состояния процесса в данный момент времени может быть выражено с помощью нечеткой метки. Нечеткая метка может быть сформирована непосредственно экспертом или получена на основе некоторого преобразования исходного временного ряда [52].

В условиях, когда моделируемым процессам присуща высокая степень неопределенности, методы прогнозирования на основе нечетких моделей временных рядов позволяют выработать наиболее

адекватную оценку будущих изменений в социально-экономических системах.

Ряд исследований [5; 27; 35], в которых изучается прогнозирование с помощью моделей нечетких временных рядов, демонстрирует положительные результаты. Точность прогнозирования улучшается за счет использования генетических алгоритмов для настройки параметров нечетких моделей прогнозирования, изменения количества нечетких множеств, используемых для описания временного ряда, использование достаточного числа производственных правил, модификации интервалов, на которые разбивается исходный ряд и др.

В то же время есть примеры неудовлетворительного применения нечеткой логики в прогнозировании временных рядов [14]. Причиной тому неопределенность при создании набора производственных правил, необходимость адекватного выбора переменных, включаемых в модель, сложность построения модели ввиду новизны, малой изученности данной темы и большой вариативности при построении модели.

2.3 Программные средства интеллектуального прогнозирования

Для использования прогнозирования в принятии решений целесообразным может оказаться интеграция этих моделей с экспертной системой или системой поддержки принятия решений. В этом случае нужно задуматься о программной реализации нечетко-логических моделей. Попробуем рассмотреть существующие библиотеки, в той или иной мере охватывающие тему нечеткой логики, нечеткого вывода и т.д.

Fuzzy Logic Toolbox для *Matlab*. Хотя данный продукт имеет все необходимые возможности, для наших задач он не подходит, т.к. функционирует на собственном языке программирования и зарегистрирован под проприетарной лицензией.

Fuzzy Logic Toolkit для *Octave*. Аналог *Fuzzy Logic Toolbox*, относящийся к категории программного обеспечения с открытым исходным кодом. Распространяется по лицензии GPLv2, ограничивающей его использование в коммерческой деятельности.

jFuzzyLogic. Позиционируется как наиболее полная библиотека по нечеткой логике и стандарт де-факто в исследовательской и прикладной деятельности. Интересен своей имплементацией языка нечеткого контроля (FCL), стандартизирующего разработку систем нечеткого вывода. Данный язык соответствует стандарту IEC 61131-7. Тем не менее, обладает таким недостатком как сложность в использовании.

fuzzylite. Еще одна альтернатива, отличительными чертами которой являются открытый исходный код, наиболее свободная лицензия, простота в использовании, наличие различных алгоритмов вывода, лингвистических переменных, операторов нечеткой логики. Также планируются дополнения в виде новых нечетких контроллеров, алгоритмов кластеризации и адаптивных нейро-нечетких систем

вывода. Уникальной особенностью является импорт из трёх и экспорт в шесть различных языков описания нечетких систем (FLL, FLD, FIS, FCL и др.). Отдельным преимуществом является доступность библиотеки в исполнении на трех языках программирования: C++, Java, Python. Кроме того, автором является доктор наук (PhD) в области искусственного интеллекта Х. Рада-Вилела [31].

frbs Пакет *frbs* написан на языке статистического программирования R. Это позволяет использовать его в тесной интеграции с другими инструментами и методами R, предназначенными для работы с временными рядами, прогнозированием, доступа к различным источникам данным, инструментами как классической математической статистики, так и с более новыми методами машинного обучения. Пакет разработан группой авторов из Гранадского университета: аспирантов и докторов наук. Помимо этого, в последней версии *frbs* добавлена возможность использовать универсальный язык для описания нечетких моделей *frbsPMML*, являющийся диалектом PMML, языка разметки предиктивных моделей. PMML — основанный на XML язык, предлагающий стандарт для описания моделей, сгенерированных алгоритмами data-майнинга и машинного обучения. Таким образом, реализована возможность импорта и экспорта нечетких моделей на языке PMML.

FuzzyEngine, *funzy* и др. либо неполны, либо заточены под чисто учебные цели и поэтому непригодны для наших задач.

По итогам обзора можно выделить *fuzzylite* и *frbs* в качестве основных претендентов к использованию. Однако, в ходе дальнейшего исследования было обнаружено, что пакет *fuzzylite* ориентирован преимущественно на программирование нечетких контроллеров, а не на аналитическую работу с временными рядами. В то время, как пакет *frbs* заточен под работу с временными рядами (регрессия) и категориальными переменными (классификация). Учитывая специфику задачи прогнозирования временных рядов по социально

Таблица 2.1: Сравнение программного обеспечения, реализующего методы нечеткой логики

Название пакета	Лицензия	Юзабилити	Платформа
Fuzzy Logic Toolbox	Проприетарная	Хорошее	Win, Linux, Mac
Fuzzy Logic Toolkit	Свободная	Хорошее	Win, Unix, Unix-like
jFuzzyLogic	Свободная	Удовл.	Java
fuzzyLite	Свободная	Отличное	Java, C++, Python
frbs	Свободная	Хорошее	R

значимой тематике, целесообразно выбрать пакет `frbs` в качестве основного инструмента для прогнозирования с помощью нечеткой логики.

2.4 Обзор литературы по нечеткому социальному прогнозированию

Первые случаи применения теории нечеткой логики связаны с созданием контроллеров для управления техническими системами. Так, в 1987 году в японском городе Сендай была открыта железнодорожная линия, на пассажирских поездах которой были установлены нечеткие системы, регулирующие ускорение, торможение и остановку поезда. С тех пор применение нечетких контроллеров значительно возросло, затрагивая, к примеру, производство бытовой техники (кондиционеры, посудомоечные машины, пылесосы), автомобилестроение (повышение энергоэффективности двигателей), распознавание образов и др.

Однако, перспективы применения нечеткой логики для моделирования и прогнозирования социально-экономических процессов остаются малоизученными. Данный небольшой обзор рассматривает работы по нечеткому прогнозированию с использованием данных государственной статистики.

В статье Н. А. Абдулаевой [43] исследуется уровень бедности в Азербайджане в зависимости от доходов населения, коэффициента безработицы, уровня инфляции и прожиточного минимума, а такжедается прогноз уровня бедности на три года. Для прогноза уровня бедности предлагается нечеткая регрессионная модель. Утверждается превосходство нечеткого регрессионного моделирования перед классическим регрессионным моделированием при прогнозировании уровня бедности, так как оно обеспечивает большую точность результатов.

Работа М. Г. Мамедовой и З. Г. Джабраиловой [46] показывает целесообразность применения нечеткой логики в моделировании демографических аспектов рынка труда на примере задачи прогнозирования численности экономически активного населения.

Предложена методика прогнозирования численности экономически активного населения с использованием модели нечетких временных рядов. Установленный горизонт прогнозирования — три года. Кроме того, метод применен для долгосрочного прогнозирования (до 2025 года). Проведен сравнительный анализ и интерпретация полученных прогнозных данных для экономически активного, общего и трудоспособного населения, позволившие определить вероятное перспективное состояния рынка труда.

В статье П. С. Пака и Г. Кима [30] предлагается основанный на теории нечеткости метод прогнозирования численности и состава населения (пол, возраст, район проживания) на долгосрочный период. На первом этапе представлена нечеткая модель, состоящая из набора правил типа «если-то», для оценки общего роста населения в 402 районах региона Кансай. Антецеденты и консеквенты правил построены с помощью реальных данных, и эти правила составляют нечеткие высказывания и регрессионные модели, соответственно. На втором этапе описан метод для оценки социального роста по полу и возрасту в каждом районе на основе метода нечеткой кластеризации. Метод ориентирован на оценку долгосрочных социоэкономических изменений в миграции населения. Точность результатов предложенной модели сравнивается с традиционной регрессионной моделью, сравнение оказывается в пользу предлагаемой нечеткой модели.

Исследование А. Сасу [33] предлагает методику прогнозирования демографических процессов на основе теории нечетких временных рядов. Своебразная особенность модели заключается в её способности прогнозировать требуемый показатель, используя неполные, нечеткие входные данные. Автор утверждает возможность установки практически неограниченного горизонта прогнозирования для данной модели.

Суммируя информацию из рассмотренных источников, заметим некоторые тенденции:

- Более высокая точность нечетких моделей в сравнении с классическими моделями для социального прогнозирования.
- Широкий диапазон горизонтов прогнозирования: возможность краткосрочного, среднесрочного, долгосрочного прогнозирования.
- Разнообразие сценариев применения нечетких моделей для социально-экономических прогнозов.

2.5 Теорема о нечеткой аппроксимации

В 1994 году Б. Коско доказал теорему о нечеткой аппроксимации (Fuzzy Approximation Theorem) [24], согласно которой, любая математическая система может быть аппроксимирована системой на нечеткой логике. Следовательно, с помощью естественно-языковых высказываний «если-то», с последующей их формализацией средствами теории нечетких множеств, можно сколько угодно точно отразить произвольную взаимосвязь «входы-выход» без использования сложного аппарата дифференциального и интегрального исчислений, традиционно применяемого в управлении и идентификации. Практические успехи нечеткого управления получили теоретическое обоснование.

Теорема 2.5.1 (Теорема о нечеткой аппроксимации). *Аддитивная нечеткая система F равномерно аппроксимирует $f : X \rightarrow Y$, если множество X компактно и f непрерывна.*

Выводы по главе 2

В работе был проведен обзор литературы, описывающей практическое применение нечетких моделей для социально-экономического прогнозирования. Характерна тенденция положительных результатов, улучшения точности в сравнении с классическими методами.

Приведена теорема, доказывающая, что нечеткие системы являются универсальными аппроксиматорами. Данное утверждение позволяет с уверенностью смотреть на будущее применения нечетких моделей.

На основе обзора методов прогнозирования, анализа динамики применения этих методов в прошедшие десятилетия, обзора программных библиотек, реализующих принципы нечеткой логики, сделаны следующие выводы:

1. Увеличение размерности и доли неопределенности в наблюдаемых данных приводят к необходимости искать новые методы анализа и прогнозирования данных.
2. Интеллектуальные методы анализа данных дают противоречивые результаты в силу малой изученности и проблемности данной области. Тем не менее, наблюдается положительная тенденция в сторону уточнения моделей.
3. Существующие свободно распространяемые программные библиотеки позволяют реализовывать системы нечеткого вывода.

Рассматриваемые нечеткие модели прогнозирования временных рядов, помимо ликвидации неопределенности и возможности работать со слабо формализуемыми входными данными, обладают преимуществом простоты для пользователя благодаря использованию продукционной модели знаний. Данная модель оперирует правилами, написанными естественным языком, что позволяет пользоваться и

совершенствовать её специалисту в предметной области, а не только разработчику метода и программной реализации прогнозирования. Это положительно сказывается на адекватности прогноза.

