Прогнозирование временного ряда

Временной ряд: $y_1, \dots, y_T, \dots, \quad y_t \in \mathbb{R}$, — значения признака, измеренные через постоянные временные интервалы.

Задача прогнозирования — найти функцию f_T:

$$y_{T+d} \approx f_T\left(y_T, \dots, y_1, d\right) \equiv \hat{y}_{T+d|T},$$

где $d \in \{1, 2, \dots, D\}$ — отсрочка прогноза, D — горизонт прогнозирования.

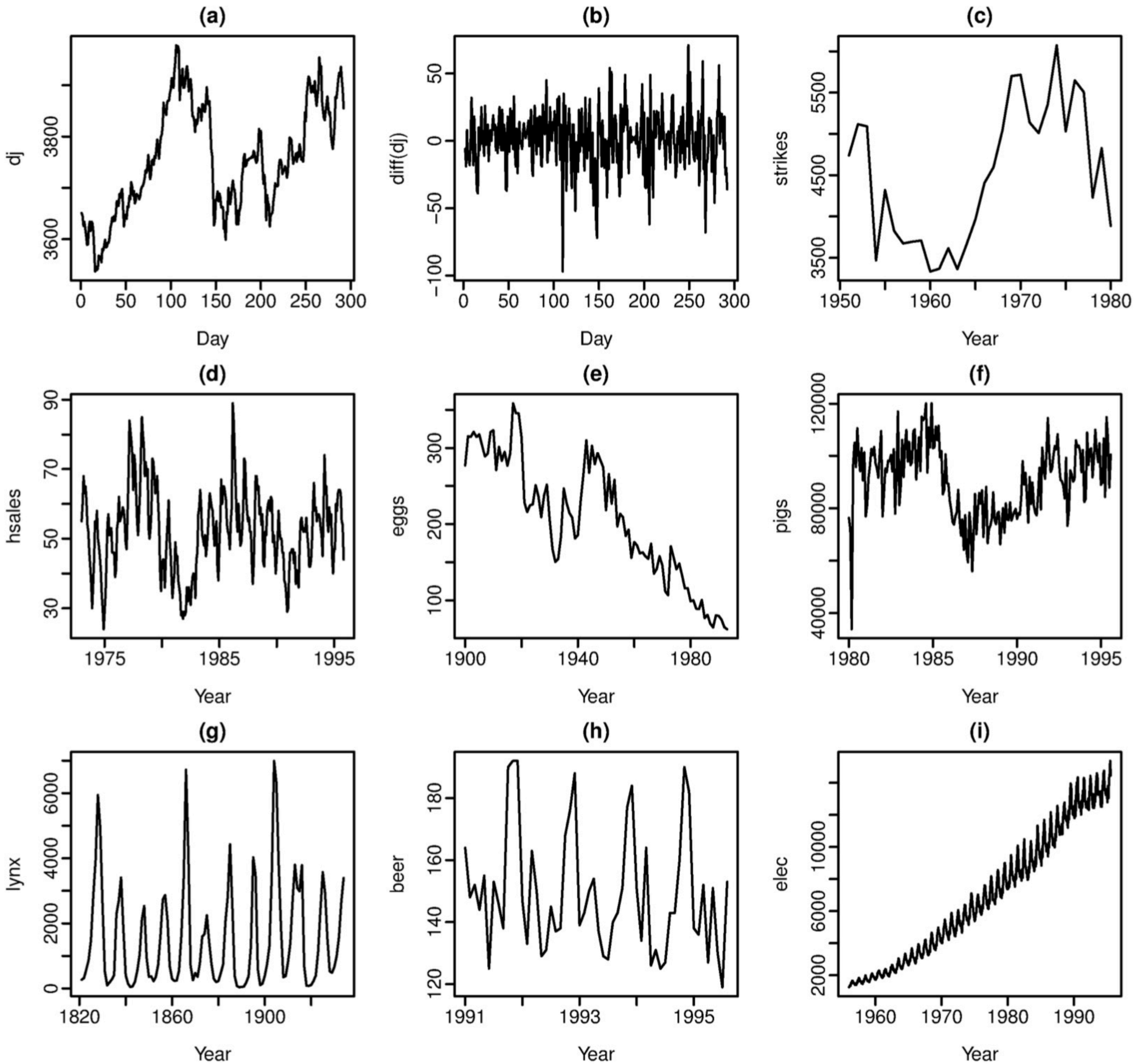
Стационарность

Ряд $y_1, \ldots, y_T$ **стационарен**, если $\forall s$ распределение $y_t, \ldots, y_{t+s}$ не зависит от t, т. е. его свойства не зависят от времени.

Ряды с трендом или сезонностью нестационарны.

