

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ  
НАБЕРЕЖНОЧЕЛНИНСКИЙ ИНСТИТУТ (ФИЛИАЛ) ФЕДЕРАЛЬНОГО  
ГОСУДАРСТВЕННОГО АВТОНОМНОГО ОБРАЗОВАТЕЛЬНОГО  
УЧРЕЖДЕНИЯ ПРОФЕССИОНАЛЬНОГО ОБРАЗОВАНИЯ  
«КАЗАНСКИЙ (ПРИВОЛЖСКИЙ) ФЕДЕРАЛЬНЫЙ УНИВЕРСИТЕТ»



**Карамышев А.Н., Сайфутдинова А.А., Исавнин А.Г., Махмутов И.И.,  
Розенцвайг А.К.**

## **ОСНОВЫ НЕЙРОСЕТЕВОГО ПРОГНОЗИРОВАНИЯ В ПАКЕТЕ "STATISTICA NEURAL NETWORKS"**

Электронный образовательный ресурс  
для студентов экономических направлений подготовки

Набережные Челны  
2017 г.

УДК 338.27  
ББК 65.054  
О-75

Основы нейросетевого прогнозирования в пакете "STATISTICA Neural Networks". Электронный образовательный ресурс для студентов экономических направлений подготовки / Составители: Карамышев А.Н., Сайфутдинова А.А., Исавнин А.Г., Махмутов И.И., Розенцвайг А.К. – Набережные Челны: Издательско-полиграфический центр НЧИ К(П)ФУ, 2017. – 42 с.

Электронный образовательный ресурс предназначен для студентов экономического отделения, изучающих дисциплины «Методы социально-экономического прогнозирования», «Методы принятия управленческих решений», «Методы оптимальных решений», «Математические методы и модели в экономике» по образовательным программам бакалавриата и магистратуры.

Рецензенты: доктор экономических наук, профессор Пуряев А.С.  
кандидат педагогических наук, доцент Титова С.В.

Печатается по решению учебно-методической комиссии экономического отделения Набережночелнинского института (филиала) К(П)ФУ (протокол №10 от 15.05.2017 г.)

© Набережночелнинский институт (филиал) ФГАОУ ВО «Казанский (Приволжский) федеральный университет», 2017 г.

## СОДЕРЖАНИЕ

ВВЕДЕНИЕ .....	4
1. ТЕОРЕТИЧЕСКИЕ ОСНОВЫ НЕЙРОННЫХ СЕТЕЙ .....	5
1.1. Модели нейронов .....	5
1.2. Архитектура сетей.....	8
1.3. Обучение нейронных сетей.....	11
2. ПРОГНОЗИРОВАНИЕ ПРИБЫЛИ ОАО «СБЕРБАНК РОССИИ» НА ОСНОВЕ НЕЙРОННЫХ СЕТЕЙ .....	12
2.1. Анализ периодичности исходного ряда.....	12
2.2. Алгоритм прогнозирования ряда на основе нейронных сетей.....	15
2.3. Выбор лучшей сети и анализ качества полученных моделей .....	21
2.4. Полученный прогноз .....	38
ЗАКЛЮЧЕНИЕ .....	40
ЛИТЕРАТУРА .....	41
ПРИЛОЖЕНИЕ 1. Прибыль банка в период с 2006 по 2015 г.г. ....	42

## ВВЕДЕНИЕ

В настоящее время нейронные сети успешно используются для принятия важных управленческих решений в государственном управлении, на предприятиях, в кредитных организациях и банках, т.е. в тех сферах деятельности, где требуется дать прогноз основных трендов и выявить факторы успеха. Нейронные сети являются дополнением классических методов анализа и применяются там, где стандартные процедуры не дают нужного эффекта [3].

В алгоритмах большинства классических (статистических) методов прогнозирования есть этапы построения моделей, основанных на тех или иных предположениях (например, что зависимость между факторами и результатом является линейной или переменные имеют нормальное распределение). Нейросетевой подход свободен от указанных модельных ограничений, подходит как для линейных, так и для сложных нелинейных зависимостей. Также он эффективен при обосновании зависимостей между переменными. Важным преимуществом нейронных сетей является их способность к самообучению.

Основным программным продуктом для проведения нейросетевых исследований является пакет STATISTICA Neural Networks [3]. К преимуществам данного программного продукта можно отнести также мощный арсенал аналитических средств, который позволяет проводить исследования практически всеми известными методами прогнозирования.

# 1. ТЕОРЕТИЧЕСКИЕ ОСНОВЫ НЕЙРОННЫХ СЕТЕЙ

В последнее время интерес к нейронным сетям существенно вырос. Это обусловлено их широкими прикладными возможностями и развитием соответствующего программного обеспечения. Область применения нейронных сетей достаточно обширна: в государственном управлении, банковской деятельности, промышленности, технике, медицине и в других областях, где требуется решить задачи прогнозирования и классификации.

## 1.1. Модели нейронов

Нейрон представляет собой единицу обработки информации в нейронной сети. На рисунке 1.1 представлена модель нейрона, лежащего в основе искусственных сетей. В этой модели можно выделить три основных элемента [2, С. 40]:

1. Набор синапсов (synapse) или связей (connecting link), каждый из которых характеризуется своим весом (weight) или силой (strength). В частности, сигнал  $x_j$  на входе синапса  $j$ , связанного с нейроном  $k$ , умножается на вес  $w_{kj}$ . Первый индекс синаптического веса  $w_{kj}$  относится к рассматриваемому нейрону, а второй – ко входному окончанию синапса, с которым связан данный вес. В отличие от синапсов мозга синаптический вес искусственного нейрона может иметь как положительные, так и отрицательные значения.

2. Сумматор (adder) складывает входные сигналы, взвешенные относительно соответствующих синапсов нейрона. Эту операцию называют линейной комбинацией. [2, С. 41]

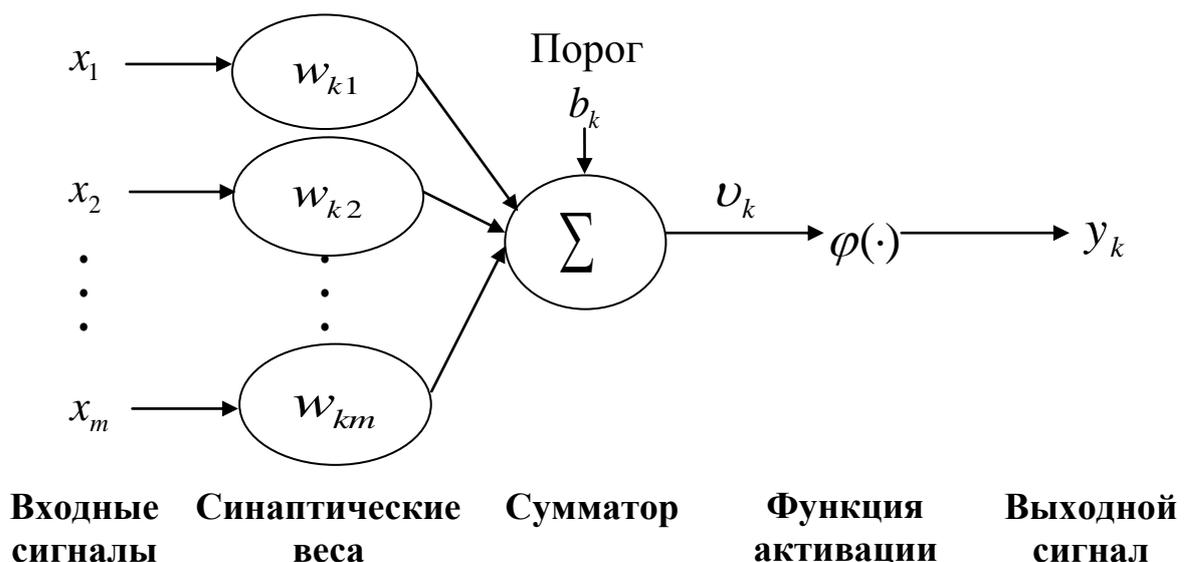


Рисунок 1.1. Нелинейная модель нейрона

1. Функция активации (activation function) ограничивает амплитуду выходного сигнала нейрона. Эта функция также называется функцией сжатия (squashing function). Обычно нормализованный диапазон амплитуд выхода нейрона лежит в интервале  $[0,1]$  или  $[-1,1]$ .

В модель нейрона, представленную на рисунке 1.1, включен пороговый элемент (bias), который обозначен  $b_k$ . Эта величина отражает увеличение или уменьшение входного сигнала, подаваемого на функцию активации [2, С. 41].

В математическом представлении функционирование нейрона  $k$  осуществляется следующим образом:

$$u_k = \sum_{j=1}^m w_{kj} x_j \quad (1.1)$$

$$y_k = \varphi(u_k + b_k) \quad (1.2)$$

где  $x_1, x_2, \dots, x_m$  – входные сигналы;  $w_{k1}, w_{k2}, \dots, w_{km}$  – синаптические веса нейрона  $k$ ;  $u_k$  – линейная комбинация входных воздействий (linear combiner output);  $b_k$  – порог;  $\varphi(\cdot)$  – функция активации (activation function);  $y_k$  – выходной сигнал нейрона. Использование порога  $b_k$  обеспечивает эффект аффинного преобразования (affine transformation) выхода линейного сумматора  $u_k$ . В модели, представленной на рисунке 1.1, постсинаптический потенциал вычисляется следующим образом:

$$v_k = u_k + b_k \quad (1.3)$$

В зависимости от того, какое значение принимает порог  $b_k$ , положительное или отрицательное, индуцированное локальное поле (induced local field) или потенциал активации (activation potential)  $u_k$  нейрона  $k$  изменяется так, как представлено на рисунке 1.2.

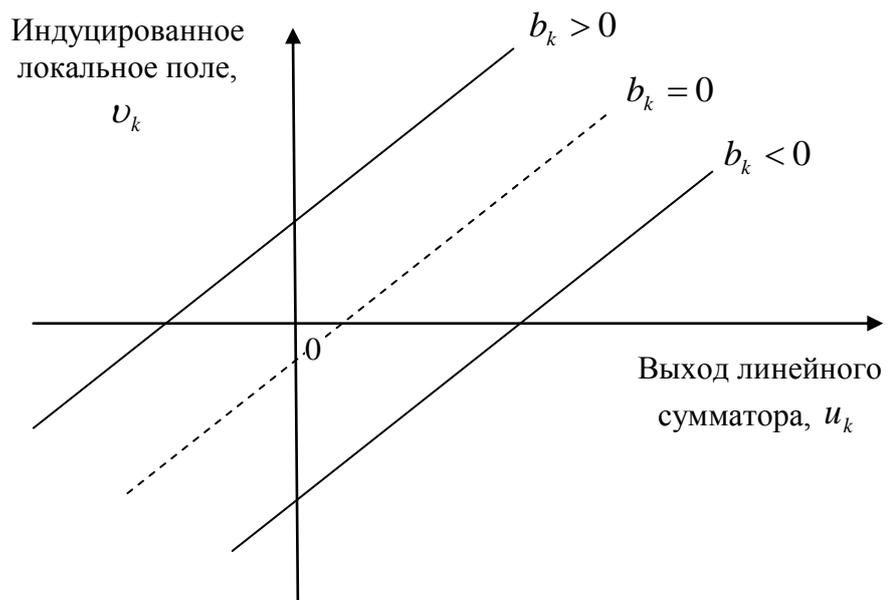


Рис. 1.2. Аффинное преобразование, вызванное наличием порога ( $u_k = 0, v_k = b_k$ )

Поскольку  $b_k \neq 0$ , то график  $v_k$  не проходит через начало координат как график  $u_k$ . Порог  $b_k$  является внешним параметром искусственного нейрона  $k$ . Принимая во внимание выражение (1.3), формулы (1.1), (1.2) можно преобразовать к следующему виду [2, С. 42]:

$$u_k = \sum_{j=0}^m w_{kj} x_j, \quad (1.4)$$

$$y_k = \varphi(v_k) \quad (1.5)$$

В выражении (1.4) добавился новый синапс. Его выходной сигнал равен:  $x_0 = +1$ , а его вес:  $w_{k0} = b_k$ .

## 1.2. Архитектура сетей

Различают следующие виды сетей по данному классификационному признаку:

**1. Однослойные сети прямого распространения.** В простейшем случае в сети существует входной слой (input layer) узлов источника, информация от которого передается на выходной слой (output layer) нейронов (вычислительные узлы) [2, С. 53]. Такая сеть называется сетью прямого распространения, которая представлена на рисунке 1.3. Нейронная сеть называется однослойной (single-layer network), если единственным слоем является слой вычислительных элементов (нейронов) [2, С. 53].

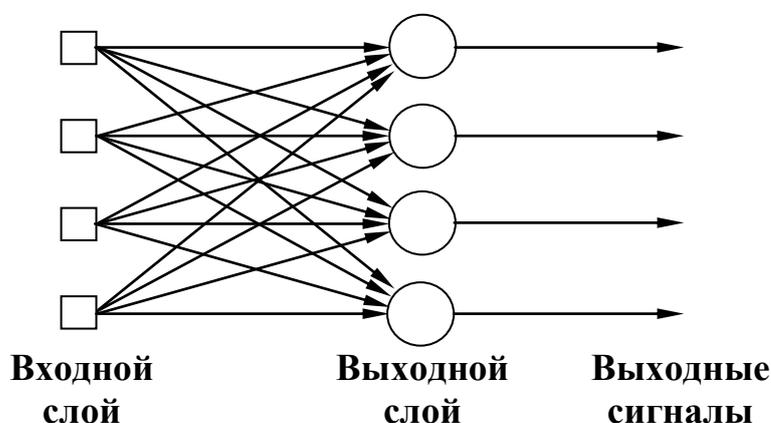


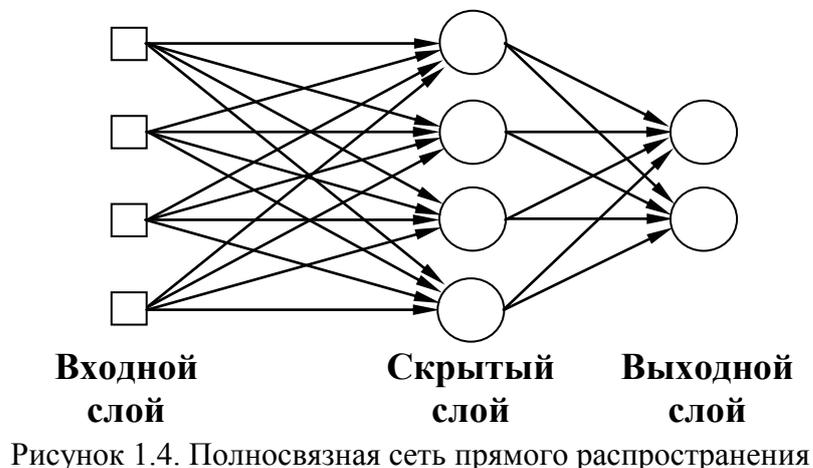
Рисунок 1.3. Сеть прямого распространения с одним слоем нейронов

**2. Многослойные сети прямого распространения.** Эти сети характеризуются наличием одного или нескольких скрытых слоев (hidden layer), т.е. узлов, которые называются скрытыми нейронами (hidden neuron) или скрытыми элементами (hidden unit). Функция скрытых нейронов заключается в посредничестве между внешним входным сигналом и выходом нейронной сети [2, С. 53].

Узлы источника входного слоя сети образуют соответствующий входной вектор, который является входным сигналом, поступающим на нейроны 2 слоя (т.е. первого скрытого слоя). Выходные сигналы второго слоя применяются в качестве входов для третьего слоя и так далее. Таким образом, исходящие сигналы нейронов каждого из слоев сети являются входящими сигналами для следующих за ними слоев.

