

# **Решение задачи прогнозирования с помощью нейронной сети**

Лекция 11 (1 час)

Емельянова М.Г.

## Сущность задачи прогнозирования

**Прогнозирование (регрессия)** – установление зависимости непрерывных выходных переменных от входных.

Регрессия используется для установления зависимостей в факторах. Например, в задаче прогнозирования зависимой величиной является **объёмы продаж**, а факторами, влияющими на эту величину, могут быть предыдущие **объёмы продаж**, **изменение курса валют**, **активность конкурентов** и т.д.

К этому же типу задач относится прогнозирование временного ряда на основе исторических данных.

## Пример решения задачи прогнозирования

У аналитика имеются данные о ежемесячном количестве проданного товара за несколько лет. Ему необходимо, основываясь на этих данных, определить, какое количество товара будет продано через месяц и через два месяца. Исходные данные по продажам находятся в файле Trade.txt.

### Фрагмент данных.

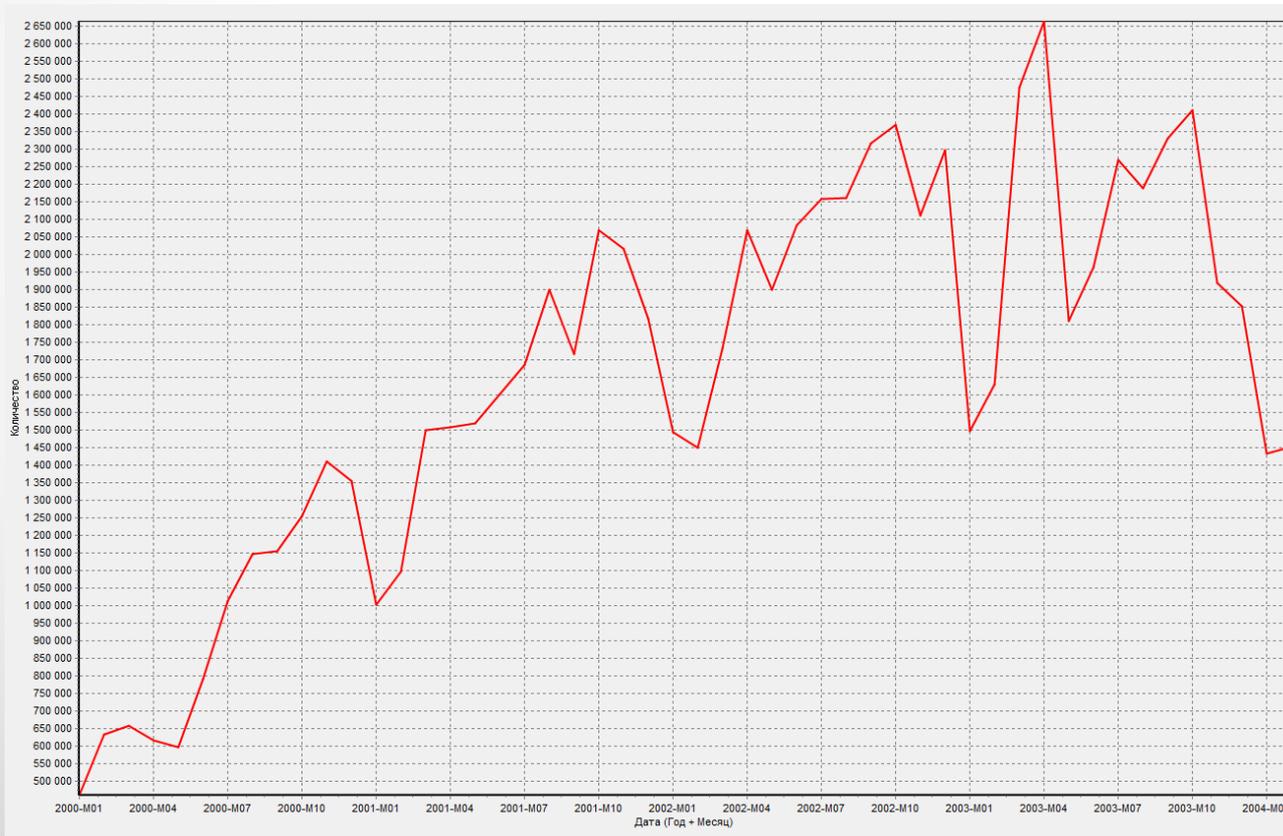
Дата (Год + Месяц)	Количество
2000-M01	462523,419
2000-M02	633208,196
2000-M03	660159,299
2000-M04	617455,3417
2000-M05	597354,4794
2000-M06	793517,4512
2000-M07	1015944,2862
2000-M08	1148052,2523
2000-M09	1156623,1715
2000-M10	1255021,9423
2000-M11	1410114,5606
2000-M12	1357230,3388
2001-M01	1003317,7317
2001-M02	1097048,6263
2001-M03	1498977,3427
2001-M04	1507696,4482
2001-M05	1520761,5589
2001-M06	1602674,5245

