ПРОГНОЗИРОВАНИЕ И АНАЛИЗ ВРЕМЕННЫХ РЯДОВ

ПРОГНОЗИРОВАНИЕ – ВСЕГДА ЛИ ОСУЩЕСТВИМО?

- *Скажите нам, что ждет нас в будущем, чтобы мы знали, что вы боги.* (Исаия 41:23)

Примеры известных предсказаний о компьютерах.

- *"Я думаю, в мире найдётся рынок от силы для 5 компьютеров«* (*Т*омас Ватсон-старший, президент IBM , 1943)
- *Вес компьютеров в будущем не будет превышать 1,5 тонны.* (Популярная Механика, 1949)
- «*Нет никаких причин для того, чтобы кто-нибудь захотел иметь компьютер у себя дома*». (Кен Олсен, президент компании Digital Equipment, 1977 г.)

ИСХОДНЫЕ СТАТИСТИЧЕСКИЕ ДАННЫЕ

- Временной ряд - последовательность наблюдаемых значений какого-либо признака, упорядоченных в неслучайные моменты времени. Отличием анализа временных рядов от анализа случайных выборок является предположение о равных промежутках времени между наблюдениями и их хронологический порядок.

Если рассматривать значения одного признака у одного объекта в равноотстоящие моменты времени, то последовательность $\boldsymbol{x}(\boldsymbol{t}_1), \boldsymbol{x}(\boldsymbol{t}_2), \ldots, \boldsymbol{x}(\boldsymbol{t}_N)$ называют *одномерным временным рядом.*

Если регистрировать значения *p* признаков у одного объекта, то говорят о статистическом анализе *многомерного временного ряда* $\boldsymbol{X}(\boldsymbol{t}) = (\boldsymbol{x}^1(\boldsymbol{t}_k), \boldsymbol{x}^2(\boldsymbol{t}_k), \ldots, \boldsymbol{x}^p(\boldsymbol{t}_k))$, ***k*=1,2, …, *N***

ПРИМЕРЫ ВРЕМЕННЫХ РЯДОВ

Примеры временных рядов:

- рыночные цены
- объемы продаж в торговых сетях
- объемы потребления и цены электроэнергии
- объемы грузовых и пассажирских перевозок
- дорожный трафик (прогнозирование пробок)

ОСНОВНЫЕ ФАКТОРЫ ВРЕМЕННЫХ РЯДОВ

1. *Долговременные*, формирующие общую тенденцию в изменении анализируемого признака $x(t)$. Обычно описывается при помощи монотонной функции $f(t)$, называемой *трендом*.

2. *Сезонные*, формирующие периодически повторяющиеся в определенное время года колебания анализируемого признака. Описывается периодической функцией $\varphi(t)$ с периодом, кратным сезонам.

3. *Циклические,* формирующие изменения анализируемого признака, обусловленные действием долговременных циклов экономической, демографической или астрономической природы. Описывается функцией $\psi(t)$.

4. *Случайные*, не поддающиеся учету и регистрации. Их воздействие обуславливает *стохастическую природу* анализируемого признака. Обозначается $\varepsilon(t)$

ОБЩИЕ МОДЕЛИ ВРЕМЕННЫХ РЯДОВ

- *Аддитивная форма*

$$\boldsymbol{x(t) = \lambda_1 f(t) + \lambda_2 \varphi(t) + \lambda_3 \psi(t) + \varepsilon(t)}, \quad \mathbf{(1)}$$