Выбранный подход к прогнозированию с использованием методов интеллектуального анализа временных рядов и нечеткой логики имеет потенциал для того, чтобы эффективно решить задачу государственного управления и оценки стратегического развития региона. Для этого соответствующее программное обеспечение должно быть внедрено в действующие информационно-аналитические системы.

Дальнейшая работа будет связана с адаптацией какой-либо нечеткой модели и её программной реализацией для опытного оценивания возможностей интеллектуального анализа временных рядов.

ГЛАВА 3

Разработка модели прогнозирования численности наркозависимых в Санкт-Петербурге на основе нечеткой модели с многими переменными

Введение

Проблема эффективного прогнозирования социально-экономических процессов в настоящее является исключительно актуальной задачей. Особенную сложность прогнозирование приобретает в задачах государственного управления. В силу того, что государственные планы и прогнозы затрагивают жизнь большого числа людей, возрастает цена ошибки. Поэтому необходимо минимизировать риски. Одним из способов минимизации рисков является научный подход в прогнозировании возможных сценариев развития общества.

Социально-экономическим процессам свойственны неопределенность, наличие скрытых факторов влияния, слабая предсказуемость. Системы, характеризующиеся такими процессами, называют нелинейными, хаотическими, случайными, неопределенными. Существуют математические теории, призванные «уточнить» неточности: теория детерминированного хаоса, теория вероятностей, нечеткая логика и др. В данной работе рассматриваются прикладные аспекты социального прогнозирования с помощью нечеткой логики.

Цель данной главы: разработка модели прогнозирования доли наркозависимых в населении Санкт-Петербурга с помощью индикаторов наркотизации.

Поставленные задачи для достижения цели главы:

- опытная проверка результатов одномерных нечетко-логических моделей и сравнение с моделями на основе теории нечетких временных рядов.

- опытная оценка результатов многомерных нечетко-логических моделей, действующих в качестве предикторов социально-экономические показатели.

3.1 Вычислительный эксперимент нечеткой авторегрессии

3.1.1 Используемые программные средства

В настоящей статье программная реализация методов нечеткого прогнозирования базируется на библиотеке «frbs» языка статистического программирования R. Библиотека разработана докторами философии и аспирантами Гранадского университета. В библиотеке реализовано более 15 методов нечеткой классификации и регрессии. В библиотеке рассматриваются системы с многими входами и единственным выходом (MISO) с данными в виде вещественных чисел.

CHB — система нечеткого вывода.

1. CHB, основанные на разбиении области определения.

- Метод Ванга и Менделя. Предназначен для решения задач регрессии [39].
- Метод Чи. Предназначен для решения задач классификации [6].
- Взвешенный метод Ишибучи. Предназначен для решения задач классификации [17].

2. CHB, основанные на искусственных нейронных сетях.

- Адаптивная нечеткая система вывода с использованием нейросети. Предназначена для решения задач регрессии [21].
- Гибридная нечеткая система вывода с использованием нейросети. Предназначена для решения задач регрессии [23].

3. CHB, основанные на алгоритмах кластеризации.

- Субтрактивная кластеризация и метод нечеткой кластеризации с-средних. Предназначена для решения задач регрессии [7].
- Динамическая эволюционирующая система вывода с использованием нейросети. Предназначена для решения задач регрессии [22].

4. СНВ, основанные на генетических алгоритмах.

- Метод Трифта. Предназначен для решения задач регрессии [37].
- Генетическая нечеткая система для обучения нечетких правил, основанная на технологии MOGUL. Предназначен для решения задач регрессии [8].
- Метод Ишибути, основанный на генетическом кооперативно-конкурентном обучении. Предназначен для решения задач классификации [18].
- Метод Ишибути, основанный на гибридизации генетического кооперативно-конкурентного обучения и Питтсбургского метода. Предназначен для решения задач классификации [19].
- Структурный обучающий алгоритм на нечеткой среде. Предназначен для решения задач классификации [13].
- Генетический алгоритм для латеральной настройки и выбора правил лингвистической нечеткой системы. Предназначен для решения задач регрессии [1].

5. СНВ, основанные на методе градиентного спуска.

- СНВ с использованием эвристик и градиентного спуска. Предназначена для решения задач регрессии [20].
- Правила нечеткого вывода по методу спуска. Предназначен для решения задач регрессии [29].

3.1.2 Методика проведения эксперимента

Основными направлениями нечеткого прогнозного моделирования, описанными в литературе, являются теория нечеткой логики [41] и теория нечетких временных рядов [36]. При этом, для прогнозирования по методу нечетких временных рядов исследователям удалось добиться существенного снижения погрешности прогнозов даже для авторегрессионного случая.

Известно [16], что до некоторых пределов математика сводима к логике. Это и будет исходной точкой нашего эксперимента. Попробуем сравнить доказанные в своей успешности методы теории нечетких временных рядов (основанные на математической теории нечетких множеств) с методами нечеткой логики, реализованными в пакете «frbs».

Объектом прогнозирования являются данные по поступившим в университет Алабамы абитуриентам за период 1971-1992 гг. (см. Рис. 3.1). Именно они были использованы в статье, впервые описавшей теорию нечетких временных рядов и с тех пор являются основой для сравнения моделей.

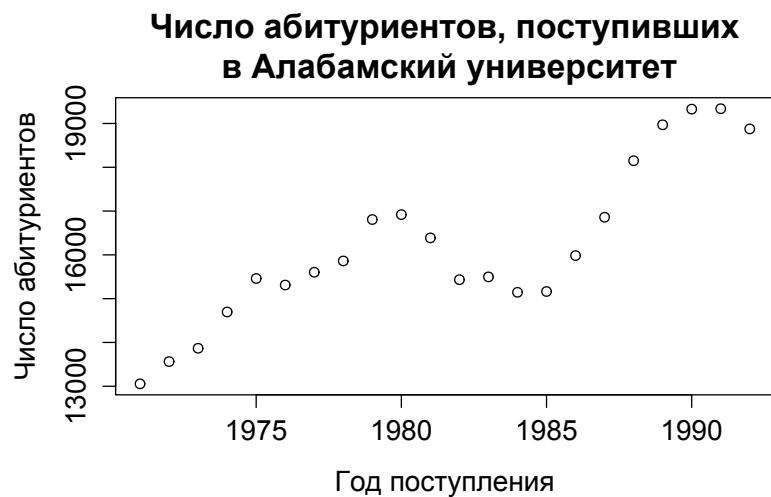


Рис. 3.1: Число поступивших в Алабамский университет, 1971-1992.

По итогам экспресс-теста регрессионных моделей решено было выбрать модель авторов Л.-Х. Ванга и Дж. М. Менделя, основанную на разбиении области определения [39]. Критериями отбора были точность прогноза на тестовых данных и сложность вычислений. Модель демонстрирует хорошую точность и низкую вычислительную сложность.

Алгоритм Ванга и Менделя состоит из пяти шагов:

1. Разбиение областей определения входных и выходных численных данных на нечеткие интервалы.
2. Генерация нечетких правил на основе предоставленных данных.
3. Присвоение степени каждому из сгенерированных правил с целью разрешения противоречий между правилами.
4. Создание комбинированной нечеткой базы правил, основанной одновременно на 1) автоматически сгенерированных правилах и 2) правилах на естественном языке, предложенных экспертами.
5. Определение отображения входного пространства на выходное, основываясь на комбинированной базе правил, с помощью процедуры дефазификации.

Доказана возможность предлагаемого отображения аппроксимировать любую непрерывную функцию вещественного переменного на компактном множестве с произвольной точностью. Примеры нечетких правил и функций принадлежности, генерируемых моделью, см. на Табл. 3.1 и Рис. 3.2, соответственно.

На выходе модели — численность поступивших в университет Алабамы за период 1971-1992 гг. (t). На входе — эта же переменная, но со сдвигом на один (t_{minus1}) и два (t_{minus2}) года назад, соответственно. Таким образом реализована импровизированная авторегрессия по переменной t .

Таблица 3.1: Пример нечетких правил.

1	IF tminus2 is v.1_a.1 and tminus1 is v.2_a.1 THEN t is c.1
2	IF tminus2 is v.1_a.7 and tminus1 is v.2_a.7 THEN t is c.5
...	
17	IF tminus2 is v.1_a.9 and tminus1 is v.2_a.6 THEN t is c.7

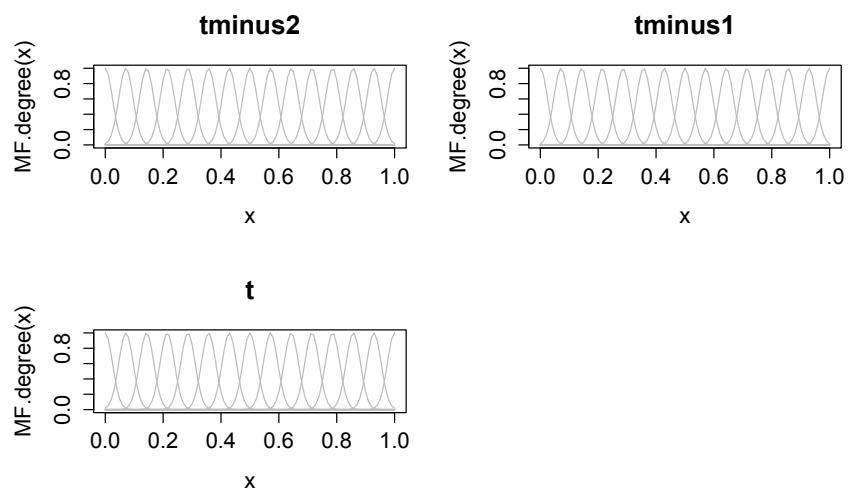


Рис. 3.2: Пример функций принадлежности для лингвистических переменных.

Аналогичным образом в тестовых примерах библиотеки «frbs» прогнозировались хаотические временные ряды Маки-Гласса, при этом, однако, сдвиг происходил на 4 и 8 позиций назад и общее число наблюдений составляло 300, а не 22, как в случае с временным рядом поступивших в университет Алабамы.

Общее множество данных разбивалось на два подмножества — данные для обучения модели и данные для сравнения результатов работы модели — в зависимости от величины горизонта прогнозирования.

3.1.3 Результаты эксперимента

В эксперименте использовались два значения горизонта прогнозирования — 5 лет (см. Рис. 3.3) и 3 года (см. Рис. 3.4). Заметим, что в обоих случаях симуляция в области данных для обучения сработала достаточно хорошо, несмотря на крайне малый объем выборки.



Рис. 3.3: Прогноз числа поступивших в Алабамский университет, горизонт 5 лет.

Прогноз числа поступивших, горизонт 3 года.

