На правах рукописи

ПАВЛОВ ВЛАДИМИР ДМИТРИЕВИЧ

ПРОГРАММНЫЙ КОМПЛЕКС, МЕТОДЫ МОДЕЛИРОВАНИЯ И ПРОГНОЗИРОВАНИЯ МНОГОКОМПОНЕНТНЫХ МОДЕЛЕЙ ДИНАМИЧЕСКИХ РЯДОВ С ИСПОЛЬЗОВАНИЕМ ФУНКЦИИ РАМСЕЯ

Специальность 05.13.18 - Математическое моделирование, численные методы и комплексы программ

АВТОРЕФЕРАТ
диссертации на соискание ученой степени кандидата технических наук

Самара 2009

Работа выполнена на кафедре математических методов в экономике государственного образовательного учреждения высшего профессионального образования «Самарский государственный аэрокосмический университет имени академика С.П. Королева»

Научный руководитель: доктор технических наук,
доктор экономических наук, профессор,
Семенычев Валерий Константинович

Официальные оппоненты: доктор физико-математических наук, профессор
Жданов Александр Иванович

кандидат технических наук, доцент
Бахарева Надежда Федоровна

Ведущая организация: Государственное образовательное учреждение высшего профессионального образования "**Самарский государственный технический университет**" (г.Самара)

2009 года в 12-00 на заседании
при государственном образовательном
ального образования «Самарский
университет имени академика С.П.
Московское шоссе, 34

иться в библиотеке Самарского
университета имени академика

09 г.

А. А. Калентьев

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

В исследованиях технических, биологических, социальных и экономических процессов широкое распространение получили процессы с логистическим характером динамики тренда анализируемых показателей: тренд сначала растет медленно, затем ускоряется, а затем снова замедляет свой рост, стремясь к некоторому уровню насыщения (рис. 1, Y_{k1}). Известны также случаи спадающего логистического тренда (рис. 1, Y_{k2}).

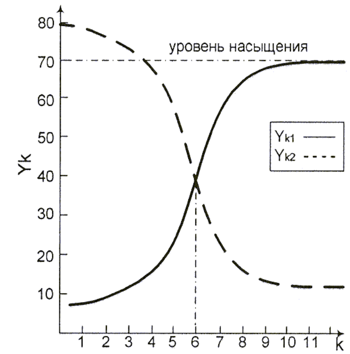

Рис. 1. Графики логистических функций

Уникальным свойством логистического тренда[1] является его способность моделировать качественные изменения в развитии динамики процессов, характеризующиеся сменой знака второй производной при сохранении знака первой производной.

Речь идет о динамике чаще во времени, хотя аргументом процесса могут быть и пространственные переменные: технические, физические, экономические и другие параметры.

Известные модели логистической динамики являются нелинейными по параметрам, а в анализируемой траектории кроме тренда присутствует и стохастическая компонента.

Актуальность темы исследования обусловлена недостаточной точностью известных методов и программных комплексов моделирования и прогнозирования логистического тренда.

Кроме того, в реальной практике в динамической траектории наряду с логистическим трендом обычно отмечается присутствие и других детерминированных компонент (полиномиальных, гармонических), что существенно усложняет задачу моделирования и прогнозирования, тем более, что компоненты модели также могут быть нелинейными по параметрам. В известной научной литературе не описаны методы и программные средства моделирования многокомпонентных временных рядов с логистическим трендом.

Теоретическую базу диссертации составили труды отечественных ученых: Айвазяна С.А., Афанасьева В.Н., Блинова А.О. Елисеевой И.И., Клейнера Г.Б., Мхитаряна В.С., Пахомовой Е.А., Прохорова С.А., Семёнычева В.К., Стерника Г.М., Тихомирова Н.П., Хачатряна С.Р., Черняка А.В., Четыркина Е.М. и др., а также зарубежных учёных - Берндта Э., Бокса Дж., Доугерти К., Дженкинса Г., Джонстона Дж., Рамсея Дж., Твисса и др.

[1] В международной практике принято название «logistic curves», в то время как в Российской практике большее распространение получило название «S-тренд» и «S-образные кривые»

На данный момент известно тринадцать аналитических выражений, для моделирования логистического тренда, наиболее распространенными из которых, и чаще всего применяемыми на практике, являются модели Верхулста и Гомперца.

Общая трудность идентификации логистических моделей заключается в том, что все известные модели являются нелинейными по независимым переменным, а порой и по параметрам.

Для отражения циклического характера процессов обычно к модели основного тренда добавляют циклические компоненты, а добавление линейной компоненты позволяет отразить и линейное изменение показателей во времени, хотя бы как приближении к нелинейности общего вида.

Обычно ряды динамики, в которых присутствуют циклические компоненты, моделируют в непараметрической форме и в несколько этапов. При этом требуемый для удовлетворительного по точности моделирования объем выборки составляет от 4 до 10 периодов колебательной компоненты, что соответствует объему выборки от 48 до 120 ежемесячных наблюдений.

Необходимость использования больших объемов выборок существенно ограничивает область применения известных способов моделирования: во-первых, не всегда есть возможность вести наблюдения за объектом в течении такого длительного промежутка времени – ретроспективные данные на несколько десятков лет назад, как правило, просто отсутствуют; а во-вторых, за такой длительный промежуток времени зачастую происходит эволюция компонент модели, как по виду, так и по параметрам, что может привести к малой точности моделирования и, особенно, прогнозирования неслучайных компонент ряда динамики. При пространственной переменной возможна ее неоднородность.

В силу этого, актуальна разработка новых моделей, способов и программ их параметризации, призванная увеличить точность моделирования реальных рядов динамики, в которых наряду с логистическим трендом присутствуют другие детерминированные компоненты, причем на коротких выборках, когда модели компонент можно считать стационарными по видам моделей и по параметрам.