# Пример решения задачи прогнозирования

## Решение в Deductor Studio.

1. Импорт данных.
2. Просмотр данных в виде таблицы и диаграммы.

Диаграмма данных



## Пример решения задачи прогнозирования

### Решение в Deductor Studio.

3. При необходимости использование обработчиков «Оценка качества данных», «Заполнение пропусков», «Редактирование выбросов», «Спектральная обработка».

[https://basegroup.ru/deductor/function/algorithm?algoritym\\_group%5B%5D=82](https://basegroup.ru/deductor/function/algorithm?algoritym_group%5B%5D=82).

4. Для используемого набора данных применим обработчик «Спектральная обработка».

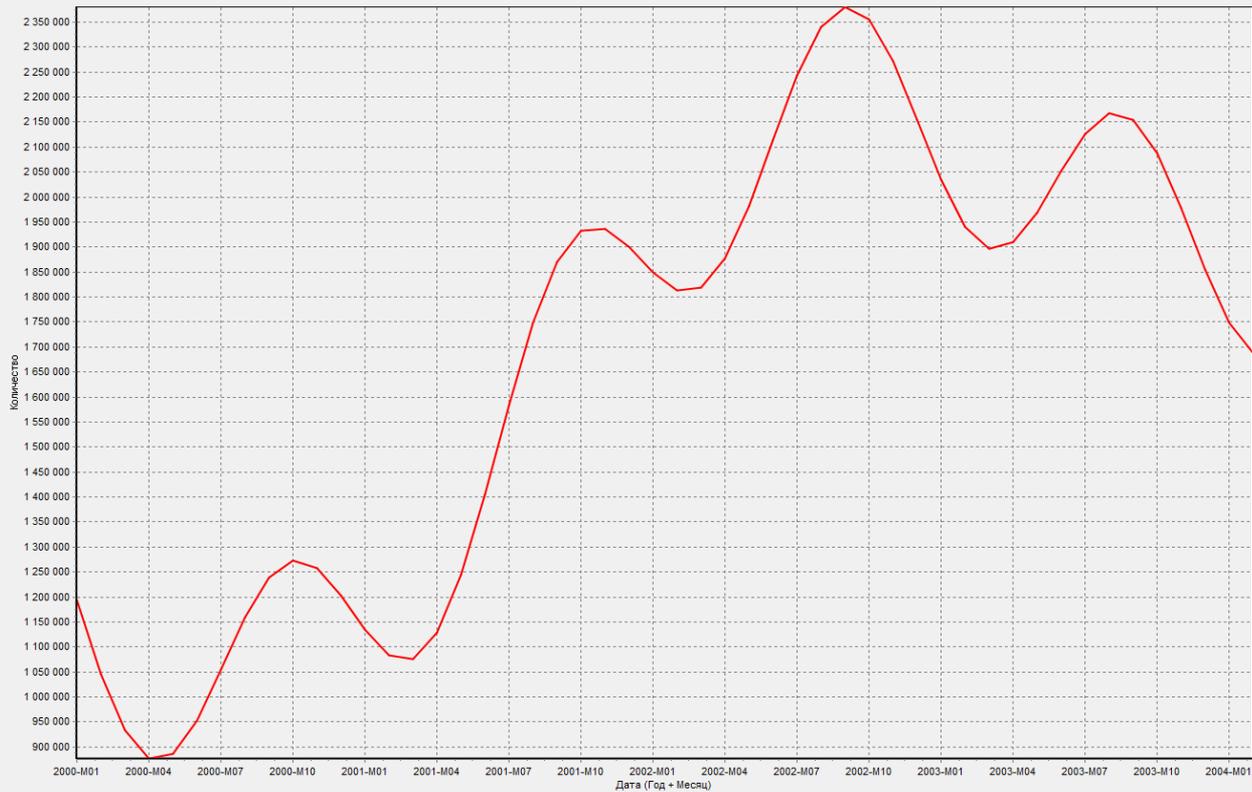
Спектральная обработка предназначена для очистки от шумовой составляющей и сглаживания рядов данных. Сглаживание необходимо в том случае, когда ряд данных оказывается неравномерным, содержит большое количество мелких структур, препятствующих исследованию более значительных объектов и закономерностей.