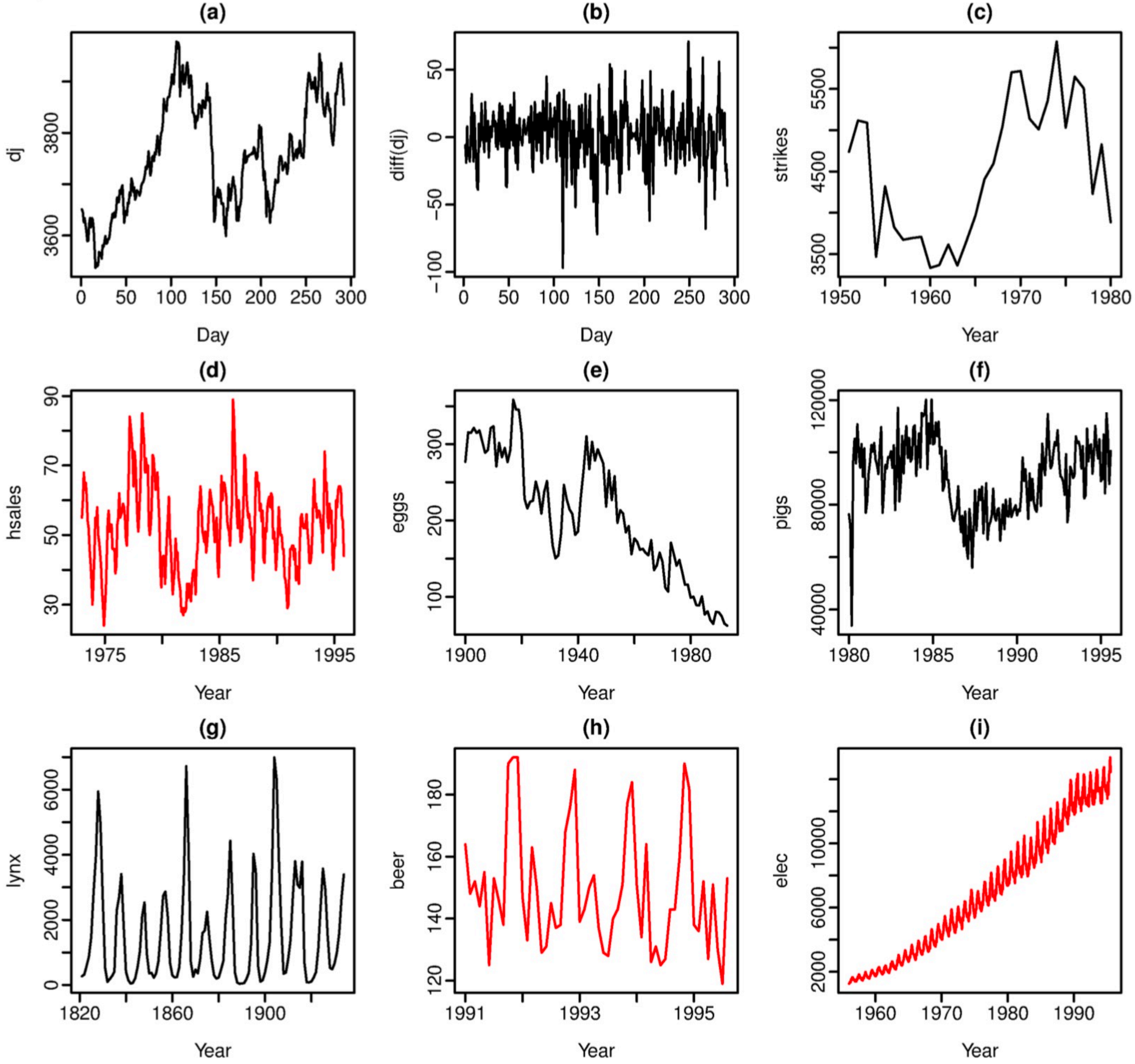
Ряды с непериодическими циклами стационарны, поскольку нельзя предсказать заранее, где будут находится максимумы и минимумы.

