M. Sc. Geodesy and Geoinformation Science

Module:

Analysis of Stochastic Processes WS 25/26

Exercise 3: Regression Analysis

GROUP 2

Pugin, Luke J.*	Reinoso Rojas, Víctor A.*	de Seriis, Jonas B.*	Li, Shuo*
411499	479301	515623	515067

December 10, 2025

*Group work completed by all group members. This assignment is therefore valid without physical signature

1. Introduction

Time series arising in geodetic and engineering applications often contain both a deterministic trend and a stochastic component. In order to interpret the underlying physical process or to perform further statistical analyses (such as correlation or spectral investigations), it is necessary to separate these two parts in a consistent way.

The aim of this assignment is to apply the concepts from the lecture *Analysis of Stochastic Processes* to several synthetic time series with different types of trends. The data sets range from almost linear behaviour, over clearly curved (polynomial) trends, to genuinely nonlinear growth laws such as exponential and logistic functions. For each series an appropriate regression model is chosen, its parameters are estimated by least squares, and the corresponding trend is removed. The resulting detrended sequences are then compared and interpreted with respect to their remaining stochastic structure.

The report is organised as follows. Task 1 deals with linear and quadratic trend models for two simple time series. Task 2 extends this idea to higher–order polynomials and investigates model selection by means of significance tests and residual analysis. Task 3 introduces nonlinear trend models (exponential and logistic) and demonstrates how they can be handled within the least–squares framework. Finally, Task 4 compares two detrended series that originate from the same underlying data and discusses the consistency of the resulting residuals.

2. Methods

The overall analysis follows the standard regression and detrending workflow presented in the lecture and can be summarised in four main steps.

2.1. Visual inspection and model choice

For each time series the raw data are first plotted as a function of time. The qualitative shape of the graph (approximately linear, parabolic, purely convex, sigmoidal with saturation, . . .) is used to formulate an initial hypothesis about a suitable functional model:

- low–order polynomials in time for approximately linear or smoothly curved trends,
- exponential functions for purely convex growth with rate proportional to the current value,
- logistic-type functions for monotone growth with upper (and, if necessary, lower) asymptotes.

This step is purely exploratory, but it provides the structural form of the regression model used in the subsequent adjustment.

2.2. Linear least-squares adjustment

Whenever the model is linear in the unknown parameters (for example for polynomials in time or after a suitable transformation such as the logarithm for an exponential law), parameter estimation is carried out using the Gauss–Markov model. The observation vector contains the measured values of the series; the design matrix is constructed from the chosen basis functions in time (constant term, powers of t, etc.). All measurements are assumed to be equally weighted and uncorrelated, so the weight matrix is taken as the identity.

For each model the normal equations are solved to obtain the parameter estimates, residuals and the empirical reference standard deviation. From the inverse normal matrix the standard deviations of the estimated parameters are computed, which are then used to form t-values and perform two-sided significance tests at the chosen confidence level.

2.3. Nonlinear regression

For the genuinely nonlinear models (logistic trends with or without offset), the regression function cannot be linearised in a simple closed form. In these cases an iterative nonlinear least-squares approach is employed. The nonlinear model $f(p, t)$ is specified in terms of a parameter vector p and the time t, and the optimal parameters are obtained as the minimiser of the sum of squared residuals

$$\min_p \sum_i \bigl(x(t_i) - f(p, t_i)\bigr)^2.$$

In the implementation this optimisation is carried out with the MATLAB routine `lsqcurvefit`, using physically reasonable initial guesses for the parameters. Convergence is checked by monitoring the stability of the parameter updates and the residual norm.

2.4. Detrending and residual analysis

Once the regression parameters have been estimated, the fitted trend $\hat{x}(t)$ is evaluated at all observation epochs and subtracted from the original series,

$$x_{\text{det}}(t_i) = x(t_i) - \hat{x}(t_i).$$

The detrended series is then analysed graphically and, where appropriate, by simple statistics such as empirical standard deviation or cross-correlation. The main diagnostic criteria are:

- absence of systematic curvature or long-term drift in the residuals,
- approximately constant variance over time (no obvious heteroscedasticity),
- statistical significance of the model parameters, especially of the highest-order term in polynomial models.