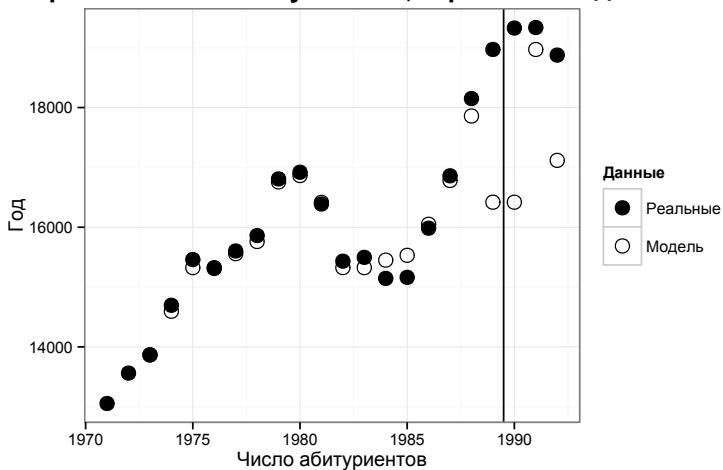


Рис. 3.4: Прогноз числа поступивших в Алабамский университет, горизонт 3 года.

В части прогноза же результаты неудовлетворительные. Это подтверждается и оценкой погрешности прогнозирования (см. Табл. 3.2). Погрешность гораздо выше, чем при авторегрессии методами нечетких временных рядов. С другой стороны, точность увеличивается при уменьшении горизонта прогнозирования. Это говорит о том, что модель адаптируется при изменении параметров и входных данных.

Таблица 3.2: Оценка точности результатов прогнозирования.

Горизонт	MSE	RMSE	SMAPE
5	6.750612e+06	2.598194e+03	3.674307e+00
3	3.897863e+06	1.974301e+03	2.330708e+00

Основной причиной низкой точности прогноза, по нашему предположению, является недостаточность исходных данных. Расширение доступной для модели информации с целью повышения качества прогноза может происходить по трем направлениям:

- Увеличение объема выборки (периода, который охватывает временной ряд).
- Увеличение количества производственных правил либо изменение существующих правил.
- Увеличение количества либо изменение структуры входных показателей, отбор наиболее релевантных задач показателей.

В процессе вычислительного эксперимента автор столкнулся с техническими трудностями:

- Невозможность оценки всех моделей из библиотеки «frbs» с варьированием всех параметров моделей ввиду нехватки вычислительных мощностей.
- Сложность доступа к данным.

3.2 Разработка модели прогнозирования численности наркозависимых в Санкт-Петербурге на основе нечеткой модели с многими переменными

В связи с расширением масштабов незаконного оборота и немедицинского потребления наркотиков в России и в ответ на усиление таких негативных тенденций, связанных с наркоситуацией, как устойчивое сокращение численности населения России, в том числе уменьшение численности молодого трудоспособного населения, была принята Стратегия государственной антинаркотической политики Российской Федерации до 2020 года [49].

В качестве генеральной цели Стратегии постулируется «существенное сокращение незаконного распространения и немедицинского потребления наркотиков». Таким образом, можно утверждать, что Российская Федерация придерживается

преимущественно политики снижения потребления. Это и будет предпосылкой к дальнейшим положениям.

3.2.1 Теоретико-методологические основы анализа наркоситуации

По мнению теоретиков анализа наркоситуации [45], с позиций социального управления приобщение части населения к употреблению психоактивных веществ целесообразно рассматривать как риск. С точки зрения управления, риск — это событие или группа однородных случайных событий, которому присущи два основных свойства — вероятность и ущерб.

Вероятность — признак, означающий возможность с той или иной степенью точности рассчитать и прогнозировать частоту наступления неблагоприятного события (в данном случае — акта потребления наркотиков) при наличии достаточного количества данных и результатов наблюдений. Вместе с тем для риска всегда характерна случайность, непредсказуемость наступления события, означающая невозможность точно определить время и место его возникновения. А поскольку риск выбора потребления наркотиков в современном российском обществе сохраняется, то, с позиций социального управления, управление антинаркотической деятельностью, а через него — и организация влияния на наркоситуацию и контроля за проявляющимися в ее рамках тенденциями — это прежде всего управление рисками.

Одним из важнейших вопросов исследования наркоситуации выступает анализ и оценка факторов риска, прямо или косвенно влияющих на тенденции наркотизации. Факторы риска можно разделить на две группы:

1. факторы, имманентно присущие субъекту наркотизации:

- наследственные

- гендер
- возраст
- социальный статус
- социокультурные факторы

2. внешние факторы:

- доступность наркотических и иных психоактивных веществ
- информационные факторы

Помимо непосредственно факторов наркотизации для анализа целесообразно рассмотреть **индикаторы воздействия**, т.е. статистические и эпидемиологические показатели, отражающие криминальную ситуацию, социальное положение и состояние здоровья населения. Основными индикаторами воздействия определены следующие:

1. Социально-демографические и экономические:

- число и доля лиц, попробовавших наркотики хотя бы раз в жизни;
- процент лиц определенной возрастной группы, употребляющих наркотики;
- спрос на наркотики среди населения;
- спектр употребляемых наркотиков;
- средняя продолжительность и качество жизни населения;
- сумма социально-экономического ущерба от наркотиков (социальная стоимость употребления наркотиков).

2. Медицинские:

- заболеваемость наркоманией;

- количество отравлений наркотиками;
- процент лиц определенной возрастной группы, зависимых от наркотиков;
- смертность, связанная с наркотиками;
- доля повторных обращений в медицинскую службу больных наркоманией;
- продолжительность и качество жизни лиц, употребляющих наркотики;
- распространенность ВИЧ и гепатитов среди потребителей наркотиков.

3. Криминальные:

- число задержанных правоохранительными органами лиц с положительным результатом освидетельствования на состояние наркотической интоксикации;
- индикатор доступности наркотиков среди населения;
- число и доля лиц, осужденных за преступления, связанные с наркотиками;
- рецидивная преступность, связанная с наркотиками.

Не все перечисленные выше индикаторы, необходимые для полноценного и качественного мониторинга наркоситуации, собираются статистическими службами и используются при проведении практических исследований, однако, попытаемся оценить эффективность прогнозирования наркоситуации с помощью той их части, которая доступна для исследования.

3.2.2 Использование индикаторов мониторинга наркоситуации для прогнозирования доли наркозависимых в населении Санкт-Петербурга

Из числа доступных для анализа показателей, хранящихся в Интегрированной системе информационно-аналитического обеспечения деятельности исполнительных органов государственной власти Санкт-Петербурга, были отобраны три комбинации для составления мультифакторного прогноза показателя «Состоит на учете больных с диагнозом «наркомания», на 100 тыс. населения», который, на наш взгляд, отражает критерии выполнения задач Стратегии антинаркотической политики в области снижения потребления.

Таблица 3.3: Оценка точности результатов прогнозирования.

№ набора	MSE	RMSE	SMAPE
(1-мерный)	6.750612e+06	2.598194e+03	3.674307e+00
1	1099	33.14	3.733
2	6460	80.38	11.14
3	9805	99.02	14.44

- Набор показателей №1. Независимые переменные:
 - Численность безработных, всего;
 - Преступления связанные с незаконным оборотом наркотиков, зарегистрировано;
 - Состоит на учете больных с диагнозом «наркомания», на 100 тыс. населения (ретроспективные данные).
- Набор показателей №2. Независимые переменные:

- В состоянии наркотического опьянения;
 - Число отравлений наркотическими веществами, Всего, Все население от 0 до 99 лет всего;
 - Состоит на учете больных с диагнозом «наркомания», на 100 тыс. населения (ретроспективные данные).
- Набор показателей №3. Независимые переменные:
 - Число лиц, осужденных за (ст. 228-233 УК РФ), возрастная структура осужденных 14-17 лет;
 - Число лиц, осужденных за (ст. 228-233 УК РФ), возрастная структура осужденных 18-24 лет;
 - Состоит на учете больных с диагнозом «наркомания», на 100 тыс. населения (ретроспективные данные).

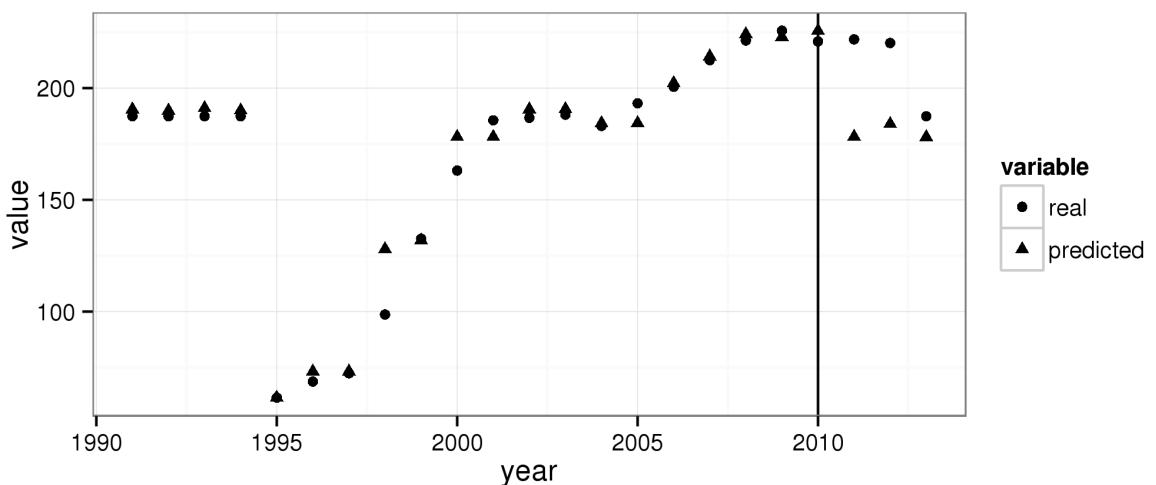


Рис. 3.5: Прогноз с набором показателей №1, горизонт 3 года.

С точки зрения точности прогноза, при использовании набора показателей №1 наблюдается существенное улучшение относительно одномерного варианта. Однако, требуемый уровень точности на данный момент не достигнут.