Спектральная обработка наиболее часто применяется для предварительной подготовки данных в задачах прогнозирования, т.к. позволяет сделать временной ряд более гладким, благодаря чему полученная прогнозная модель обладает высокими обобщающими качествами.

# Пример решения задачи прогнозирования

## Решение в Deductor Studio.

Диаграмма после спектральной обработки.



## Пример решения задачи прогнозирования

### Решение в Deductor Studio.

5. Трансформация данных к скользящему окну (обработчик «Скользящее окно»).

При решении задачи прогнозирования используются данные, в которых есть входные столбцы – **факторы** и есть выходные столбцы – **результат**.

Набор данных Trade.txt содержит столбец с выходными значениями.

Строить прогноз на будущее будем, основываясь на данных прошлых периодов, т. е. предполагая, что количество продаж на следующий месяц зависит от количества продаж за предыдущие месяцы. Это значит, что входными факторами для модели могут быть продажи за текущий месяц, продажи за месяц ранее и т.д., а результатом должны быть продажи за следующий месяц.

## Обработчик «Скользящее окно»

**Обработчик «Скользящее окно»** преобразует последовательность значений ряда в таблицу, где соседние записи представлены как соседние поля данных (окно – поскольку выделяется только некоторый непрерывный участок данных, скользящее – поскольку это окно «перемещается» по всему набору).

Потребность в такой таблице часто возникает при построении моделей, анализе и прогнозировании временных рядов, когда требуется подавать на вход модели значения нескольких смежных отсчётов из исходного набора данных (<https://basegroup.ru/deductor/function/algorithm/sliding-window>).

## Обработчик «Скользящее окно»

Значения в одном из полей записи будут относиться к текущему отсчёту, а в других – смещены от текущего отсчёта «в будущее» или «в прошлое». Таким образом, преобразование скользящего окна имеет два параметра:

- глубина погружения – количество «прошлых» отсчётов, попадающих в окно;
- горизонт прогнозирования – количество «будущих» отсчётов.

## Обработчик «Скользящее окно»

Пример применения скользящего окна.

Есть история продаж за половину года по месяцам, представленная таблицей:

Первый день месяца	Объем продаж (тыс. руб.)
01.01.2014	1000
01.02.2014	1160
01.03.2014	1210
01.04.2014	1130
01.05.2014	1250
01.06.2014	1300

## Обработчик «Скользящее окно»

Для граничных (относительно начала и конца всей выборки) положений окна будут формироваться неполные записи, т.е. записи, содержащие пустые значения для отсутствующих прошлых или будущих отсчётов.

Алгоритм преобразования позволяет исключить такие записи из выборки (тогда для нескольких граничных отсчётов записи формироваться не будут) либо включить их (тогда формируются записи для всех имеющихся отсчётов, но некоторые из них будут неполными).

## Обработчик «Скользящее окно»

Если задать глубину погружения 2 и горизонт прогнозирования 1, то получим следующую таблицу с неполными записями.

Первый день месяца	Объем продаж 2 месяца назад	Объем продаж месяц назад	Объем продаж в текущий месяц	Объем продаж на следующий месяц
				1000
01.01.2014			1000	1160
01.02.2014		1000	1160	1210
01.03.2014	1000	1160	1210	1130
01.04.2014	1160	1210	1130	1250
01.05.2014	1210	1130	1250	1300
01.06.2014	1130	1250	1300	
	1250	1300		
	1300			

Или следующую таблицу с полными записями.

Первый день месяца	Объем продаж 2 месяца назад	Объем продаж месяц назад	Объем продаж в текущий месяц	Объем продаж на следующий месяц
01.03.2014	1000	1160	1210	1130
01.04.2014	1160	1210	1130	1250
01.05.2014	1210	1130	1250	1300

## Обработчик «Скользящее окно»

### Решение в Deductor Studio.

В Мастере обработки «Скользящее окно» указать необходимые параметры.