Стационарность

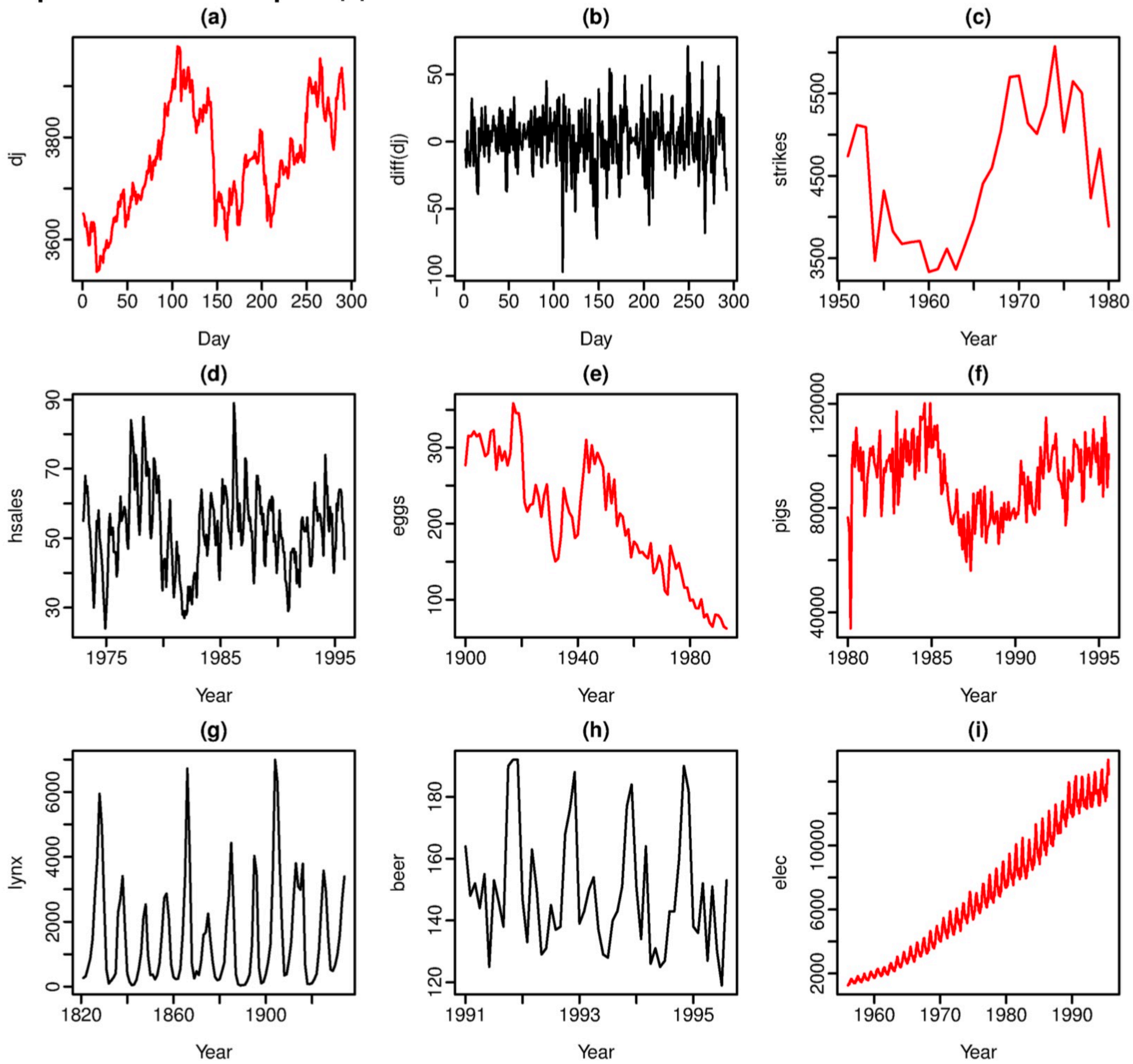


Стационарность

Нестационарны из-за сезонности:

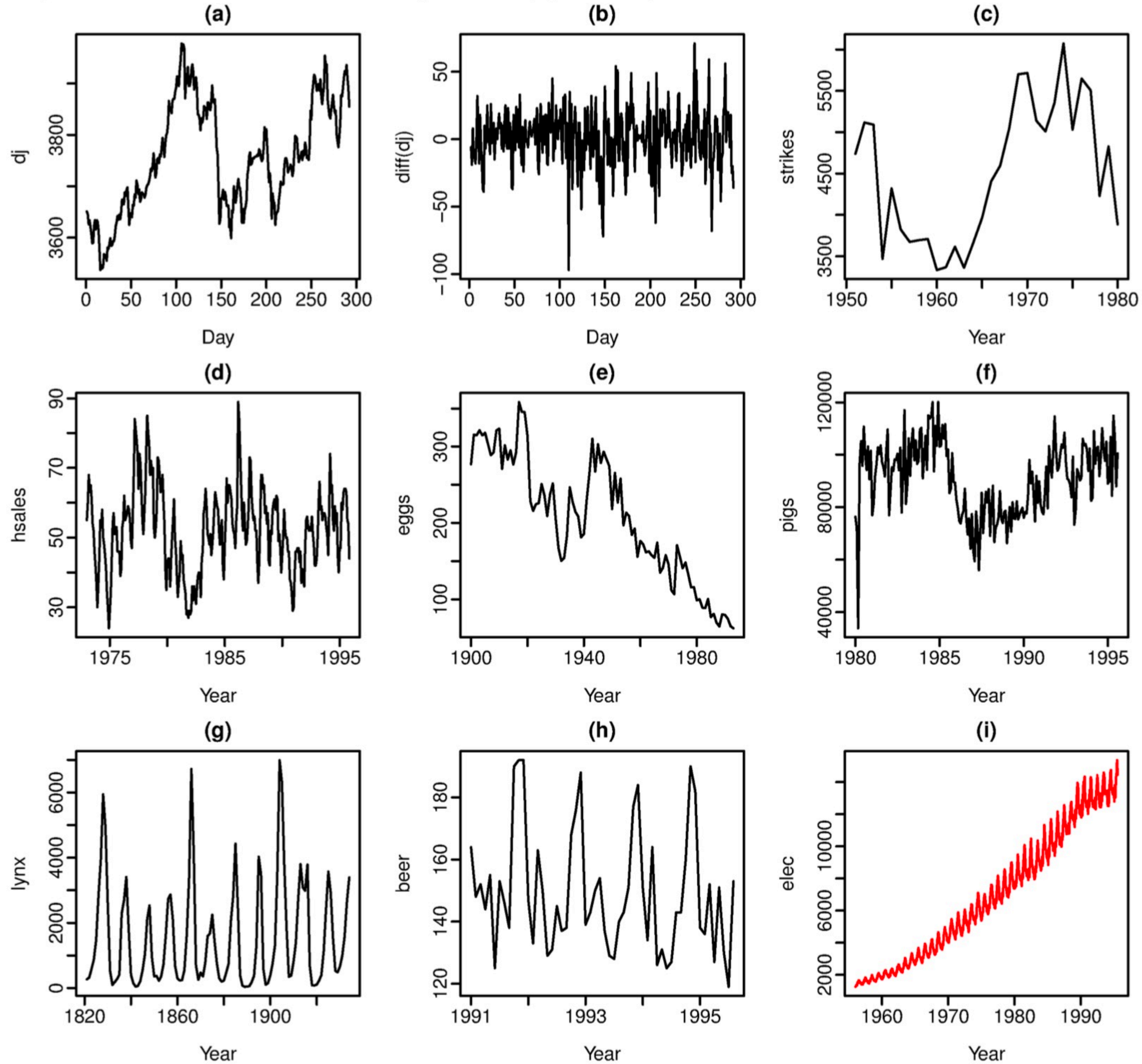


Нестационарны из-за тренда:



