

ПРОГНОЗИРОВАНИЕ ПОТОКОВ ПРЯМЫХ ИНОСТРАННЫХ ИНВЕСТИЦИЙ В РЕГИОНЫ РФ

УДК 339.727.22

Лилия Валериевна Матраева,
к.э.н., доцент кафедры Финансы и Банковское дело, Российский государственный торгово-экономический университет, Тульский филиал
Тел.: 8 (903) 841-89-49
Эл. почта: matraeva@rambler.ru

Данная статья посвящена проблеме прогнозирования потоков прямых иностранных инвестиций (далее в тексте статьи – ПИИ) в регионы РФ. Оценка прогнозных значений ПИИ проводится как с помощью классических методов прогнозирования (например, АРПСС), так и с помощью новых, которым в настоящее время отдают предпочтение многие исследователи – нейронных сетей.

В статье также проводится сравнительный анализ эффективности использования вышеназванных методов. Дается оценка ситуаций, когда современные методы прогнозирования превосходят классические, что иллюстрируется на реальных примерах.

Ключевые слова: прямые иностранные инвестиции, прогнозирование, АРПСС, нейронные сети, региональная инвестиционная политика.

Lilia V. Matraeva,

PhD in economics, associate professor of chair «the Finance and Banking», Russian State University of Trade and Economics, Tula branch
Тел.: 8 (903) 841-89-49
E-mail: matraeva@rambler.ru

FORECASTING OF FOREIGN DIRECT INVESTMENTS INTO THE REGIONS OF RUSSIAN FEDERATION

This article is dedicated to the problem of forecasting of Foreign Direct Investments (later in article – FDI) into the regions of Russian Federation. Assessment of FDI forecasted levels is done using not only classical forecasting methods (e.g. ARIMA) but with modern methods used by researches currently – neural networks.

Comparative analysis of efficiency of above mentioned methods is done in this article as well. Assessment is done on when modern forecasting methods are more suitable than classical ones which is illustrated on real life examples.

Keywords: foreign direct investments, forecasting, ARIMA, neural networks, regional investment policy.

1. Введение

В настоящее время существует достаточно большой арсенал методов моделирования и прогнозирования рядов динамики. К наиболее популярным среди экономистов можно отнести: методы скользящей средней, методы экспоненциального сглаживания, прогнозирование случайной компоненты по регрессионной модели, методы направленные на выявление сезонной и циклической компоненты и др. Тем не менее, основная масса методов достаточно требовательна к однородности и стандартности данных, что встречается не часто в экономических исследованиях. Кроме того, сами по себе экономические процессы очень часто нельзя однозначно отнести к стационарным, что также существенно затрудняет процедуру прогнозирования. В рамках региональной экономики задача может осложняться еще и ограниченностью выборки, так как многие региональные показатели являются ежегодными, а статистика собственно Российской Федерации ведет отсчет с 1992 года.

В рамках полного исследования были построены прогнозы объемов ПИИ для 80 субъектов РФ за период с 1996 по 2011 годы. Для примера проиллюстрируем полученные результаты на примере Москвы и Московской области.

2. Разведывательный этап анализа рядов динамики

Следует отметить, что в рамках анализируемого периода присутствуют два момента, характеризующихся серьезным изменением макросреды — кризисы 1998 и 2002.

Процесс моделирования и анализа временных рядов целесообразно начинать с описания общей тенденции в целом, то есть идентификации направления и характера роста.

Гипотеза об отсутствии тенденции проверялась с помощью кумулятивного t -критерия, метода сравнения средних уровней временного ряда. Результаты расчетов позволяют сделать вывод о том, что не все анализируемые ряды динамики показали отсутствия тенденций в изменении параметров модели, следовательно, только часть рядов можно идентифицировать как стационарные.