Целью исследования является разработка математических методов и инструментальных средств, обеспечивающих повышение точности моделирования и прогнозирования технических и экономических процессов, в классе многокомпонентных моделей рядов динамики с логистическим трендом.

В соответствии с целью исследования были поставлены следующие **задачи**:

1. Предложить виды моделей, которые отвечали бы многообразию технических и экономических процессов с логистическим трендом.
2. Для предложенных моделей разработать и реализовать методы, допускающие идентификацию на коротких выборках.
3. Разработать расширяемый программный комплекс, осуществляющий моделирование и прогнозирование временных рядов предложенными методами, обладающий интуитивно понятным интерфейсом, что сделает возможным его использование не только профессионалами, но и людьми, обладающими начальными навыками работы с компьютером.
4. Разработанным программным комплексом провести исследование точности и области применения предложенных моделей и методов их параметризации.
5. Провести апробирование программного комплекса на реальных статистических данных.

Методологической основой исследования послужили методы и модели теории вероятностей и математической статистики, численные методы математики, известные результаты в теории авторегрессии, положения теории функций комплексного переменного, дифференциального исчисления, теории рядов, эконометрики и теории управления.

В процессе исследования проанализированы работы отечественных и зарубежных авторов, специалистов в области математического моделирования, информационных технологий а также обработаны данные по реальным техническим и экономическим процессам.

Использовался пакет Microsoft Office Excel 2003 и программный комплекс «Logistic», разработанный при участии автора в среде Borland Delphi 7.0.

Научная новизна. К числу основных результатов, полученных лично соискателем и определяющих научную новизну диссертаций, можно отнести следующее:

1. Предложено использовать модель Рамсея как основу для моделирования и прогнозирования девяти моделей временных многокомпонентных рядов, включающих в себя основной логистический тренд и дополнительные колебательную и линейную компоненты, которые адекватны многим реальным техническим и экономическим процессам.
2. Для рассматриваемых моделей сконструированы параметрические модели авторегрессии - скользящего среднего (ARMA-модели), в которых основой параметризации моделей является решение линейных уравнений из корреляционных моментов и параметров модели.
3. Проведено исследование точности предложенных моделей и методов их параметризации на тестовых выборках, с использованием разработанного программного комплекса, при различных соотношениях сигнал/шум и в широком диапазоне изменения параметров моделей.
4. Испытание разработанного программного комплекса на реальных данных подтвердило возможность использования предложенных моделей и методов идентификации для моделирования и прогнозирования различных технических и экономических процессов, на примере двенадцати применений в реальной практике.

На защиту выносятся следующие основные результаты в области разработки и развития математических методов и моделей, методологии, расширения области применения моделирования и прогнозирования многокомпонентных рядов динамики с логистическим трендом и их инструментальная (программная) поддержка:

1. Применение модели Рамсея и аддитивных компонент во временной области в виде многокомпонентных моделей.
2. Расширения области их реального применения.
3. «Перепараметризация» нелинейных моделей рядов динамики с линейными и колебательными компонентами, с использованием Z – преобразования.
4. Разработанные, испытанные на тестовых и реальных выборках программные средства моделирования и прогнозирования многокомпонентных рядов динамики.

Практическая ценность проведенных в диссертационной работе исследований, заключается в возможности использования полученных результатов и разработанного программного комплекса для моделирования и прогнозирования широкого класса реальных технических, экономических, социальных и биологических процессов с высокой точностью и на малых выборках, в широком динамическом диапазоне значений параметров, при многокомпонентности детерминированной составляющей. Важно, что для определения параметров логистического тренда не нужны априорные данные о уровне насыщения логистического тренда.

Апробация и внедрение результатов исследования.

Полученные теоретические, методологические и практические результаты работы обсуждались на 8-ми семинарах и конференциях: V Всероссийской научно-практической конференции «Электронный бизнес: проблемы, развитие и перспективы» (г. Воронеж, ВГУ, 23-24 ноября 2006г.); Научно-практической конференции СГАУ (г. Самара, 15-16 декабря 2006г.); Научно-практической конференции ПГУ (г. Пенза, ПГУ, 2007г.); II Всероссийской научно-практической конференции молодых ученых. (г. Самара, СМИУ, 1 марта 2007г.); IV Международной научно-практической конференции «Экономическое прогнозирование: модели и методы» (г. Воронеж, ВГУ, 2008г.); III Всероссийской научно-практической конференции молодых ученых. (г. Самара, СМИУ, 26-27 апреля 2008г.); V Международной заочной научно-практической конференции «Интеллектуальные технологии в образовании, экономике и управлении» (ИТОЭУ-2008) (г.Воронеж, 2008г.); XIV Международной открытой научной конференции "Современные проблемы информатизации" (г.Воронеж, 2008г.).

Результаты проведенных исследований и разработанный программный комплекс использованы в лекционных курсах и при проведении лабораторных работ по курсам «Математическое моделирование в экономике» и «Эконометрика» в Самарском муниципальном институте управления и при обучении по специальности «Математические методы в экономике» в Самарском государственном аэрокосмическом университете, что подтверждено актами внедрения.

Публикации. По теме диссертационного исследования было опубликовано 11 научных работ, в том числе в ведущем рецензируемом научном журнале, определенном Высшей аттестационной комиссией - одна, издано одно методическое пособие, получено одно свидетельство о государственной регистрации программы для ЭВМ.

Структура диссертации. Диссертационное исследование изложено на 173 страницах, состоит из введения, трех глав, заключения, библиографического списка и приложений.

СОДЕРЖАНИЕ РАБОТЫ

Во введении показана актуальность темы диссертации, сформулированы цели и задачи работы, раскрыта научная новизна, изложена практическая значимость полученных результатов, сформулированы основные положения, выносимые на защиту.