Сеть, которая представлена на рисунке 1.4, называется сетью 4-4-2, потому что имеет 4 входа, 4 скрытых нейрона, 2 выходных нейрона. В общем случае, сеть прямого распространения с  $m$  входами,  $h_1$  нейронами первого скрытого слоя,  $h_2$  нейронами второго скрытого слоя и  $q$  нейронами выходного слоя называется сетью  $m - h_1 - h_2 - q$  [2, С. 54].

Нейронная сеть, представленная на рисунке 1.4, считается полносвязной (fully connected), поскольку все узлы каждого конкретного слоя соединены со всеми узлами смежных слоев [2, С. 54]. Если некоторые из синаптических связей в сети отсутствуют, то она называется неполносвязной (partially connected) [2, С. 54].



**3. Рекуррентные сети.** Рекуррентная нейронная сеть (recurrent network) отличается от сети прямого распространения наличием, по крайней мере, одной обратной связи (feedback loop) [2, С. 54]. Например, рекуррентная сеть может состоять из единственного слоя нейронов, каждый из которых направляет свой выходной сигнал на входы всех остальных нейронов слоя [2, С. 54]. Архитектура такой сети представлена на рисунке 1.5.

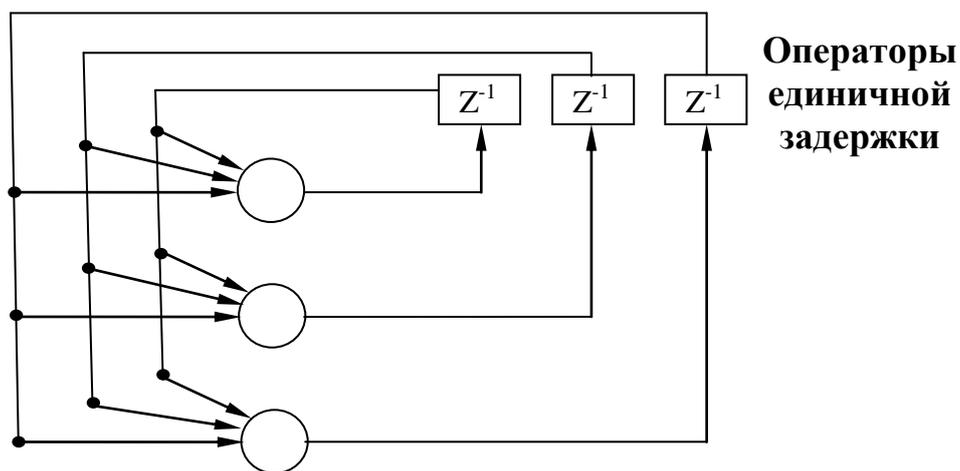


Рисунок 1.5. Рекуррентная сеть без скрытых нейронов и обратных связей нейронов с самими собой

Представим на рисунке 1.6 рекуррентную сеть со скрытыми нейронами.

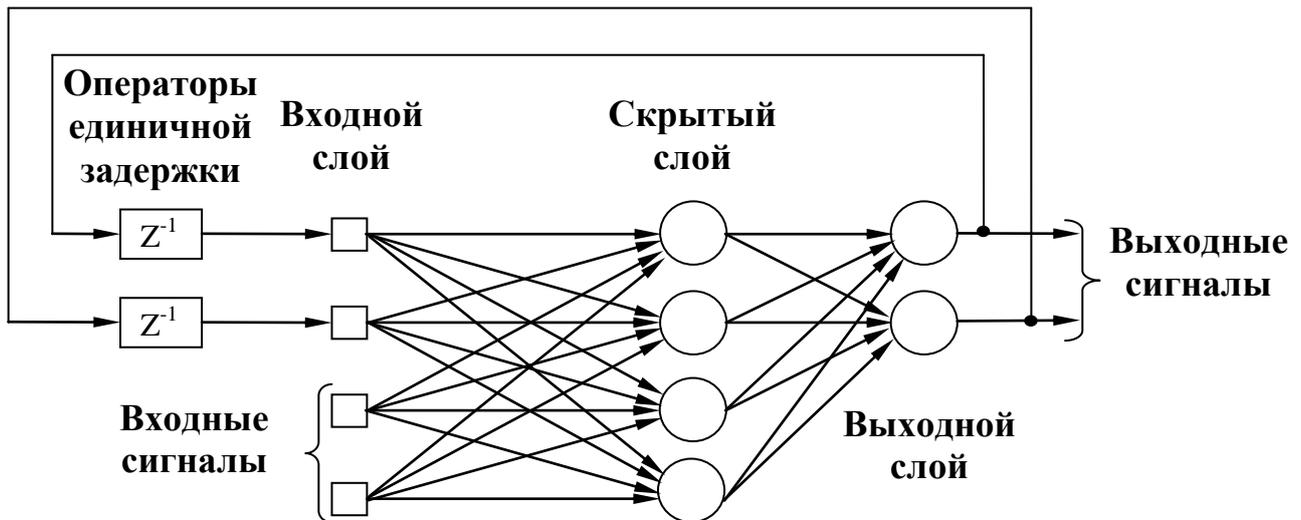


Рисунок 1.6. Рекуррентная сеть со скрытыми нейронами

Наличие обратных связей в сетях 1.5. и 1.6., оказывает непосредственное влияние на способность таких сетей к обучению и их производительность. Более того, обратная связь подразумевает исполнение элементов единичной задержки (unit-delay element) ( $z^{-1}$ ), что приводит к нелинейному динамическому поведению, если в сети содержатся нелинейные нейроны [2, С. 56].

### ***1.3. Обучение нейронных сетей***

Можно выделить алгоритмы обучения с учителем и без учителя.

#### **Обучение с учителем**

Обучение с учителем предполагает, что для каждого входного вектора существует целевой вектор, представляющий собой требуемый выход. Вместе они называются обучающей парой. Обычно сеть обучается на некотором числе таких обучающих пар. Предъявляется входной вектор, вычисляется выход сети и сравнивается с соответствующим целевым вектором. Разность (ошибка) с помощью обратной связи подается в сеть, и веса изменяются в соответствии с алгоритмом, стремящимся минимизировать ошибку. Векторы обучающего множества предъявляются последовательно, ошибки вычисляются и веса подстраиваются для каждого вектора до тех пор, пока ошибка по всему обучающему массиву не достигнет приемлемого низкого уровня [3].

#### **Обучение без учителя**

Обучение без учителя осуществляется без вмешательства внешнего учителя, контролирующего процесс обучения [2]. Существует лишь независимый от задачи уровень качества представления, которому должна научиться нейронная сеть, и свободные параметры сети оптимизируются по отношению к этому уровню.

Обучение без учителя является намного более правдоподобной моделью обучения для биологической системы. Развита Кохоненом Т. и многими другими, она не нуждается в целевом векторе для выходов и, следовательно, не требует сравнения с predetermined идеальными ответами. Обучающее множество состоит лишь из входных векторов. Обучающий алгоритм подстраивает веса сети так, чтобы получались согласованные выходные векторы, т.е. чтобы предъявление достаточно близких входных векторов давало одинаковые выходы. Процесс обучения, следовательно, выделяет статистические свойства обучающего множества и группирует сходные векторы в классы. Предъявление на вход вектора из данного класса даст определенный выходной вектор, но до обучения невозможно предсказать, какой выход будет производиться данным классом входных векторов. Следовательно, выходы подобной сети должны трансформироваться в некоторую понятную форму, обусловленную процессом обучения. На практике достаточно легко идентифицировать установленную сеть связь между входом и выходом.

## 2. ПРОГНОЗИРОВАНИЕ ПРИБЫЛИ ОАО «СБЕРБАНК РОССИИ» НА ОСНОВЕ НЕЙРОННЫХ СЕТЕЙ

Решим задачу прогнозирования прибыли организации на основе автоматизированных нейронных сетей. Исходные данные для прогнозирования представлены в приложении 1.

### 2.1. Анализ периодичности исходного ряда

Для анализа периодичности исходного ряда воспользуемся спектральным анализом Фурье, который позволяет выявить циклически повторяющиеся значения, определить периоды этих циклов. Он часто применяется при обработке результатов исследований (при анализе данных, выявления характерных частот, в целях подавления шума и т.д.). Это бывает удобно для определения, например, сезонных компонент.

В меню «*Statistic*» выберем «*Advanced linear/Nonlinear models*». В выпадающем списке выбираем «*Time Series/Forecasting*» в соответствии с рисунком 2.1.

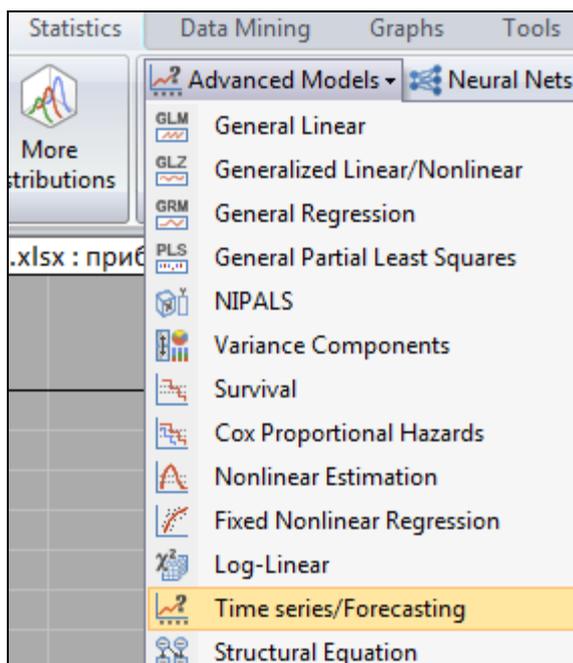


Рисунок 2.1. Модуль «Statistics»

В появившемся диалоговом окне «*Анализ временных рядов (Time series analysis)*» нажимаем на кнопку «*Spectral (Fourier) analysis*» (рисунок 2.2).

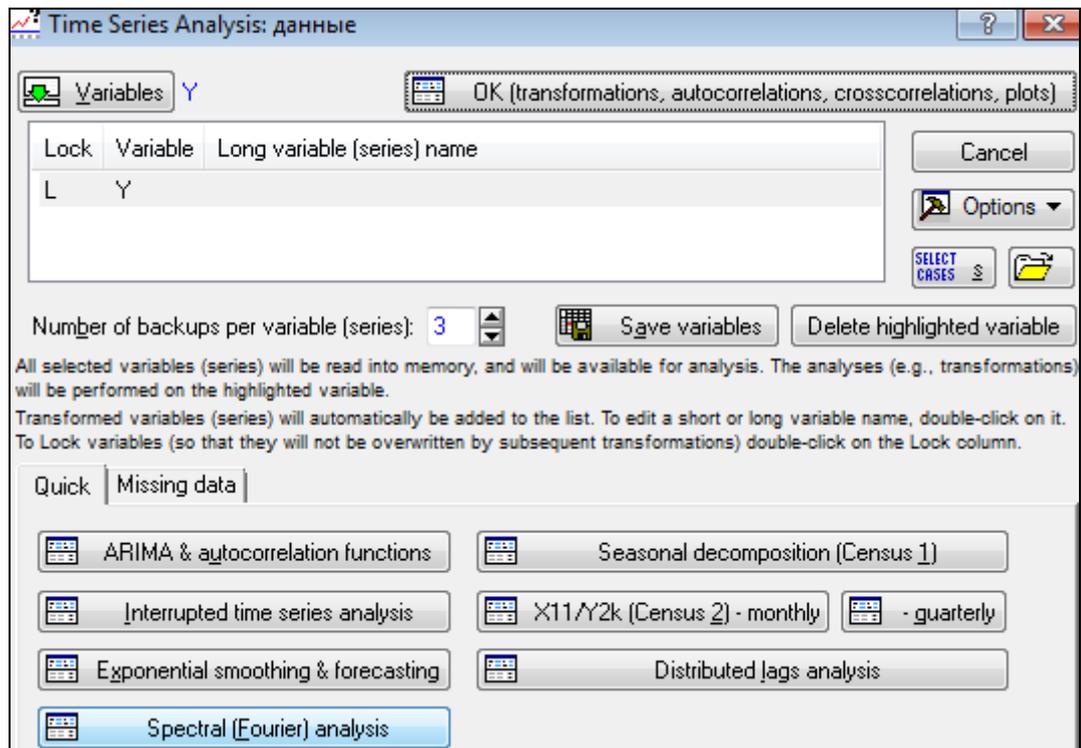


Рисунок 2.2. Диалоговое окно «Анализ временных рядов»

В окне спектрального анализа нажимаем на кнопку «*Single series Fourier analysis*» (см. рисунок 2.3). В открывшемся окне анализа выбираем «*Period*» и нажимаем на кнопку «*Periodogram*» (как на рисунке 2.4).

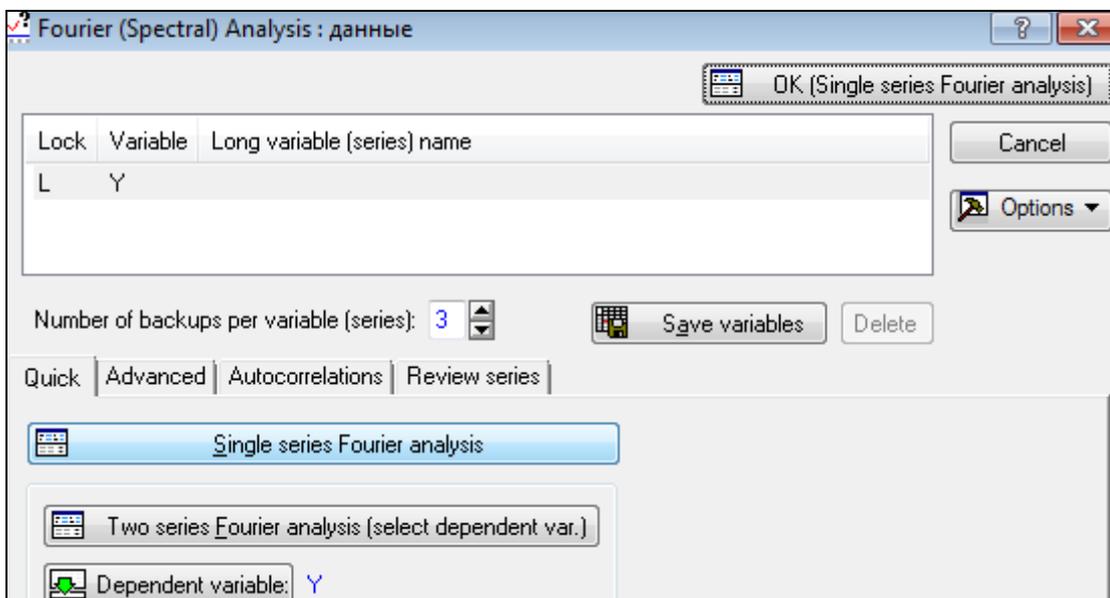


Рисунок 2.3. Диалоговое окно «Анализ Фурье»