Перед процессом построения прогнозных моделей проведем предварительный этап в рамках которого упорядочим объекты и систематизируем регионы в зависимости от того типа распределения, который в наибольшей степени соответствует исходным данным. Это даст дополнительную информацию в процессе прогнозирования и обеспечит улучшение качества оценки прогнозных значений показателей. Этот этап в последнее время некоторые авторы все чаще рассматривают в качестве обязательного приема анализа экономических и социальных явлений, так как социально-экономические показатели далеко не всегда подчиняются нормальному распределению. [1, 2]. Использование данного этапа в процессе анализа позволит сгруппировать группы регионов по сценариям динамического развития и сопоставить эти сценарии с базисными условиями, предпосылками этого развития.

Исследование динамических рядов проводилось по функциям: линейная, логарифмическая, экспоненциальная, S -функция, полиномиальная (степень от 2 до 6).

Оценка качества и достоверности проводилась с помощью коэффициента детерминации (R^2) и F -критерия Фишера. Из множества возможных уравнений было выбрано то, которое, при прочих равных условиях, имело наименьшую меньшую ошибку оценки.

Описанные сценарии развития отражают основные тенденции, свойственные конкретной территории и на основании их можно в дальнейшем проводить прогнозирование. Функции трендов рассчитывались для всех регионов РФ. Для примера приведем информацию для г. Москвы: экспоненциальная функция тренда с $R^2 = 0,654$, Московской области: кубическая с $R^2 = 0,948$, г. Санкт-Петербурга: кубическая с $R^2 = 0,696$ и Ленинградской области: кубическая с $R^2 = 0,774$. В табл. 1 представлены агрегированные данные по сценариям потоков ПИИ в субъекты РФ.

Таблица 1

Сценарии потоков ПИИ в субъекты РФ

Функция тренда	Количество регионов
Кубическая	40
Экспоненциальная	9
Степенная	7
Квадратичная	4
S-функция	1
Не достаточно положительных значений	11
Не определено	8
Итого	80

Исследовав динамические ряды по 80 субъектам Федерации по обозначенным выше функциям, можно сделать следующие заключения: восходящие тренды продемонстрировали 76,25% регионов РФ; устойчивые тренды продемонстрировали только 33 субъекта РФ (41,25%); основную тенденцию описывают функции полинома 3-ей степени (50,0%).

Очень часто в современном региональном анализе встречается подход, основанный на прогнозировании временных рядов по выделенным группам по инвестиционному потенциалу и климату. Авторы подобных подходов придерживаются того мнения, что поскольку такие группировки объединяют регионы со схожими условиями и предпосылками развития, то и их временные тренды также должны быть схожи.

Проверим эту гипотезу методом от обратного, оценив каким регионам присущи самые устойчивые тренды, характеризующиеся высоким коэффициентом детерминации (более 0,8). К ним относятся Архангельская область, Московская область, Амурская область, Нижегородская область, Республика Адыгея, Псковская область, Калужская область, Еврейская ОА, Белгородская область. Визуального анализа списка областей достаточно, чтобы заключить, что данные регионы относятся к разным группам по инвестиционному потенциалу и климату.

Таким образом, первоначальный разведывательный этап анализа рядов динамики, позволил прийти к заключению, что прогнозирование в рамках групп регионов не представляется целесообразным: прогнозирование следует проводить по каждому региону в отдельности.

Выбор методов прогнозирования для решения поставленной задачи,

должен базироваться на следующих предпосылках:

1. Методы должны учитывать нестационарность описываемых процессов;
2. Методы должны работать с нестандартными данными;
3. Методы должны учитывать ограниченность выборки;
4. Результаты должны с заданной точностью описывать изменение заданных показателей.