If these criteria are not satisfied, the functional model is revised (e.g. by increasing the polynomial degree or by adopting a different nonlinear form) and the procedure is repeated.

The detailed application of this methodology to each dataset, together with the numerical results and figures, is presented in the subsequent task sections.

3. Task 1: Linear and quadratic trend models

Task 1 provides two time series: the first contains acceleration measurements, while the second represents observation data that follows a clearly curved trend.

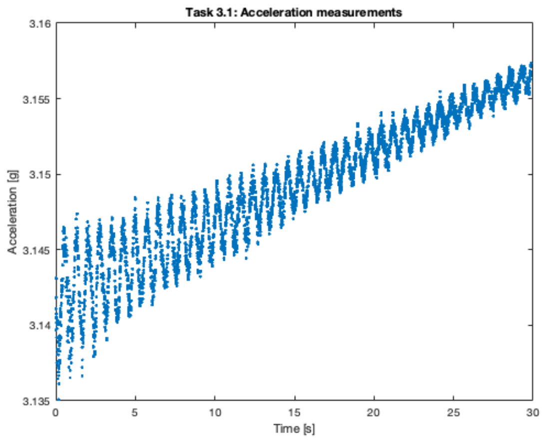

Figure 1: Acceleration data (Exercise3-1.txt).

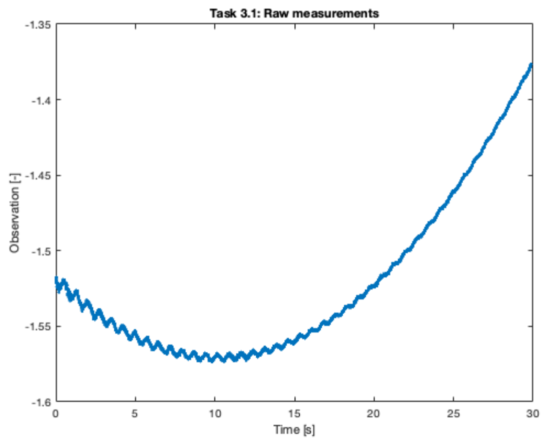

Figure 2: Observation data (Exercise3-2.txt).

We see that the first time series oscillates around an approximately constant value with a very weak linear trend, whereas the second time series exhibits a pronounced curved (parabolic) behaviour.

3.1. Results

To perform a regression analysis, we need to specify an appropriate functional model for each series and estimate its parameters by least squares. The resulting regression functions can later be used for de-trending and further analysis.

3.1.1 Acceleration data from `Exercise3-1.txt`

For the first dataset, the trend can be reasonably approximated by a **linear** model. We therefore use the functional model

$$\mathrm{reg}_1(t) = a\,t + b, \tag{1}$$

with unknown parameters a and b.

The parameters are estimated by a Gauss–Markov least-squares adjustment[1], assuming an identity covariance matrix (equally weighted, uncorrelated observations). The resulting regression line is shown in Fig. 3.

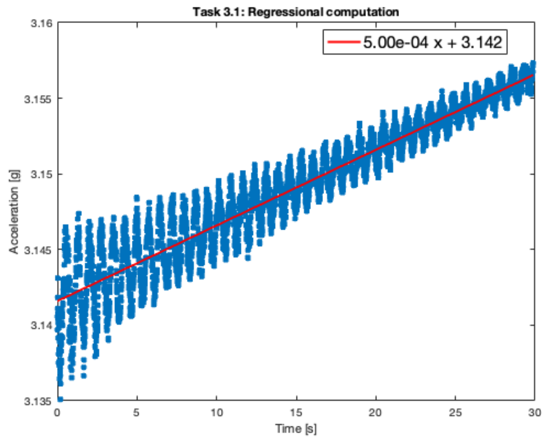

Figure 3: Acceleration data and estimated linear regression line.