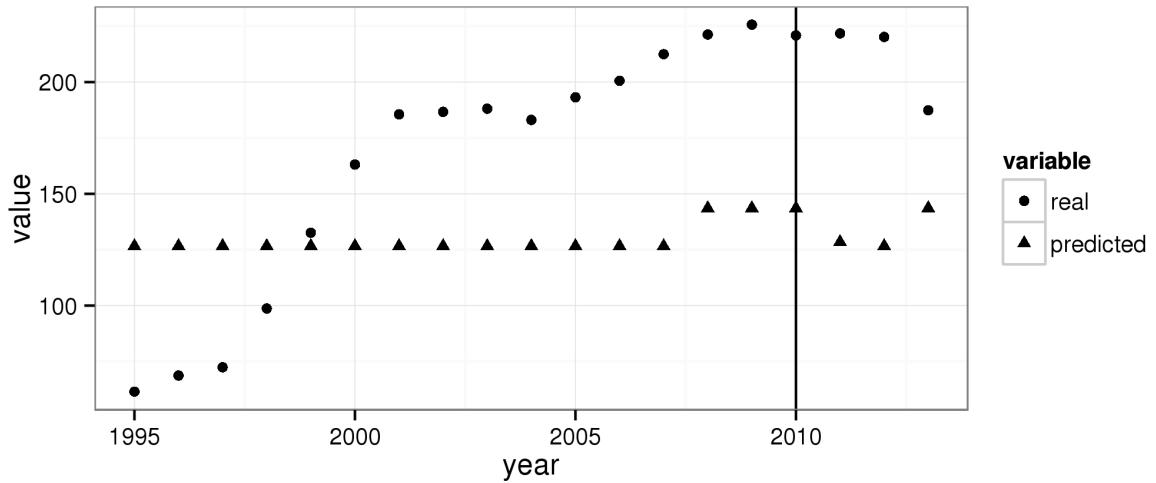


Рис. 3.6: Прогноз с набором показателей №2, горизонт 3 года.

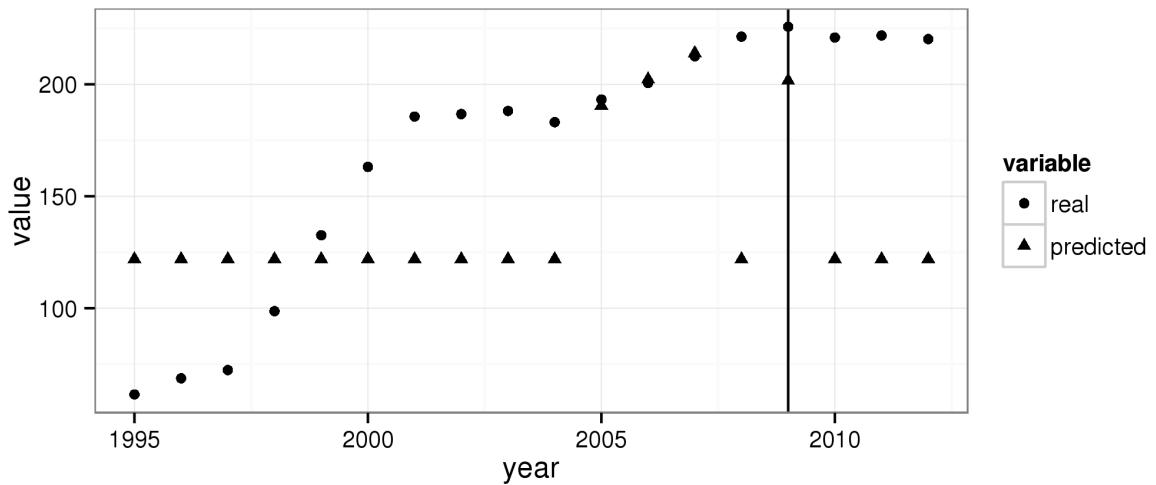


Рис. 3.7: Прогноз с набором показателей №3, горизонт 3 года.

Низкие оценки точности прогноза при использовании наборов показателей №2 и №3 объясняются, по мнению автора, неполнотой исходных данных, которая выражается в наличии пропущенных значений в независимых переменных, требующих их экстраполяции (в нашем случае для заполнения пропусков использовалась функция медианы), что заведомо снижает количество доступной для модели информации, и искажает характеристики случайных величин, тем самым нарушая механизмы прогнозирования.

3.3 Формальное описание модели

Предлагаемая модель прогнозирования численности наркозависимых на основе нечеткой логики состоит из ряда последовательных аналитических шагов:

1. Определение исходных данных;
2. Заполнение пропусков в исходных данных (imputation);
3. Выбор горизонта прогнозирования;
4. Обучение модели с помощью алгоритма нечетко-логического вывода;
5. Получение результатов модели (численных значений временных рядов, оценок ошибки прогнозирования, связей между переменными и др.);
6. Интерпретация результатов модели

Остановимся на некоторых из данных шагов подробнее.

Определение исходных данных.. Модель относится к категории моделей с множеством входных переменных и одной выходной переменной (MISO). Это означает, что прежде всего определиться с целевой переменной, прогнозирование которой и будет осуществляться. В нашем случае, это показатель "Состоит на учете больных с диагнозом «наркомания», на 100 тыс. населения полученный из Информационной системы информационно-аналитического обеспечения органов государственной власти Санкт-Петербурга. Однако, в «сыром» виде использовать показатели оценки наркоситуации не эффективно, т.к. ситуация, наблюдаемая органами государственной статистики и объективная ситуация, как правило, расходятся на несколько порядков ввиду нелегальной природы самого феномена наркотизма, наркорынков и пр. Для объективной оценки

феномена наркотизации вводится понятие коэффициента латентности [47], определяемого как отношение фактического значения исследуемой величины (в нашем случае — числа наркозависимых) к наблюдаемому значению. Оценка коэффициента латентности выходит за рамки данного исследования, добавим лишь, что как правило латентный анализ используется в сопряжении с установкой и анализом пороговых уровней, позволяющих судить о степени кризисности ситуации в зависимости от принадлежности тех или иных целевых величин соответствующим уровням. Кроме того, для приведения модели к стандартизированному виду, возможности её сопоставления с аналогами из Банка моделей и для внедрения элементов вероятностного моделирования целесообразно переводить абсолютные значения исследуемой величины в вероятность или долю (в нашем случае — вероятность заболевания наркозависимостью для отдельно взятого человека). Таким образом, общая формула манипуляций, производимых с исследумой переменной до непосредственно процесса прогнозирования выглядит следующим образом:

$$p_{addiction} = \frac{\rho_{drug}}{\rho_{population}} \cdot k \quad (3.1)$$

Где $p_{addiction}$ — вероятность заболевания наркозависимостью для отдельно взятого человека, ρ_{drug} — численность наркозависимых по данным официальной статистики, $\rho_{population}$ — численность населения города Санкт-Петербурга, k — коэффициент латентности.

Далее необходимо определиться с составом входных переменных. Модель базируется на предположении о существовании явной или неявной связи между входными переменными и выходной переменной. Социальные процессы не связаны такими же жесткими законами, как физические процессы, но можно предложить методы подбора входных переменных, способных обеспечить максимальную точность прогнозирования. К таким методам можно отнести: использование экспертного знания (исследований, опросов, метода коллегиальной

оценки, гипотез и др.), обращение к аппарату теоретико-методологических основ в исследуемой предметной области, проведение корреляционного анализа. Помимо этого, модель, в силу своей универсальности, предполагает возможность проведения сценарного анализа для учета влияния не только наркоспецифических, но и общеэкономических и социальных факторов для обнаружения и потенциального устранения непредвиденного влияния изменений в какой-либо из сфер жизнедеятельности города на состояние наркоситуации.

Заполнение пропусков в исходных данных. Основой моделирования наркоситуации являются данные государственной статистики. Однако, на практике, данные статистических служб имеют свои особенности и ограничения, которые необходимо учитывать при моделировании. К ним относятся: изменение год от года методик сбора данных, наличие переходных периодов, когда данные не собираются, разная частота и периодичность сбора данных в зависимости от конкретных показателей. Это приводит к неоднородности исходных данных, необходимости их приведения к нормализованному виду для дальнейшего анализа. Необходимость решения данной задачи обусловлена ещё и тем, что зачастую требуется использовать максимально длинный временной ряд ввиду того, что точность моделей на основе алгоритмов машинного обучения прямо пропорциональна объему доступной для обучения информации. И если взять за основу максимально длинный временной ряд, неизбежно возникают пропуски в случаях его совместного использования с сопутствующими временными рядами. Для решения этой задачи используются методы т.н. «imputation» или заполнения пропусков в данных. К таким методам относятся функции агрегирования (например, месячные, годовые средние, медианы и т.п.), аппроксимирования (интерполяции, сплайны и др.), метод LOCF («последнее наблюдаемое значение переносится вперёд»), фильтры (например, сезонный фильтр Калмана), наконец, ручное заполнение

пропущенных значений и обрезание пропущенных значений в конце и начале временного ряда. Выбор конкретного метода зависит от задач исследования, но в целом следует понимать, что заполнение пропусков может привести к искажениям в исходных данных.

Выбор горизонта прогнозирования. Горизонт прогнозирования — это количество единиц времени в будущем, для которых требуется осуществить прогноз. В анализе накроситуации применяются как среднесрочные прогнозы, так и долгосрочные, в зависимости от актуальных задач. Предлагаемая модель ориентирована на краткосрочные и среднесрочные прогнозы ввиду ограниченности исходных данных для построения более длинных предсказаний.

Обучение модели с помощью нечетко-логического вывода. В качестве базового алгоритма обучения в работе используется алгоритм Ванга-Менделя, основанный на равномерном разбиении входного и выходного пространств и построения отображения, учитывающего взаимное изменение уровней переменных. Алгоритм подробнее описан в разделе 3.1.2 «Методика проведения эксперимента».

Получение результатов модели. К результатам модели можно отнести ряд аналитических продуктов. Это, прежде всего, численные значения временных рядов на период, определённый установленным горизонтом прогнозирования, соответствующие им оценки ошибки прогнозирования. Отдельно следует затронуть генерируемую моделью базу нечетких правил. Нечеткие правила типа «ЕСЛИ, ТО» описывают связи между переменными, в зависимости от их уровней, описываемых лингвистическими переменными. Например, правило может звучать так: Если уровень безработицы высок, то численность наркозависимых выше среднего. При этом понятия «высоко» и «выше среднего» могут быть уточнены путем сопоставления отдельных термов лингвистических переменных конкретным значениям функций принадлежности.

Интерпретация результатов модели. Учитывая, что основных аналитических продуктов модели два, каждый из них интерпретируется по-своему. Прогнозные значения могут оцениваться как сами по себе, так и контексте определённых пороговых уровней. В зависимости от попадания значений в те или иные уровни и от общей направленности тренда можно говорить о степени кризисности наркоситуации и планировать соответствующие контрмеры. База нечетких правил позволяет исследовать связи между переменными, использую обнаруженные зависимости для создания предложений о том, как можно повлиять на ситуацию для её наилучшего развития.

Выводы по главе 3

Экспериментальная часть данного исследования была направлена на сравнение нечетко-логического и теоретико-множественного (теория нечетких временных рядов) подходов на примере авторегрессии по одной переменной. Согласно результатам эксперимента, нечетко-логический подход показал неудовлетворительные результаты на заданных данных. Приведены варианты разрешения этой трудности.