Мастер обработки - Скользящее окно (2 из 4)

**Скользящее окно**  
Транспонирование столбца с использованием скользящего окна

Дата (Год + Месяц)  
✓ Количество

Имя столбца: COL2  
Тип данных: Вещественный  
Назначение: ✓ Используемое

Глубина погружения: 12  
Горизонт прогнозирования: 0

Оставлять неполные записи

< Назад    **Далее >**    Отмена

## Пример решения задачи прогнозирования

### Решение в Deductor Studio.

5. Используем узел «Нейросеть».

Установим в качестве входных поля «Количество – 12», «Количество – 11», «Количество – 2» и «Количество – 1», а в качестве выходного – «Количество». Остальные поля сделаем информационными.

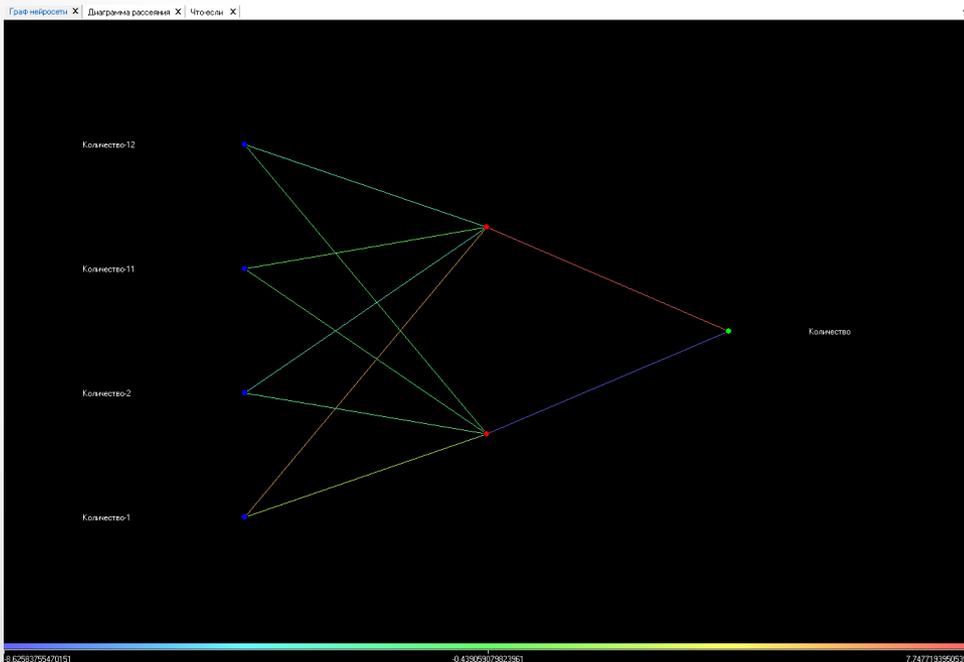
**Настройка параметров нейросети – предыдущая лекция.**