В первой главе «Обзор существующих методов идентификации моделей временных рядов с логистическим характером динамики тренда» рассмотрены:

- теоретические аспекты моделирования и прогнозирования;
- модели рядов и существующие методы их идентификации;
- модели, используемые для идентификации логистической динамики.

Тем самым в данной главе заложена теоретическая база работы, показана актуальность темы исследования.

В **п. 1.1** принято определение **математической модели,** как образа реальной системы (процесса) в форме математических соотношений, отражающих существенные свойства моделируемой системы и замещающих её в ходе исследования и управления.

В настоящее время при моделировании и прогнозировании чаще используют непараметрические методы в силу сложности аналитической параметризации нелинейных по параметрам рядов. В непараметрических методах не высока точность прогнозирования, не строится теоретическая модель явления или процесса.

К достоинствам параметрических методов можно отнести следующее:

- результаты представимы в виде удобных аналитических выражений;
- параметрические методы демонстрируют «общую картину» развития процесса, дают, как правило, большую точность прогноза.

В **п. 1.2** рассмотрена структура ряда динамики, а также осуществлен выбор критериев точности прогноза и моделирования.

Временным рядом называется ряд наблюдений $Y_1, Y_2, ..., Y_N$ анализируемой случайной величины ξ_k, произведенных в последовательные моменты времени $t_1, t_2, ..., t_N$, где $t_k =$ (kΔ), $Y_k = Y(t_k) = Y(k\Delta)$, Δ– шаг дискретизации, N- объем выборки.

В общем виде модель временного ряда может быть представлена в виде суммы двух компонент:

$$Y_k = D_k + \varepsilon_k, \qquad (1)$$

где D_k - детерминированная (систематическая) компонента ряда; ε_k - случайная (стохастическая) компонента ряда.

Детерминированная компонента может быть разложена на отдельные, детерминированные компоненты (составляющие):

1. **Долговременная (вековая) составляющая**, формирующая общую в длительной перспективе тенденцию в изменении анализируемого признака L_k.
2. **Сезонная составляющая** – S_k, формирующаяся под влиянием сезонных колебаний технического или экономического показателя в течение заданного периода времени менее года.
3. **Циклическая (конъюнктурная) оставляющая** – C_k, формирующая изменения анализируемого признака в связи с действием долговременных циклов – технической, астрофизической, экономической или демографической природы, например, волны Кондратьева, демографические «ямы» и пики, циклы солнечной активности и т.п.

Показана целесообразность при идентификации ориентироваться на ARMA- модель.

Приведен обзор основных показателей, характеризующих точность модели и прогноза. В качестве критерия оценки точности моделирования выбран коэффициента детерминации:

$$R^2 = \frac{\sum_{k=1}^{N}\left(Y_k^o - M\left[Y_k\right]\right)^2}{\sum_{k=1}^{N}\left(Y_k - M\left[Y_k\right]\right)^2} = 1 - \frac{\sum_{k=1}^{N}\left(Y_k^o - Y_k\right)^2}{\sum_{k=1}^{N}\left(Y_k - M\left[Y_k\right]\right)^2}, \qquad (2)$$

где $M\{...\}$ - оператор математического ожидания.

Для оценки погрешности прогноза использовалась MAPE-оценка:

$$\bar{\gamma} = \frac{1}{l}\sum_{k=1}^{l}\frac{\left|Y_k - Y_k^*\right|}{Y_k}100\%. \qquad (3)$$

В **п. 1.3** приведено двадцать пять различных технических, экономических, биологических и социальных процессов, изменяющихся по логистическим законам.

Жизнь технических систем (как, впрочем, и других систем, например биологических) можно изобразить в виде логистической (S-образной) кривой (рисунки 2, 3), показывающей, как меняются во времени главные характеристики системы (мощность, производительность, скорость, число выпускаемых систем, популяции животных, социальная активность и т.д.). Видим, что на рисунке 3 помимо логистического тренда в исходных данных присутствует и колебательная компонента.

Проведенный анализ многочисленных данных по рынку недвижимости Самарской области показал, что по логистическим законам происходит и изменение стоимости однокомнатных, двухкомнатных, трехкомнатных квартир как на первичном так и на вторичном рынках недвижимости.

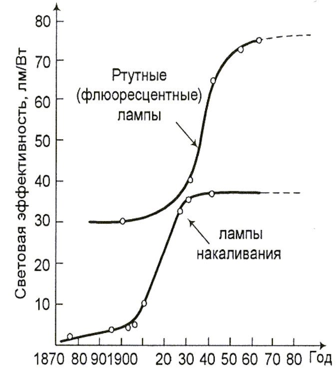

Рис. 2. Увеличение эффективности двух различных источников света. Моделирование функцией Верхулста (Перла-Рида).

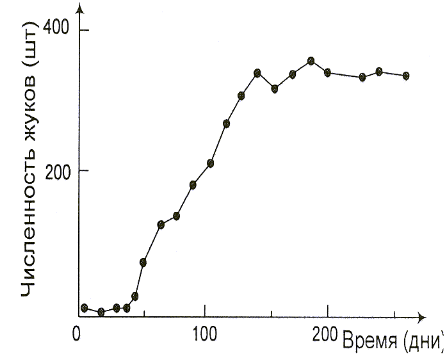

Рис. 3. Динамика численности жука Rhizopertha dominica в 10-граммовой порции пшеничных зерен, пополняемых каждую неделю.

Приведены тринадцать моделей, используемых для моделирования логистической динамики. Наиболее популярными из них являются функкция

Верхулста (Перла-Рида): $$D_k = \frac{1}{A_0 + A_1\exp\left(-\alpha_1 k\Delta\right)}, \qquad (4)$$

и функция Гомперца: $$D_k = A \cdot B^{C^{k\Delta}}. \qquad (5)$$

Обзор известных методов моделирования и прогнозирования логистической динамики показал, что все они ориентированы исключительно на идентификацию логистического тренда. Но и при этом они обладают невысокой точностью и обычно требуют для идентификации дополнительных априорных знаний о параметре тренда (об уровне насыщения).