Numerically, we obtain for example

$$a \approx 0.0005, \qquad b \approx 3.142, \tag{2}$$

so that

$$\mathrm{reg}_1(t) \approx 0.0005\,t + 3.142. \tag{3}$$

The detrended series is obtained by subtracting the regression function from the raw observations,

$$x_1^{\mathrm{det}}(t_i) = x_1(t_i) - \mathrm{reg}_1(t_i), \tag{4}$$

leading to a time series that fluctuates around zero with a harmonic pattern (Fig. 4).

3.1.2 Observation data from `Exercise3-2.txt`

In contrast, the second dataset clearly follows a **parabolic** trend. A linear model would not be sufficient, so we adopt a quadratic functional model

$$\mathrm{reg}_2(t) = a\,t^2 + b\,t + c, \tag{5}$$

with unknowns a, b and c.

[1] As introduced in Adjustment Theory I for polynomial functional models.

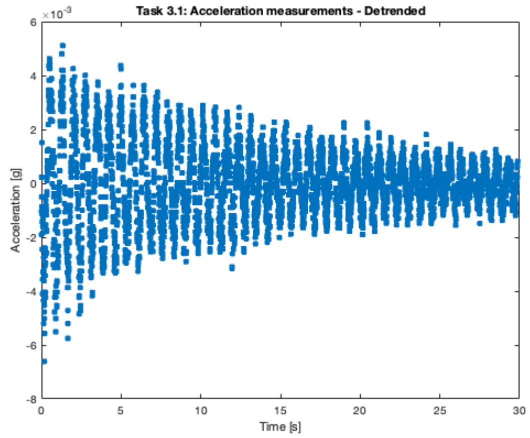

Figure 4: Detrended acceleration data after removing the linear trend.

Again, the parameters are estimated by a Gauss–Markov least-squares adjustment using equally weighted and uncorrelated observations. The resulting polynomial fit is illustrated in Fig. 5.

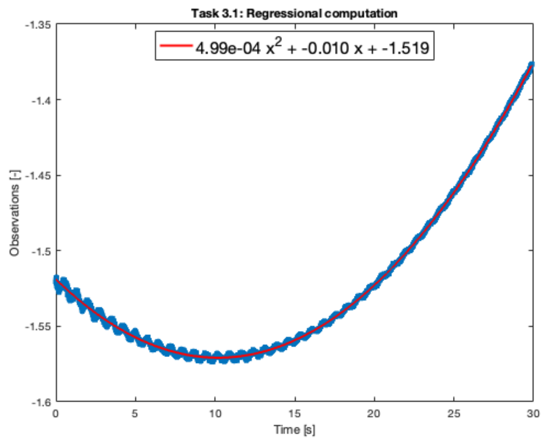

Figure 5: Quadratic regression function fitted to the observation data.

The adjustment yields, for example,

$$a \approx 4.99 \times 10^{-4}, \qquad b \approx -0.010, \qquad c \approx -1.519, \tag{6}$$

so that the regression function can be written as

$$\mathrm{reg}_2(t) \approx 4.99 \times 10^{-4}\, t^2 - 0.010\, t - 1.519. \tag{7}$$

The detrended series is then obtained in complete analogy to the first dataset:

$$x_2^{\mathrm{det}}(t_i) = x_2(t_i) - \mathrm{reg}_2(t_i), \tag{8}$$

resulting in a time series that fluctuates around zero with a nearly harmonic pattern (Fig. 6).

3.2. Interpretation

After applying a Gauss–Markov least-squares adjustment with a linear trend model for the first dataset

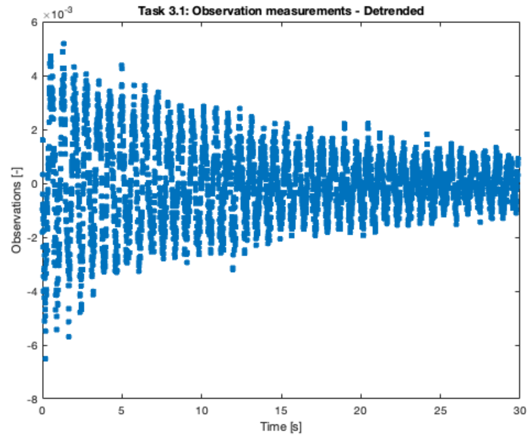

Figure 6: Detrended observation data after removing the quadratic trend.