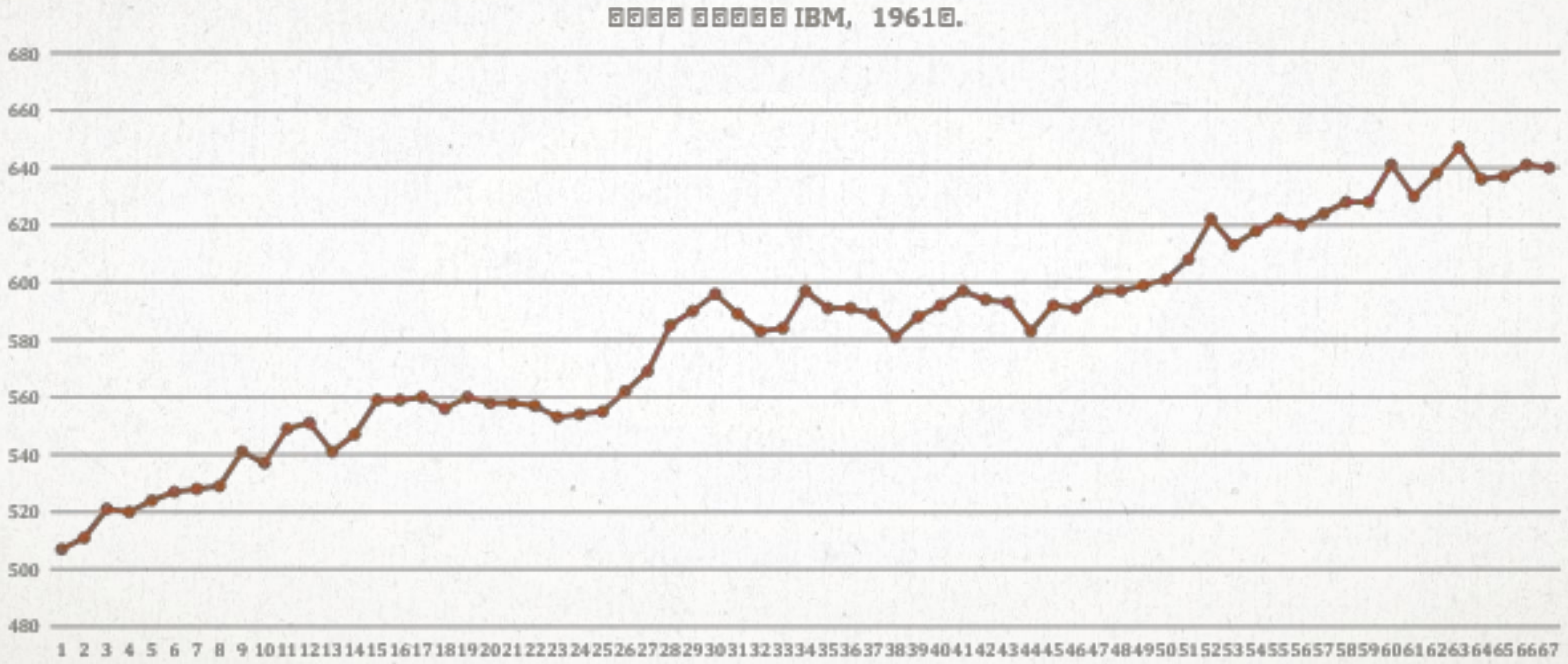
$$\lambda_i = \begin{cases} 1, & \text{фактор участвует в формировании уровней ряда } x(t) \\ 0, & \text{в противном случае} \end{cases}$$