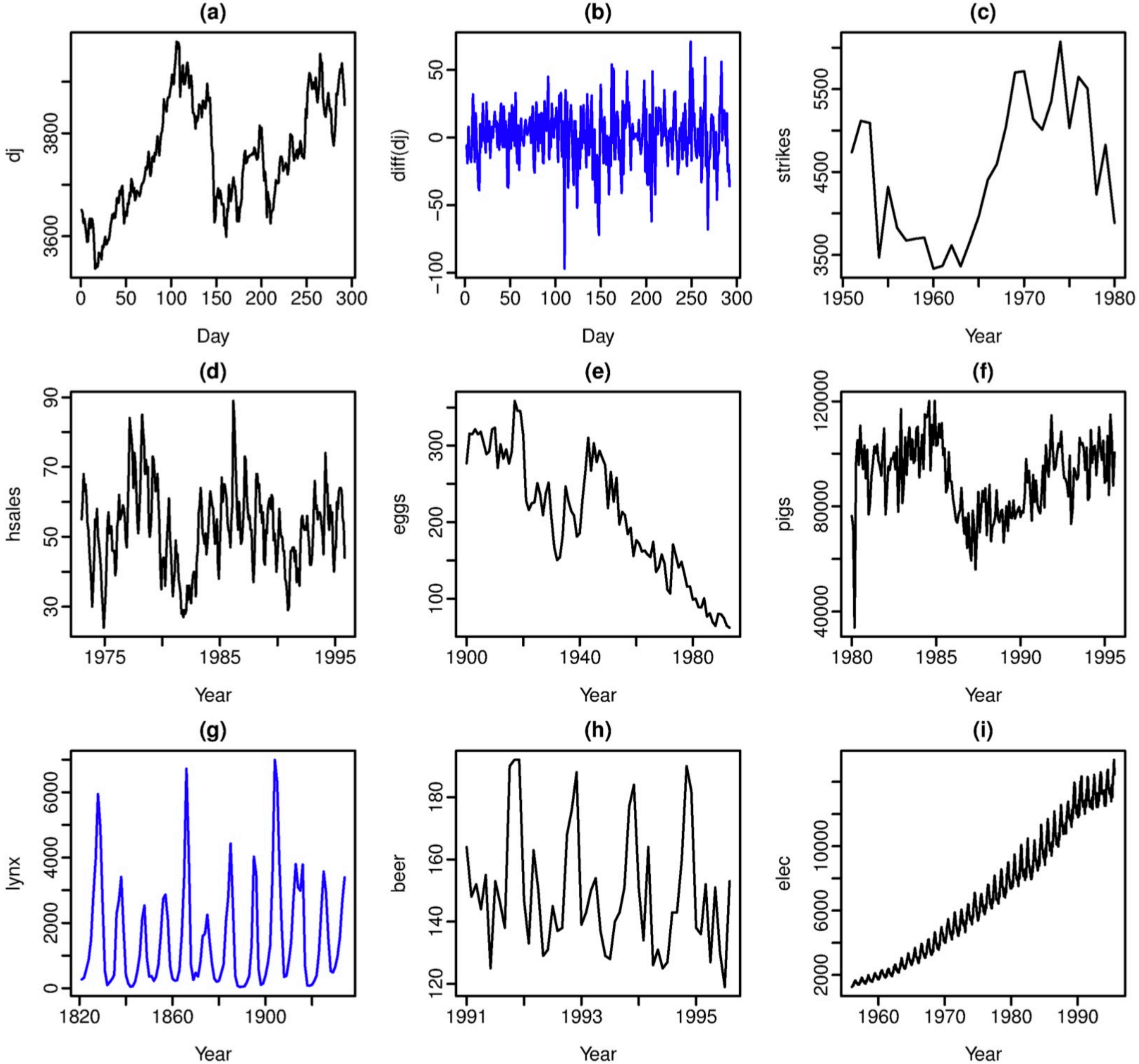
Стационарность

Нестационарны из-за меняющейся дисперсии:



Стационарность

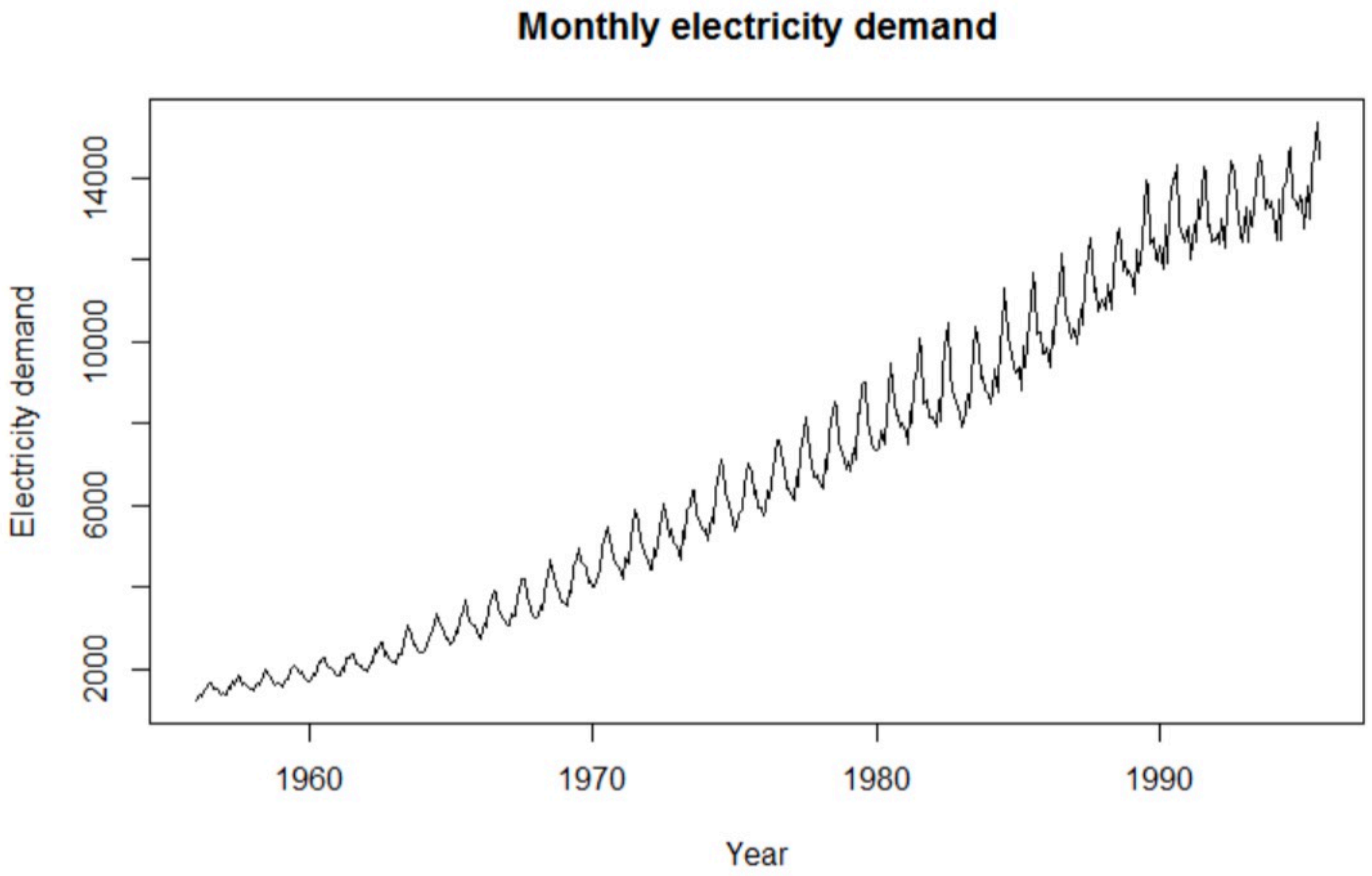
Стационарны:

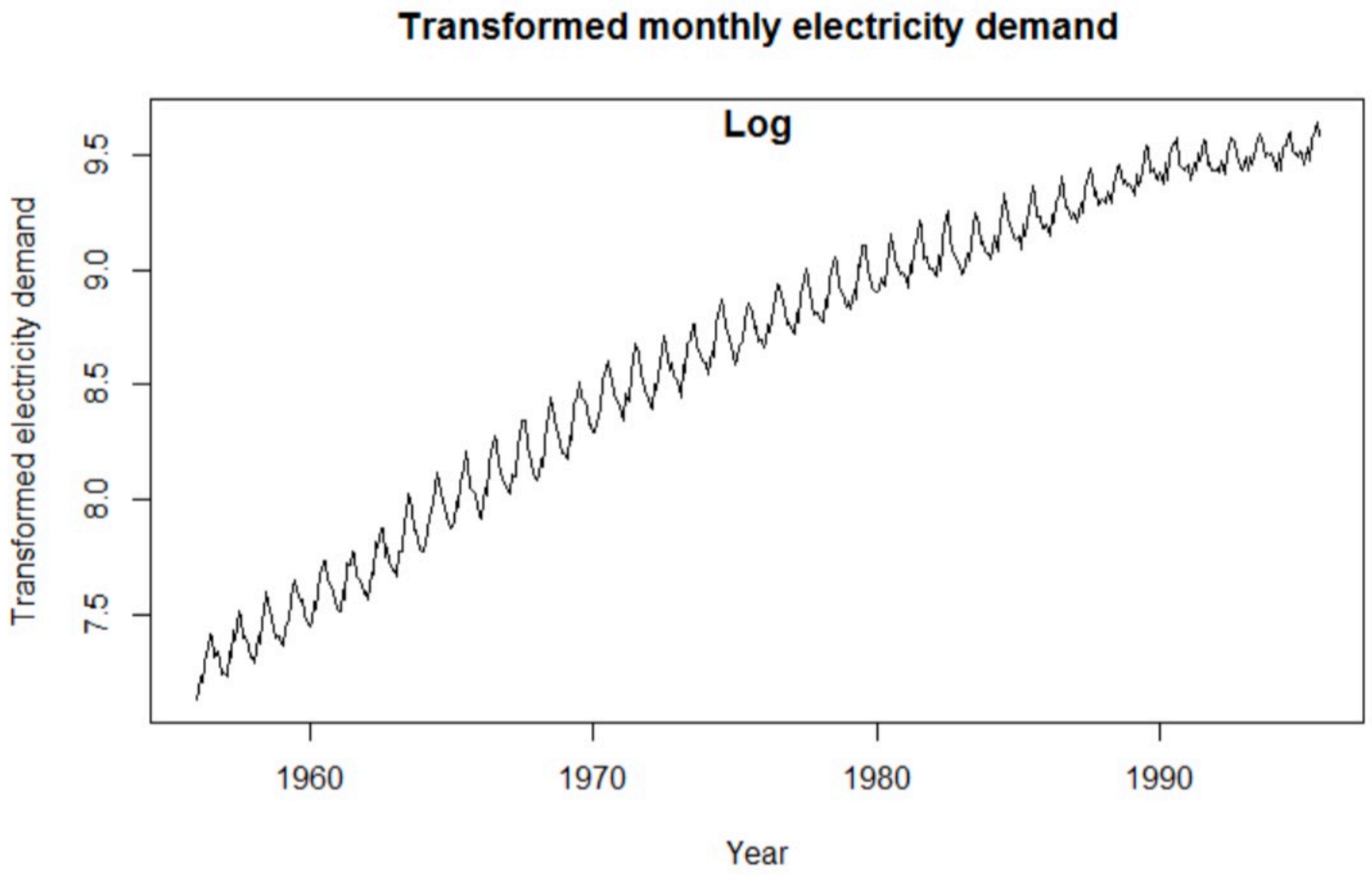


Стабилизация дисперсии

Для рядов с монотонно меняющейся дисперсией можно использовать стабилизирующие преобразования.