(Section 3.1.1) and a quadratic model for the second dataset (Section 3.1.2), and subsequently removing these trends, the two detrended time series look very similar.

To quantify this, we compute the sample cross-correlation between the two detrended series. As shown in Fig. 7, the maximum of the cross-correlation function is

$$r_{\mathrm{cross}} \approx 0.999, \tau_{shift} = 0,$$

which indicates that both detrended time series exhibit an almost identical noise and harmonic observation behavior.

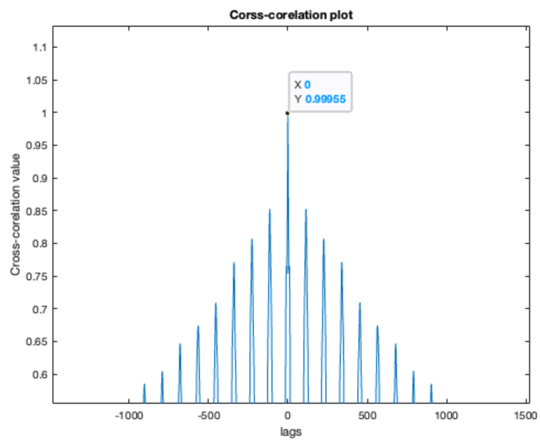

Figure 7: Cross-correlation between the two detrended time series.

Further time series analysis or de-trending is not possible, as the remaining time series is non-stationary (variable variance, with a constant mean), meaning that performing a least-squares adjustment for a sine wave fitting is not possible, as the amplitude can not be definitively estimated. However when switching from the time domain to the freuqency domain, provides new possibility to further de-trend or analyze the already LSA-detrended time series.

4. Task 2: Linear trend models

The time series from `Exercise3-3.txt` consists of n equally spaced measurements $x(t_i)$, where the first column contains the time t_i and the second column the observation values $x_i = x(t_i)$.[2]

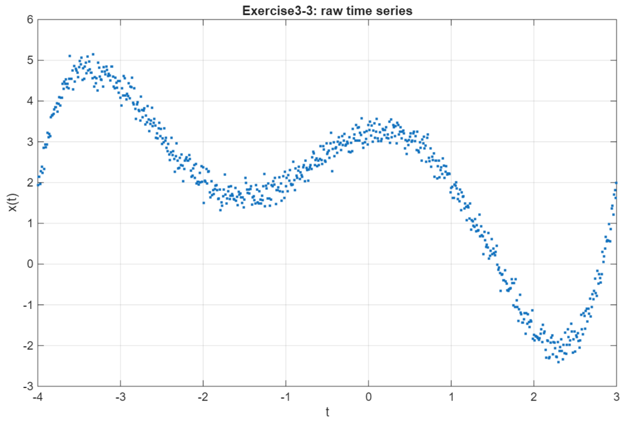

Figure 8: Raw time series from `Exercise3-3.txt`.

4.1. Functional and stochastic model

Following the lecture, we assume a *linear regression model in the parameters*, where the trend is approximated by a polynomial of degree k in time:

$$x(t_i) = \beta_0 + \beta_1 t_i + \beta_2 t_i^2 + \cdots + \beta_k t_i^k + v_i, \tag{9}$$

with unknown coefficients β_j and residuals v_i.

The corresponding design matrix is

$$A = \begin{bmatrix} 1 & t_1 & t_1^2 & \dots & t_1^k \\ 1 & t_2 & t_2^2 & \dots & t_2^k \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & t_n & t_n^2 & \dots & t_n^k \end{bmatrix}, \tag{10}$$

and the observation vector is

$$\mathbf{l} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}. \tag{11}$$

The stochastic model follows the Gauss–Markov assumptions: the x_i are equally weighted and uncorrelated, and the time t is regarded as error free. Hence we set

$$\mathbf{P} = \mathbf{I}_n, \qquad \mathrm{Cov}(\mathbf{l}) = \sigma_0^2 \mathbf{P}. \tag{12}$$