В части многомерного моделирования данное исследование было направлено на сравнение точности прогнозной модели при использовании разных входных индикаторов наркоситуации. Согласно результатам эксперимента, наблюдается значительное улучшение точности при условии достаточного объема и низкой зашумленности исходных данных.

Приоритетными направлениями дальнейшей деятельности являются:

- Анализ предметной области «Наркоситуация в Санкт-Петербурге» с целью выделения релевантных показателей для прогнозирования.
- Использование API информационно-аналитической системы «Антинар» для упрощения доступа к данным государственной статистики.
- Создание графического интерфейса пользователя (предположительно с помощью библиотеки Shiny) с тем, чтобы обеспечить возможность использования моделей более широким кругом пользователей.
- Более тщательный и полный выбор показателей для прогнозирования. (Возможно использование данных в разрезе по полу, возрасту; разбивка временных рядов на месяца вместо годов и т.п.);

- Обеспечение стабильно-высокой точности прогноза;
- Интеграция модели прогнозирования в «ИС ИАО».

ГЛАВА 4

Применение модели прогнозирования численности наркозависимых для информационно-аналитической поддержки принятия решений органами государственной власти

4.1 Текущая модель организации антинаркотического мониторинга и анализа наркоситуации на региональном уровне на примере Санкт-Петербурга

Мониторинг и анализ наркоситуации на данный момент являются одними из самых актуальных вопросов Государственной антинаркотической политики и деятельности по противодействию незаконному обороту наркотиков и распространению наркомании.

В Санкт-Петербурге мониторинг наркоситуации осуществляется с 2003 года. На основе ранее организованного комплексного мониторинга социально-экономического развития города – как субъекта Федерации, на первом этапе анализировалось исполнение мероприятий целевых программ Санкт-Петербурга по противодействию злоупотреблению наркотиками и их незаконному обороту.

В настоящее время целевое предназначение мониторинга гораздо шире. С одной стороны – это информационно-аналитическая поддержка процесса формирования надведомственных решений, обеспечивающих выполнение государственной антинаркотической политики на территории субъекта, а с другой – информационное обеспечение на уровне каждого ведомства, решающего обозначенную проблему с учетом деятельности других ведомств.

Правовой основой организации мониторинга наркоситуации, последовательного создания межведомственной Базы данных, а затем АИС мониторинга является Стратегия государственной антинаркотической политики, а также Постановление правительства Российской Федерации «Об утверждении Положения о

государственной системе мониторинга наркоситуации в Российской Федерации».

На уровне региона вступило в силу Постановление Губернатора Санкт-Петербурга от 20.09.2012 N 60-пг "О мониторинге наркоситуации в Санкт-Петербурге".

Автоматизированные средства проведения мониторинга наркоситуации. В Санкт-Петербурге мониторинг наркоситуации осуществляется с применением автоматизированной информационной системы, являющейся по сути информационно-аналитической системой АИС «Антинар СПб».

АИС «Антинар СПб» - автоматизированная система, ориентированная на комплексный анализ наркоситуации во всех сферах жизнедеятельности города, выявление важнейших тенденций и закономерностей ее развития, информационно-аналитическую поддержку процесса принятия управленческих решений руководством Санкт-Петербурга.

Среди основных задач, решаемых за счет автоматизации процессов мониторинга, анализа и прогнозирования выделяют:

- сбор, предварительная обработка, структурирование, интеграция и хранение информации о наркоситуации (фильтрация, проверка на полноту и достоверность, верификация, резервирование, восстановление и т.д.);
- поддержание в актуальном состоянии Базы данных (БД АИС «Антинар СПб»), содержащей разностороннюю информацию о наркоситуации в Санкт-Петербурге; аналитическая обработка информации, решение расчетных задач, получение агрегированных текущих и прогнозных оценок развития наркоситуации в городе на основе применения современных информационных технологий;
- моделирование и прогнозирование процессов наркоситуации в Санкт-Петербурге; разработка предложений для принятия

соответствующих управленческих решений; анализ исполнения мероприятий целевых программ по противодействию злоупотреблению наркотиками и их незаконному обороту;

- информационно-аналитическая и информационно-справочная поддержка управленческой деятельности руководства города.

Таким образом, основное предназначение АИС «Антинар СПб» – информационно-аналитическая поддержка процесса принятия управленческих решений Губернатором Санкт-Петербурга, Правительством Санкт-Петербурга, руководителями ИОГВ по определению приоритетных направлений антинаркотической деятельности в Санкт-Петербурге.

Создание межведомственного информационного ресурса:

- определение круга отраслевых и территориальных структур (ведомств), осуществляющих деятельность по противодействию злоупотреблению наркотиками и их незаконному обороту и разработку Положения об информационном взаимодействии между ними в процессе мониторинга;
- разработка, согласование и утверждение Перечня показателей наркоситуации и Регламента подготовки и представления данных в единую базу данных (БД).

Перечень составляют показатели по различным направлениям мониторинга наркоситуации: незаконный оборот наркотиков, наркомания, окружающая среда, профилактика наркозависимости, показатели мониторинга наркоситуации в Российской Федерации по методике ГАК. Наличие в Перечне показателей, относящихся к районному уровню, позволяет осуществлять мониторинг наркоситуации с учетом особенностей ее развития в районах города, там, где непосредственно проживают группы риска населения, имеющие отношение к незаконному обороту наркотиков.

В результате проделанной работы сформирована единая межведомственная База данных по показателям наркоситуации, обеспечивающая возможность их анализа, сопоставления и прогноза.

Разработанный алгоритм мониторинга базируется на принципах использования единых и обязательных для всех участников методологических подходов и критериев оценки показателей наркотической ситуации, открытости используемых для мониторинга данных, безвозмездного обмена информацией. Результат – возможность проводить на уровне города комплексный объективный анализ динамики ситуации.

На втором этапе при тесном взаимодействии с Комитетом по вопросам законности, правопорядка и безопасности и с Комитетом по информатизации и связи создана АИС мониторинга наркоситуации «Антинар СПб», которая позволяет наблюдать и анализировать срезы состояния тех или иных взаимоувязанных показателей по итогам квартала, полугодия и года. Система реализована с учетом всех требований по работе с конфиденциальной информацией и аттестована по классу 1Г. Сформирован работоспособный коллектив специалистов по анализу данной предметной области.

Работа с данными по показателям наркоситуации, поступившим из разных ведомств, позволяет не только увидеть картину в целом, но и выявлять скрытые процессы или намечающиеся тенденции, изучать состояние и динамику наркоситуации, ее отдельные составляющие, например:

- наркорынок;
- заболеваемость, в том числе по районам;
- преступность, связанная с незаконным оборотом наркотиков и т.д.

Информационная поддержка принятия управленческих решений. Важнейшей характеристикой статистической информации, определяющей ее пригодность для решения задач социального

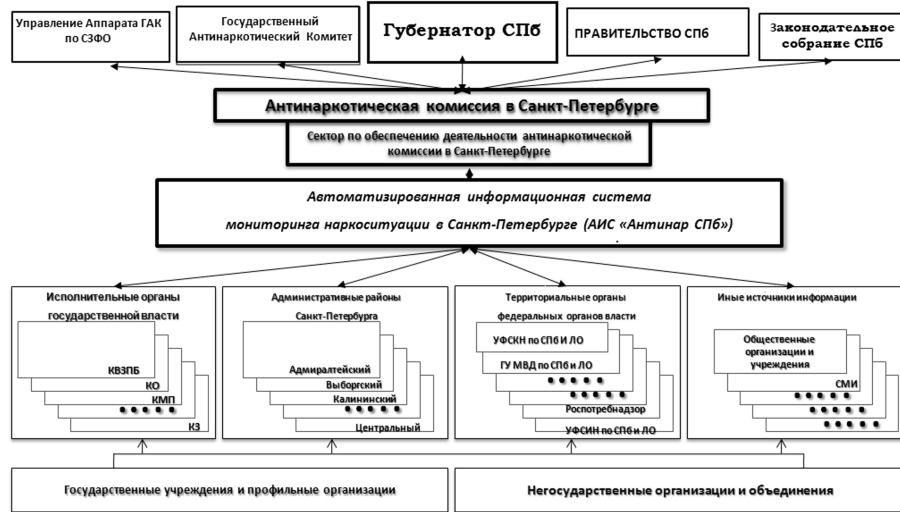


Рис. 4.1: Организация межведомственного мониторинга наркоситуации с использованием АИС «Антинар СПб»

управления, является достоверность, под которой понимается степень адекватности отображения информацией описываемых явлений, событий или процессов. Достоверность данных в базе АИС «Антинар СПб» обеспечивается документальным сопровождением и указанием источников всех загружаемых значений.

Наличие межведомственной системы позволяет координировать деятельность соответствующих исполнительных органов государственной власти с учетом общегородских задач (рис. 4.4). Например, модернизация системы учета больных в государственных наркологических учреждениях здравоохранения или профилактических программ, осуществляемых заинтересованными в решении своих проблем администрациями районов.

Комплексный анализ наркоситуации на средствах АИС «Антинар СПб» включает:

- анализ динамики и прогнозирование процессов распространения и последствий незаконного оборота наркотиков;
- оценку влияния наркоситуации на социально-экономические аспекты жизнедеятельности города;

- оценка численности потребителей психоактивных веществ, неучтенных в официальной статистике (Оценка величины латентности);
- оценка кризисности наркоситуации в регионе на основе латентностных характеристик;
- анализ выполнения целевых программ по противодействию злоупотреблению наркотиками и их незаконному обороту.

Злоупотребление психоактивными веществами является сложным социальным процессом, скрытым от непосредственного наблюдения и оценивания. Поэтому его распространение нельзя ограничивать только фиксированным перечнем статистических показателей, учитываемых государственной и ведомственной статистикой. В силу специфики проблемы наркотизации населения, прежде всего ее высокой латентности, классические методы изучения ведомственной статистики недостаточны для объективной оценки состояния наркообстановки. Поэтому обязательной составляющей мониторинга являются данные социологических исследований и экспертных оценок по целому ряду вопросов наркоситуации.

В связи с этим следует отметить, что в городе до настоящего времени не сложилась система комплексного учета данных тематических оценочных исследований, проводимых государственными структурами и общественными организациями или фондами в различных слоях общества и по сути своей мониторирующих происходящие в обществе процессы. Никто из специалистов ИОГВ не может точно сказать кто, сколько, как и зачем исследует на подведомственных им территориях, какие результаты получает и как их использует в дальнейшем. А ведь только объединение усилий государства и общества обеспечит успех в противодействии наркомании. Эта проблема требует своего решения в рамках обсуждаемого мониторинга.