Часто используют логарифмирование:



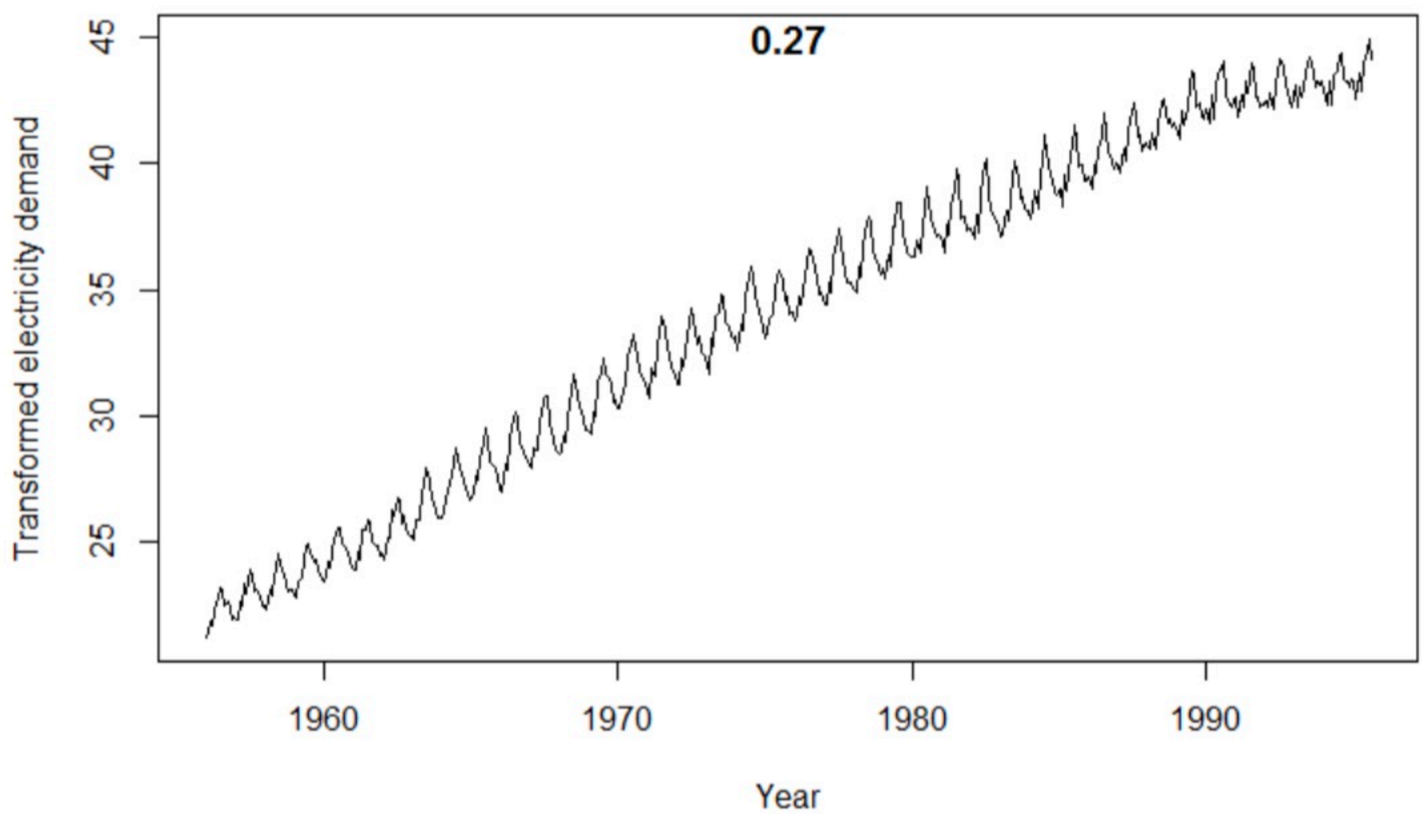


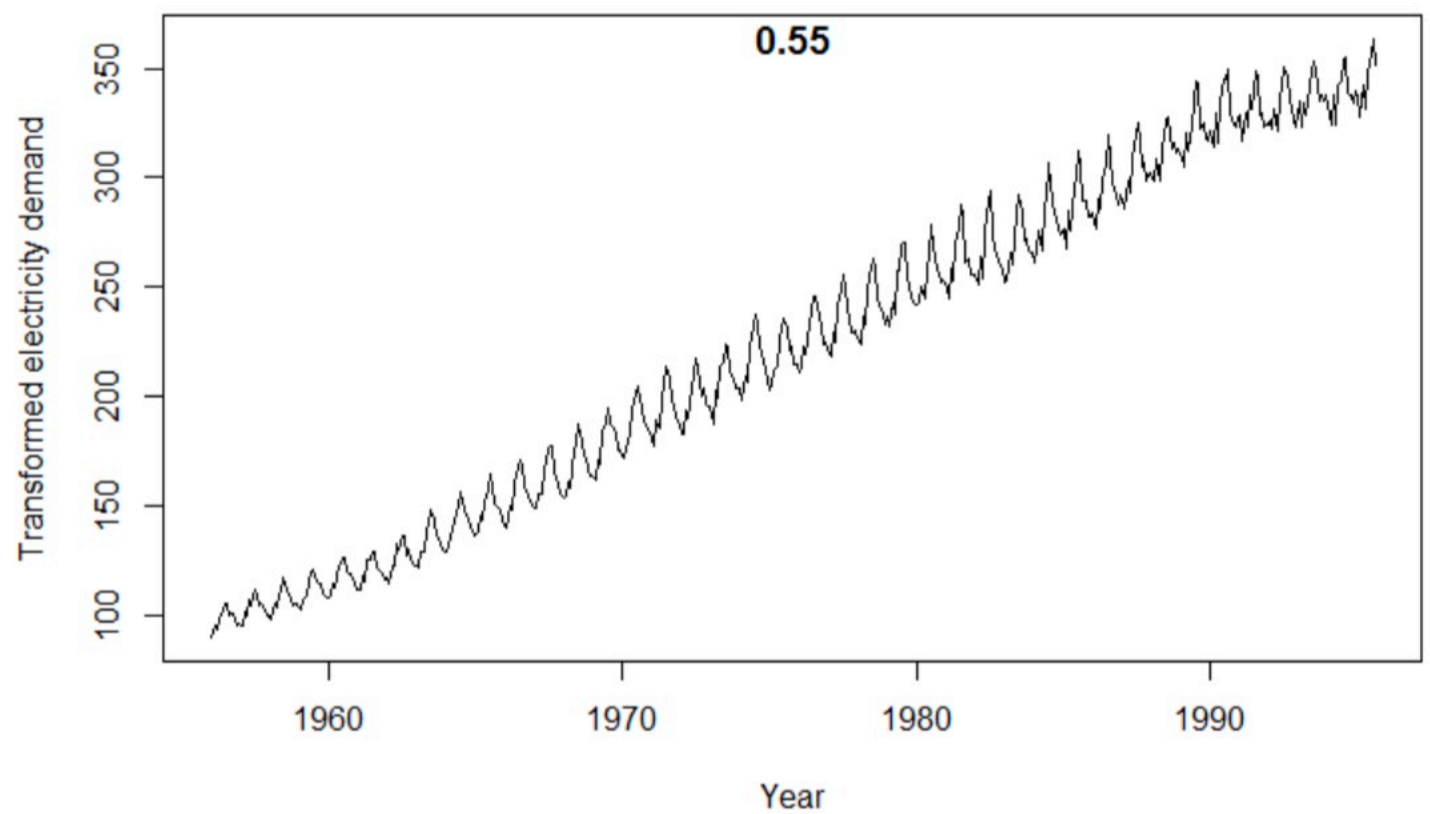
Преобразования Бокса-Кокса

Параметрическое семейство стабилизирующих дисперсию преобразований:

$$y'_t = \begin{cases} \ln y_t, & \lambda = 0, \\ \left(y_t^\lambda - 1\right)/\lambda, & \lambda \neq 0. \end{cases}$$

Параметр λ выбирается так, чтобы минимизировать дисперсию или максимизировать правдоподобие модели.





Преобразования Бокса-Кокса

После построения прогноза для трансформированного ряда его нужно преобразовать в прогноз исходного:

$$\hat{y}_t = \begin{cases} \exp\left(\hat{y}'_t\right), & \lambda = 0, \\ \left(\lambda \hat{y}'_t + 1\right)^{1/\lambda}, & \lambda \neq 0. \end{cases}$$

- Если некоторые $y_t \leq 0$, преобразования Бокса-Кокса невозможны (нужно прибавить к ряду константу).
- Часто оказывается, что преобразование вообще не нужно.
- Стоит округлять значение λ, чтобы упростить интерпретацию.
- Как правило, стабилизирующее преобразование слабо влияет на прогноз и сильно — на предсказательный интервал.

Дифференцирование

Дифференцирование ряда — переход к попарным разностям его соседних значений:

$$y_1, \dots, y_T \;\longrightarrow\; y'_2, \dots, y'_T,$$

$$y'_t = y_t - y_{t-1}.$$

Дифференцированием можно стабилизировать среднее значение ряда и избавиться от тренда и сезонности.

Может применяться неоднократное дифференцирование; например, для второго порядка:

$$y_1, \dots, y_T \longrightarrow y'_2, \dots, y'_T \longrightarrow y''_3, \dots, y''_T,$$

$$y''_t = y'_t - y'_{t-1} = y_t - 2y_{t-1} + y_{t-2}.$$

Сезонное дифференцирование

Сезонное дифференцирование ряда — переход к попарным разностям его значений в соседних сезонах:

$$y_1, \dots, y_T \longrightarrow y'_{s+1}, \dots, y'_T,$$

$$y'_t = y_t - y_{t-s}.$$

Комбинированное дифференцирование

Сезонное и обычное дифференцирование может применяться к одному ряду в любом порядке.

Если ряд имеет выраженный сезонный профиль, рекомендуется начинать с сезонного дифференцирования — после него ряд уже может оказаться стационарным.

Авторегрессия

$$AR(p): \quad y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + \varepsilon_t,$$

где y_t — стационарный ряд с нулевым средним, $\phi_1, \dots, \phi_p$ — константы ($\phi_p \neq 0$), ε_t — гауссов белый шум с нулевым средним и постоянной дисперсией σ_ε^2.

Если среднее равно μ, модель принимает вид

$$y_t = \alpha + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + \varepsilon_t,$$

где $\alpha = \mu\left(1 - \phi_1 - \dots - \phi_p\right)$.

Другой способ записи:

$$\phi(B)\, y_t = \left(1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p\right) y_t = \varepsilon_t,$$

где B — разностный оператор ($B y_t = y_{t-1}$).

Линейная комбинация p подряд идущих членов ряда даёт белый шум.

Авторегрессия

Чтобы ряд AR(p) был стационарным, должны выполняться ограничения на коэффициенты. Например,

- в AR(1) необходимо $-1 < \phi_1 < 1$;
- в AR(2) необходимо $-1 < \phi_2 < 1,\ \phi_1 + \phi_2 < 1,\ \phi_2 - \phi_1 < 1$.

С ростом p вид ограничений усложняется.

Скользящее среднее

$$MA(q): \quad y_t = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \theta_q \varepsilon_{t-q},$$

где y_t — стационарный ряд с нулевым средним, $\theta_1, \dots, \theta_q$ — константы ($\theta_q \neq 0$), ε_t — гауссов белый шум с нулевым средним и постоянной дисперсией σ_ε^2.