Методы идентификации при предположении других компонент в выборке в известной научной литературе не описаны.

Проведенный в первой главе диссертации анализ существующих моделей логистической динамики и методов их параметризации позволяет сделать следующие **выводы**:

- анализ существующего состояния методов моделирования и прогнозирования указывает на актуальность разработки новых математических методов и инструментария для научной и практической деятельности широкого круга специалистов. Методы, предлагаемые для идентификации моделей логистической динамики, сложны, требуют большого объема статистических данных и обладают низкой точностью прогнозирования, что сдерживает их широкое практическое применение;
- актуальна разработка методов, позволяющих осуществлять моделирование и прогнозирование на коротких выборках, когда протекающие

процессы можно считать стационарными по видам моделей и их параметрам. Данная задача особенно актуальна для описания эволюционирующей динамики, когда применение больших выборок непригодно, т.к. за время сбора информации сильно изменяются внешние условия, а также для процессов, по которым отсутствуют статистические данные за длительный период времени.

- для реальной практики целесообразнее использовать многокомпонентные модели, в состав которых помимо основного тренда входят линейные и колебательные компоненты.

Далее результаты первой главы используются при построении и параметризации многокомпонентных логистических рядов.

Во второй главе «Инструментарий построения моделей для рядов логистической динамики» описан метод параметризации, позволяющий выполнять идентификацию моделей на коротких выборках и осуществлять их параметризацию, с использованием функции Рамсея:

$$D_k = C\left(1-\left(1+\alpha(k\Delta)\right)\cdot \exp\left(-\alpha(k\Delta)\right)\right) \qquad (6)$$

Предложено ее применение в качестве логистического тренда при построении многокомпонентных моделей. На основе функции Рамсея предложено девять моделей, а также проведена их параметризация.

В **п. 2.1** описан предложенный В.К. Семенычевым метод параметризации на основе Z-преобразования. К недостаткам метода для наиболее часто используемой функции Верхулста отнесены гетероскедастический характер стохастической компоненты ε_k и нелинейное преобразование исходной выборки Y_k (переход к обратным значениям), ведущее к параметризации не по значениям выборки Y_k, а по $Z_k = 1/Y_k$. Показана невозможность применения модели Верхулста для построения и последующей параметризации многокомпонентных моделей логистической динамики.

Для моделирования логистического тренда предложено принять относительно мало известную функцию Рамсея, которая, во-первых, использовалась лишь для однокомпонентной модели пространственной динамики, а во-вторых, известный метод ее параметризации не обеспечивает высокую точность на малых выборках.

В **п. 2.2** проведено исследование функции Рамсея и многокомпонентных динамических моделей, построенных на ее основе.

На рисунке 4 приведены графики функций Верхулста и Рамсея. Ряд 1 отображает данные функции Верхулста со следующими параметрами: $A_0 = 0{,}1$, $A_1 = 0{,}25$, $\alpha = 0{,}1$. В ряде 3 представлены данные функции Рамсея с параметрами $C = 7$, $\alpha = 0{,}11$. Видим, что ряд 3 практически полностью совпадает с первым рядом.

Справедливо и более общее утверждение: для различных моделей Верхулста можно подобрать модель Рамсея, что позволяет рассчитывать на ее широкое применение на практике. Проведенные исследования функции Верхулста и Рамсея при различных сочетаниях параметров показали близость графиков этих двух функций, что свидетельствует о возможности использования функции Рамсея при моделировании логистической динамики, в том числе и в случаях, когда для этих целей традиционно использовалась функция Верхулста

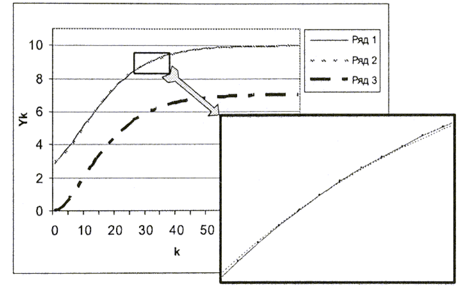

Рис. 4. Сравнение графиков функций Верхулста и Рамсея

На основе логистической функции Рамсея было построено девять многокомпонентных моделей:

$$Y_1^k = C\left(A-\left(B+\alpha(k\Delta)\right)\cdot \exp\left(-\alpha(k\Delta)\right)\right)+\varepsilon_k \qquad (7)$$

$$Y_2^k = C(1-(1+\alpha k\Delta)\exp(-\alpha k\Delta))+\varepsilon_k \qquad (8)$$

$$Y_3^k = C(1-(1+\alpha k\Delta)\exp(-\alpha k\Delta))+A_1 k\Delta+\varepsilon_k \qquad (9)$$

$$Y_4^k = C(1-(1+\alpha k\Delta)\exp(-\alpha k\Delta))+A_1\sin(\omega k\Delta+\varphi)+\varepsilon_k \qquad (10)$$

$$Y_5^k = C(1-(1+\alpha k\Delta)\exp(-\alpha k\Delta))+A_1\sin(\omega k\Delta+\varphi_1)+A_2 k\Delta+\varepsilon_k \qquad (11)$$

$$Y_6^k = C(1-(1+\alpha k\Delta)\exp(-\alpha k\Delta))+A_1\sin(\omega_1 k\Delta+\varphi_1)+A_2\sin(\omega_2 k\Delta+\varphi_2)+\varepsilon_k \qquad (12)$$

$$Y_7^k = C(1-(1+\alpha k\Delta)\exp(-\alpha k\Delta))+A_1\sin(\omega_1 k\Delta+\varphi_1)+A_2\sin(\omega_2 k\Delta+\varphi_2)+A_3 k\Delta+\varepsilon_k \qquad (13)$$

$$Y_8^k = C(1-(1+\alpha k\Delta)\exp(-\alpha k\Delta))+A_1 k\Delta\sin(\omega_1 k\Delta+\varphi_1)+A_2 k\Delta+\varepsilon_k \qquad (14)$$

$$Y_9^k = C(1-(1+\alpha k\Delta)\exp(-\alpha k\Delta))+A_1 e^{\alpha k\Delta}\sin(\omega_1 k\Delta+\varphi_1)+A_2 k\Delta+\varepsilon_k \qquad (15)$$

Модели, содержащие две гармонических компоненты (модели 12-13), позволяют одновременно описывать как сезонные, так и циклические колебания, характерные для реальных процессов.