Безусловно, выдвинутые требования являются достаточно противоречивыми. И каждое из данных условий в лучшей степени выполняют различные методы. Считается, что нестационарные процессы в лучшей степени описывают авторегрессионные и адаптивные модели. Для нелинейных процессов, имеющих либо хаотическую, либо квазипериодическую, либо смешанную основу в наибольшей степени подходят специальные искусственные сети [2]. Поэтому остановиться однозначно на выборе одного метода в рамках данной задачи не представляется возможным. Для достижения поставленной цели воспользуемся несколькими методами, наиболее полно удовлетворяющими выдвинутым требованиям и сравним качество полученных прогнозов. Поскольку в нашем случае наиболее важной представляется динамика развития исследуемого показателя на конец анализируемого периода, а не средний общий тренд его развития, то в этом случае наиболее целесообразно воспользоваться адаптивными методами прогнозирования, так как «адаптивные модели и методы имеют механизм автоматической настройки на изменение исследуемого показателя. Инструментом прогноза является модель, первоначальная оценка параметров которой производится по нескольким первым наблюдениям. На ее основе делается прогноз, ко-

торый сравнивается с фактическими наблюдениями. Далее модель корректируется в соответствии с величиной ошибки прогноза и вновь используется для прогнозирования следующего уровня, вплоть до исчерпания всех наблюдений» [2]. То есть такие модели направлены на то, чтобы отображать в большей степени тенденцию, сложившуюся на конечный момент анализа.

В качестве базовых моделей в данном случае предлагается воспользоваться являются моделью авторегрессии и проинтегрированного скользящего среднего (модель АР-ПСС) и моделью экспоненциального сглаживания. Обе выбранные модели способны учитывать последние изменения условий прогноза и различную информационную ценность уровней ряда. Свойства этих моделей и особенности их применения достаточно полно описаны в работах [3, 4, 5].

3. Прогнозирование методом АРПСС

В рассматриваемом случае модели АРПСС показали лучшие параметры по сравнению с моделями экспоненциального сглаживания для всех исследуемых временных рядов.

Процесс прогнозирования с помощью модели АРПСС включал несколько классических этапов [5]:

– сначала был проведен анализ функции (АКФ) и частной автокорреляционной функции (ЧАКФ) с целью проверки ряда на стационарность, если к однозначному заключению в данном случае прийти не удалось, использовался расширенный тест Дики-Фуллера или ADF-тест. В случае нестационарного ряда необходимо использовать оператор взятия последовательных разностей, тем самым определяется значение параметра d (порядка разности), это позволит перейти к стационарному ряду.

– на втором этапе эмпирически были подобраны параметры p (порядок авторегрессии) и q (порядок скользящего среднего). Если параметр сезонности не удавалось однозначно определить эмпирически, то проводилась проверка ряда на сезонность (цикличность) на основании соответствия ряда последовательности Фурье. В результате был сформирован набор моделей, подлежащих дальнейшему анализу.

– на заключительном этапе сравнивались прогностические способности

Таблица 2

Основные характеристики прогнозных моделей, описывающих потоки ПИИ для субъектов РФ с R-квадрат больше 98%

Субъект РФ	Тип модели	R-квадрат	Средний относительный модуль ошибки
Алтайский край	ARIMA(0,0,0)	0,999	21,123
Воронежская область	ARIMA(0,0,0)	0,987	123,231
г. Москва	ARIMA(0,0,0)	0,980	24,100
Ивановская область	ARIMA(0,0,0)	1,000	38,882
Иркутская область	ARIMA(0,0,0)	0,993	64,624
Калининградская область	ARIMA(1,1,0)	0,983	48,264
Калужская область	ARIMA(0,1,0)	0,987	133,558
Кемеровская область	ARIMA(0,0,0)	0,993	60,509
Кировская область	ARIMA(0,1,0)	0,995	214,465
Красноярский край	ARIMA(0,0,0)	0,981	64,733
Курганская область	ARIMA(0,1,0)	0,998	11967,680
Липецкая область	ARIMA(0,1,0)	0,998	164,936
Магаданская область	ARIMA(0,0,0)	0,994	20647,512
Нижегородская область	ARIMA(0,1,0)	0,996	47,123
Омская область	ARIMA(0,1,0)	0,998	101,760
Приморский край	ARIMA(0,0,0)	0,985	68,978
Республика Адыгея	ARIMA(0,2,0)	0,999	107,006
Республика Мордовия	ARIMA(0,0,0)	0,996	22,248
Саратовская область	ARIMA(0,0,0)	0,990	79,706
Тамбовская область	ARIMA(0,0,0)	0,999	23,541
Томская область	ARIMA(0,1,0)	1,000	1640,510
Тульская область	ARIMA(0,1,0)	0,981	54,036
Тюменская область	ARIMA(1,1,0)	1,000	4,905
Удмуртская Республика	ARIMA(1,0,0)	1,000	39,203
Челябинская область	ARIMA(0,0,0)	0,996	319,966

трендов представлены на рис. 1 и 2 на примере Москвы и Московской области.