АИС «Антинар СПб» как региональный сегмент государственной системы мониторинга наркоситуации в Российской Федерации

Согласно постановлению Правительства №485, методике мониторинга наркоситуации ГАК и приложениям к ней определен порядок мониторинга по показателям наркоситуации на региональном и федеральном уровне. Согласно данным постановлениям, в базе данных АИС «Антинар СПб» создан раздел «Мониторинг наркоситуации по методике ГАК». Тем самым АИС «Антинар СПб» является региональным сегментом государственной системы мониторинга наркоситуации в Российской Федерации.

На основе программно-технических средств АИС «Антинар СПб» организовано функционирование единой базы данных по вопросам наркоситуации в субъектах СЗФО в рамках организации мониторинга наркоситуации на территории СЗФО согласно методике ГАК. С этой целью реализовано:

1. создано программно-техническое ядро единой БД.
2. в единой БД «Антинар СПб – СЗФО» размещены:
 - показатели мониторинга наркоситуации согласно приложениям №№ 1-53 к методике мониторинга наркоситуации ГАК;
 - данные по показателям для оценки степени тяжести наркоситуации в субъектах СЗФО.

Источниками информации являются исполнительные органы государственной власти и органы местного самоуправления субъектов СЗФО и УФСКН субъектов СЗФО. Передача данных с мест на первом этапе может обеспечиваться по электронной почте, на бумажных носителях, либо с помощью разработанного типового модуля. При этом информационные потоки в предлагаемом формате и с установленной периодичностью будут поступать из регионов СЗФО в Единую БД АИС «Антинар СПб - СЗФО».

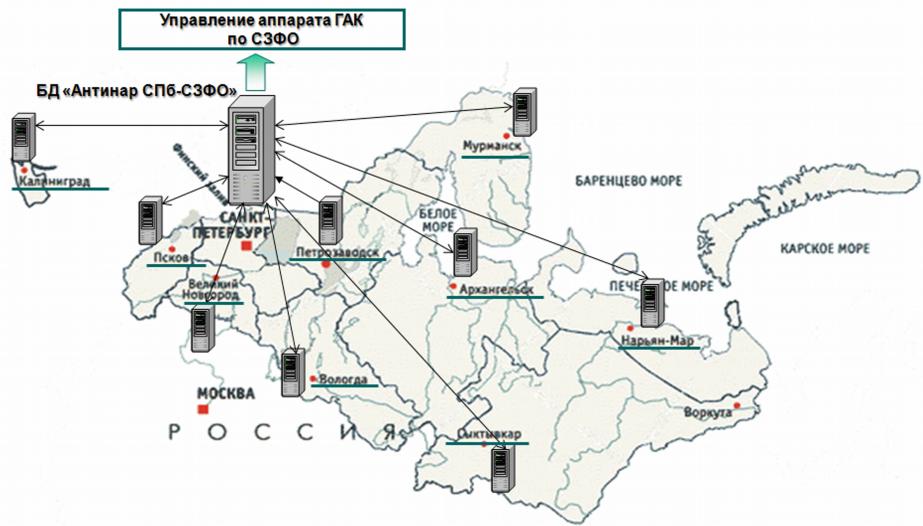


Рис. 4.2: Схема формирования информационно-коммуникационной основы мониторинга наркоситуации в СФЗО

Геоинформационный сегмент АИС «Антинар СПб» В геоинформационном сегменте АИС «Антинар СПб» рассматривается анализ информации, имеющей жесткую привязку к географическим. В АИС «Антинар СПб» представлены ГИС-слои по направлениям:

1. мониторинг сообщений населения о местах распространения наркотиков по телефонам горячей линии;
2. мониторинга деятельности ФСКН – места задержаний распространителей наркотиков, лабораторий и пр.;
3. мониторинг токсикологической ситуации, включая места отравлений наркотическими средствами и психотропными веществами, отравлений с целью опьянения и пр.;
4. мониторинг профилактической деятельности в Санкт-Петербурге, включая места проведения мероприятий антинаркотической направленности;
5. и другие слои, включающий сопутствующую информацию о территории.

Достоинством данного подхода является его наглядность и простота восприятия, а также жесткая привязка к реальным объектам, что упрощает интерпретацию результатов и выработку управлеченческих решений. Вследствие описанных особенностей рассматриваемые методы применимы в задаче оценки эффективности работы с населением, в частности, в задаче организации мероприятий первичной профилактики. Организация профилактической работы подразумевает определение места проведения и целевую аудиторию ряда мероприятий, что позволяет осуществлять ее анализ посредством ГИС.

Портал «Антинаркотическая политика в Санкт-Петербурге». Пространство Интернет является важнейшим ресурсом, посредством которого реализуется масса различных информационно-пропагандистских мероприятий. Поэтому особенно ценной, на наш взгляд, явилась возможность освещения государственной антинаркотической политики со страниц интернет-сайта.

Поскольку информационные ресурсы базы данных Автоматизированной информационной системы мониторинга наркоситуации в Санкт-Петербурге (АИС «Антинар СПб»), информационно-аналитические материалы, разработанные на инструментальных средствах системы доступны только для руководителей и специалистов ИОГВ (система защиты информации АИС классифицирована по классу 1Г), то возникла необходимость в создании информационного инструмента для более широкого круга пользователей. Это и обусловило создание сайта.

В рамках решения задачи организации электронного правительства по направлению создания сети сайтов и порталов, направленных на иллюстрацию деятельности органов государственной власти населению в Интернет, разработан и функционирует интернет-ресурс «Антинаркотическая политика в Санкт-Петербурге».

Главная цель создания данного сайта – объективное и полное отражение государственной антинаркотической политики,

реализуемой исполнительными органами государственной власти на территории Санкт-Петербурга и его районов, формирование у широкого круга пользователей сайта представления о деятельности ИОГВ и их возможностях в области противодействия злоупотреблению наркотическими средствами и их незаконному обороту.

Сайт позволяет избежать длительных процедур ознакомления с деятельностью органов власти в поисках тематической информации по противодействию распространению наркозависимости. Сайт дает возможность предоставления любому заинтересованному контингенту пользователей документов, фотоматериалов и другой информации в больших объемах и с высокой оперативностью.

С помощью сайта можно мгновенно известить большое количество пользователей о главных мероприятиях и важнейших событиях антинаркотической деятельности в Санкт-Петербурге и его районах.

С учетом практически неограниченного охвата аудитории мы получили мощное и эффективное информационно-пропагандистское средство, несущее пользователям актуальную информацию о деятельности отраслевых и территориальных ИОГВ Санкт-Петербурга по реализации государственной антинаркотической политики. Одной из главных задач сайта является обеспечение координации деятельности территориальных органов исполнительной власти и антинаркотических комиссий в районах Санкт-Петербурга, а также организация их взаимодействия между собой, с органами исполнительной власти города, органами местного самоуправления муниципальных образований, общественными объединениями и организациями.

Методологические подходы к анализу наркоситуации. Для комплексной оценки наркоситуации в системе реализован известный методический подход, основанный на сопоставлении значения того или иного индикативного показателя наркоситуации с устанавливаемым для него пороговым (нормативным) уровнем. Это

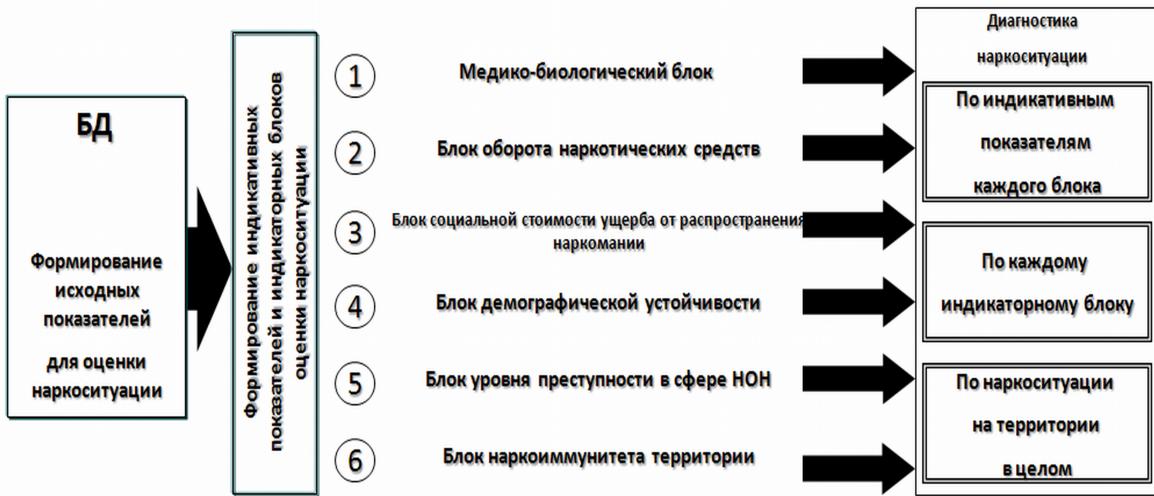


Рис. 4.3: Методические основы анализа наркоситуации с использованием АИС «Антинар СПб»

позволяет определять степень кризисности ситуации, складывающейся в регионе (рис. 4.3).

Анализ наркоситуации по данной методике позволяет дать оценку социальной стоимости ущерба от распространения наркомании, количеству наркозависимых на территории, числу лиц, занимающихся незаконным оборотом наркотиков, уровню преступности в сфере незаконного оборота наркотиков, уровню наркотизации различных групп населения (женщины, подростки) и демографической устойчивости территории, а также выйти на интегральную оценку наркоситуации в регионе.

По каждому блоку индикативных показателей производится оценка пороговых уровней наркоситуации. В используемой методике определения пороговых значений индикативных показателей мы исходили из положений о демографической безопасности территории. Под демографической безопасностью понимаем состояние защищенности общества, проживающего на территории, его образа жизни, культурно-исторического наследия и др. Именно сохранение коренного населения территории мы видим основной целью мероприятий укрепления национальной безопасности, в частности

борьбы с распространением наркозависимости.

Следует отметить, что использование пороговых уровней показателей наркоситуации позволяет получить нижнюю оценку кризисности наркоситуации, таким образом, возникает дополнительная задача определения оптимального направления деятельности по ликвидации кризисности наркоситуации.

Пороговые значения показателей служат для оценки степени кризисности ситуации в рассматриваемой сфере. В литературе встречается определение уровней кризисности порогов индикаторов как степеней угроз нормальному функционированию системы. Предполагается, что по выбранному перечню индикативных показателей должны быть определены некие критические значения, в случае превышения которых следуют изменения, требующие оперативного вмешательства в рассматриваемый процесс. Нами рассматривается следующая система пороговых уровней:

ПК1 – предкризисное состояние первого уровня, характеризуется ожиданием наступления кризиса с 95% вероятностью в течение пяти лет;

ПК2 – предкризисное состояние второго уровня, характеризуется ожиданием наступления кризиса с 95% вероятностью в течение трех лет;

ПК1 – кризисное состояние, характеризуется наличием явных угроз демографической безопасности региона;

ПК2 – кризисное состояние второго уровня, характеризуется завершающей стадией эпидемического процесса развития наркомании и значительной угрозой демографической безопасности региона;

ПК3 – кризисное состояние третьего уровня, характеризуется практическим уничтожением этноса, массовым оттоком коренного

населения, замещением населения трудовыми мигрантами.