4.2. Least-squares adjustment and t-tests

For a given degree k, the least-squares solution is obtained from the normal equations

$$\hat{\boldsymbol{\beta}} = (A^\top A)^{-1} A^\top \mathbf{l}, \tag{13}$$

[2]The dataset was provided as part of the course material.

with residuals

$$\mathbf{v} = \mathbf{l} - A\hat{\boldsymbol{\beta}}. \tag{14}$$

The empirical reference standard deviation is

$$\hat{s}_0 = \sqrt{\frac{\mathbf{v}^\top \mathbf{v}}{n-u}}, \tag{15}$$

where $u = k+1$ is the number of unknown parameters.

The cofactor matrix of the estimated parameters is given by

$$Q_{\hat{\boldsymbol{\beta}}} = (A^\top A)^{-1}, \tag{16}$$

and the standard deviations of the individual coefficients read

$$\sigma_{\hat{\beta}_j} = \hat{s}_0 \sqrt{(Q_{\hat{\boldsymbol{\beta}}})_{jj}}. \tag{17}$$

For each coefficient we compute a t-value

$$t_j = \frac{\hat{\beta}_j}{\sigma_{\hat{\beta}_j}}, \tag{18}$$

and perform a two-sided t-test with significance level $\alpha = 5\%$ and $\nu = n - u$ degrees of freedom. A coefficient is considered statistically significant if

$$|t_j| > t_{\alpha/2,\nu}. \tag{19}$$

4.3. Model selection by increasing the polynomial degree

We followed the strategy described in the assignment: starting from a straight line ($k = 1$), we increased the polynomial degree step by step, evaluated the significance of all coefficients and inspected the residuals. The main numerical results are:

- **Degree** $k = 1$ (straight line):

  $$x(t) = \beta_0 + \beta_1 t.$$

  The adjustment yields an empirical standard deviation of $\hat{s}_0 = 1.1650$. The estimated coefficients, their standard deviations, t-values, and significance decisions are listed in Table 1. The critical value for the two-sided t-test at $\alpha = 5\%$ and $\nu = 699$ degrees of freedom is $t_{0.975,699} \approx 1.963$.

Table 1: Estimated coefficients, standard deviations, t-values, and significance (degree $k = 1$).

Coefficient	$\hat{\beta}_j$	$\sigma_{\hat{\beta}_j}$	t_j	Significant?
β_0	$+1.60036$	0.04533	$+35.308$	Yes
β_1	-0.73071	0.02175	-33.604	Yes

Both coefficients are clearly significant ($|t_j| \gg 1.963$). However, the residual plot in Fig. 9 displays a strong, almost sinusoidal systematic pattern, indicating that a straight-line trend is insufficient to adequately model the time series.

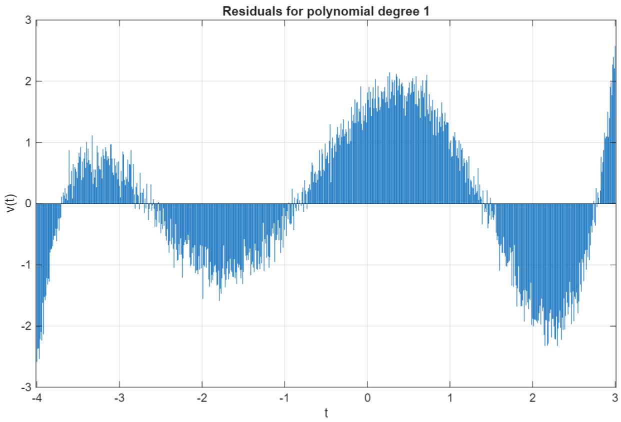

Figure 9: Residuals for polynomial degree $k = 1$.

- **Degree** $k = 2$:

$$x(t) = \beta_0 + \beta_1 t + \beta_2 t^2.$$

The adjustment yields an empirical standard deviation of $\hat{s}_0 = 1.1045$. The estimated coefficients, their standard deviations, t-values, and significance decisions are listed in Table 2. The critical value for the two-sided t-test at $\alpha = 5\%$ and $\nu = 698$ degrees of freedom is $t_{0.975,698} \approx 1.963$.