Мультипликативная форма

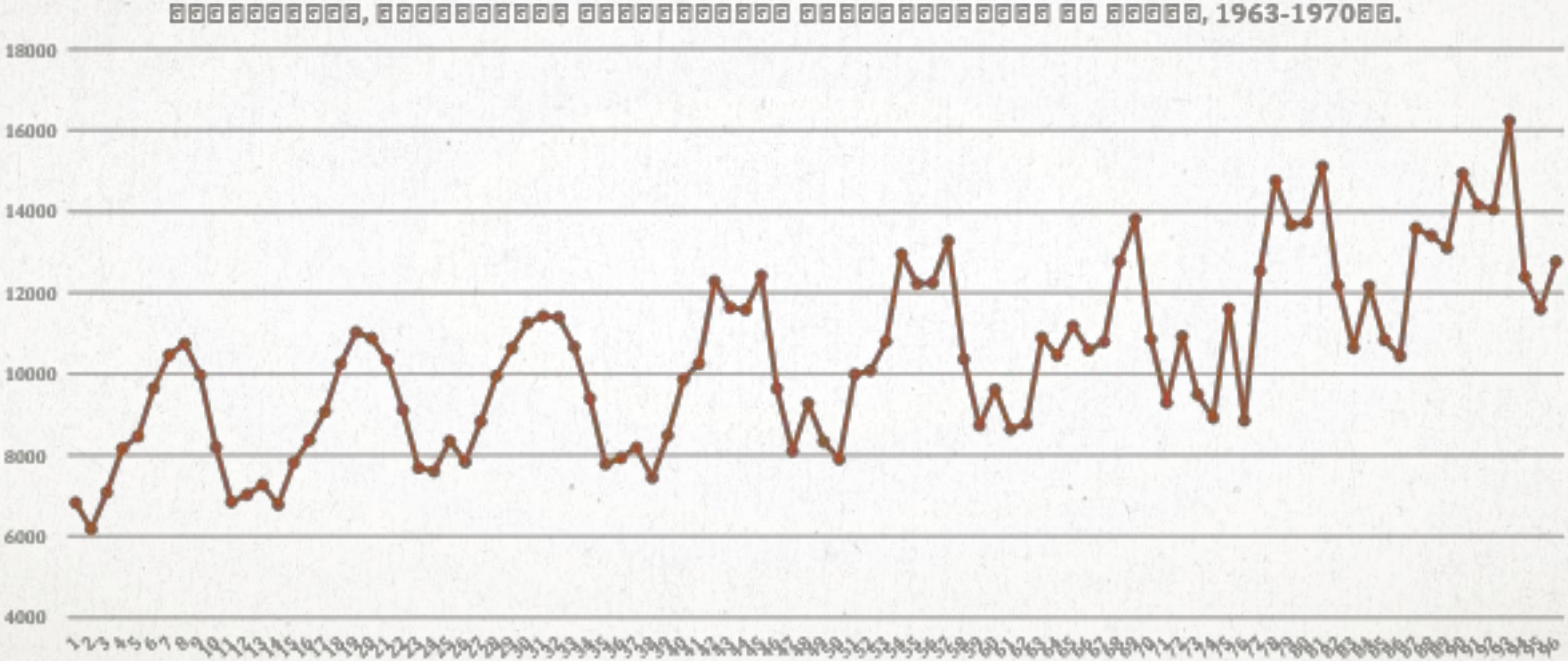
$$\boldsymbol{x(t) = f(t)^{\lambda_1} * \varphi(t)^{\lambda_2} * \psi(t)^{\lambda_3} * \varepsilon(t)}$$

$$\boldsymbol{\ln x(t) = \lambda_1 \ln f(t) + \lambda_2 \ln \varphi(t) + \lambda_3 \ln \psi(t) + \ln \varepsilon(t)}$$