Отнесение ситуации в регионе к какому-либо из перечисленных состояний аналогична проверке гипотезы наличия в регионе кризисной ситуации. Уменьшение вероятностей ошибок первого и второго рода в данном случае достигается за счет выбора соответствующего критерия отнесения ситуации к соответствующему состоянию.

В данной методике определение пороговых уровней опирается на величину латентности наркоситуации. Основное положение состоит в том, что реальное число наркозависимых не должно превысить определенного критического значения, когда развитие наркомании приведет к деградации общества и развалу экономики региона. Таким образом, определение пороговых уровней учитывает необходимость снижения фактической преступности и заболеваемости в сфере незаконного оборота наркотиков, а также уменьшение теневой доли данных процессов.

Прогнозирование наркоситуации в Санкт-Петербурге
Прогнозирование наркоситуации является важнейшей составляющей деятельности в сфере информационной поддержки принятия управлеченческих решений, поскольку позволяет не только оценить эффективность принимаемых мер противодействия, но и моделировать развитие наркоситуации в зависимости от сценария противодействия с целью нахождения наилучшего решения.

Получение прогноза требует разработки специфических подходов в силу скрытого характера явления наркомании. Для организации эффективных мер противодействия существует необходимость прогнозировать развитие наркоситуации в зависимости от общей социальной, экономической, психологической и политической обстановки на территории.

Прогнозирование наркоситуации в АИС «Антинар СПб» осуществляется на основе модуля анализа и ситуационного прогнозирования для автоматизированной информационной системы

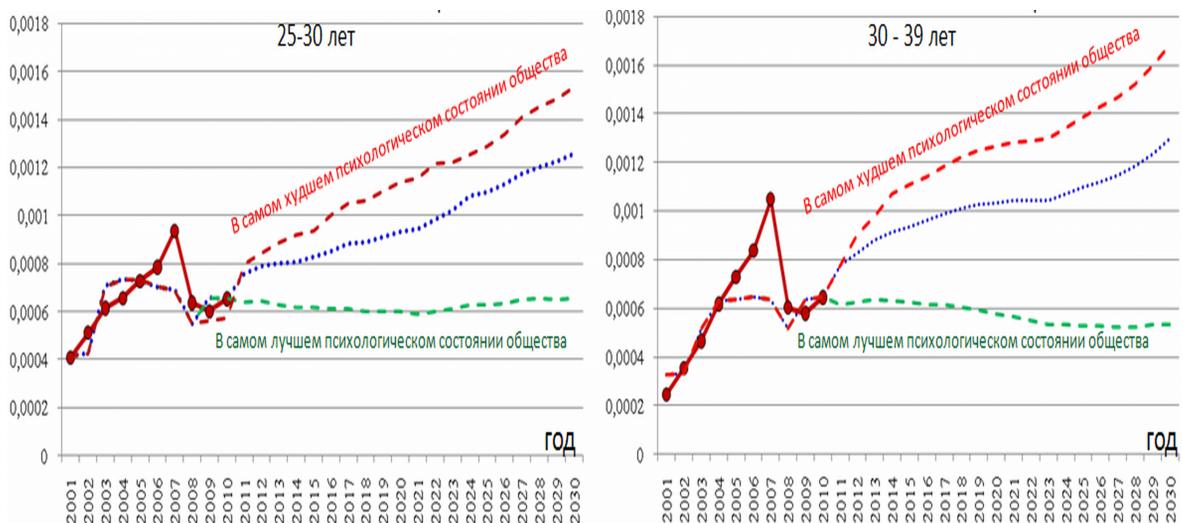


Рис. 4.4: Прогнозирование вероятности заболеваемости наркоманией в лучшем и худшем эмоционально-психологическом сценарии развития Санкт-Петербурга

мониторинга наркоситуации в Санкт-Петербурге (МАСП АИС «Антинар СПб»).

Комплекс предназначен для ситуационного прогнозирования развития наркоситуации на территории субъекта Федерации. По задаваемым сценариям прогнозируется динамика численности наркозависимых с распределением по полу, возрасту и тяжести заболевания, экономический ущерб от распространения наркомании, демографические потери от наркомании и снижение демографической устойчивости территории.

Создание автоматизированной информационной системы подразумевает организацию информационного обмена, как данными, так и документами, что позволяет унифицировать и упростить взаимодействие между ведомствами и организациями, участниками мониторинга. И в конечном итоге перейти к информационному документообороту.

Но широкий функционал определяет строгие критерии выбора организации, ответственной за создание и сопровождение автоматизированной информационной системы, так как ошибка в

данном вопросе может привести к развалу всей системы мониторинга.

В Санкт-Петербурге отлажена система информационного взаимодействия субъектов антинаркотической деятельности, создана и действует многоуровневая система профилактики наркомании. Однако, продолжающийся рост незаконного распространения и немедицинского потребления наркотиков обуславливают необходимость оперативного реагирования на наркоситуацию всех участников антинаркотической деятельности. Необходима выработка новых путей и возможностей для повышения эффективности противодействия наркомании, поднятия уровня доверия населения города к работе исполнительных органов государственной власти, правоохранительным и контролирующими органам. Одним из таких путей является организация комплексного мониторинга наркоситуации на основе использования современных информационно-коммуникационных технологий.

4.2 Сравнение эффективности применяемых на практике моделей прогнозирования наркоситуации с предлагаемой новой моделью на основе нечеткой логики

На данный момент в Банке моделей кратко- и среднесрочного прогнозирования социально-экономического развития Санкт-Петербурга задействован ряд достаточно конвенциональных моделей прогнозирования: ARIMA (оценивание по методу максимального правдоподобия, комбинированному методу, методу наименьших квадратов), линейный, степенной, полиномиальный и логарифмический тренд, экспоненциальное сглаживание (двойное), линейная регрессия, Метод Хольта (аддитивный тренд), мультипликативная модель Хольта-Винтерса (аддитивный тренд, длина цикла 12). С помощью них прогнозируются показатели по различным блокам: демография, промышленное производство,

строительство, торговля и услуги населению, внешнеэкономическая деятельность, инвестиции, финансы, уровень жизни населения, труд и занятость, здравоохранение, преступность.

Хотя данные методы не применяются в Банке моделей для прогнозирования наркоситуации, в силу их большой распространенности и разнообразия сфер жизнедеятельности города и показателей, которые прогнозируются с их помощью, попробуем протестировать их на показателе «Состоит на учете больных с диагнозом наркомания на 100 тыс. населения, чел.» для сравнения точности результатов.

Условия для сравнения: горизонт прогнозирования — 3 года, данные в разрезе по годам, исходный датасет расширен от границ 1995-2012 до 1991-2013 ввиду использования методов заполнения пропусков при нечетком моделировании для охвата большего диапазона в независимых переменных.

Таблица 4.1: Сравнение ошибок прогнозирования предлагаемой модели и аналогов из Банка моделей кратко- и среднесрочного прогнозирования. Данные в разрезе по годам.

Название модели	MSE	RMSE	SMAPE
ARIMA	507.743352	22.533161	1.753912
Нечеткая логика	1098.56	33.14	3.73
Экспоненциальное сглаживание	374.527341	19.352709	1.427809
Метод Хольта	157.591757	12.553556	1.368674
Линейная модель	1925.323490	43.878508	3.907447

По результатам сравнения, нечеко-логический подход демонстрирует чувствительность к количеству входной информации, как и предполагалось ранее. Однако, даже в условиях ограниченности исходных данных нечеткая модель более верно, чем аналоги, указывает общую направленность тренда. Таким образом, можно утверждать необходимость применения временных рядов с

наибольшей частотой измерений. Например, при использовании данных в разрезе по месяцам, а не по годам, мы получаем в 12 раз больше информации для обработки моделью, что позитивно сказывается на точности конечного результата.

Заключение

Прогнозирование социально-экономических процессов в целом и наркоситуации в частности — один из наиболее важных инструментов для поддержки принятия решений при управлении региональным развитием, позволяя планировать действия и ресурсы в соответствии с гипотетическим состоянием социальной системы в будущем.

Немаловажным обстоятельством является включенность системы информационно-аналитической поддержки контроля за наркоситуацией в общегосударственную стратегию антинаркотической политики, что требует строгой подчиненности данной системы целям и задачам, выдвигаемым на данном этапе реализации антинаркотической политики. Поэтому применяемые методы должны максимально эффективно обеспечивать достижение целей мониторинга и анализа наркоситуации.

Основные теоретические результаты работы носят следующий характер:

- Изучены используемые в мировой практике антинаркотические политики;
- Изучены методы и модели прогнозирования временных рядов;
- Проведено сравнение моделей на предмет их эффективности в региональном управлении;
- Сформулированы требования к требуемой модели прогнозирования численности наркозависимых на территории Санкт-Петербурга;
- Спроектирован программный модуль прогнозирования для информационно-аналитической системы.

В области практических результатов удалось достичь следующего:

- Модифицирована модель прогнозирования на основе нечеткой логики под нужды анализа наркоситуации;

- Разработан модуль прогнозирования для информационно-аналитической системы;
- Разработаны рекомендации к применению новой модели и его программного инструмента;
- Продемонстрированы характеристики интеллектуального анализа данных как средства регионального управления.

В связи с тем, что с каждым годом наркоситуация сопровождается всё новыми вызовами, одними из последних для примера можно назвать распространение синтетических наркотиков, требуются ответы на эти вызовы. В том числе это касается подсистемы информационно-аналитического обеспечения принятия решений. Разработанная модель и сопутствующий программный инструмент в силу своей универсальности могут быть не только использованы для решения своей прямой задачи — прогнозирования численности наркозависимых, но и могут быть адаптированы для других задач, в особенности имеющих отношение к сложным, многосоставным связям между явлениями.

Разработанный инструмент, учитывая рекомендуемые условия его применения, можно рекомендовать к внедрению в качестве компонента информационно-аналитической системы. Следует отметить обнаружившееся в ходе работы обстоятельство наличия у интеллектуальных методов прогнозирования специфических аналитических продуктов, удобных инструментов анализа которых пока что не обнаружено. В случае нечеткой логики таким продуктов является база нечетких правил. Таким образом, можно утверждать о необходимости развития области интеллектуального моделирования и прогнозирования для наилучшего применения теоретических достижений на практике.