## Пример решения задачи прогнозирования

### Решение в Deductor Studio.

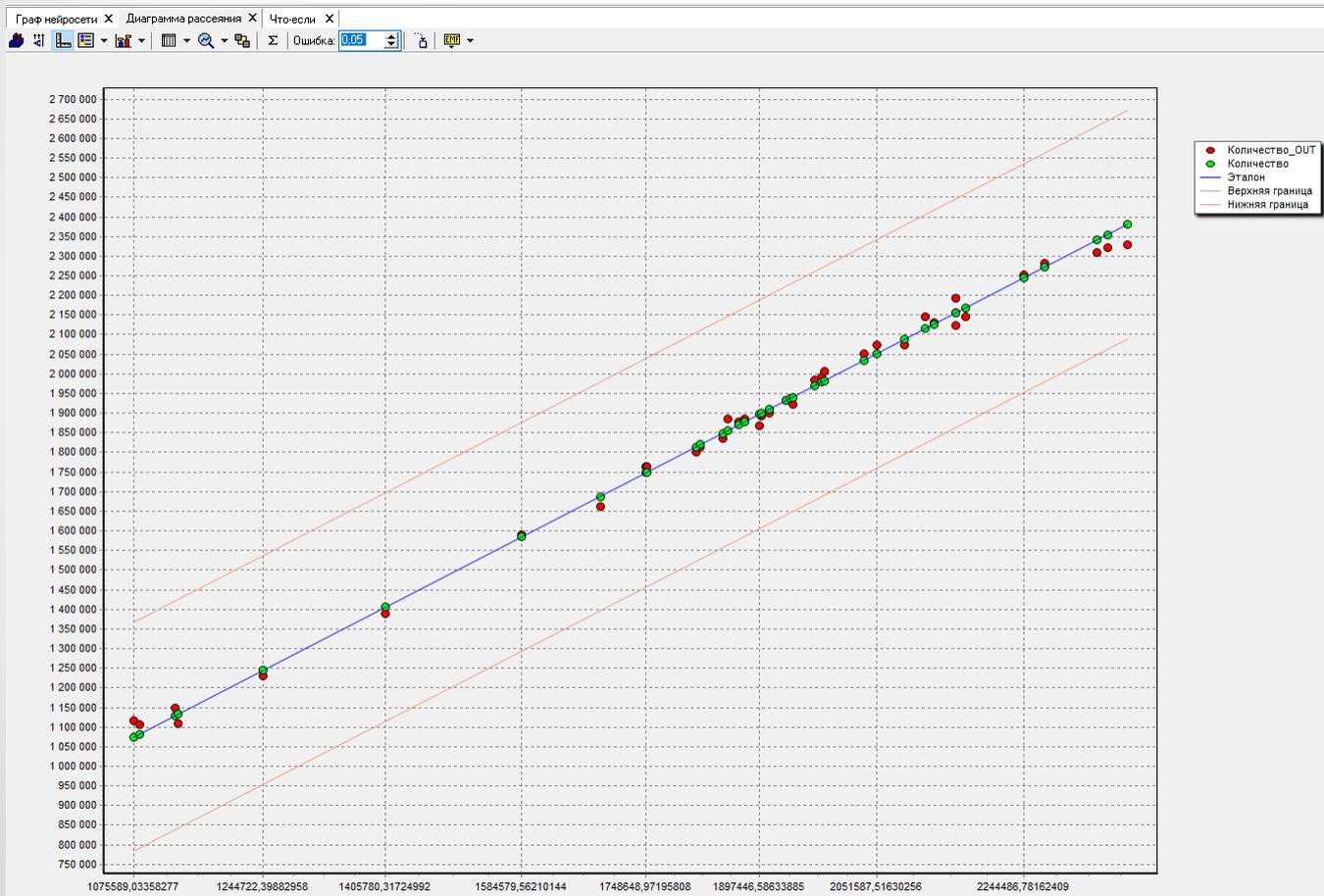
6. Просмотр результатов в визуализаторах «Граф нейросети», «Диаграмма рассеяния», «Что-если».

### Визуализатор «Граф нейросети».



# Диаграмма рассеяния

Диаграмма рассеяния наглядно показывает качество обучения. На этой диаграмме отображается отклонение прогнозного значения величины от её истинного значения.



На диаграмме рассеяния отображаются выходные значения каждого для каждого из примеров обучающей выборки, координаты которых по оси X – это значение выхода на обучающей выборке (эталон), а по оси Y – значение выхода, рассчитанное обученной моделью на том же примере.

Прямая диагональная линия представляет собой ориентир (линию идеальных значений). Чем ближе точка к этой линии, тем меньше ошибка модели.

## Пример решения задачи прогнозирования

### Решение в Deductor Studio.

7. Применяем обработчик «Прогнозирование».

Настройка параметров прогнозирования.

Визуализатор «Диаграмма прогноза».

Мастер обработки - Прогнозирование (2 из 4)

**Настройка прогнозирования временного ряда**  
Настройте связи столбцов для прогнозирования временного ряда

Столбец	При очередном шаге брать значения из
Дата (Год + Месяц)	
Количество-12	Количество-11
Количество-11	Количество-10
Количество-10	Количество-9
Количество-9	Количество-8
Количество-8	Количество-7
Количество-7	Количество-6
Количество-6	Количество-5
Количество-5	Количество-4
Количество-4	Количество-3
Количество-3	Количество-2
Количество-2	Количество-1
Количество-1	Количество
Количество	
Шаг прогноза	