Особый интерес представляет возможность моделирования эволюционирующих рядов. Использование в моделях (14)-(15) гармоник с эволюционирующими амплитудами позволяет расширить область применения предложенных моделей.

Многочисленные графики построенных моделей с различными параметрами, их представлено в данной главе двадцать, показали широкие и, зачастую, неожиданные по формам кривых возможности моделирования рядов динамики логистами в сочетании с линейными и гармоническими компонентами.

Можно сделать следующие выводы:

1. Функция Рамсея по своему виду очень близка к функции Верхулста, что делает возможным ее применение в случаях, когда для моделирования традиционно использовалась функция Верхулста.

2. Аддитивно добавленные к функции Рамсея линейные и синусоидальные компоненты позволили существенно расширить область применения моделей логистической динамики.
3. Приведенные графики по своему виду совпадают с графиками реальных статистических данных (например, рисунок 5), иллюстрируя возможности применения предложенных моделей для их аппроксимации.

Приведенный «атлас моделей» позволит исследователю легче ориентироваться при выборе того или иного типа модели, лучше понять возможности применения каждой из предложенных моделей.

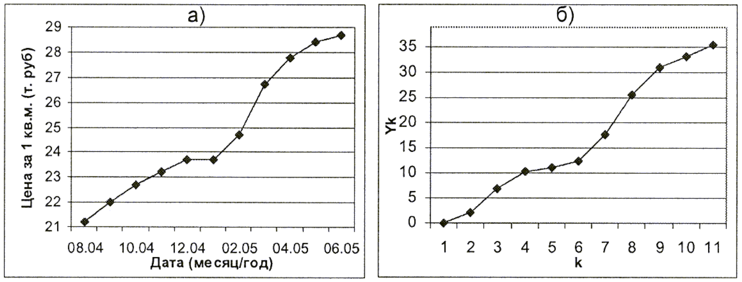

Рис. 5. а) Стоимость 1 кв. метра жилья (г. Самара) в среднем ценовом диапазоне; б) график многокомпонентной модели (11)

В **п. 2.3** предложены методы параметризации девяти моделей (формулы 7-15), построенных на основе логисты Рамсея.

Покажем процесс параметризации модели (11), в состав которой помимо основного, логистического тренда входят также линейная и гармоническая компоненты. Для данной модели получена следующая ARMA-модель шестого порядка:

$$\begin{aligned} Y_k &= 2Y_{k-1} - 2Y_{k-2} + 2Y_{k-3} - Y_{k-4} + m_1(2Y_{k-1} - 4Y_{k-2} + 4Y_{k-3} - 4Y_{k-4} \\ &+ 2Y_{k-5}) - m_2(Y_{k-2} - 2Y_{k-3} + 2Y_{k-4} - 2Y_{k-5} + Y_{k-6}) + \\ &+ m_3(Y_{k-1} - 2Y_{k-2} + Y_{k-3}) - m_4(2Y_{k-2} - Y_{k-3} + 2Y_{k-4}) + \\ &+ m_5(Y_{k-3} - 2Y_{k-4} + Y_{k-5}) + \varepsilon_k, \end{aligned} \qquad (16)$$

где: $m_1 = \lambda_1$, $m_2 = \lambda_1^2$, $m_3 = \lambda_2$, $m_4 = \lambda_1\lambda_2$, $m_5 = \lambda_1^2\lambda_2$, $\lambda_1 = \exp(-\alpha\Delta)$, $\lambda_2 = 2\cos(\omega\Delta)$, ξ_k - новая стохастическая компонента, по своей структуре аналогичная Y_k.

$$\begin{aligned} \varepsilon_k &= 2\varepsilon_{k-1} - 2\varepsilon_{k-2} + 2\varepsilon_{k-3} - \varepsilon_{k-4} + m_1(2\varepsilon_{k-1} - 4\varepsilon_{k-2} + 4\varepsilon_{k-3} - 4\varepsilon_{k-4} \\ &+ 2\varepsilon_{k-5}) - m_2(\varepsilon_{k-2} - 2\varepsilon_{k-3} + 2\varepsilon_{k-4} - 2\varepsilon_{k-5} + \varepsilon_{k-6}) + \\ &+ m_3(\varepsilon_{k-1} - 2\varepsilon_{k-2} + \varepsilon_{k-3}) - m_4(2\varepsilon_{k-2} - \varepsilon_{k-3} + 2\varepsilon_{k-4}) + \\ &+ m_5(\varepsilon_{k-3} - 2\varepsilon_{k-4} + \varepsilon_{k-5}), \end{aligned}$$