Для оценки адекватности модели были проанализированы ряды остатков. Автокорреляционная функция (АКФ) и частная автокорреляционная функция (ЧАКФ) остатков в построенных функциях попадают в доверительный интервал, что позволяет сделать заключение, что автокорреляция остатков практически отсутствует и, следовательно, остатки имеют случайный характер, визуальный анализ гистограмм распределения остатков подтвердил этот факт.

Как видно из данных, приведенных в таблице 2, среди классических методов прогнозирования наиболее продуктивным применительно к целям исследования оказался метод АР-ПСС (ARIMA), так как в изучаемых динамических рядах присутствует большая зависимость от предыдущих значений и имеет место некоторая цикличность.

Учитывая небольшое количество наблюдений, на основании которых проводилось прогнозирование, резуль-

данных моделей и проводилась проверка их на адекватность на основе анализа остатков. В хорошей модели взаимосвязь остатков низкая, графики автокорреляционной и частной автокорреляционной функции не превышают допустимых границ. [6].

Расчеты производились с использованием прикладных пакетов Statistica 8.0 и SPSS 20.0. Среди осуществленных методов прогноза, исходя из параметров адекватности и точности прогноза, оптимальной моделью представляется прогноз, представленный в табл. 2.

Прогнозирование не проводилось по следующим субъектам РФ в связи с тем, что в более чем 30% наблюдений в этих регионах потоки ПИИ отсутствовали: Кабардино-Балкарская республика, Карачаево-Черкесская республика, Республика Алтай, Республика Бурятия, Республика Дагестан, Республика Игушетия, Республика Калмыкия, Республика Марий Эл, Республика Северная Осетия-Алания, Республика Тыва, Республика Хакасия, Ульяновская область, Чеченская Республика, Чукотский АО.