Горизонт прогноза  12

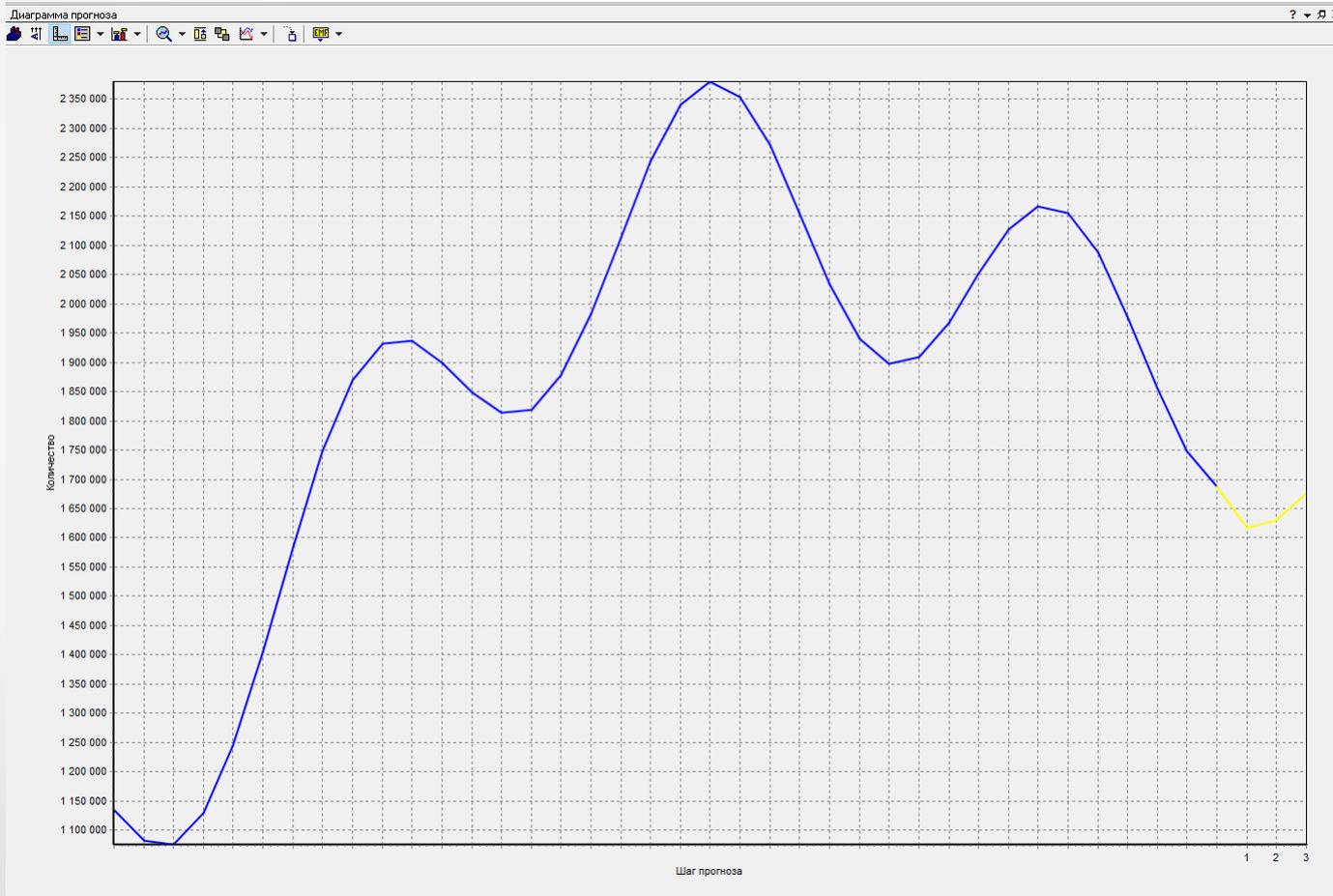
Добавлять поле "Шаг прогноза"  
 Исходные данные

< Назад **Далее >** Отмена

# Пример решения задачи прогнозирования

Визуализатор «Диаграмма прогноза».

Прогнозирование на три шага вперёд.



## Вопросы для проверки

1. В чём суть задачи прогнозирования?
2. Для чего предназначен обработчик «Спектральная обработка»?
3. Для чего предназначен обработчик «Скользящее окно»?
4. Что отображается на диаграмме рассеяния?
5. С помощью какого обработчика в Deductor можно осуществить прогнозирование на указанное количество шагов вперёд?