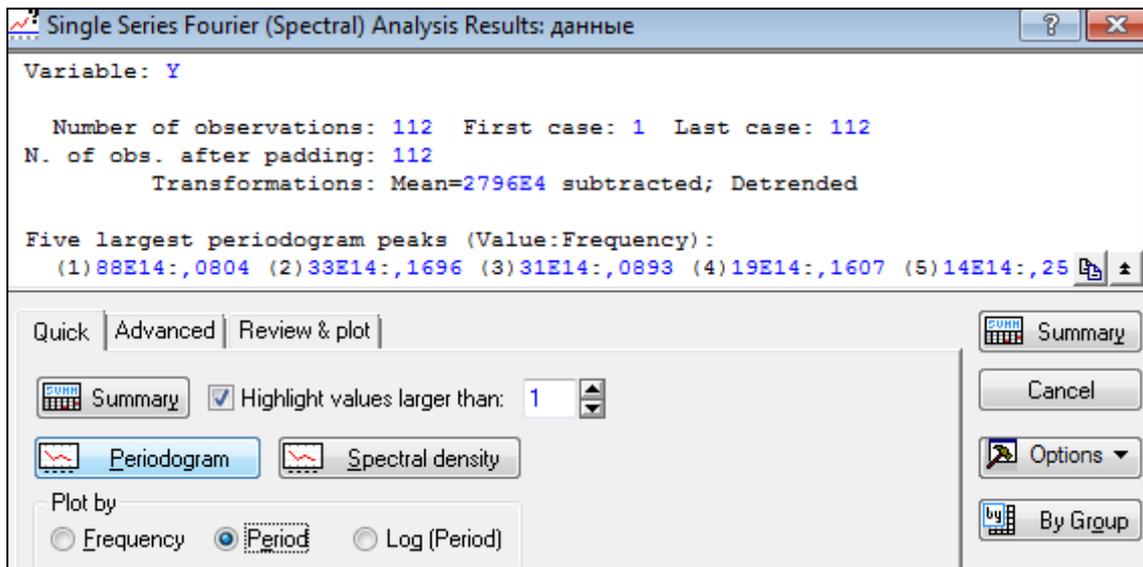


Рисунок 2.4. Спектральный анализ Фурье

В результате получаем следующий график (рисунок 2.5):

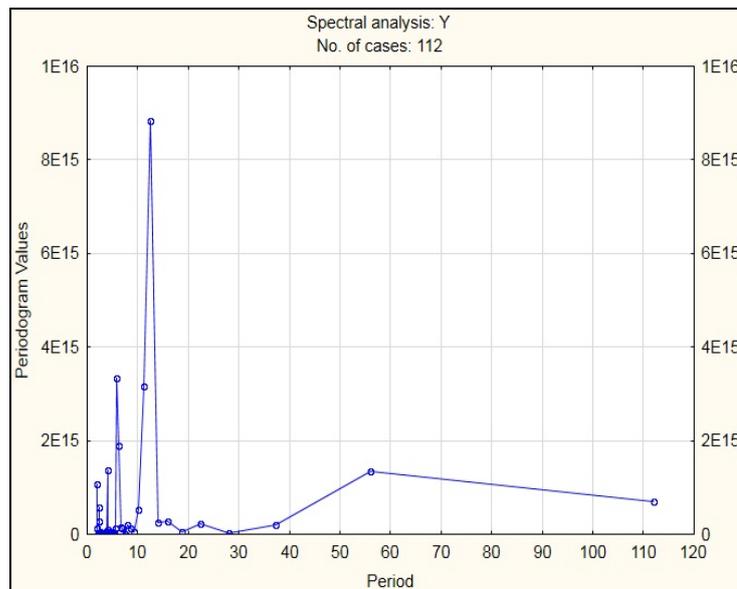


Рисунок 2.5. График «Спектральный анализ Фурье»

Выбросы есть на периодах 2, 4, 6, 12. В данном случае наибольший пик соответствует значению равному 12, т.е. цикл повторяется каждые 12 мес.

## 2.2. Алгоритм прогнозирования ряда на основе нейронных сетей.

Для прогнозирования ряда заходим во вкладку «Statistic» выбираем модуль «Neural Networks». Из предложенного списка выбираем «Time series (regression)» и нажимаем на кнопку «OK» (см. рис. 2.6).

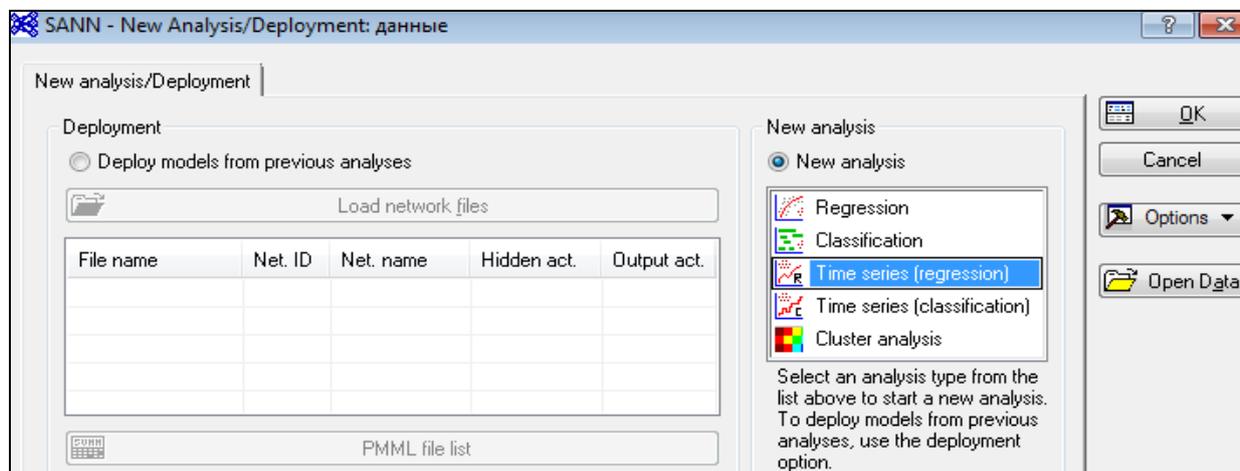


Рисунок 2.6. Задание параметров для анализа временных рядов

### Выбор переменной для анализа

Поскольку у нас осуществляется прогнозирование прибыли на основе отчета о прибыли за прошлые годы, то для дальнейших расчетов выбираем одну переменную «Прибыль (Y)» (см. рис. 2.7). После чего нажимаем на кнопку «OK».

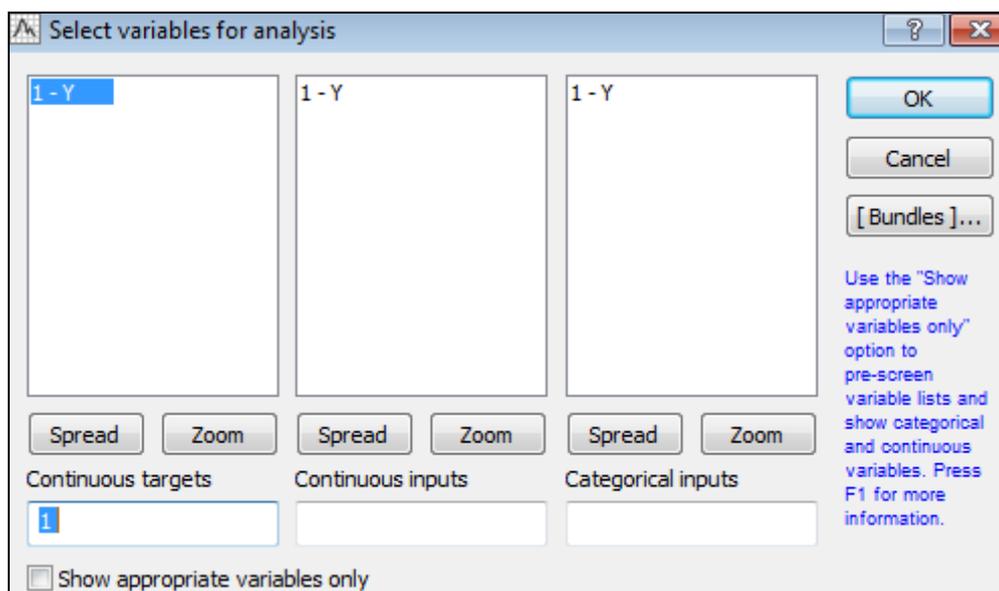


Рисунок 2.7. Окно выбора переменной

### ***Задание условия для кросс-проверки***

Кросс-проверка необходима для оценки точности прогнозирования путем сравнения данных из тестовой и обучающей выборок. Сама модель строится на основе обучающей выборки, а точность прогноза определяется по тестовой выборке. Модель, показывающая результаты близко приближенные к тестовой выборке, считается успешно прошедшей кросс-проверку.

Прогнозирование ряда будем осуществлять на основе 100 месяцев (с 2006 г. по апрель 2014 г. включительно), задаем условия в соответствии с рисунком 2.8.

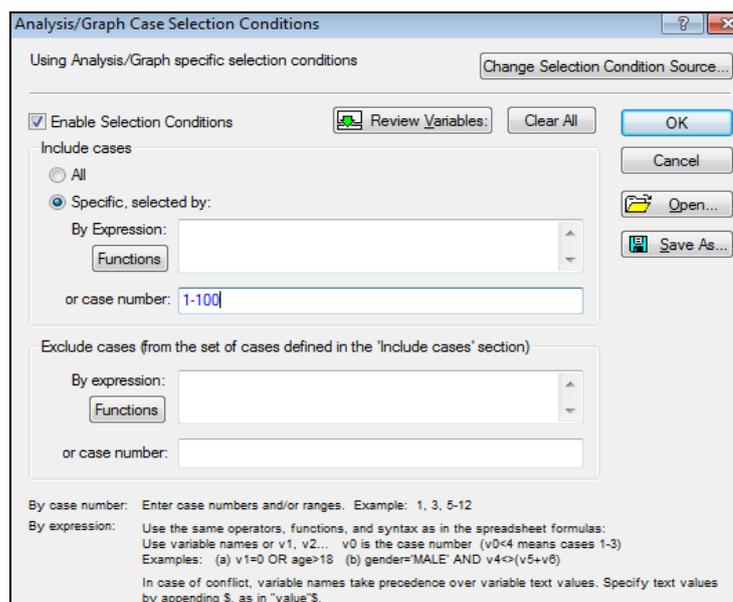


Рисунок 2.8. Условие отбора («Case Selection»)

### ***Выбор стратегии прогнозирования и указание необходимых параметров ее реализации***

Выбираем стратегию «Автоматизированная сеть» (см. рис. 2.9). Во вкладке «Подвыборка (ПНС и АНС) (Sampling (CNN and ANS))» задаем размеры обучающей, контрольной, проверочной подвыборок в соответствии с рисунком 2.10.

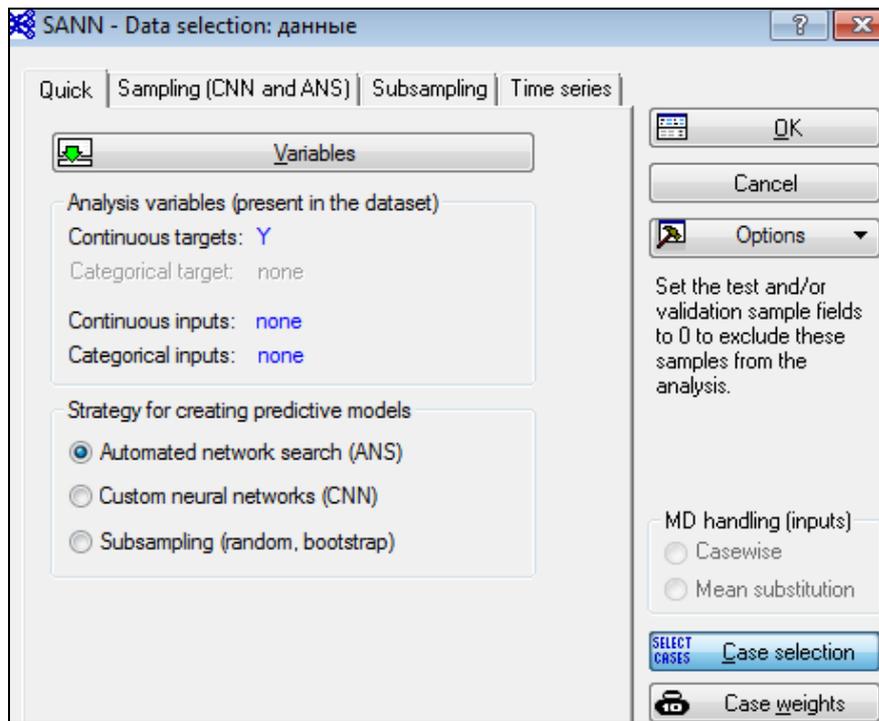


Рисунок 2.9. Выбор стратегии

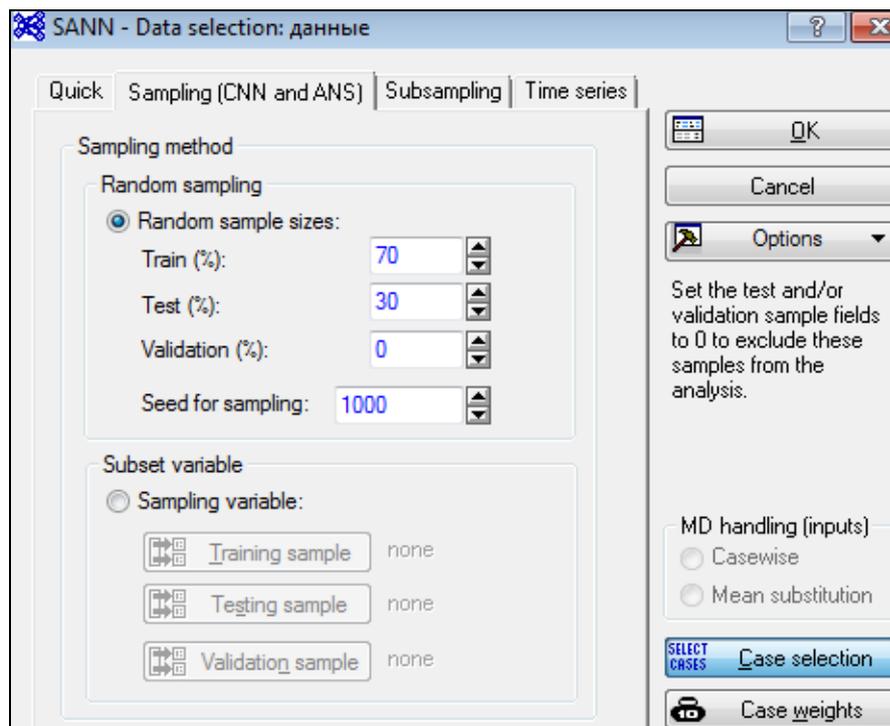


Рисунок 2.10. Вкладка «Подвыборка (ПНС и АНС)»

В данном случае указываем нулевой объем проверочной подвыборки (*Validation*).

Далее указываем размер окна (поскольку при спектральном анализе Фурье был использован шаг 12 месяцев, то здесь возьмем такую же размерность шага) на вкладке «*Time series*», который равен 12 (в соответствии с рисунком 2.11), и нажимаем на кнопку «*OK*». Необходимо отметить, что размер окна должен быть не меньше, чем период ряда.