Прогнозные значения показателей ПИИ для регионов РФ и графики их

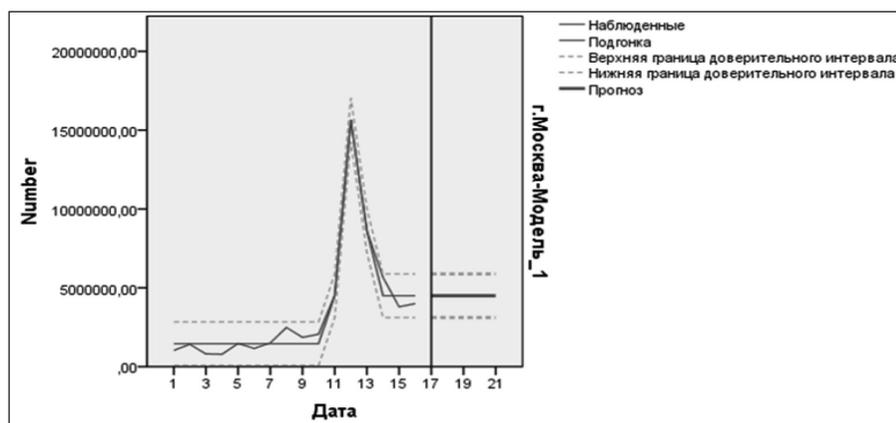


Рис. 1. График тренда ПИИ по г. Москва, 1996–2011 годы

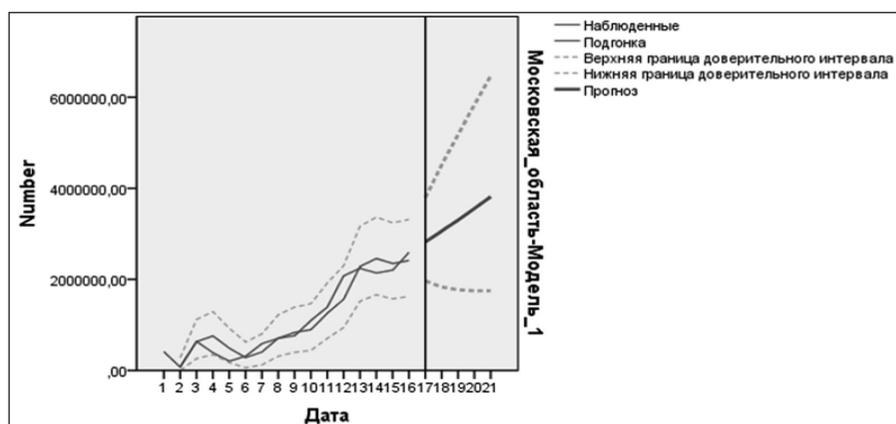


Рис. 2. График тренда ПИИ по Московской области, 1996–2011 годы

таты прогнозных моделей, полученные для большинства регионов РФ, можно признать удовлетворительными, тем более что в данном случае интересовал характер прогнозного тренда, то есть погрешность прогноза являлась не критичным фактором. Тем не менее, для 19 регионов прогностические способности полученных функций не являются достаточно высокими ($R^2 < 0,9$). Попробуем улучшить прогноз, используя инструментальный нейронных сетей.

4. Прогнозирование с использованием методов нейронных сетей

В последнее время в исследованиях методы прогнозирования с помощью нейронных сетей достаточно часто стали встречаться в исследованиях и многие авторы стали критично относиться и пренебрегать классическими методами прогнозирования. Однако, здесь следует особо отметить, что далеко не во всех случаях использование этих методов приводит к существенному улучшению качества моделей.

Последние исследования в данной области показывают, что при наличии явной линейности в задаче, способность нейронных сетей к обобщению оказывается более слабой по отношению к классическим методам. Это объясняется как раз нелинейностью сетей по своей сути [2, 7]. То есть нейронные сети чаще всего имеет смысл использовать при решении таких задач, когда в явном виде определить тип зависимости не представляется возможным или классические методы показали не достаточно удовлетворяющие исследователя результаты.

Теоретические основы и специфика построения моделей достаточно хорошо рассмотрена в работах [2, 8, 9]. В нашем случае основное внимание было сосредоточено на возможности применения методов нейросетевого прогнозирования именно к тем временным рядам, где классические методы показали недостаточно хорошие результаты ($R^2 < 90$). В качестве инструмента реализации был выбран пакет Statistica Neural Networks, поскольку он обладает достаточным для решения данной задачи инструментарием типовых сетей и методов их обучения. При решении поставленной задачи наиболее целесообразно использовать инструмен-

Сравнительные характеристики качества прогнозных моделей региональных потоков ПИИ АРПСС и нейронных сетей (RBF)

Субъект РФ	АРПСС		Нейронные сети – RBF	
	R-квадрат	Средний относительный модуль ошибки	R-квадрат	Средний относительный модуль ошибки
Астраханская область	0,636	1751,4	0,578	774,83
Белгородская область	0,541	60,14	0,733	76,79
Волгоградская область	0,682	189,66	0,531	247,90
г.Санкт Петербург	0,793	42,74	0,453	97,69
Забайкальский край	0,683	161	0,580	3239,00
Краснодарский край	0,759	73,62	0,656	87,20
Курская область	0,527	193,94	0,712	107,91
Ленинградская область	0,634	30,84	0,854	21,73
Мурманская область	0,771	95,7	0,445	136,36
Новгородская область	0,580	69,1	0,629	113,66
Новосибирская область	0,743	15,85	0,649	269,99
Оренбургская область	0,729	466,06	0,379	476,79
Орловская область	0,659	107,66	0,038	119,00
Республика Карелия	0,807	61,69	0,407	82,84
Самарская область	0,724	27,64	0,626	26,07
Сахалинская область	0,892	20,13	0,639	129,55
Свердловская область	0,829	116,14	0,370	160,98
Ставропольский край	0,642	54,34	0,492	90,08
Хабаровский край	0,879	57,59	0,627	62,95