Список литературы

1. *Alcala R., Alcala-Fdez J., Herrera F.* A Proposal for the genetic lateral tuning of linguistic fuzzy systems and its interaction with rule selection // IEEE Transactions on Fuzzy Systems. — 2007. — Vol. 15, no. 4. — Pp. 616–635.
2. *Caulkins J. P.* Estimating Elasticities of Demand for Cocaine and Heroin With Data from the Drug Use Forecasting System: Final Report / United States Department of Justice, National Institute of Justice. — Washington, DC, 1995.
3. *Chaiken J. M., Chaiken M. E.* Understanding the Drug Use Forecasting (DUF) Sample of Adult Arrestees / National Institute of Justice. — 1993. — P. 59.
4. *Chapman M. [et al.]* ADAM II 2009 Annual Report: Arrestee Drug Abuse Monitoring Program II / Office of National Drug Control Policy. — 2010.
5. *Chen S.* Forecasting enrollments based on fuzzy-time series // Fuzzy sets and systems. — 1996.
6. *Chi Z., Yan H., Pham T.* Fuzzy algorithms: with applications to image processing and pattern recognition. — Singapore : World Scientific, 1996. — 227 pp.
7. *Chiu S.* Method and software for extracting fuzzy classification rules by subtractive clustering // Fuzzy Information Processing Society. — 1996. — Pp. 461–465.
8. *Cordon O. [et al.]* MOGUL: A methodology to obtain genetic fuzzy rule-based systems under the iterative rule learning approach // International Journal of Intelligent Systems. — 1999. — Vol. 14. — Pp. 1123–1153.
9. *Council N. R.* Informing America's Policy on Illegal Drugs: What We Don't Know Keeps Hurting Us / ed. by C. F. Manski, J. V. Pepper, C. V. Petrie. — Washington, D.C. : National Academy Press, 2001. — P. 424.

10. *Davenport T. H., Patil D. J.* Data Scientist: The Sexiest Job of the 21st Century // Harvard Business Review. — 2012. — Oct.
11. *EMCDDA* Modelling drug use: methods to quantify and understand hidden processes. — Luxembourg : Office for Official Publications of the European Communities, 2001. — 246 pp. — URL: http://www.emcdda.europa.eu/attachements.cfm/att_80903_EN_Monograph6.pdf ; EMCDDA Scientific Monograph Series; No 6.
12. *Gantz J., Reinsel D.* Extracting Value from Chaos // IDC iView «Extracting Value from Chaos». — 2011.
13. *Gonzalez A., Perez R.* Selection of relevant features in a fuzzy genetic learning algorithm // IEEE Transactions on Systems, Man, and Cybernetics. — 2001. — Vol. 31, no. 3. — Pp. 417–425.
14. *Hoekstra M.* Hotel booking prediction by means of fuzzy logic prediction: MA thesis / Hoekstra M. — Amsterdam : VU University Amsterdam, 2010. — 25 pp.
15. *Hunt D.* Arrestee Drug Abuse Monitoring II in the United States: Technical Documentation Report / Inter-university Consortium for Political & Social Research. — 2013. — P. 50.
16. *Irvine A. D.* Principia Mathematica // The Stanford Encyclopedia of Philosophy / ed. by E. N. Zalta. — Winter 2013. — 2013. — URL: <http://plato.stanford.edu/archives/win2013/entries/principia-mathematica/>.
17. *Ishibuchi H., Nakashima T.* Effect of rule weights in fuzzy rule-based classification systems // IEEE Transactions on Fuzzy Systems. — 2001. — Vol. 1. — Pp. 59–64.
18. *Ishibuchi H., Nakashima T., Murata T.* Performance evaluation of fuzzy classifier systems for multidimensional pattern classification problems // IEEE transactions on Systems, Man, and Cybernetics. — 1999. — Vol. 29, no. 5. — Pp. 601–618.

19. *Ishibuchi H., Yamamoto T., Nakashima T.* Hybridization of fuzzy GBML approaches for pattern classification problems // IEEE transactions on Systems, Man, and Cybernetics. — 2005. — Vol. 35, no. 2. — Pp. 359–365.
20. *Ishibuchi H.* [et al.] Empirical study on learning in fuzzy systems by rice taste analysis // Fuzzy Sets and Systems. — 1994. — Vol. 64, no. 2. — Pp. 129–144.
21. *Jan J. S. R.* ANFIS: adaptive-network-based fuzzy inference system // IEEE Transactions on Systems, Man, and Cybernetics. — 1993. — Vol. 23, no. 3. — Pp. 665–685.
22. *Kasabov N., Song Q.* DENFIS: Dynamic evolving neural-fuzzy inference system and its application for time-series prediction // IEEE Transactions on Fuzzy Systems. — 2002. — Vol. 10, no. 2. — Pp. 144–154.
23. *Kim J., Kasabov N.* HyFIS: Adaptive neuro-fuzzy inference systems and their application to nonlinear dynamical systems // Neural Networks. — 1999. — Vol. 12, no. 9. — Pp. 1301–1319.
24. *Kosko B.* Fuzzy Systems as Universal Approximators // IEEE Transactions on Computers. — 1994. — Nov. — Issue 11. — Pp. 1329–1333. — URL: <http://sipi.usc.edu/~kosko/FuzzyUniversalApprox.pdf>.
25. *MacCoun R. J., Reuter P.* Drug War Heresies: Learning from Other Vices, Times, and Places. — Cambridge University Press, 2001. — 479 pp.
26. *Meehl P. E.* Clinical versus statistical prediction: a Theoretical Analysis and a Review of the Evidence. — University of Minnesota Department of Psychology. — URL: <http://www.psych.umn.edu/people/meehlp/032ClinstixBook.pdf>.

27. *Melike S., Degtiarev K. Y.* Forecasting Enrollment Model Based on First-Order Fuzzy Time Series // World Academy of Science, Engineering and Technology. — 2008.
28. National Drug Control Strategy / the Whitehouse [сайт]. — URL: http://www.whitehouse.gov/sites/default/files/ondcp/policy-and-research/ndcs_2014.pdf (visited on 10/10/2014).
29. *Nomura H., Hayashi I., Wakami N.* A learning method of fuzzy inference rules by descent method // IEEE International Conference on Fuzzy Systems. — 1992. — Pp. 203–210.
30. *Pak P. S., Kim G.* Long-Range Prediction of Population by Sex, Age and District Based on Fuzzy Theories // Computational Intelligence for Modelling and Prediction. — 2005.
31. *Rada-Vilela J.* fuzzylite: a fuzzy logic control library. — 2014. — URL: <http://www.fuzzylite.com>.
32. *Robinson M. B., Scherlen R. G.* Lies, Damned Lies, and Drug War Statistics: A Critical Analysis of Claims Made by the Office of National Drug Control Policy. — Albany : State University of New York Press, 2007. — 286 pp.
33. *Sasu A.* An application of fuzzy time series to the romanian population // Bulletin of the Transilvania University of Brasov. — 2010. — Vol. 3, issue 52. — Pp. 125–132.
34. *Sawyer J.* Measurement and prediction, clinical and statistical // Psychological Bulletin. — 1966. — Vol. 66, no. 3. — Pp. 178–200.
35. *Saxena P., Sharma K., Easo S.* Forecasting Enrollments based on Fuzzy Time Series with Higher Forecast Accuracy Rate // Computer Technology & Applications. — 2012. — No. 3. — Pp. 957–961.
36. *Song Q., Chissom B. S.* Fuzzy time series and its models // Fuzzy Sets and Systems. — 1993. — No. 54. — Pp. 269–277.

37. *Thrift P.* Fuzzy logic synthesis with genetic algorithms // In Proceedings of the Fourth International Conference on Genetic Algorithms. — San Diego, 1991. — Pp. 509–513.
38. *Travis J., Feucht T. E.* Drug Use Forecasting: 1994 Annual Report On Adult And Juvenile Arrestees / National Institute of Justice. — 1995. — P. 52.
39. *Wang L.-X., Mendel J. M.* Generating fuzzy rules by learning from examples // IEEE Transactions on Systems, Man and Cybernetics. — 1992. — Nov. — Vol. 22, no. 6. — Pp. 1414–1427.
40. *Wiessing L.* Pilot project to develop a model of geographical spread of drug misuse in the European Union / European Monitoring Centre for Drugs and Drug Addiction. — Lisbon, Apr. 1999. — URL: <http://www.emcdda.europa.eu/html.cfm/index1374EN.html>.
41. *Zadeh L.* An Outline of a New Approach to the Analysis of Complex Systems and Decision Processes // IEEE Transactions on Systems, Man, and Cybernetics. — 1973. — No. 1. — Pp. 28–44.
42. *Zadeh L. A.* Fuzzy Sets // Information and control. — 1965. — Vol. 8, no. 3. — Pp. 338–353. — URL: www.asianscientist.com/books/wp-content/uploads/2013/06/2895_chap01.pdf.
43. *Абдуллаева Н. А.* Нечеткий подход к прогнозированию уровня бедности // Управление большими системами. — М., 2010. — Вып. 30. — С. 98—114.
44. *Гегедюш Н. С., Масленникова Е. В., Мокеев М. М. [и др.]* Государственное и муниципальное управление: конспект лекций. — Москва, 2008. — 192 с.
45. *Карпец А. В., Махров И. Е., Виноградова Э. М.* Мониторинг и оценка наркоситуации. — М. : ЗАО Юстицинформ, 2010. — 224 с.

46. *Мамедова М. Г., Джабраилова З. Г.* Нечеткая логика в прогнозировании демографических аспектов рынка труда // Искусственный интеллект. — 2005. — Вып. 3. — С. 450—460.
47. Методологические основы мониторинга и комплексного анализа развития наркоситуации на территории региона / под ред. Ю. Н. Захаров, В. В. Иваненков, А. А. Коржик. — СПб : Нестор-История, 2012. — 200 с.
48. О Стратегии национальной безопасности Российской Федерации до 2020 года: указ Президента РФ от 12.05.2009 № 537 / Президент России [сайт]. — 2009. — URL: <http://graph.document.kremlin.ru/page.aspx?1012112> (дата обр. 22.12.2014).
49. Об утверждении Стратегии государственной антинаркотической политики Российской Федерации до 2020 года: указ Президента РФ от 9.06.2010 № 690 / Президент России [сайт]. — 2010. — URL: <http://graph.document.kremlin.ru/page.aspx?1;1285491> (дата обр. 22.12.2014).
50. *Чучуева И. А.* Классификация методов и моделей прогнозирования. — URL: <http://habrahabr.ru/post/177633/>.
51. *Чучуева И. А.* Модель прогнозирования временных рядов по выборке максимального правдоподобия: дис. ... канд. техн. наук / Чучуева И. А.. — М : МГТУ им. Н. Э. Баумана, 2012. — 153 с.
52. *Ярушкина Н. Г., Афанасьева Т. В., Перфильева И. Г.* Интеллектуальный анализ временных рядов: учебное пособие. — Ульяновск : УлГТУ, 2010. — 320с с.