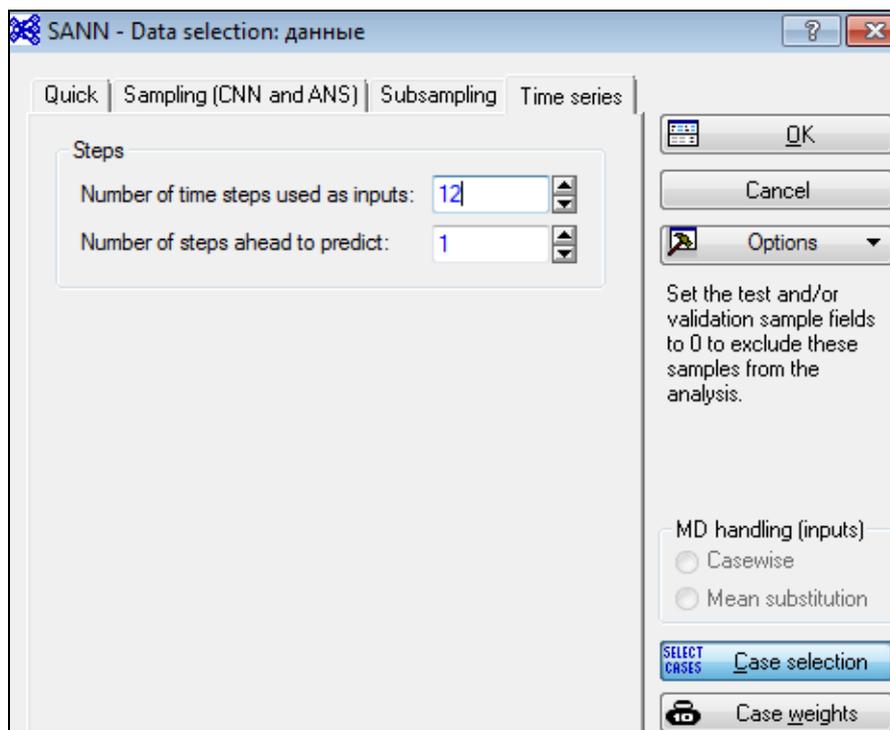


Рисунок 2.11. Выбор размера окна

После задания всех необходимых настроек нажимаем на кнопку «OK».

### ***Выбор типа сети и его параметров***

В открывшемся диалоговом окне «Автоматизированный поиск сети (*Automated network search*)» выбираем тип сети. Обычно для временных рядов, имеющих тренд, применяют тип сети «Многослойный перцептрон (*MLP*)».

Далее задаем на вкладке «*Quick*» (см. рис. 2.12):

- произвольное количество сетей для обучения (в нашем случае -100 ед.);
- указываем для перцептронов интервал скрытых нейронов (1-10);
- количество сохраненных сетей оставляем по умолчанию равным 5 ед.

На вкладке «*Функции активации*» задаем функции для скрытых нейронов в соответствии с рисунком 2.13.

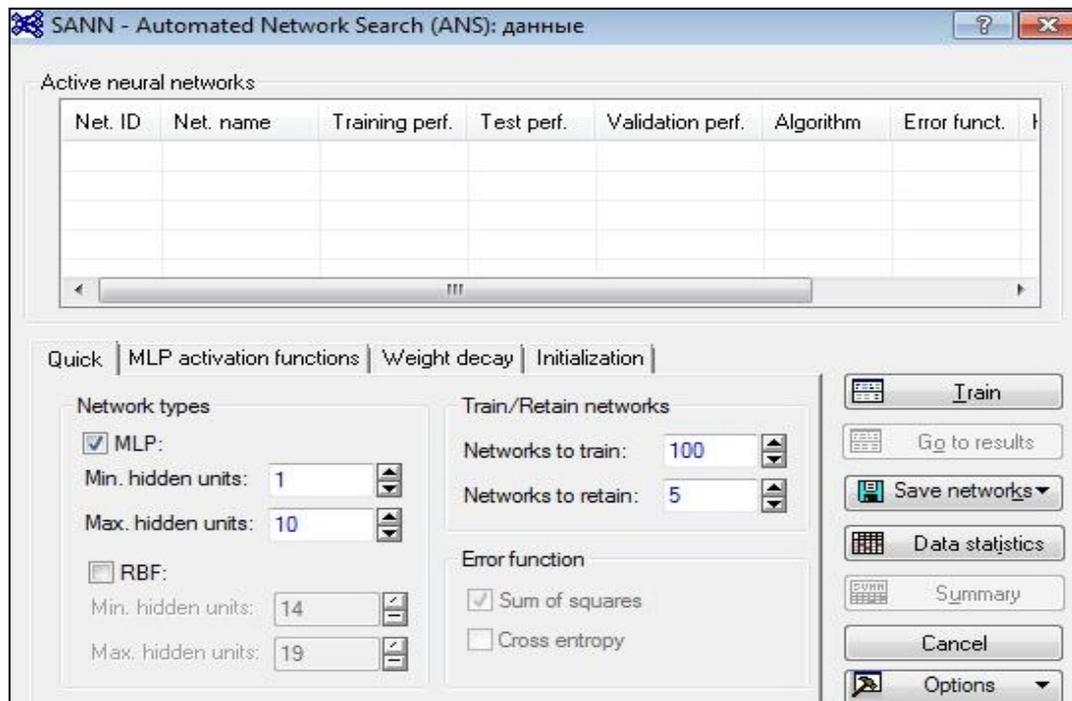


Рисунок 2.12. – Определение параметров сети «Многослойный перцептрон (MLP)»

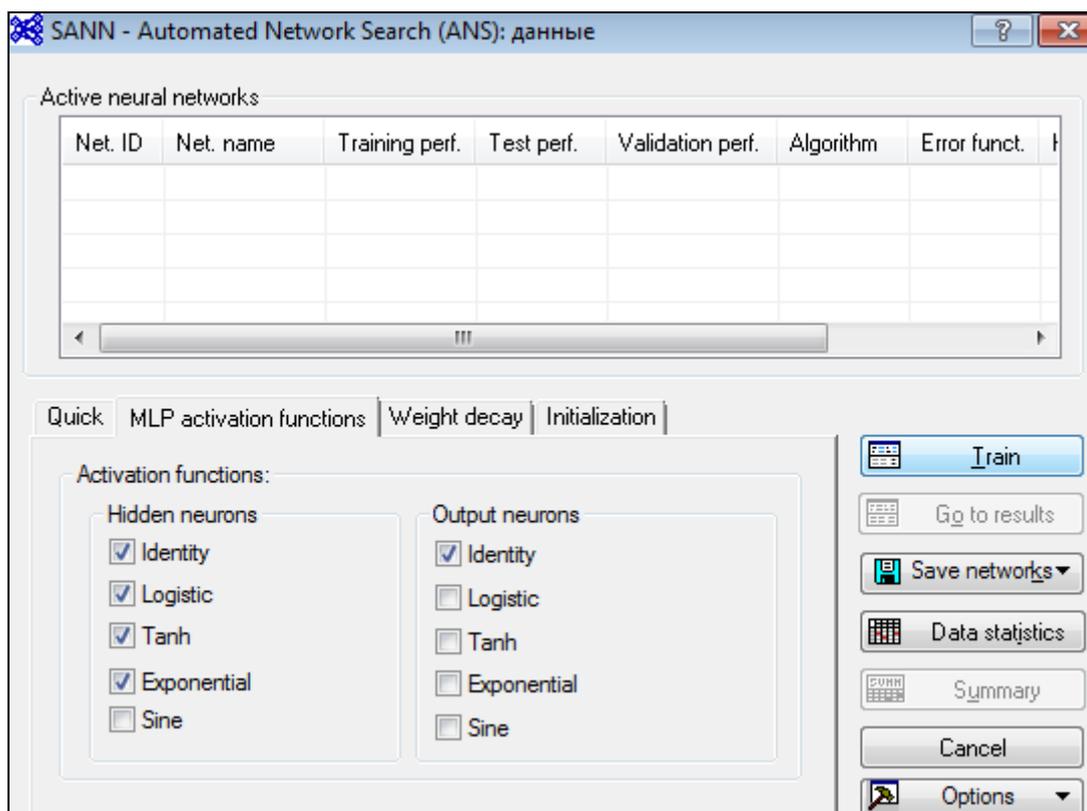


Рисунок 2.13. – Функции активации

На вкладке «Функции активации (MLP activation functions)» необходимо указать функции для входных и выходных нейронов. На практике наиболее часто для активации

входных нейронов используют тождественную и логистическую функции, а для выходных нейронов – тождественную функцию.

### Обучение сети

Для запуска процесса обучения сети нажимаем на кнопку «Train».

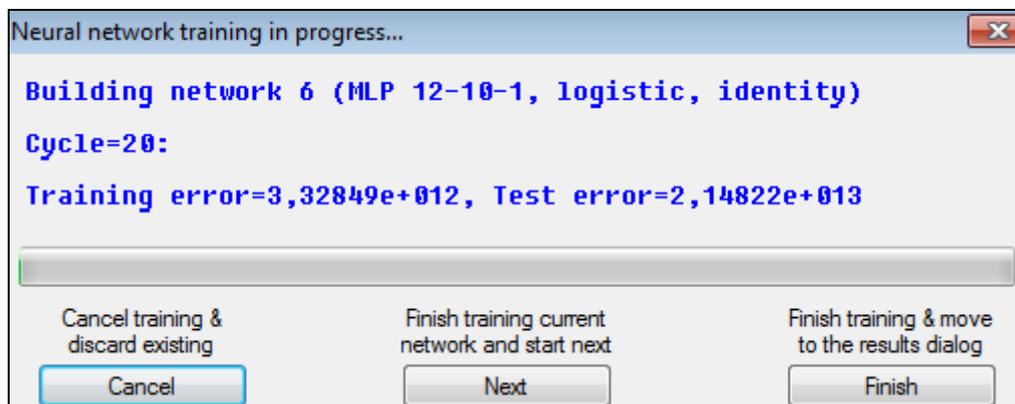


Рисунок 2.14.– Диалоговое окно «Процесс обучения сетей»

В открывшемся окне обучения (рис.2.14) приведена информация об архитектуре сетей (на рисунке 2.14. - сеть 6, модель: MLP 12-10-1), цикле обучения (на рисунке 2.14. – 20 цикл), ошибке на обучающей и на контрольной выборках (на рисунке 2.14. - 3,32... и 2,14... соответственно). В нашем примере мы будем рассчитывать ошибки после проведения кросс-проверки.

Кнопка «Next» предназначена для обучения следующих сетей. Нажатием кнопки «Finish» переходим к окну результатов; кнопки «Cancel» - закрываем данное окно.

### 2.3. Выбор лучшей сети и анализ качества полученных моделей

Окно результатов обучения выглядит следующим образом (см. рисунок 2.15):

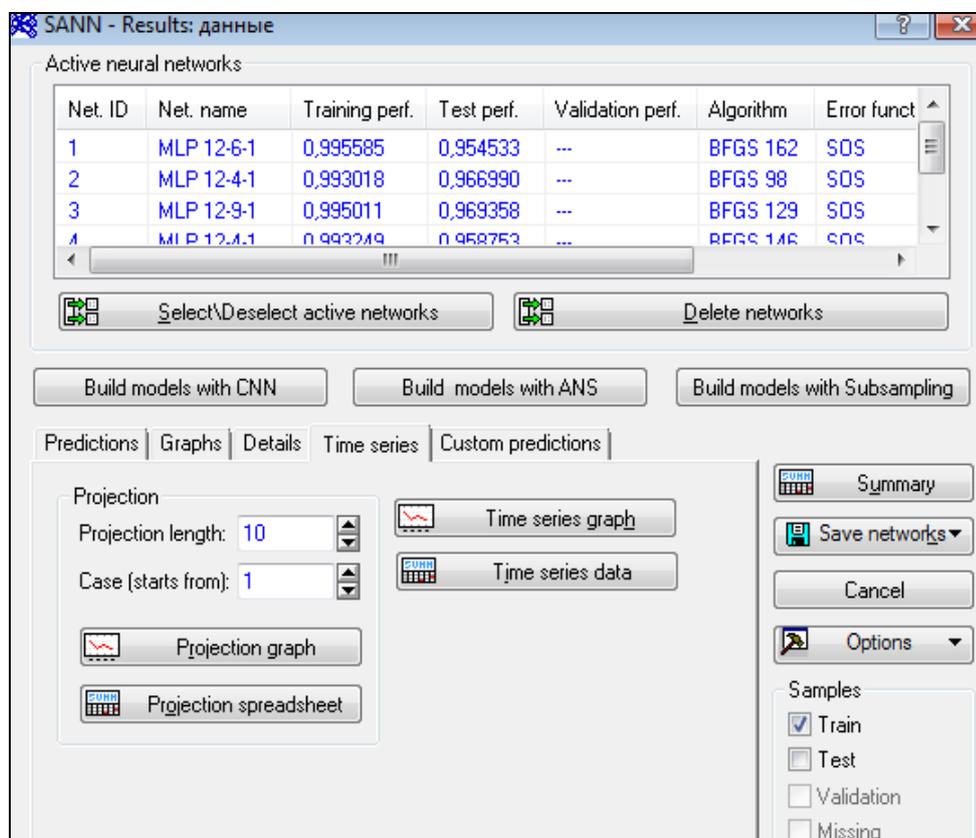


Рисунок 2.15. Диалоговое окно «Результаты»

Для удобства пользователя предусмотрена возможность представления результатов обучения в табличном виде с помощью нажатия кнопки «Summary», кликом на нее. Полученные результаты (см. рисунок 2.16) необходимо подвергнуть анализу. С целью выбора наилучшей сети рассмотрим производительность сетей и распределение остатков.

Summary of active networks (данные)											
Include cases: 1:100											
Index	Net. name	Training perf.	Test perf.	Validation perf.	Training error	Test error	Validation error	Training algorithm	Error function	Hidden activation	Output activation
1	MLP 12-6-1	0,995585	0,954533		7,406488E+10	1,184107E+13		BFGS 162	SOS	Tanh	Identity
2	MLP 12-4-1	0,993018	0,966990		5,842654E+11	9,579460E+12		BFGS 98	SOS	Tanh	Identity
3	MLP 12-9-1	0,995011	0,969358		1,903153E+11	8,224500E+12		BFGS 129	SOS	Logistic	Identity
4	MLP 12-4-1	0,993249	0,958753		5,384272E+11	1,069068E+13		BFGS 146	SOS	Logistic	Identity
5	MLP 12-7-1	0,995218	0,957442		1,470372E+11	1,098427E+13		BFGS 125	SOS	Tanh	Identity

Рисунок 2.16. Таблица результатов обучения

Архитектура сетей соответствует *MLP* (многослойному персептрон), количество входов равно 12, скрытых нейронов 6, 4, 9, 4, 7 для сетей 1-5 соответственно (см. рис. 2.16, столбец «Net.name») и один выход. Производительность сетей достаточно высокая, т.к. зна-

чения, содержащиеся в столбцах 3 и 4, приближены к 1. Если значения в обучающем и контрольном столбцах имеют приблизительно одинаковые значения, то данная модель хорошо прогнозирует исходный ряд.

Также в таблице на рисунке 2.26. представлены контрольные ошибки и ошибки обучения, функции ошибки и активации.

Для анализа качества прогнозов для исходного ряда построенными сетями необходимо сформировать гистограмму распределения остатков и диаграмму рассеяния целевых и выходных переменных.

### ***Распределение остатков***

Во вкладке «*Graphs*» отмечаем по оси X – *Residual (Остатки)*, а Y – *Output (Выходные переменные)*, после чего нажимаем на кнопку «*Histograms of X*» (как на рисунке 2.17).