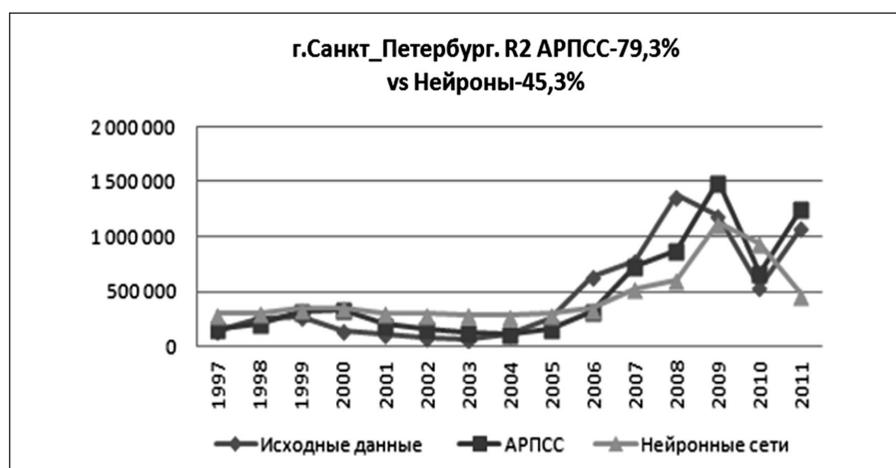


Рис. 3. Прогнозные значения потоков ПИИ, полученные с помощью нейронных сетей, Санкт-Петербург.



Рис. 4. Прогнозные значения потоков ПИИ, полученные с помощью нейронных сетей, Ленинградская область.

тарий радиальных базисных сетей (RBF), так как в данном случае они имеют следующие преимущества перед также реализованными в данном пакете сетями типа многослойных перцептрон (MLP) [2]:

- Более быстрая обучаемость по сравнению с MLP, что является очень важным фактором в наших условиях недостаточности данных;

- Моделирование реализуется с помощью одного слоя (то есть отсутствует задача поиска оптимального количества слоев);

- В выбранном прикладном пакете конструктор сети автоматически достаточно успешно решает одну из самых сложных задач при построении сетей типа RBF – определение числа радиальных элементов, положение их центров и величины отклонений.

- Лучшее качество прогнозирования при небольшом количестве данных.

Сети радиальных базисных функций представляют собой специальный тип нейронных сетей с прямыми связями. Основное их назначение – аппроксимация и интерполяция многомерных функций для решения, в частности, задач прогнозирования. Их математическую основу составляет теория аппроксимации и интерполяции многомерных функций. Сколь угодно точная аппроксимация функций достигается при этом путем комбинации радиально симметричных функций [10].

В таблице 3 приведены сравнительные характеристики моделей АРПСС и RBF для областей, для которых планировалось улучшить качество моделей.

Как видно из приведенных в таблице 3 данных, нейронные сети показали лучшие результаты только для четырех областей из 19 построенных моделей. Высокие ошибки прогнозирования в данном случае объясняются недостаточным количеством данных, в результате построенные нейронные сети оказались «недообученными».

На рис. 3 и 4 показаны примеры того, где нейронные сети показали лучший результат, чем классические методы прогнозирования и наоборот (г. Санкт-Петербург и Ленинградская область).

5. Заключение

Тем не менее, в целом, классические методы прогнозирования при решении поставленной задачи показали более адекватные результаты. Здесь особенно хотелось бы отметить, что в некоторых эконометрических задачах, когда количество входных данных невелико, нейронные сети уступают классическим методам прогнозирования. Их целесообразно использовать только в случае достаточного для качественного обучения количества входных данных, когда тип зависимости определить достаточно сложно.