Далее в два этапа определим коэффициенты модели (16). На первом этапе, применяя тот или иной метод сглаживания, например, МНК, находим оценки параметров $m_1^0, m_2^0, m_3^0, m_4^0$ и m_5^0 из условия:

$$\begin{aligned} m_1^0, m_2^0, m_3^0, m_4^0, m_5^0 &= \underset{m_1, m_2, m_3, m_4, m_5}{\arg\min} \sum_{k=5}^{N} (Y_k - \{2Y_{k-1} - 2Y_{k-2} + 2Y_{k-3} - Y_{k-4} + \\ &+ m_1(2Y_{k-1} - 4Y_{k-2} + 4Y_{k-3} - 4Y_{k-4} + 2Y_{k-5}) - m_2(Y_{k-2} - 2Y_{k-3} + 2Y_{k-4} - \\ &- 2Y_{k-5} + Y_{k-6}) + m_3(Y_{k-1} - 2Y_{k-2} + Y_{k-3}) - m_4(2Y_{k-2} - Y_{k-3} + 2Y_{k-4}) + \\ &+ m_5(Y_{k-3} - 2Y_{k-4} + Y_{k-5})\})^2. \end{aligned} \qquad (17)$$

Решая соответствующую нормальную СЛАУ 5-го порядка, найдем оценки параметров m_1, m_2, m_3, m_4 и m_5. Учитывая, что $m_1 = \lambda_1$, $m_3 = \lambda_2$, получим следующие формулы для оценок параметров α и ω:

$$\alpha^0 = -\frac{\ln m_1^0}{\Delta}, \quad \omega^0 = \frac{\arccos(m_3^0 / 2)}{\Delta}.$$

На втором этапе, для определения C^0, A_1^0, A_2^0 и φ^0 воспользуемся условием:

$$\begin{aligned} C^0, A_2^0, A_3^0, A_4^0 &= \underset{C, A_2, A_3, A_4}{\arg\min} \sum_{k=0}^{N} (Y_k - (C(1 - (1 + \alpha^0 k\Delta)\exp(-\alpha^0 k\Delta)) + \\ &+ A_3 \sin\omega^0 k\Delta + A_4 \cos\omega^0 k\Delta + A_2 k\Delta))^2, \end{aligned} \qquad (18)$$

где $A_3 = A_1\cos\varphi$, $A_4 = A_1\sin\varphi$.

Данное условие приводит к СЛАУ 4-го порядка, решая которую, находим оценки параметров C^0, A_2^0 и рассчитываем $\varphi^0 = arctg(A_4^0 / A_3^0)$ и $A_1^0 = A_3^0 / \cos(\varphi^0)$.

Как показали исследования, рассмотренный метод параметризации оказывается работоспособен и на коротких выборках, в два-три раза больших порядка ARMA-модели, что значительно расширяет область применения предложенных моделей. Аналогичные ARMA-модели получены и для моделей (8)-(15).

В третьей главе «Программный комплекс моделирования и прогнозирования динамических рядов с использованием функции Рамсея» приведено описание программного комплекса «Logistic», позволяющего осуществлять моделирование и прогнозирование временных рядов, с помощью девяти моделей, предложенных во второй главе. С помощью разработанного программного комплекса проведено исследование точности моделирования и прогнозирования предложенными моделями, а также определена область их применения. Также проведено аппробирование программного комплекса на реальных статистических данных.

В **п. 3.1** дано описание разработанного программного комплекса «Logistic», с помощью которого можно осуществлять моделирование и прогнозирование временных рядов с логистическом трендом. Программа разработана в среде Borland Delphi 7.0. Структурная схема программного комплекса представлена на рисунке 6.

Условно программный комплекс можно разделить на два модуля: модуль тестирования предложенных моделей для оценки их метрологических характеристик и модуль моделирования и прогнозирования временных трендов, на основе реальных статистических данных.

Первый модуль используется при оценке работоспособности и точности программного комплекса на тестовых выборках, с назначаемыми параметрами модели и шума. По результатам тестирования автоматически создается отчет в виде файла формата «Excel», в который заносятся все необходимые для анализа данные, а также результаты автоматической обработки полученных данных.

Второй модуль осуществляет моделирование и прогнозирование реальных временных рядов. Исходные данные заносятся в таблицы Excel, после чего они загружаются в программу, для дальнейшей обработки. Результаты моделирования также могут быть экспортированы в файл Excel.

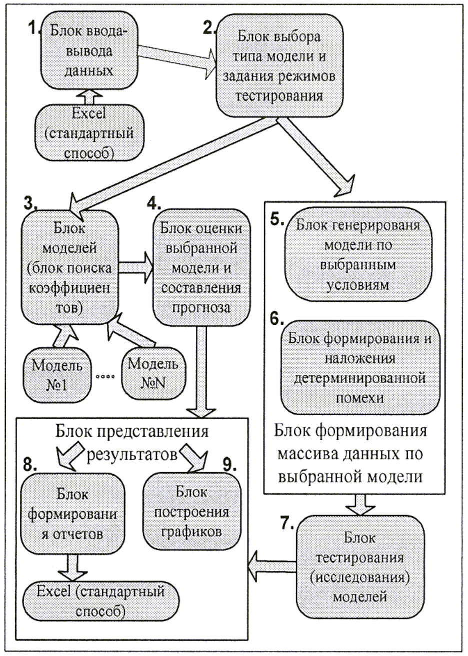

Рис. 6. Структурная схема программного комплекса «Logistic»

Также в программном комплексе заложена возможность предварительной обработки данных: 1) существует возможность обработки не самих исходных данных, а их относительных значений, что позволяет снизить вычислительные погрешности округления в случае обработки очень больших значений, или значений, близких к нулю; 2) существует возможность прореживания исходных данных, т.е. на основе исходного ряда строится ряд, состоящий только из k-ых отсчетов, где k задается пользователем. Как показали исследования, точность моделирования повышается при назначении k, равном порядку авторегрессии, т.к. в этом случае мы получаем выборку, состоящую из некоррелированных величин.

Помимо этого существует возможность задания значения периода гармонической компоненты вручную. На практике встречаются случаи, когда значение периода колебательной компоненты нам известно, исходя из некоторых априорных данных. Ручное задание значения периода позволяет

повысить точность вычислений за счет сокращения количества неизвестных параметров модели, и как следствие, уменьшения размерности СЛАУ при ее параметризации.