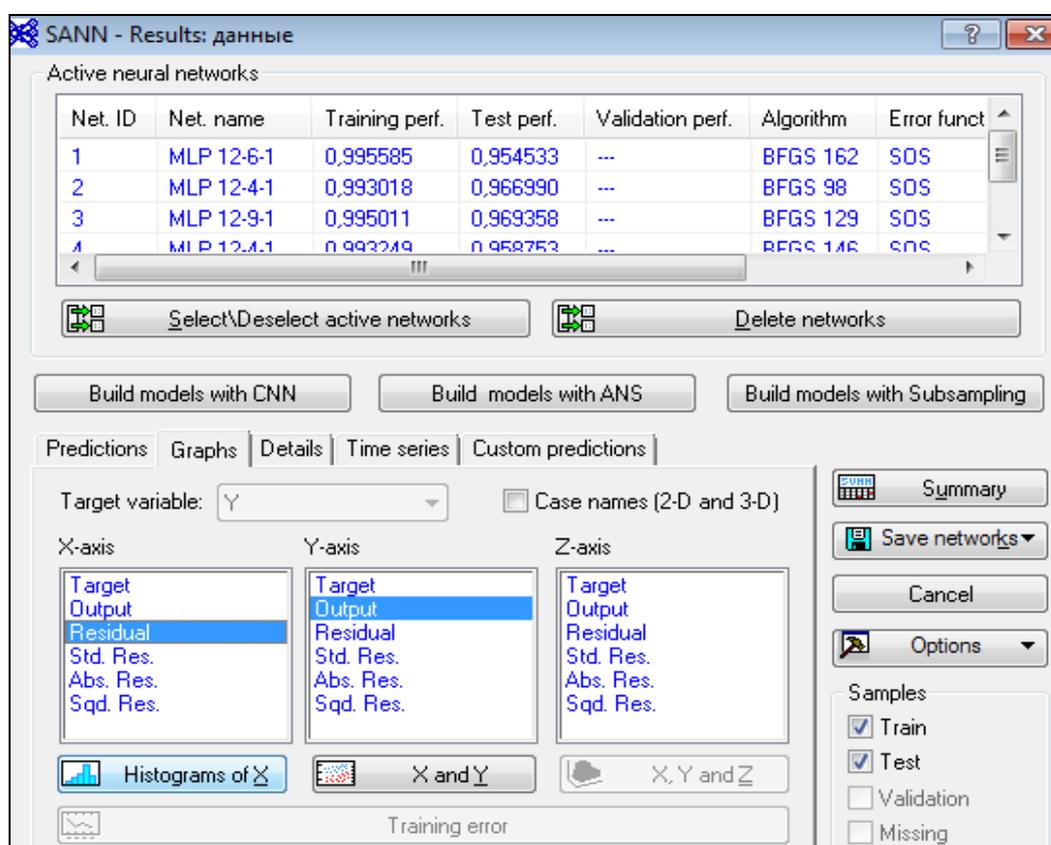


Рисунок 2.17. Задание переменных для построения гистограммы остатков

Необходимо обратить внимание на то, что на панели *Выборка (Samples)* поставлены флажки на обучающую и контрольную подвыборки.

Получаем следующие остатки (представлены на рисунках 2.18. – 2.22):

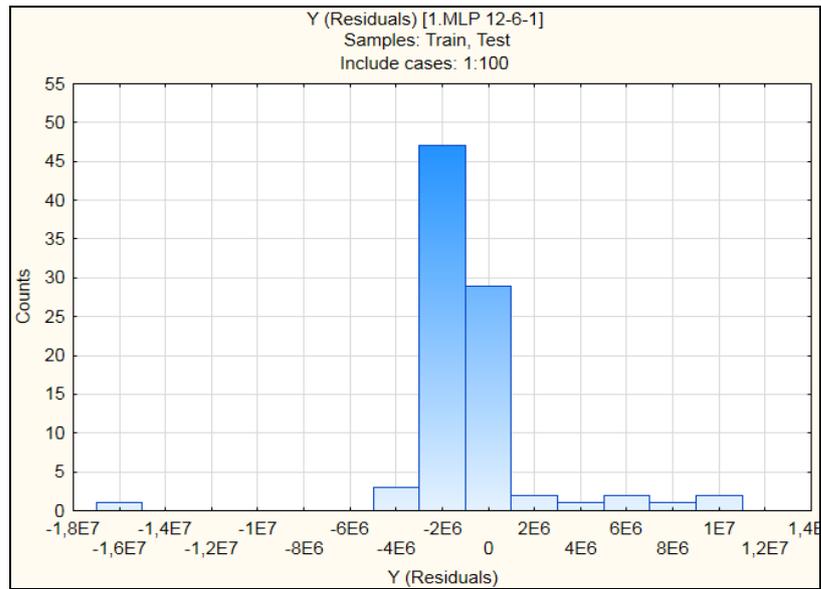


Рисунок 2.18. – Гистограмма распределения остатков сети MLP 12-6-1

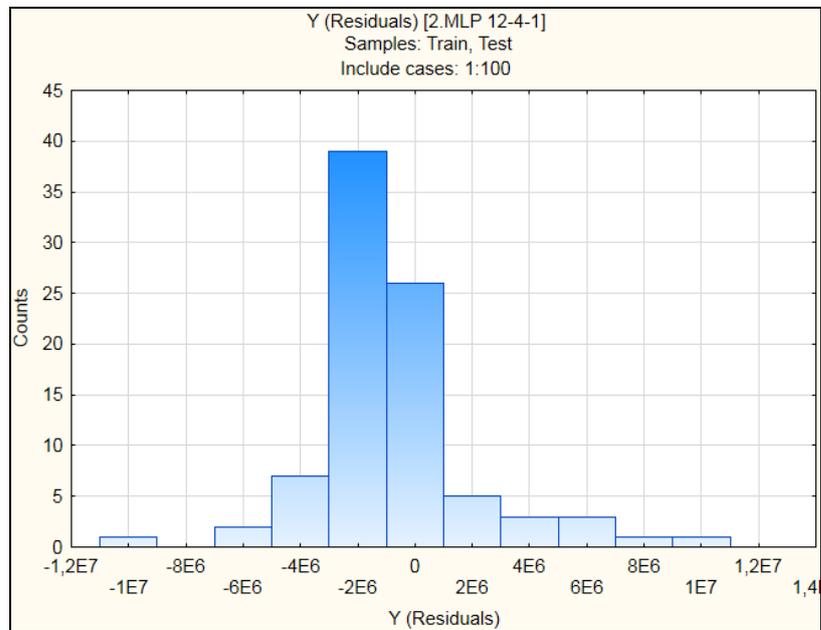


Рисунок 2.19. Гистограмма распределения остатков сети MLP 12-4-1

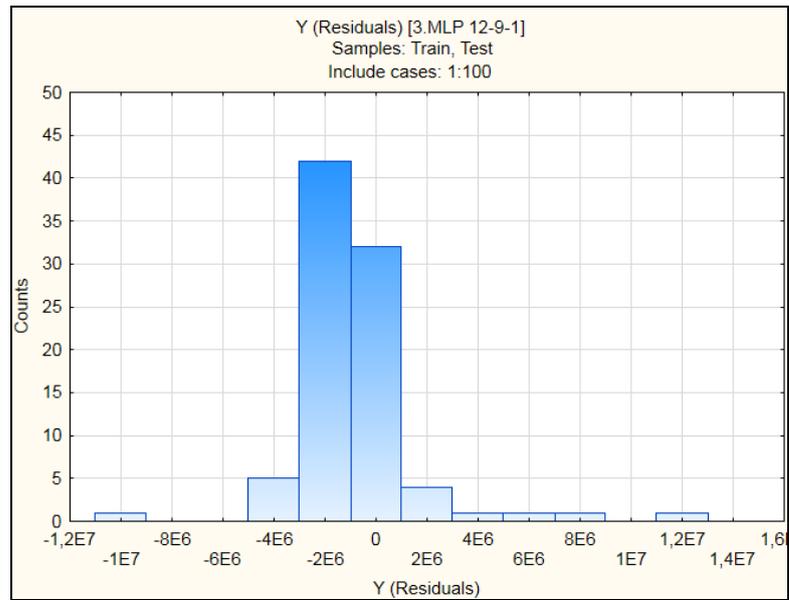


Рисунок 2.20. Гистограмма распределения остатков сети MLP 12-9-1

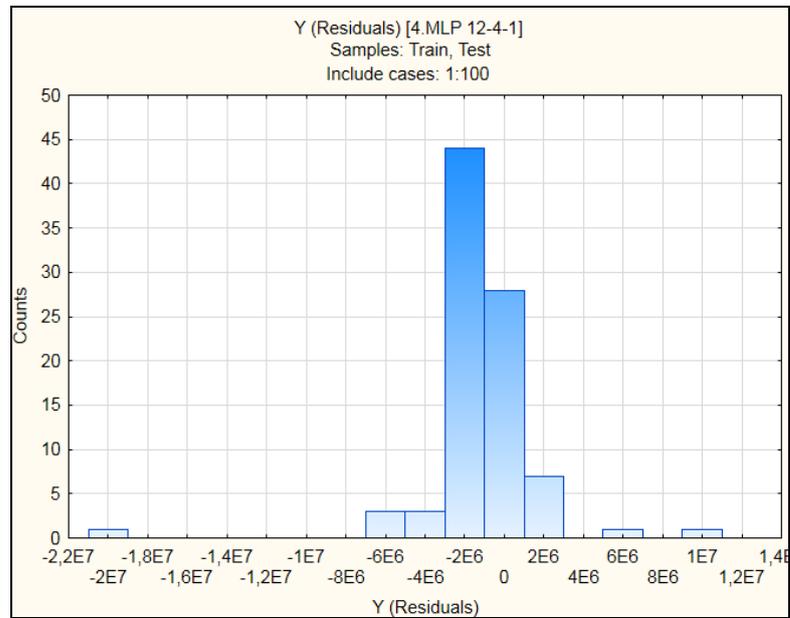


Рисунок 2.21. Гистограмма распределения остатков сети MLP 12-4-1

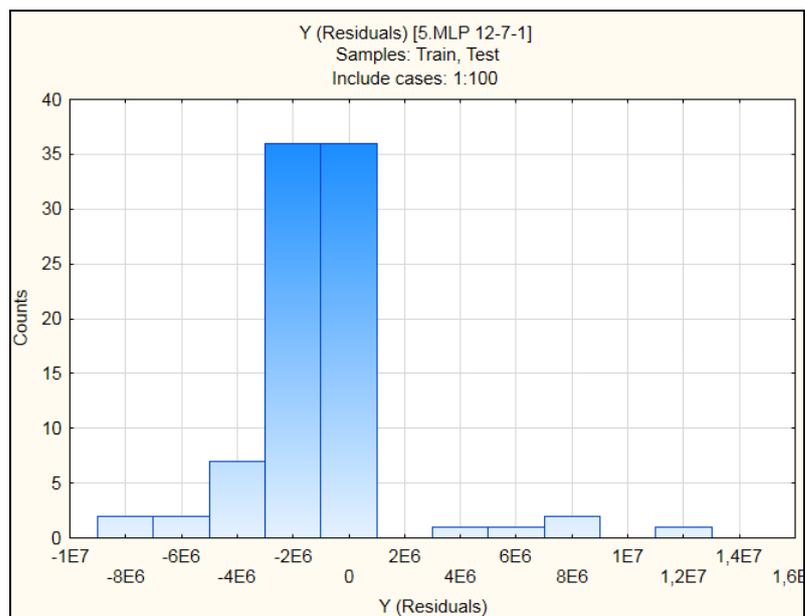


Рисунок 2.22. Гистограмма распределения остатков сети MLP 12-7-1

Показателем хорошего качества построенных моделей является нормальное распределение остатков, т.е. расположение блоков должны соответствовать нормальному распределению Гаусса.

Полученные гистограммы распределения остатков необходимо подвергнуть анализу с целью оценки степени соответствия нормальному распределению. Рассмотрим, в частности, гистограмму распределения остатков сети MLP 12-7-1 (рис. 2.22).

Для проведения этого анализа необходимо получить значения остатков для рассматриваемых сетей. Переходим во вкладку «*Predictions (Прогнозные значения)*», отмечаем ячейку остатков «*Residuals*» (в соответствии с рисунком 2.23.) и нажимаем на кнопку «*Predictions*».

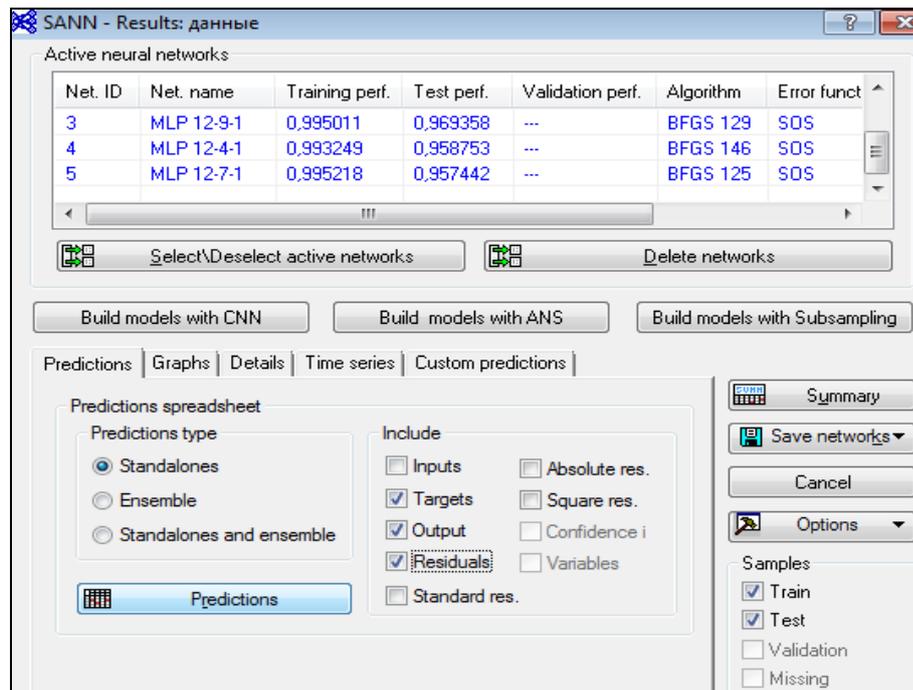


Рисунок 2.23. Вкладка «Predictions»

В результате будет сформирована таблица с прогнозными значениями и остатками для каждой сети (рис.2.24).

В данной таблице первые 12 значений отсутствуют, поскольку они используются для кросс-проверки.

В дереве *Statistica* выбираем нашу таблицу «*Predictions spreadsheet for Y (данные)*» и делаем ее активной для последующей работы только с данными из этой таблицы. Правой клавишей мыши вызываем контекстное меню и выбираем из представленного списка опцию «*Use as Active Input*» (см. рисунок 2.25). После этого значок таблицы «*Predictions spreadsheet for Y*» будет обведен красной рамкой (см. рис. 2.26).

После этого выделяем столбец остатков сети 12-7-1 (*Y – Residual 5.MLP 12-7-1*), вызываем контекстное меню и выбираем строку «*Graphs of Block Data*» / «*Histogram: Block columns*» (как на рис.2.26).