Для 14 субъектов РФ прогнозные модели не были построены в связи с отсутствием необходимого количества входных данных, что является одной из проблем, препятствующих построению адекватной оценки развития показателя и служит барьером для иностранных инвесторов в их желании осуществлять инвестиции на территории подобных регионов.

Литература

1. Заварина Е.С., Озерова К.А. Межрегиональная дифференциация доходов населения России: возможности анализа и регулирования // Вопросы статистики. – 2010. – №5. – С. 30–39.
2. Морозова Т.Г., Пиккулькин А.В., Тихонов В.Ф., и др. Прогнозирование и планирование в условиях рынка. Учеб. Пособие для вузов. – М.: ЮНИТИ-ДАНА, 1999. – 318 с.
3. Айвазян С.А., Мхитарян В.С. Прикладная статистика и основы эконометрии. – М.: ЮНИТИ, 1998. – 465 с.
4. Елисеева И.И., Курышева С. В., Костеева Т. В., Бабаева И. В., Михайлов Б. А. Эконометрика. Учебник под ред. И. И. Елисеевой. М.: Финансы и статистика, 2003. – 344 с.
5. Мхитарян В. С., Архипова М. Ю., Балаш В. А. Эконометрика: Учебник. – М.: Проспект, 2011. – 384 с.
6. Электронный учебник по статистике. Москва, StatSoft, 2001. – WEB: / home/textbook/default.htm. (вход свободный, дата обращения 02.06.2013)
7. Анисимов В.Н., Соломахо К.Л. Об эффективности модели ARIMA при прогнозировании экономических процессов // Известия Челябинского научного центра, вып. 2 (44), 2009. – С.44–48.

8. Боровиков В. П. Нейронные сети. STATISTICA Neural Networks: Методология и технологии современного анализа данных. – М.: Горячая линия-Телком, 2008. – 392 с.

9. Калан Р. Основные концепции нейронных сетей. – М.: ИД «Вильямс», 2001. – 290 с.

10. Нейросети и нейрокомпьютеры // сайт удаленного обучения Российского государственного социального университета. – URL: <http://21.net78.net/lesson.php?glava=6> (вход свободный, дата обращения 07.07.2013)

References

1. Zavarina E.S. Ozerov K.A. Inter-regional differentiation of incomes in Russia: Opportunities Analysis and Control // Questions of Statistics. – 2010. – № 5. – Pp.30–39.
2. Morozova T.G., Pikulkin A.V., Tikhonov, V.F., etc. Forecasting and Planning in the market. Textbook. Manual for high schools. – Moscow: UNITY –DANA, 1999. – 318 p.
3. Aivazyan S.A., Mkhitarian V.S. Applied statistics and econometrics basis. – Moscow: UNITY, 1998. – 465 p.
4. Eliseev, I.I., Kurysheva S.V., Kosteeva T.V., Babaev I.V., Mikhailov B.A. Econometrics. Textbook ed. I.I. Eliseevoj. Moscow: Finance and Statistics, 2003. – 344 p.
5. Mkhitarian V.S., Arkhipova M.Yu, Balazs V.A. Econometrics: A Textbook. – М.: Prospect, 2011. – 384 p.
6. Electronic textbook on statistics. Moscow, StatSoft, 2001. – WEB: / home / textbook / default.htm. (free admission , date of access 02.06.2013)
7. Anisimov V.N., Solomakho K.L. About the effectiveness of the ARIMA model in forecasting economic processes // Proceedings of the Chelyabinsk Scientific Center, vol. 2 (44), 2009. – P. 44–48.
8. Mushrooms V.P. neural network. STATISTICA Neural Networks: Methodology and technology of modern data analysis. – Moscow: Hotline Telcom 2008. – 392p.
9. Kалан R. Basic concepts of neural networks. - Moscow: Publishing House "Williams", 2001. – 290 p.
10. Neural networks and neurocomputers / site / remote training of the Russian State Social University. – URL: <http://21.net78.net/lesson.php?glava=6> (free admission , date of access 07.07.2013)