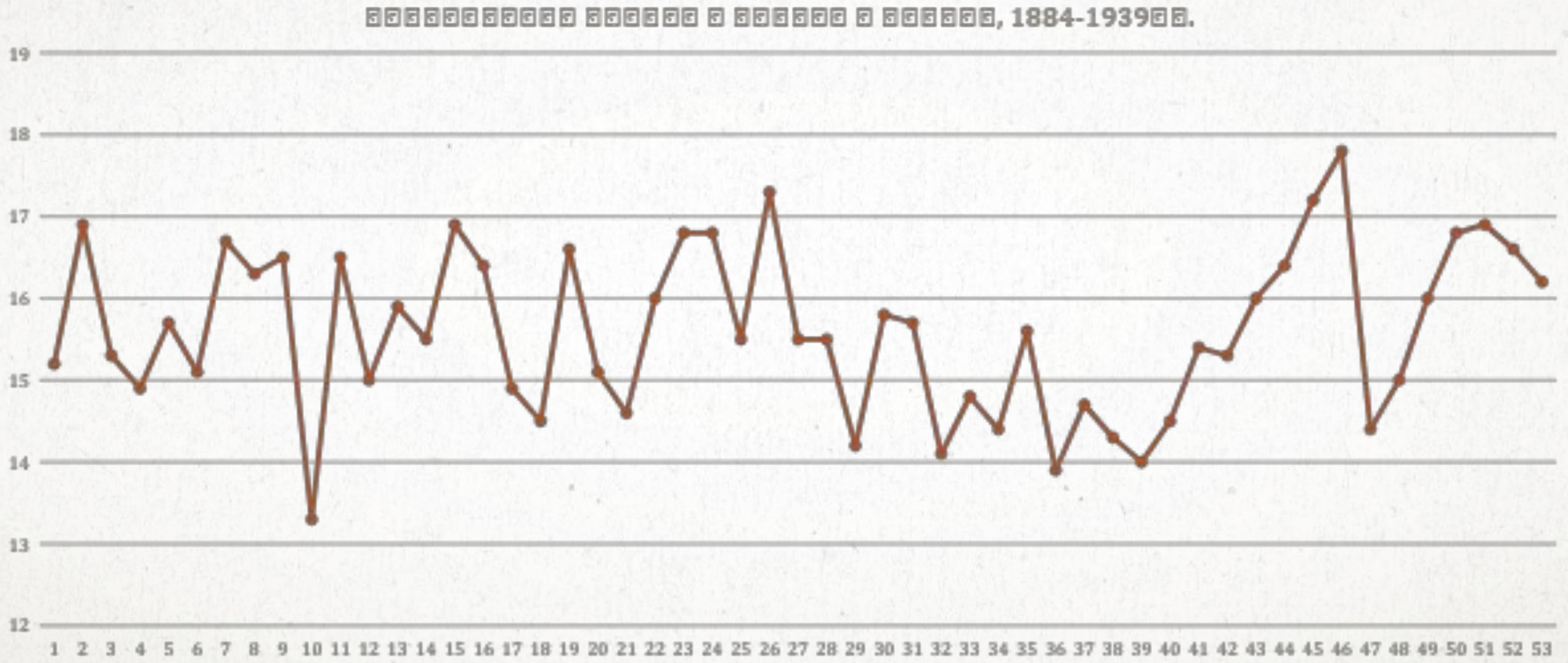
ПРИМЕРЫ ВРЕМЕННЫХ РЯДОВ



ПРИМЕРЫ ВРЕМЕННЫХ РЯДОВ



ПРИМЕРЫ ВРЕМЕННЫХ РЯДОВ



ЗАДАЧИ АНАЛИЗА ВРЕМЕННЫХ РЯДОВ

По имеющейся траектории анализируемого временного ряда $x(t)$ требуется:

- Определить, какие из неслучайных составляющих $f(t)$, $\varphi(t)$ и $\psi(t)$ присутствуют в разложении (1)
- Построить оценки для тех неслучайных функций, которые присутствуют в разложении (1)
- Подобрать модель, адекватно описывающую поведение «случайной составляющей $\varepsilon(t)$, и статистически оценить параметры этой модели

ТЕСТИРОВАНИЕ НАЛИЧИЯ/ОТСУТСТВИЯ НЕСЛУЧАЙНОЙ СОСТАВЛЯЮЩЕЙ

- Если неслучайные составляющие отсутствуют (ряд является стационарным), то ряд состоит из статистически независимых наблюдений, случайно варьирующихся около некоторого постоянного уровня a, т.е.

$$\boldsymbol{H_0}:\ \mathbf{M}\big(\boldsymbol{x}(\boldsymbol{t})\big) = \boldsymbol{a} = \boldsymbol{const}$$

Иначе, существует зависимость от времени неслучайной составляющей анализируемого временного ряда (ряд является нестационарным)

$$\boldsymbol{H_1}:\ \mathbf{M}\big(\boldsymbol{x}(\boldsymbol{t})\big) \neq \boldsymbol{const}$$

ВИДЫ ПРОГНОЗОВ

- Прогноз может быть краткосрочным, среднесрочным и долгосрочным.
- **Краткосрочный прогноз** представляет собой прогноз на несколько шагов вперед, т.е. осуществляется построение прогноза не более чем на 3% от объема наблюдений или на 1-3 шага вперед.
- **Среднесрочный прогноз** - это прогноз на 3-5% от объема наблюдений, но не более 7-12 шагов вперед; также под этим типом прогноза понимают прогноз на один или половину сезонного цикла. Для построения краткосрочных и среднесрочных прогнозов вполне подходят статистические методы.
- **Долгосрочный прогноз** - это прогноз более чем на 5% от объема наблюдений.