Predictions spreadsheet for Y (данные)													
Samples: Train, Test													
Include cases: 1:100													
Case name	Y Target	Y - Output 1. MLP 12-6-1	Y - Residuals 1. MLP 12-6-1	Y - Output 2. MLP 12-4-1	Y - Residuals 2. MLP 12-4-1	Y - Output 3. MLP 12-9-1	Y - Residuals 3. MLP 12-9-1	Y - Output 4. MLP 12-4-1	Y - Residuals 4. MLP 12-4-1	Y - Output 5. MLP 12-7-1	Y - Residuals 5. MLP 12-7-1	Y - Output	Y - Residuals
Jan-06	9654780												
Feb-06	9254471												
Mar-06	9225123												
Apr-06	13540789												
May-06	13621457												
Jun-06	13620145												
Jul-06	19256047												
Aug-06	19365478												
Sep-06	19874540												
Oct-06	25254708												
Nov-06	25250450												
Dec-06	24000017												
Jan-07	10550478	11065498	-515020	9219325	1331153	10704588	-154110	9349079	1201399	10407131	143347		
Feb-07	10995642	10409731	585911	11081092	-85450	10875067	120575	11986344	-990702	10168929	826713		
Mar-07	11000000	11231096	-231096	13066373	-2066373	11921092	-921092	12302531	-1302531	11642516	-642516		
Apr-07	15000125	15835152	-835027	16027133	-1027008	16091718	-1091593	15378860	-378735	16137643	-1137518		
May-07	15001254	14925512	75742	15630064	-628810	15155114	-153860	15183739	-182485	16300528	-1299274		
Jun-07	15024700	16738093	-1713393	17433263	-2408563	16990585	-1965885	16664378	-1639678	17050646	-2025946		
Jul-07	21001457	24908550	-3907093	25701876	-4700419	23946583	-2945126	24186710	-3185253	24712956	-3711499		
Aug-07	21012540	21068206	-55666	20929383	83157	20631602	380938	21143689	-131149	21188087	-175547		
Sep-07	21584602	22358912	-774310	24191157	-2606555	23997869	-2413267	22898881	-1314279	24282002	-2697400		
Oct-07	30000478	30171478	-171000	29605735	394743	30483025	-482547	28120041	1880437	29510261	490217		
Nov-07	30000784	29254016	746768	29191454	809330	30366740	-365956	30398138	-397354	30780065	-779281		
Dec-07	30001254	29675507	325747	31067342	-1066088	29981762	19492	31287503	-1286249	29773010	228244		

Рисунок 2.24. Таблица «Predictions»

Predictions spreadsheet for Y (данные)												
Samples: Train, Test												
Include cases: 1:100												
Y Target	Y - Output 1. MLP 12-6-1	Y - Residuals 1. MLP 12-6-1	Y - Output 2. MLP 12-4-1	Y - Residuals 2. MLP 12-4-1	Y - Output 3. MLP 12-9-1	Y - Residuals 3. MLP 12-9-1	Y - Output 4. MLP 12-4-1	Y - Residuals 4. MLP 12-4-1	Y - Output 5. MLP 12-7-1	Y - Residuals 5. MLP 12-7-1		
9654780												
9254471												
9225123												
13540789												
13621457												
13620145												
19256047												
19365478												
19874540												
25254708												
25250450												
24000017												
10550478	11065498	-515020	9219325	1331153	10704588	-154110	9349079	1201399	10407131	143347		
10995642	10409731	585911	11081092	-85450	10875067	120575	11986344	-990702	10168929	826713		
11000000	11231096	-231096	13066373	-2066373	11921092	-921092	12302531	-1302531	11642516	-642516		
15000125	15835152	-835027	16027133	-1027008	16091718	-1091593	15378860	-378735	16137643	-1137518		
15001254	14925512	75742	15630064	-628810	15155114	-153860	15183739	-182485	16300528	-1299274		
15024700	16738093	-1713393	17433263	-2408563	16990585	-1965885	16664378	-1639678	17050646	-2025946		
21001457	24908550	-3907093	25701876	-4700419	23946583	-2945126	24186710	-3185253	24712956	-3711499		
21012540	21068206	-55666	20929383	83157	20631602	380938	21143689	-131149	21188087	-175547		
21584602	22358912	-774310	24191157	-2606555	23997869	-2413267	22898881	-1314279	24282002	-2697400		
30000478	30171478	-171000	29605735	394743	30483025	-482547	28120041	1880437	29510261	490217		
30000784	29254016	746768	29191454	809330	30366740	-365956	30398138	-397354	30780065	-779281		
30001254	29675507	325747	31067342	-1066088	29981762	19492	31287503	-1286249	29773010	228244		

Рисунок 2.25. Опция «Use as Active Input»

Workbook

- SANN (данные)
- STATISTICA NN Results Dialog
- Summary of active networks (данные)
- Y (Residuals) [1.MLP 12-6-1] Samples: Train
- Y (Residuals) [2.MLP 12-4-1] Samples: Train
- Y (Residuals) [3.MLP 12-9-1] Samples: Train
- Y (Residuals) [4.MLP 12-4-1] Samples: Train
- Y (Residuals) [5.MLP 12-7-1] Samples: Train
- Predictions spreadsheet for Y (данные)

Predictions spreadsheet for Y (данные)

Samples: Train, Test  
Include cases: 1:100

Case name	Y - Target	Y - Output 1. MLP 12-6-1	Y - Residuals 1. MLP 12-6-1	Y - Output 2. MLP 12-4-1	Y - Residuals 2. MLP 12-4-1	Y - Output 3. MLP 12-9-1	Y - Residuals 3. MLP 12-9-1	Y - Output 4. MLP 12-4-1	Y - Residuals 4. MLP 12-4-1	Y - Output 5. MLP 12-7-1	Y - Residuals 5. MLP 12-7-1
Jan-06	9654780	11065498	-515020	9219325	1331153	10704588	-15411				
Feb-06	9254471	10995642	10409731	585911	-85450	10875067	120575	#####	-990702	10168929	82
Mar-06	9225123	11000000	11231096	-231096	#####	11921092	-921092	#####	-1302531	11642516	-64
Apr-06	13540789	15000125	15835152	-835027	#####	16091718	-1091593	#####	-378735	16137643	-113
May-06	13621457	15001254	14925512	75742	#####	15155114	-153860	#####	-182485	16300528	-129
Jun-06	13620145	15024700	16738093	-1713393	#####	16990585	-1965885	#####	-1639678	17050646	-202
Jul-06	19256047	21001457	24908550	-3907093	#####	23946583	-2945126	#####	-3185253	24712956	-371
Aug-06	19365478	21012540	21068206	-56666	#####	20631602	380938	#####	-131149	21188087	-17
Sep-06	19874540	21584602	22358912	-774310	#####	23997869	-2413267	#####	-1314279	24282002	-263
Oct-06	25254708										
Nov-06	25250450										
Dec-06	24000017										
Jan-07	10550478										
Feb-07	10995642										
Mar-07	11000000										
Apr-07	15000125										
May-07	15001254										
Jun-07	15024700										
Jul-07	21001457										
Aug-07	21012540										
Sep-07	21584602										

Statistics of Block Data

Graphs of Block Data

- Graphs of Input Data
- Cut Ctrl+X
- Copy Ctrl+C
- Copy with Headers
- Paste
- Paste Special...
- Add Variables...
- Delete Variables...
- Move Variables...
- Copy Variables...
- Variable Specs...
- Bundle Manager...
- Fill/Standardize Block
- Clear
- Format
- Marking Cells

Histogram: Block Columns

- Histogram: Entire Columns
- Line Plot: Entire Columns
- Line Plot: Block Rows
- Box Plot: Block Columns
- Normal Prob. Plot: Block Columns
- Custom Graph from Block by Column
- Custom Graph from Block by Row
- Custom Graph for Entire Column
- Custom Graph for Entire Row
- Customize List...

Рисунок 2.26. Остатки сети MLP 12-7-1

В результате получаем гистограмму остатков сети 12-7-1 и график нормального распределения, представленный красной линией (рис.2.27).

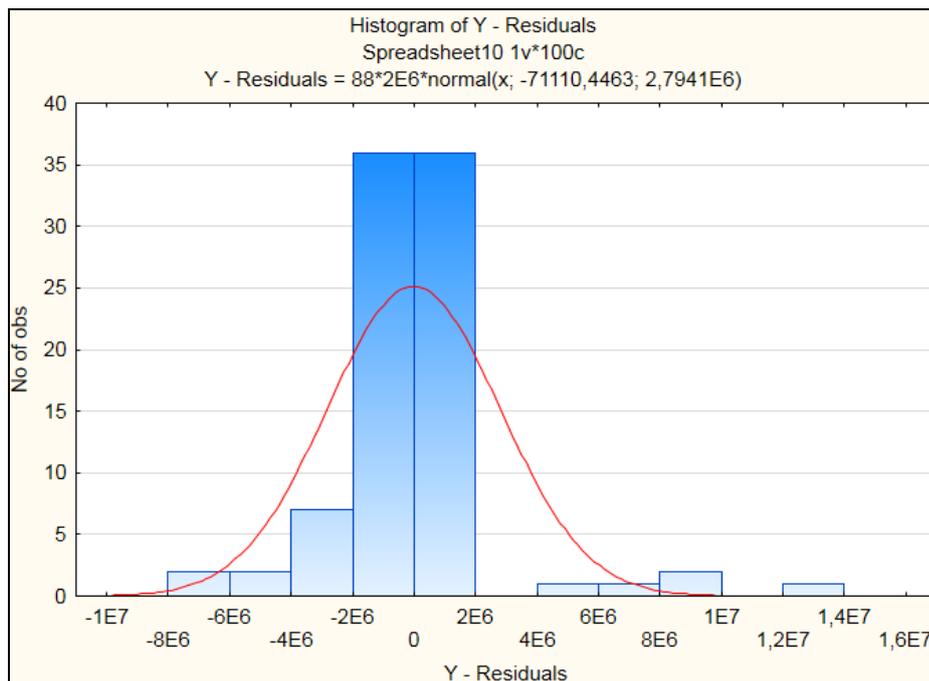


Рисунок 2.27. Гистограмма остатков сети 12-7-1 с кривой нормального распределения

Остатки приближенно повторяют линию Гаусса, поэтому будем считать их распределение нормальным.

### ***Анализ рассеяния***

Для построения диаграммы рассеяния заходим во вкладку «*Graphs (Графики)*» и задаем по оси X значения целевых переменных (*Target*), а по оси Y – значения выходных переменных (*Output*) (см. рисунок 2.28).

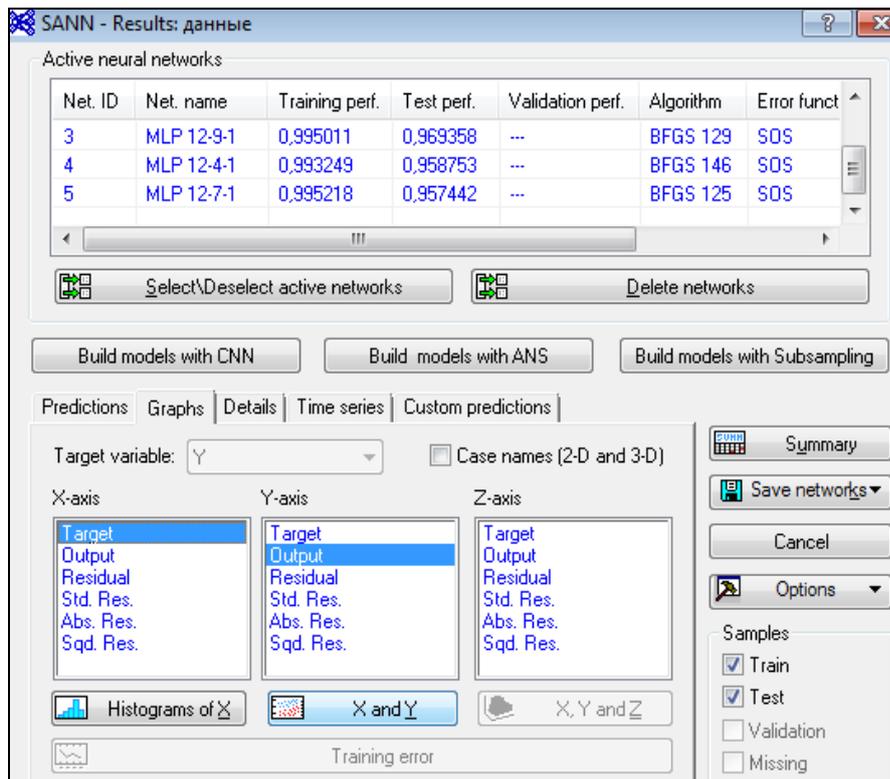


Рисунок 2.28. Вкладка «Графики»

Нажимаем на кнопку «X and Y» и получаем следующий график (рис.2.29):

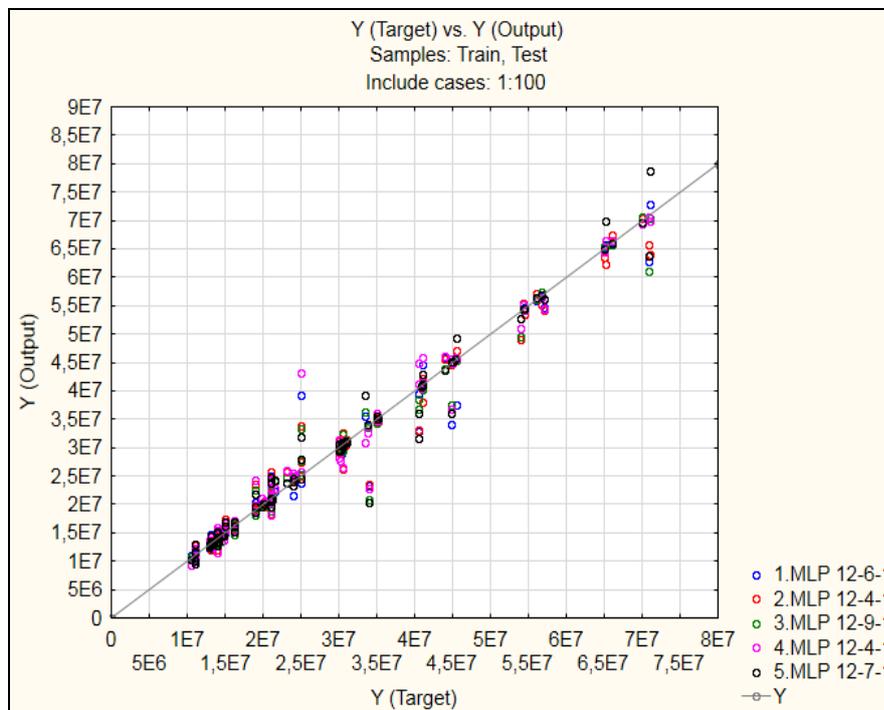


Рисунок 2.29. – Диаграмма рассеяния

Все точки приблизительно расположены на прямой, т.е. можно говорить о достаточно хорошем качестве построенных моделей.

## Прогнозирование для исходного временного ряда

Проведем анализ качества прогнозирования для исходного ряда полученными моделями. Переходим во вкладку «Time series» нажимаем на кнопку «Time series graph» (см. рисунок 2.30).

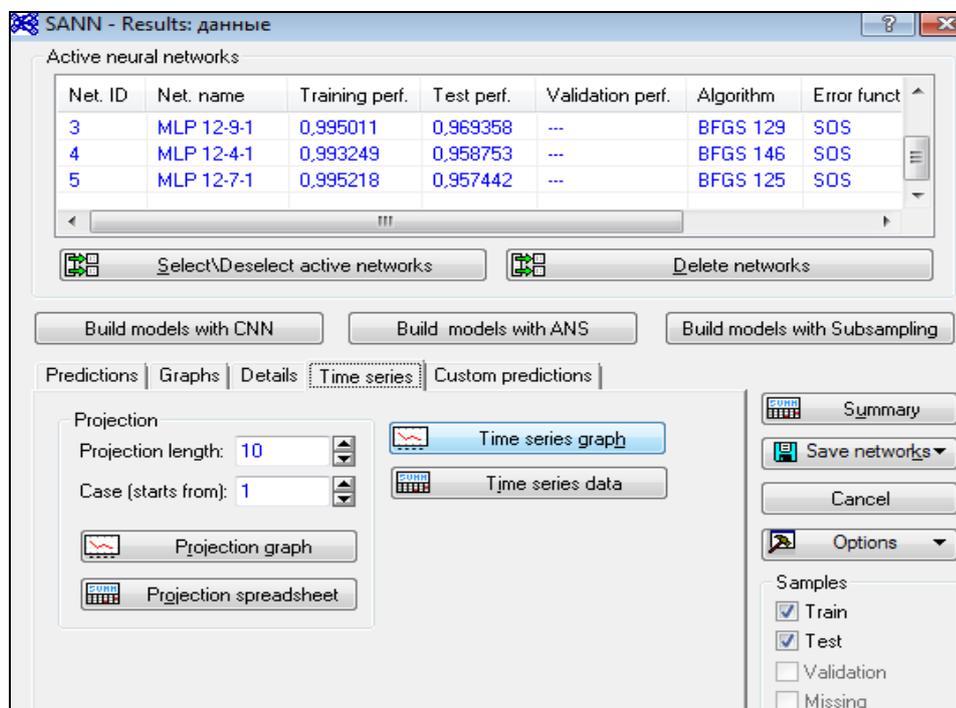


Рисунок 2.30. Вкладка «Time series» окна «Results»

Полученные графики временных рядов представлены на рисунке 2.31.