Table 2: Estimated coefficients, standard deviations, t-values, and significance (degree $k = 2$).

Coefficient	$\hat{\beta}_j$	$\sigma_{\hat{\beta}_j}$	t_j	Significant?
β_0	$+1.99123$	0.06136	$+32.453$	Yes
β_1	-0.83236	0.02355	-35.340	Yes
β_2	-0.10166	0.01139	-8.925	Yes

All three coefficients are statistically significant ($|t_j| \gg 1.963$), and $\hat{s}_0$ decreases slightly compared to the linear model. Nevertheless, the residuals in Fig. 10 still exhibit clear systematic curvature, indicating that a second-degree polynomial is not yet sufficient to fully remove the trend.

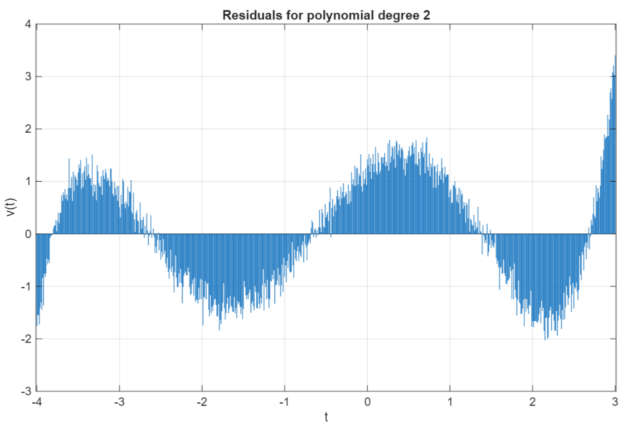

Figure 10: Residuals for polynomial degree $k = 2$.

- **Degree** $k = 3$:

$$x(t) = \beta_0 + \beta_1 t + \beta_2 t^2 + \beta_3 t^3.$$

The adjustment yields an empirical standard deviation of $\hat{s}_0 = 1.0702$. The estimated coefficients, their uncertainties, t-values, and significance decisions are given in Table 3. The critical value for the two-sided t-test at $\alpha = 5\%$ and $\nu = 697$ degrees of freedom is $t_{0.975,697} \approx 1.963$.

Table 3: Estimated coefficients, standard deviations, t-values, and significance (degree $k = 3$).

Coefficient	$\hat{\beta}_j$	$\sigma_{\hat{\beta}_j}$	t_j	Significant?
β_0	$+2.14192$	0.06343	$+33.769$	Yes
β_1	-0.55214	0.04702	-11.742	Yes
β_2	-0.16514	0.01444	-11.435	Yes
β_3	-0.04232	0.00621	-6.816	Yes

All four coefficients are clearly significant ($|t_j| \gg 1.963$), and $\hat{s}_0$ decreases compared to the quadratic model. However, the residuals in Fig. 11 still exhibit a visible oscillatory structure, indicating that additional higher-order terms may be required.

All coefficients are clearly significant. The residuals (Fig. 11) are reduced, but still show a visible oscillatory structure.

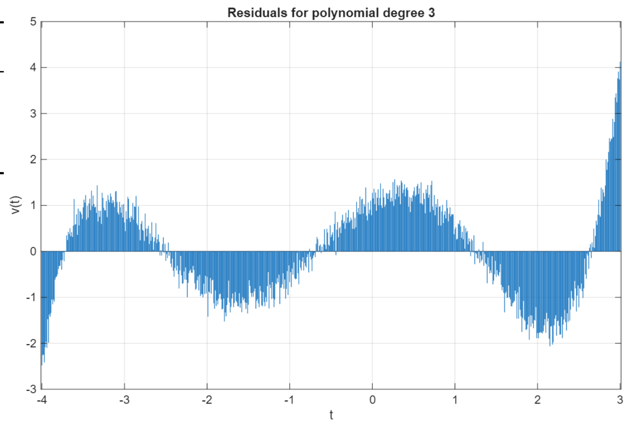

Figure 11: Residuals for polynomial degree $k = 3$.