КРИТЕРИИ ОЦЕНКИ ПРОГНОЗА

Среднее абсолютное отклонение

$$MAD = \frac{\sum_{\mathrm{i}=1}^{\mathrm{n}} |\hat{y}_i - y_i|}{\mathrm{n}}$$

Среднеквадратичная ошибка

$$MSE = \frac{\sum_{\mathrm{i}=1}^{\mathrm{n}} |\hat{y}_i - y_i|^2}{\mathrm{n}}$$

Максимальное абсолютное отклонение

$$MAXD = \max|\hat{y}_i - y_i|$$

Абсолютное среднее процентной ошибки

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{\hat{y}_i - y_i}{y_i}\right| * 100\%$$

МЕТОДЫ СГЛАЖИВАНИЯ ВРЕМЕННОГО РЯДА

- *Простая скользящая средняя*

$$s_t = \frac{y_{t-1} + y_{t-2} + y_{t-3} + \cdots + y_{t-n}}{n}$$

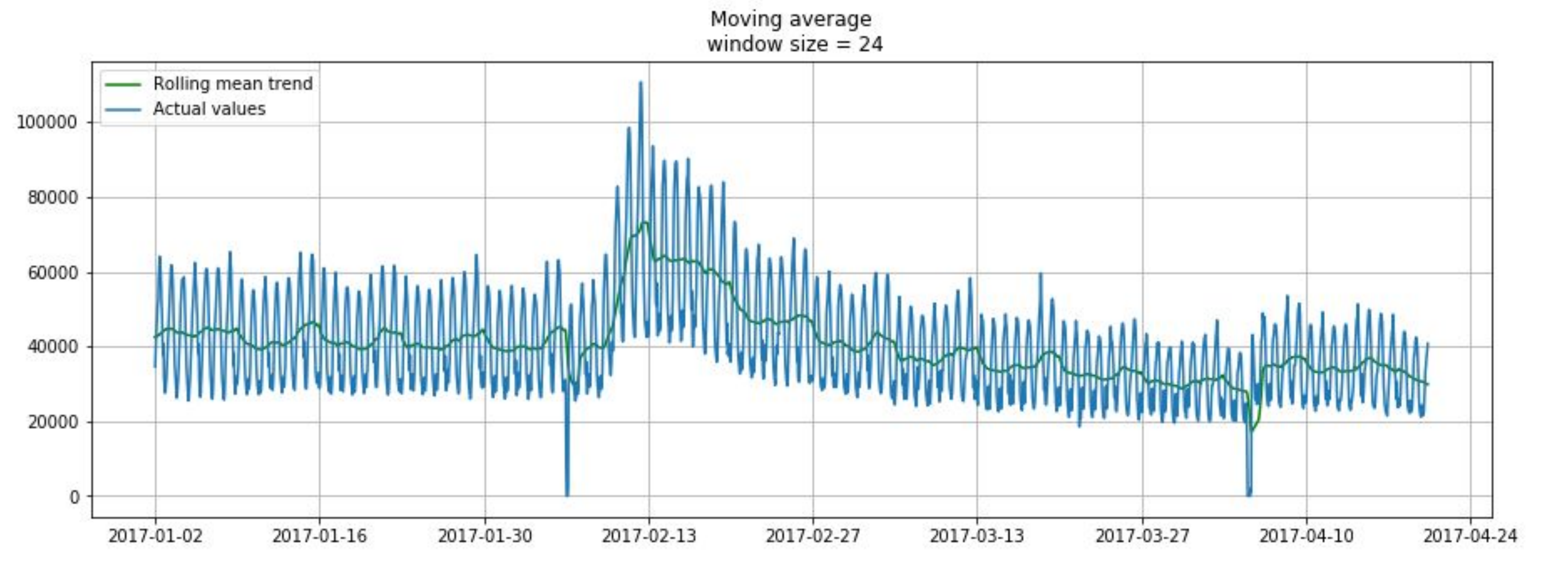
Экспоненциальная скользящая средняя (ЭСС)

$$s_t = \alpha y_{t-1} + (1 - \alpha) s_{t-1}$$

Где

- $\alpha \in [0,1]$– параметр сглаживания,
- y_{t-1}- фактическое значение исследуемого показателя за предшествующий период
- s_{t-1}- экспоненциальная скользящая средняя для предшествующего периода.