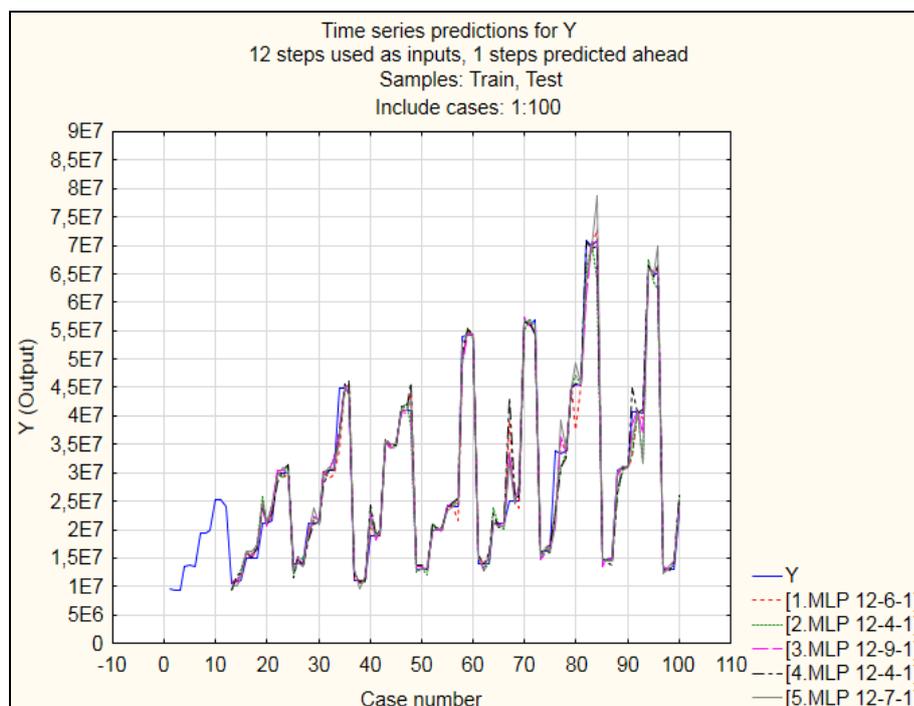


Рисунок 2.31. График временного ряда

Далее необходимо выяснить, как исследуемый временной ряд будет прогнозироваться дальше, т.е. построить проекцию временного ряда. Для этого задаем во вкладке «*Time series*» следующие параметры: «начать с наблюдения (*Case (starts from)*)», «длина проекции («*Projection length*»)» (в соответствии с рисунком 2.32) и нажимаем на кнопку «*Projection graph*».

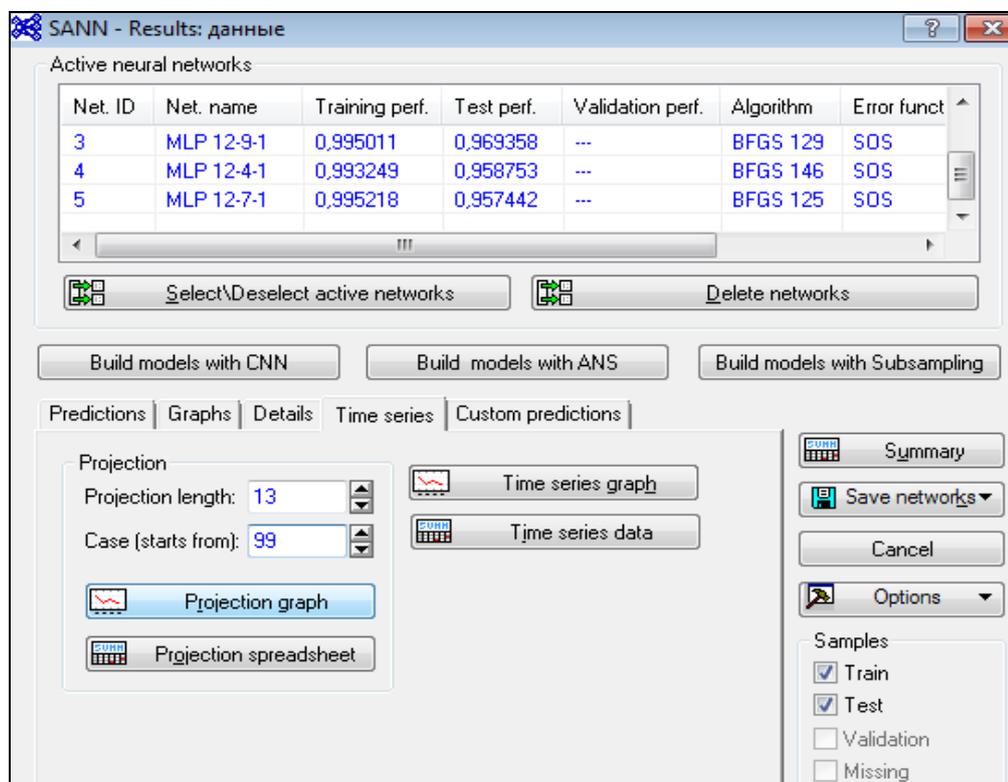


Рисунок 2.32. Проекция временного ряда

Проекцию мы будем строить на тринадцать месяцев вперед (один год вперед (двенадцать неиспользованных последних значений прибыли из таблицы Приложения 1) плюс один месяц). По этой причине на параметре «*Начать с наблюдения*» устанавливаем значение 99.

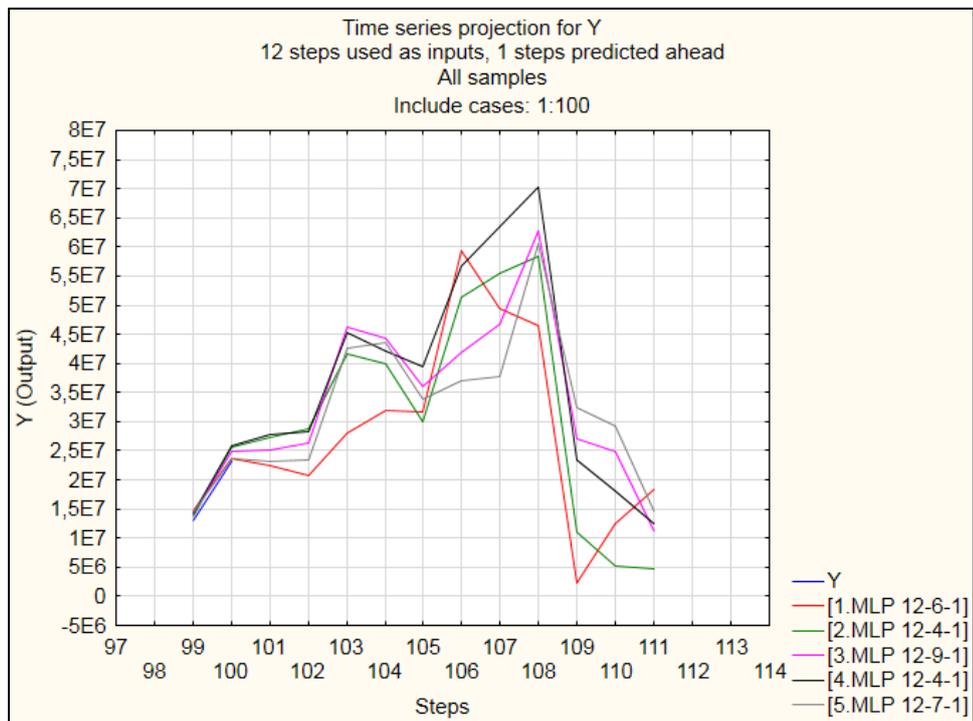


График 2.33. График проекции

Длину проекции установим в размере 144 для того, чтобы посмотреть, какие модели сформируют прогноз (см. рисунок 2.34). Результат представлен на рисунке 2.35.

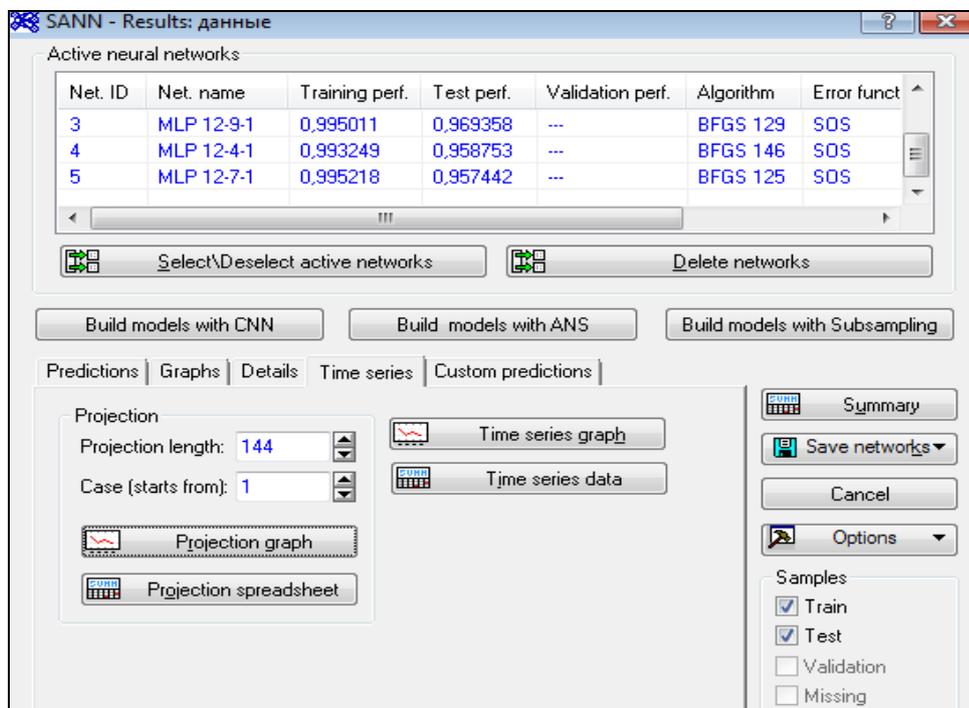


Рисунок 2.34. Проекция до 144 наблюдения

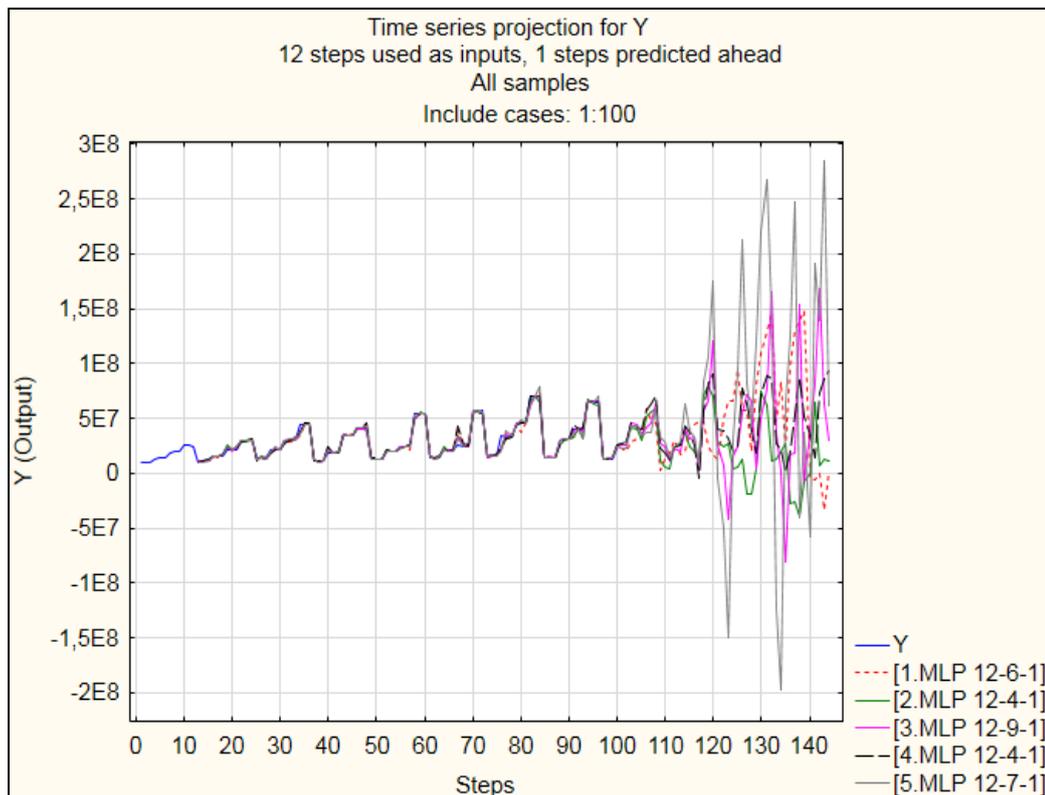


График 2.35. График проекции от 1 до 144 наблюдения

Для выбора наилучшей модели, которая будет использована в дальнейшем анализе, можно построить дополнительные проекции на интересных пользователю интервалах.

На основе анализа моделей, представленных на рис.2.35, можно удалить те из них, которые плохо прогнозируют исходный ряд. К ним относятся модели: 1, 2, 3, 5. Наилучшие результаты показала модель под номером 4.

Далее проанализируем качество выбранной модели. Для этого сравним построенный временной ряд с исходными данными. Для этого во вкладке «*Time series*» нажимаем на «*Projection spreadsheet*», получаем таблицу проекции (представлена на рисунке 2.36). Далее в этой таблице заполним для переменной Y оставшиеся значения (со 101 по 112 из таблицы Приложения 1), т.е. последние 12 месяцев исходного ряда. И построим линейный график, который представлен на рисунке 2.37.