- **Degree** $k = 4$:

$$x(t) = \beta_0 + \beta_1 t + \beta_2 t^2 + \beta_3 t^3 + \beta_4 t^4.$$

The adjustment yields an empirical standard deviation of $\hat{s}_0 = 1.0315$. The estimated coefficients, their uncertainties, t-values, and significance decisions are listed in Table 4. The two-sided t-critical value at $\alpha = 5\%$ with $\nu = 696$ degrees of freedom is $t_{0.975,696} \approx 1.963$.

The coefficient β_3 is statistically not significant, while all other coefficients remain highly significant. The empirical standard deviation decreases again, and the residuals in Fig. 12 show further

Table 4: Estimated coefficients, standard deviations, t-values, and significance (degree $k = 4$).

Coefficient	$\hat{\beta}_j$	$\sigma_{\hat{\beta}_j}$	t_j	Significant?
β_0	+2.40087	0.07050	+34.053	Yes
β_1	−0.80269	0.05665	−14.170	Yes
β_2	−0.39072	0.03361	−11.624	Yes
β_3	+0.00764	0.00904	+0.845	No
β_4	+0.02498	0.00339	+7.373	Yes

improvement, although weak systematic effects are still visible.

Here β_3 is statistically *not* significant ($t_3 \approx 0.845$), whereas all other coefficients remain highly significant. The residuals (Fig. 12) improve compared to lower degrees, but still show a weak remaining structure.

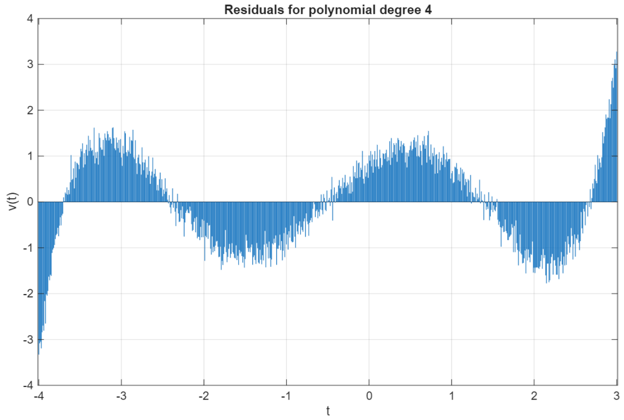

Figure 12: Residuals for polynomial degree $k = 4$.

- **Degree** $k = 5$:

$$x(t) = \beta_0 + \beta_1 t + \beta_2 t^2 + \beta_3 t^3 + \beta_4 t^4 + \beta_5 t^5.$$

This model yields a strong reduction of the empirical standard deviation to $\hat{s}_0 = 0.1974$. The estimated coefficients, their uncertainties, t-values, and significance decisions are summarized in Table 5. For $\alpha = 5\%$ and $\nu = 695$ degrees of freedom, the two-sided critical value is $t_{0.975,695} \approx 1.963$.

Table 5: Estimated coefficients, standard deviations, t-values, and significance (degree $k = 5$).

Coefficient	$\hat{\beta}_j$	$\sigma_{\hat{\beta}_j}$	t_j	Significant?
β_0	+3.21241	0.01477	+217.573	Yes
β_1	+0.49286	0.01446	+34.078	Yes
β_2	−1.34838	0.00956	−140.993	Yes
β_3	−0.54778	0.00445	−122.978	Yes
β_4	+0.14952	0.00113	+132.810	Yes
β_5	+0.04981	0.00037	+135.317	Yes

All six coefficients are extremely significant (all $|t_j| \gg 1.963$), and the residuals in Fig. 13 behave like random noise with no visible systematic structure. This indicates that a fifth-degree polynomial is capable of absorbing nearly all deterministic trend components in the series.