ГРАФИК ПРОСТОЙ СКОЛЬЗЯЩЕЙ СРЕДНЕЙ

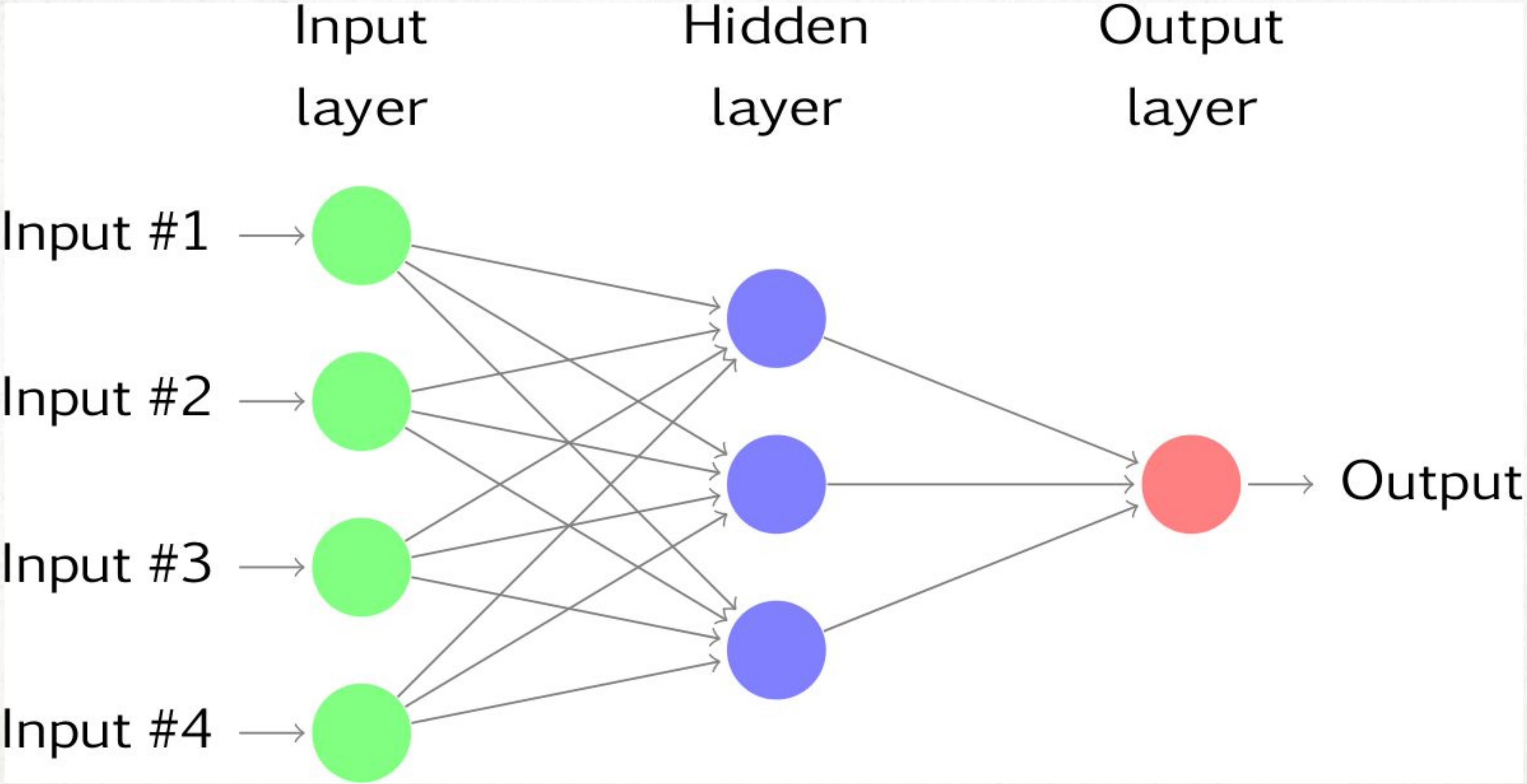


ЛИНЕЙНАЯ МОДЕЛЬ АВТОРЕГРЕССИИ

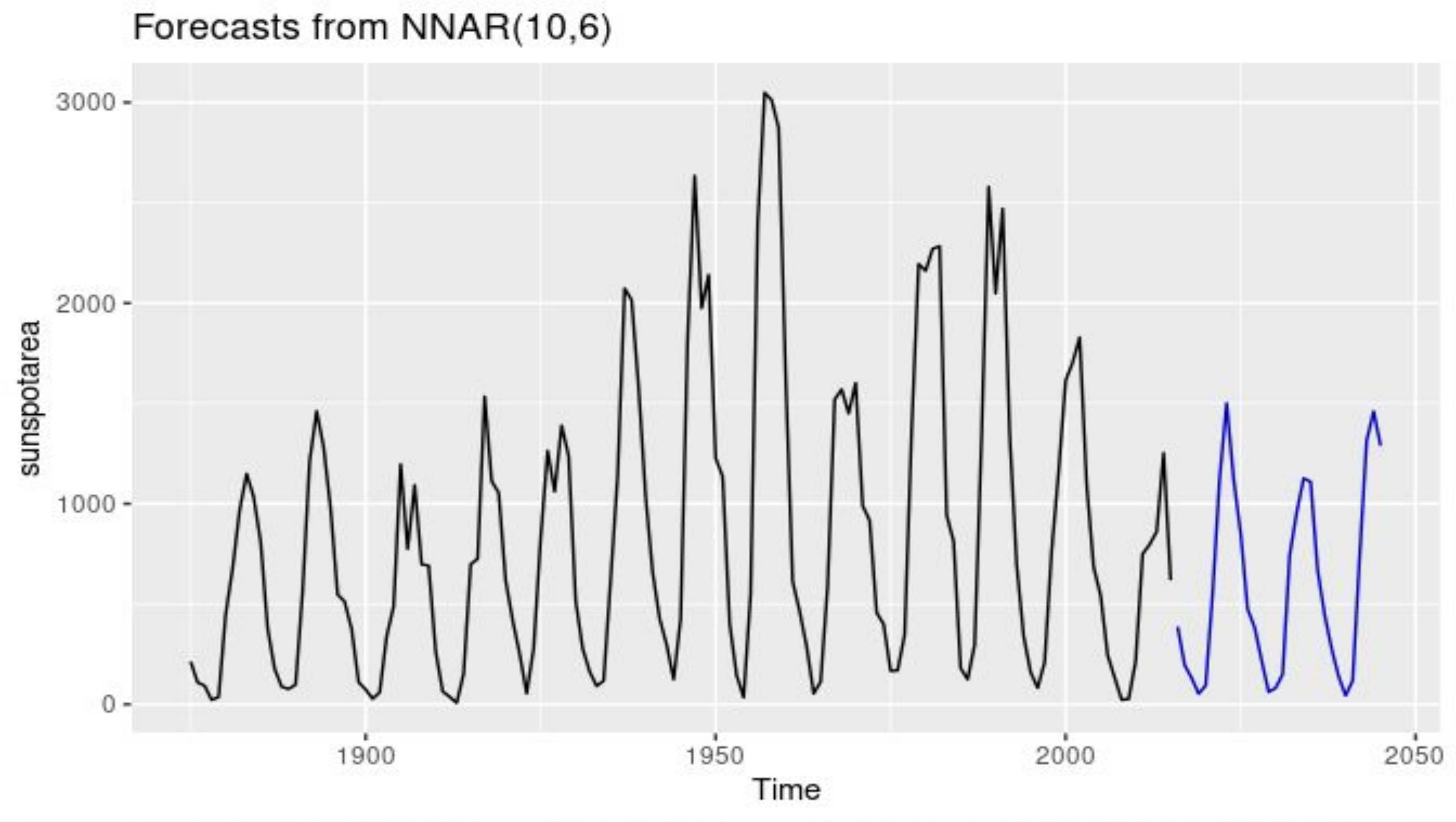
- В роли признаков - n предыдущих наблюдений ряда:

$$\hat{y}_{t+1}(w) = \sum_{j=1}^{n} w_j y_{t-j+1}, \quad w \in \mathbb{R}^n$$

НЕЙРОСЕТЕВЫЕ МОДЕЛИ ПРОГНОЗИРОВАНИЯ



ПРОГНОЗ ЧИСЛА ПЯТЕН НА СОЛНЦЕ



ЛИТЕРАТУРА

Прогнозирование: принципы и практика

Роб Джей Хайндман и Джордж Афанасопулос

https://otexts.com/fpp2/