Time series projection for Y (Predictions spreadsheet for Y (данные)) 12 steps used as inputs, 1 steps predicted ahead All samples			Time series projection for Y (Predictions spreadsheet for Y (данные)) 12 steps used as inputs, 1 steps predicted ahead All samples								
Case name	Y Target	Y(Output) MLP 12-4-1	Case name	Y Target	Y(Output) MLP 12-4-1	Y	MLP 12-4-1	Y	MLP 12-4-1	Y	MLP 12-4-1
Jan-06	9654780		Feb-08	14078095	14943229	13040051	12999104	33928333	22920164		27874059
Feb-06	9254471		Mar-08	14154780	13545842	19781621	21001940	33376544	30940940		28312428
Mar-06	9225123		Apr-08	21003065	18248961	19981754	20019305	33843278	32718090		45305592
Apr-06	13540789		May-08	21012547	21050229	20214500	20494218	45050984	44838834		42214057
May-06	13621457		Jun-08	21250740	20874502	24002547	23432169	45543255	45717875		39536695
Jun-06	13620145		Jul-08	30123000	27814973	24004650	24901598	45379654	45899497		56800515
Jul-06	19256047		Aug-08	30321450	30315912	24007453	25580648	70823006	70690060		63545094
Aug-06	19365478		Sep-08	30478000	30773734	54025478	50952812	70006023	69288196		70365132
Sep-06	19874540		Oct-08	44800121	36840508	54254789	55105565	70996207	69888606		23352622
Oct-06	25254708		Nov-08	44798254	45694311	54484100	54222377	14471666	14624034		18019856
Nov-06	25250450		Dec-08	44000478	46095207	14002054	15465853	14671666	14244579		12599151
Dec-06	24000017		Jan-09	11000475	11901088	14000244	13876509	14899768	13817895		23804198
Jan-07	10550478	9349079	Feb-09	11000325	10889644	14001470	15971855	30432569	26448291		
Feb-07	10995642	11986344	Mar-09	11000564	11124743	21000254	22803143	30876543	31194624		
Mar-07	11000000	12302531	Apr-09	18905478	24351529	21012050	20413767	30998765	30824464		
Apr-07	15000125	15378860	May-09	19000012	18976054	21020003	21328404	40599674	44942134		
May-07	15001254	15183739	Jun-09	19001654	19943269	25012547	43211703	40785879	40401856		
Jun-07	15024700	16664378	Jul-09	35000564	36108702	25001788	25681719	40554567	41114823		
Jul-07	21001457	24186710	Aug-09	35050012	34891327	25000147	25770867	65998968	66375376		
Aug-07	21012540	21143689	Sep-09	35124780	34777778	56807545	56523163	65076059	64534017		
Sep-07	21584602	22898881	Oct-09	41000256	41814073	56024511	56415599	65094521	66515646		
Oct-07	30000478	28120041	Nov-09	41000514	41359476	57000785	54434065	13011000	12863631		
Nov-07	30000784	30398138	Dec-09	41001110	45848678	16176600	15537725	13062417	12525771		
Dec-07	30001254	31287503	Jan-10	13005001	13749476	16198725	16890783	13085400	14193161		
Jan-08	14021547	11423676	Feb-10	13002500	13621389	16200146	17101113	23108333	25925572		

Case name	Y Target	Y(Output) MLP 12-4-1
May-13	30876543	31194624
Jun-13	30998765	30824464
Jul-13	40599674	44942134
Aug-13	40785879	40401856
Sep-13	40554567	41114823
Oct-13	65998968	66375376
Nov-13	65076059	64534017
Dec-13	65094521	66515646
Jan-14	13011000	12863631
Feb-14	13062417	12525771
Mar-14	13085400	14193161
Apr-14	23108333	25925572
101	23218333	27874059
102	23489650	28312428
103	32101245	45305592
104	32392546	42214057
105	32280567	39536695
106	54125678	56800515
107	54089799	63545094
108	54067854	70365132
109	9987451	23352622
110	10065478	18019856
111	10917550	12599151
112	15500550	23804198

Рисунок 2.36. Таблица проекции

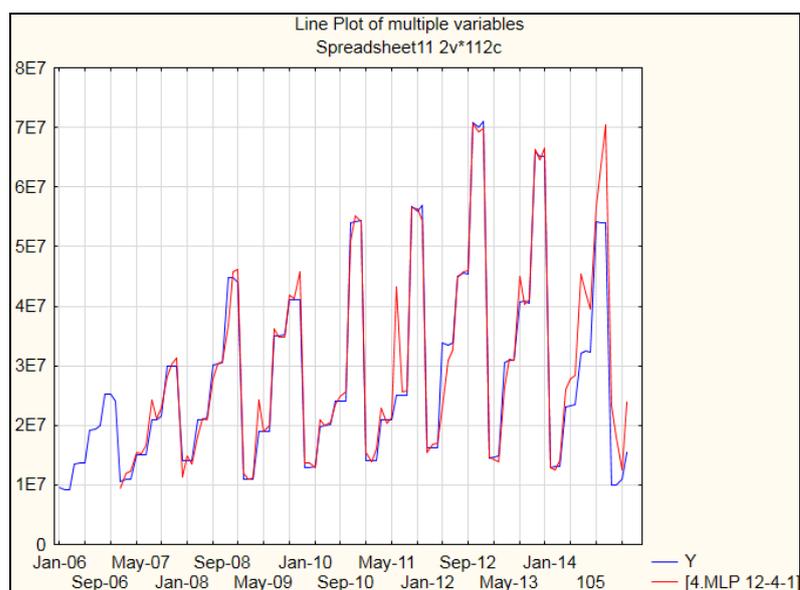


Рисунок 2.37. Линейный график

Для проведения кросс-проверки копируем последние 12 наблюдений из исходного и прогнозного рядов. Вставляем скопированный материал в новую книгу (предварительно создаем ее). Далее высчитываем относительную ошибку по формуле  $ABS(v2 - v1) / ABS(v1)$ , где  $v2$  – прогнозные значения,  $v1$  – наблюдаемые значения. Расчет ошибки представлен на рисунке 2.38. После этого рассчитываем среднее значение. Для этого выделяем столбец «Ошибка» и вызываем контекстное меню правой клавишей мыши, в котором выбираем строку «Statistics of Block Data», далее «Block Columns» – «Means». Полученный результат представлен на рисунке 2.39.

Data: Spreadsheet5* (10v by 12c)			
	1	2	3
	Y	MLP 12-4-1	ошибка
1	23218333	27874059	0,2005194
2	23489650	28312428	0,20531502
3	32101245	45305592	0,41133442
4	32392546	42214057	0,30320282
5	32280567	39536695	0,22478316
6	54125678	56800515	0,049419
7	54089799	63545094	0,17480737
8	54067854	70365132	0,3014227
9	9987451	23352622	1,33819644
10	10065478	18019856	0,79026328
11	10917550	12599151	0,15402731
12	15500550	23804198	0,53570022

Рисунок 2.38. Относительная ошибка

	1
	ошибка
MEAN case 1-12	0,390749262

Рисунок 2.39. Среднее значение

## 2.4. Полученный прогноз

Построим прогноз на период май 2015 г. – апрель 2016 г. (рисунок 2.40). Для этого во вкладке «*Time series*» указываем следующие параметры: «начать наблюдения (*Case (starts from)*)» с временного периода 1, а длина проекции («*Projection length*») составит 124 временных периода. Далее нажимаем на кнопку «*Time series graph*».

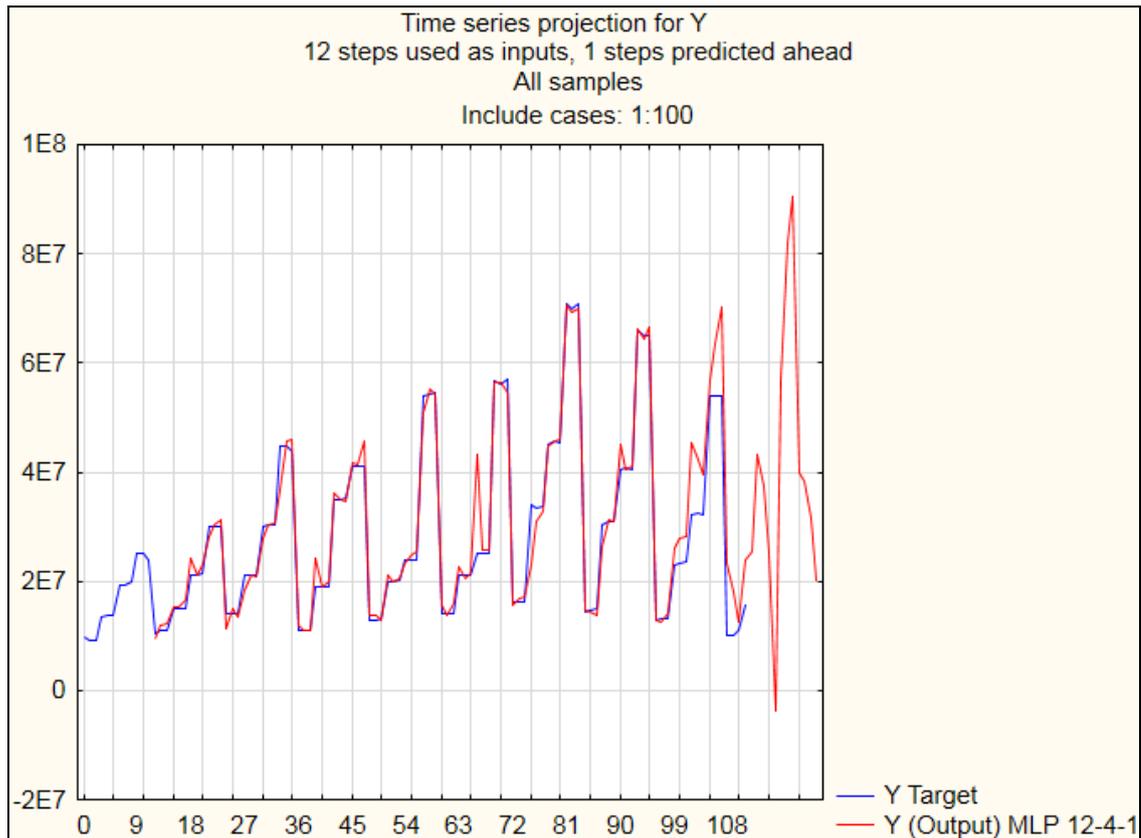


Рисунок 2.40. – График прогноза

Полученные прогнозные значения представим в таблице 2.1.

Таблица 2.1 – Прогноз на период май 2015 г. – апрель 2016 г.

<b>Месяц</b>	<b>Прогноз</b>
Май	25397063
Июнь	43291512
Июль	37536883
Август	25128160
Сентябрь	-3804539
Октябрь	56885032
Ноябрь	82011642
Декабрь	90544681
Январь	39786997
Февраль	38284194
Март	31595728
Апрель	20174145,4

На основе прогнозных значений можно сделать следующие выводы: прогнозная совокупная прибыль за год превышает фактически полученную за аналогичный предшествующий период прибыль на 38%. Однако фактические результаты, вероятнее всего, будут отличаться от прогнозных в сторону уменьшения. Это связано с макроэкономической ситуацией, а также особенностями полученной прогнозной модели (поскольку средняя относительная ошибка, выявленная кросс-проверкой, достаточно большая).

## ЗАКЛЮЧЕНИЕ

В современных реалиях факторов успеха бывает большое множество, а основные из них зачастую бывают неявными. В этих условиях исходными для анализа бывают большие массивы данных (так называемые «big data» и «data mining»), обработка которых требует применения искусственного интеллекта и вычислительных систем большой мощности.

Одним из самых перспективных направлений в изучении, анализе и прогнозировании динамики результирующих факторов являются нейронные сети. Нейронные сети построены на алгоритме функционирования биологического мозга и способны к самообучению. Знания и навыки применения таких технологий являются важнейшим требованием к специалистам в области управления.

В представленном учебном пособии рассмотрены теоретические и прикладные основы методов прогнозирования на основе нейронных сетей. В первой главе были рассмотрены теоретические основы нейронных сетей, их архитектуры и обучения. Во второй главе проводится анализ периодичности исходного ряда, осуществляется прогнозирование на основе нейронных сетей, а также выбор лучшей из них, формируется прогноз на плановый период.

Достоинством данного пособия является подробное описание алгоритмов прогнозирования на основе нейронных сетей с учетом специфических особенностей их применения. С помощью этого пособия любой пользователь с минимальным уровнем подготовки может освоить навыки регрессионного анализа и построить прогноз на основе нейронных сетей в программном пакете STATISTICA Neural Networks. Также формированию указанных навыков способствуют подробно рассмотренные практические примеры анализа прибыли банка и ее прогнозирования.

## ЛИТЕРАТУРА

1. Бизнес-прогнозирование, 7-е издание, пер. с англ. – М.: Издательский дом «Вильямс», 2003. 656 с.
2. Хайкин С. Нейронные сети. Полный курс. М.: Вильямс, 2006. 1104 с.
3. Мазманова Б.Г. Основы теории и практики прогнозирования: учебное пособие. Екатеринбург: изд. ИПК УГТУ, 2004. 456 с.
4. Бородич С.А. Эконометрика: Учеб. Пособие / С.А. Бородич. – М.:Новое знание, 2001. 408 с.
5. О.Г. Туровец, Б.Ю. Сербиновский. Организация производства на предприятии: Учебник для технических и экономических специальностей. Ростов-на-Дону: МарТ, 2002. 462 с.
6. Ковалев В.В. Управление финансами. Учебное пособие. М.: ФБК-ПРЕСС, 2005 г. 272 с.
7. Ковалева А.М. Финансы и кредит. Учебное пособие. М.: Финансы и статистика, 2005 г. 512 с.
8. Смирнова Г.Н., Сорокин А.А., Тельнов Ю.Ф. Проектирование экономических информационных систем: Учебник. М.: Финансы и статистика, 2004. 312 с.
9. Яхьяева Г. Основы искусственных нейронных сетей [Электронный ресурс]. Режим доступа: <http://www.intuit.ru>.
10. Автоматизированные нейронные сети [Электронный ресурс]. Режим доступа: <http://statsoft.ru>.
11. Первые шаги в STATISTICA. Automated Neural Networks (SANN) [Электронный ресурс]. Режим доступа: <http://www.statsoft.ru/products/>

**ПРИЛОЖЕНИЕ 1. Прибыль банка в период с 2006 по 2015 г.г.**

Месяц	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015
<b>январь</b>	9654780	10550478	14021547	11000475	13005001	14002054	16176600	14471666	13011000	9987451
<b>февраль</b>	9254471	10995642	14078095	11000325	13002500	14000244	16198725	14671666	13062417	10065478
<b>март</b>	9225123	11000000	14154780	11000564	13040051	14001470	16200146	14899768	13085400	10917550
<b>апрель</b>	13540789	15000125	21003065	18905478	19781621	21000254	33928333	30432569	23108333	15500550
<b>май</b>	13621457	15001254	21012547	19000012	19981754	21012050	33376544	30876543	23218333	-
<b>июнь</b>	13620145	15024700	21250740	19001654	20214500	21020003	33843278	30998765	23489650	-
<b>июль</b>	19256047	21001457	30123000	35000564	24002547	25012547	45050984	40599674	32101245	-
<b>август</b>	19365478	21012540	30321450	35050012	24004650	25001788	45543255	40785879	32392546	-
<b>сентябрь</b>	19874540	21584602	30478000	35124780	24007453	25000147	45379654	40554567	32280567	-
<b>октябрь</b>	25254708	30000478	44800121	41000256	54025478	56807545	70823006	65998968	54125678	-
<b>ноябрь</b>	25250450	30000784	44798254	41000514	54254789	56024511	70006023	65076059	54089799	-
<b>декабрь</b>	24000017	30001254	44000478	41001110	54484100	57000785	70996207	65094521	54067854	-



---

Подписано в печать 07.06.2017 г.  
Формат 60x84/15 Бумага офсетная Печать ризографическая  
Уч.изд.л. 2,8 Усл.-печ.л. 2,8 Тираж 50 экз.  
Заказ №865

Издательско-полиграфический центр  
Набережночелнинского института  
Казанского (Приволжского) федерального университета

---

423810, г. Набережные Челны, Новый город, проспект Мира, 68/19  
Тел./факс (8552) 39-65-99 e-mail: [ic-nchi-kpfu@mail.ru](mailto:ic-nchi-kpfu@mail.ru)