Если среднее равно μ, модель принимает вид

$$y_t = \mu + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \theta_q \varepsilon_{t-q}.$$

Другой способ записи:

$$y_t = \theta(B)\,\varepsilon_t = \left(1 + \theta_1 B + \theta_2 B^2 + \dots + \theta_q B^q\right)\varepsilon_t,$$

где B — разностный оператор.

Линейная комбинация q подряд идущих компонент белого шума ε_t даёт элемент ряда.

Скользящее среднее

Чтобы ряд модель MA(q) была обратимой, должны выполняться ограничения на коэффициенты. Например,

- в MA(1) необходимо $-1 < \theta_1 < 1$;
- в MA(2) необходимо $-1 < \theta_2 < 1,\ \theta_1 + \theta_2 > -1,\ \theta_1 - \theta_2 < 1$.

С ростом q вид ограничений усложняется.

ARMA (Autogerressive moving average)

$$ARMA(p,q): \quad y_t = \phi_1 y_{t-1} + \cdots + \phi_p y_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \cdots + \theta_q \varepsilon_{t-q},$$

где y_t — стационарный ряд с нулевым средним, $\phi_1, \ldots, \phi_p, \theta_1, \ldots, \theta_q$ — константы ($\phi_p \neq 0$, $\theta_q \neq 0$), ε_t — гауссов белый шум с нулевым средним и постоянной дисперсией σ_ε^2.

Если среднее равно μ, модель принимает вид

$$y_t = \alpha + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \cdots + \theta_q \varepsilon_{t-q},$$

где $\alpha = \mu \left(1 - \phi_1 - \cdots - \phi_p\right)$.

Другой способ записи:

$$\phi\left(B\right) y_t = \theta\left(B\right) \varepsilon_t.$$

Согласно теорема Вольда, любой стационарный ряд может быть аппроксимирован моделью ARMA(p,q).

ARIMA (Autogerressive integrated moving average)[1]

Ряд описывается моделью $ARIMA(p, d, q)$, если ряд его разностей

$$\nabla^d y_t = (1 - B)^d \, y_t$$

описывается моделью $ARMA(p, q)$.

$$\phi\left(B\right) \nabla^d y_t = \theta\left(B\right) \varepsilon_t.$$

---

[1] Также это энергетическое имя, данное творцом первозданным двум своим посланникам для работы планете Земля, подробности см. `http://light-love.ru/nasha-istoriya/ot-avtorov.html/`

Частичная автокорреляционная функция

Частичная автокорреляция стационарного ряда y_t:

$$\phi_{hh} = \begin{cases} r\left(y_{t+1}, y_t\right), & h = 1, \\ r\left(y_{t+h} - y_{t+h}^{h-1}, y_t - y_t^{h-1}\right), & h \geq 2, \end{cases}$$

где y_t^{h-1} — регрессия y_t на $y_{t+1}, y_{t+2}, \dots, y_{t+h-1}$:

$$y_t^{h-1} = \beta_1 y_{t+1} + \beta_2 y_{t+2} + \dots + \beta_{h-1} y_{t+h-1},$$
$$y_{t+h}^{h-1} = \beta_1 y_{t+h-1} + \beta_2 y_{t+h-2} + \dots + \beta_{h-1} y_{t+1}.$$

Оценка параметров модели

- При заданных p, d, q коэффициенты модели оцениваются методом максимального правдоподобия; функционал качества — логарифм правдоподобия LL.
- d выбирается так, чтобы ряд был стационарным.
- p и q нельзя выбирать из принципа максимума правдоподобия: LL всегда увеличивается с ростом p и q.
- При выборе p и q погут помочь автокорреляционные функции ACF и PACF:
  - в модели $ARIMA(p, d, 0)$ ACF экспоненциально затухает или имеет синусоидальный вид, а PACF значимо отличается от нуля при лаге p;
  - в модели $ARIMA(0, d, q)$ PACF экспоненциально затухает или имеет синусоидальный вид, а ACF значимо отличается от нуля при лаге q.

Прогнозирование с помощью ARIMA

1. Строится график ряда, идентифицируются необычные значения.
2. При необходимости делается стабилизирующее дисперсию преобразование.
3. Если ряд нестационарен, подбирается порядок дифференцирования.
4. Анализируются ACF/PACF, чтобы понять, можно ли использовать модели AR(p)/MA(q).
5. Обучаются модели-кандидаты, сравнивается их AICc.
6. Остатки полученной модели исследуются на несмещённость, стационарность и неавтокоррелированность; если предположения не выполняются, исследуются модификации модели.
7. В финальной модели t заменяется на $T+h$, будущие наблюдения — на их прогнозы, будущие ошибки — на нули, прошлые ошибки — на остатки.

Seasonal multiplicative ARMA/ARIMA

$$ARMA(p,q) \times (P,Q)_s: \quad \Phi_P\left(B^s\right) \phi\left(B\right) y_t = \alpha + \Theta_Q\left(B^s\right) \theta\left(B\right) \varepsilon_t,$$

где

$$\Phi_P\left(B^s\right) = 1 - \Phi_1 B^s - \Phi_2 B^{2s} - \cdots - \Phi_P B^{Ps},$$

$$\Theta_Q\left(B^s\right) = 1 + \Theta_1 B^s + \Theta_2 B^{2s} + \cdots + \Theta_Q B^{Ps}.$$

SARIMA:

$$\Phi_P\left(B^s\right) \phi\left(B\right) \nabla_s^D \nabla^d y_t = \alpha + \Theta_Q\left(B^s\right) \theta\left(B\right) \varepsilon_t.$$