В **п. 3.2** с помощью разработанного программного комплекса проведены метрологические исследования точности девяти предложенных многокомпонентных моделей логистической динамики и способов их параметризации.

При этом использовалась следующая методика:

-осуществлялась генерация уровней T_k детерминированных компонент моделей (7)-(15), с которыми суммировалась генерируемая стохастическая компонента ξ_k с нормальным законом распределения и регулируемой (назначаемой) мощностью. Для уменьшения зависимости исследования точности от конкретных выборок ξ_k использовалось усреднение по десяти тестовым выборкам;

-в качестве критерия точности моделирования использовался коэффициент детерминации, а точность прогноза характеризовалась МАРЕ – оценкой прогноза.

-точность исследовалась в функции коэффициента «шум/сигнал» при различных объемах выборки:

$$K_{ns} = \frac{D(\xi)}{D(D_k)}, \qquad (19)$$

где $D(\xi)$ - дисперсия стохастической компоненты с назначаемой (путем центрирования, нормирования и умножения на соответствующий коэффициент) дисперсией, $D(D_k)$ - дисперсия детерминированной компоненты модели (сумма логисты Рамсея, линейного тренда и гармоники).

Таблица 1 – Диапазон изменения параметров моделей

Параметр	MIN	MAX
C	10	140
α	0,05	0,5
A_1	0,3	30
A_2	0,01	15
A_3	0,01	2
ω_1	0,175	1,047
ω_2	0,175	1,047

При тех же выборках Y_k и соответствующих параметрах моделирования определялась зависимость МАРЕ– оценки от длины исходной выборки (глубина прогноза, как это обычно максимально допускают, составляла 1/3 от длины выборки).

Эксперименты проводились в широком диапазоне изменения параметров моделей – например, данные по диапазонам изменения параметров приведены в таблице № 1.

Результаты исследования точностных характеристик моделей, полученные при имитационном моделировании, свидетельствуют о возможности широкого использования предложенных моделей для моделирования и прогнозирования на реальных статистических данных.

Проведен сравнительный анализ точности моделирования тестовых данных различными моделями: 1) функцией Верхулста, с применением известных методов идентификации; 2) функцией Рамсея, с применением предложенного метода идентификации; 3) многокомпонентными рядами, построенными на основе функции Рамсея, с применением предложенных методов их идентификации.

В качестве моделируемой выборки использовался ряд данных, смоделированный на основе функции Верхулста, с добавлением 5% шума. Результаты моделирования данного ряда функцией Верхулста приведены в книге Е.М.Четыркина «Статистические методы прогнозирования», при этом диапазон изменения коэффициента детерминации для различных методов идентификации (их приведено шесть) составляет от 0,520 до 0,997.

Таблица 2 содержит результаты моделирования тестовых данных тремя моделями, построенными на основе логисты Рамсея (модели 7, 8 и 9) а также результаты моделирования функцией Верхулста. При этом приведены данные моделирования как выборки без шума, так и выборки с шумом.

Таблица 2 – сравнение различных методов моделирования на тестовой выборке без шума и с 5 % шумом. Объем выборки – 14-ть отсчетов

Модель	R^2 **(без шума)**	R^2 **(с шумом)**
Верхулста (метод Родса)	1	0,9970
Логиста Рамсея (модель 7)	0,9945	0,9304
Логиста Рамсея + лин. комп-та. (модель 8)	0,9914	0,9897
Логиста Рамсея + лин. комп-та. +синус. комп-та. (модель 10)	0,9999	0,9982

Таблица 3 – сравнение различных методов моделирования на короткой выборке (выборка 8-мь отсчетов с 5% шумом)

Модель	R^2
Верхулста (метод Родса)	0,8892
Логиста Рамсея (модель 7)	0,9014
Многокомпонентная логиста Рамсея (модель 8)	0,9932
Многокомпонентная логиста Рамсея (модель 10)	0,9936

Моделирование функцией Рамсея является заведомо менее точным: исходная выборка строилась на основе логисты Верхулста, поэтому даже при моделировании данных без шума коэффициент детерминации заведомо меньше единицы, что нашло отражение в таблице 3. Однако даже при таких, заведомо неравных в ущерб функции Рамсея условиях сравнения, применение функции Рамсея дает результаты, по своей точности уступающие результатам моделирования логистой Верхулста только в случае использования методов Родса и Нейра. Применение модели (7) позволило достичь точности, превышающей аналогичные показатели модели Верхулста.

Также проведено исследование точности моделирования различными методами для более коротких выборок, результаты которого, для N=8, приведены в таблице 3. Из таблицы видно, что в случае короткой выборки

предложенные модели и методы их идентификации обладают более высокой точностью, по сравнению с известными методами моделирования функцией Верхулста.

В **п. 3.3** проведено исследование программного комплекса на реальных статистических данных. Результаты одного из таких исследований приведены на рисунке 7

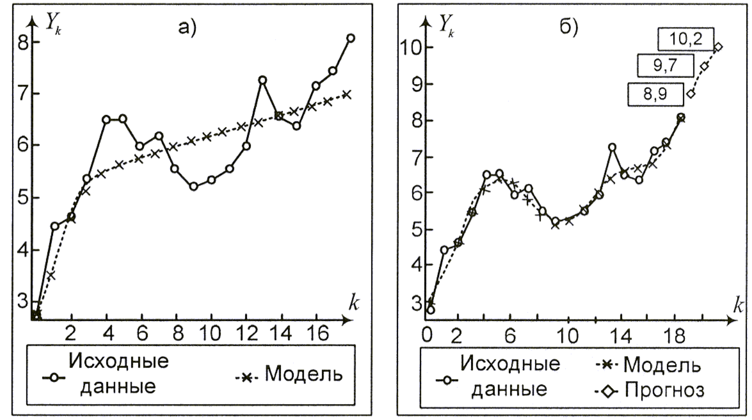

Рис.7. Результаты моделирования динамики роста цен на дизельное топливо: а) логистой с линейным трендом (модель 6) $R^2 = 0,7861$; б) логистой с линейным и двумя синусоидальными трендами (модель 10) $R^2 = 0,9400$.
Кол-во отсчетов N=19. MAPE оценка = 2% (модель 10) при глубине прогноза k=3.

На практике также востребованы и более простые модели (7)-(8) в случаях, когда в исходных данных отсутствуют колебательные или линейно изменяющиеся процессы, например, для пространственных рядов динамики.

ОСНОВНЫЕ ВЫВОДЫ И РЕЗУЛЬТАТЫ

Проведенное исследование позволяет сделать следующие выводы:

1. Предложенная для моделирования и прогнозирования многокомпонентных рядов логистической динамики функция Рамсея позволила на коротких выборках и с большей точностью осуществлять моделирование и прогнозирование.
2. Как показали исследования, предложенные типы многокомпонентных рядов логистической динамики, адекватны реальным экономическим процессам.
3. Для всех рассматриваемых типов многокомпонентных рядов логистической динамики, сконструированы параметрические ARMA-модели.

4. Высокая точность моделирования и прогнозирования рядов логистической динамики обеспечена за счет снижения ряда методических погрешностей и малых вычислительных погрешностей, а также уменьшения необходимых априорных сведений о характеристиках компонент, например, об уровне насыщения.
5. С использованием разработанного программного комплекса проведено исследование построенных моделей и предложенного метода их параметризации на тестовых выборках. Исследование на тестовых выборках показали высокую точность моделирования и прогнозирования в широком диапазоне значений параметров модели, соотношения «шум/сигнал».
6. Проведенное тестирование разработанного программного комплекса на реальных данных, подтвердило высокую точность предложенных методов и средств и возможность их использования для моделирования и прогнозирования экономических процессов различного содержания.
7. Разработанный программный комплекс, может быть использован для моделирования и прогнозирования динамических рядов, изменяющихся по логистическим законам, при малой квалификации пользователя программы.

Перечень публикаций, по теме диссертации

в журналах, рекомендованных ВАК

1. Павлов В.Д. Моделирование и прогнозирование временного ряда суммой логистической, линейной и гармонической компонент на основе ARMA–модели. [Текст] /Семенычев В.К., Павлов В.Д., Семенычев В.В. // Известия Уральского государственного экономического университета. - Екатеринбург, 2009 - № 1(23). – С. 128 - 139.

в других изданиях

2. Павлов В.Д. Эконометрическое моделирование многокомпонентных рядов динамики с использованием логисты Рамсея. [Текст] /Павлов В.Д., Семенычев В.К. // Материалы пятой всероссийской научно-практической конференции «Электронный бизнес: проблемы, развитие и перспективы». - Воронеж, 2006.- С. 125-131.
3. Павлов В.Д. Инструментарий моделирования многокомпонентных рядов динамики с использованием функции Рамсея. [Текст] / Павлов В.Д. // Материалы научно-практической конференции СГАУ,. – Самара: СГАУ, 2006. – С. 81-84.
4. Павлов В.Д. Программный комплекс для моделирования логистической динамики временных рядов. [Текст] / Семенычев В.К., Павлов В.Д..// Материалы научно-практической конференции.- Пенза, 2007.- С. 46-51.
5. Павлов В.Д. Использование математических методов и комплекса программ для моделирования и прогнозирования стоимостных характеристик строительства. . [Текст] / Семенычев В.К., Павлов В.Д. // Сборник статей II Всероссийской научно-практической конференции молодых ученых 1 марта 2007г. - Самара: СМИУ, 2007. - С. 132-143.
6. Павлов В.Д. Компьютерный анализ моделей нелинейной динамики. . [Текст] /Павлов В.Д., Семенычев В.К. // Материалы четвертой Международной научно-практической конференции «Экономическое прогнозирование: модели и методы». - Воронеж: ВГУ, 2008. - С. 255-260.
7. Павлов В.Д. Моделирование и прогнозирование логистических трендов. [Текст] / Павлов В.Д. // Сборник статей III Всероссийской научно-практической конференции молодых ученых 26-27 апреля 2008г. - Самара: СМИУ, 2008.- С. 98-104.
8. Павлов В.Д. Моделирование экономических процессов, изменяющихся по логистическим законам. [Текст] /Павлов В.Д., Семенычев В.В. // Сборник материалов V Международной заочной научно-практической конференции «Интеллектуальные технологии в образовании, экономике и управлении» (ИТОЭУ-2008).- Воронеж, 2008.- С. 39-44.
9. Павлов В.Д. Программный комплекс моделирования и прогнозирования экономических процессов логистической динамики. [Текст] / Павлов В.Д. // Современные проблемы информатизации в экономике и обеспечении безопасности. Сборник трудов. Выпуск 14 (по итогам XIV Международной открытой научной конференции).- Воронеж, 2008. - С. 64-70.

Методические указания

10. Павлов В.Д. Моделирование и прогнозирование нелинейной экономической динамики с логистическим трендом. [Текст] / Семенычев В.К., Павлов В.Д. Семенычев В.В. // Методические указания к выполнению лабораторных работ.- Самара: СМИУ, 2008. - 20 С.

Свидетельство о регистрации программы для ЭВМ

11. Свидетельство № 2009611129 о государственной регистрации программы для ЭВМ. Моделирование и прогнозирование многокомпонентных динамических рядов «Logistic». / Павлов В.Д., Семенычев В.К., Семенычев В.В. - № 2009611129 от 20 февраля 2009.