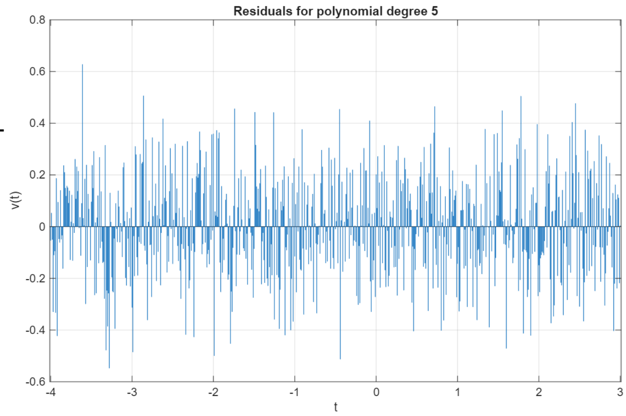

Figure 13: Residuals for polynomial degree $k = 5$.

- **Degree** $k = 6$:

$$x(t) = \beta_0 + \beta_1 t + \beta_2 t^2 + \beta_3 t^3 + \beta_4 t^4 + \beta_5 t^5 + \beta_6 t^6.$$

The adjustment yields an empirical standard deviation of $\hat{s}_0 = 0.1975$. The estimated coefficients, their uncertainties, t-values, and significance results are listed in Table 6. For $\alpha = 5\%$ and $\nu = 694$ degrees of freedom, the critical value is $t_{0.975,694} \approx 1.963$.

Table 6: Estimated coefficients, standard deviations, t-values, and significance (degree $k = 6$).

Coefficient	$\hat{\beta}_j$	$\sigma_{\hat{\beta}_j}$	t_j	Significant?
β_0	+3.21228	0.01561	+205.809	Yes
β_1	+0.49320	0.01923	+25.653	Yes
β_2	−1.34813	0.01335	−101.009	Yes
β_3	−0.54796	0.00788	−69.560	Yes
β_4	+0.14944	0.00295	+50.701	Yes
β_5	+0.04983	0.00073	+68.404	Yes
β_6	+0.000006	0.000209	+0.027	No

The new coefficient β_6 is *not* statistically significant, and the reduction in $\hat{s}_0$ from degree 5 to degree 6 is negligible. Furthermore, the residual plot in Fig. 14 is virtually identical to that of degree 5.

According to the t-test and the assignment instructions, the model selection stops at degree $k = 5$.

Since the highest-degree coefficient for $k = 6$ is not significant and the residuals do not improve visually compared to $k = 5$, we follow the criterion of the assignment and stop the procedure at degree $k = 5$.

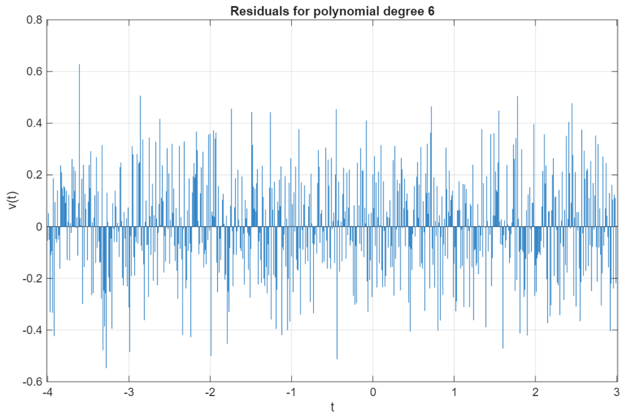

Figure 14: Residuals for polynomial degree $k = 6$.

4.4. Interpretation

The analysis shows that the time series in `Exercise3-3.txt` cannot be adequately described by a simple linear trend. Increasing the polynomial degree from $k = 1$ to $k = 4$ gradually reduces the variance factor and improves the fit, but the residuals still exhibit visible systematic structure. A degree-5 polynomial is required to remove the remaining systematic behaviour.

The final trend model adopted is the degree-5 polynomial

$$x(t) \approx 3.2124 + 0.4929\,t - 1.3484\,t^2 - 0.5478\,t^3 + 0.1495\,t^4 + 0.04981\,t^5, \tag{20}$$

for which all coefficients are highly significant at the 5 % level and the residuals behave like random noise.

Figure 15 illustrates the final polynomial trend together with the raw observations, and Fig. 16 shows the detrended series obtained by subtracting this degree-5 trend from the data.

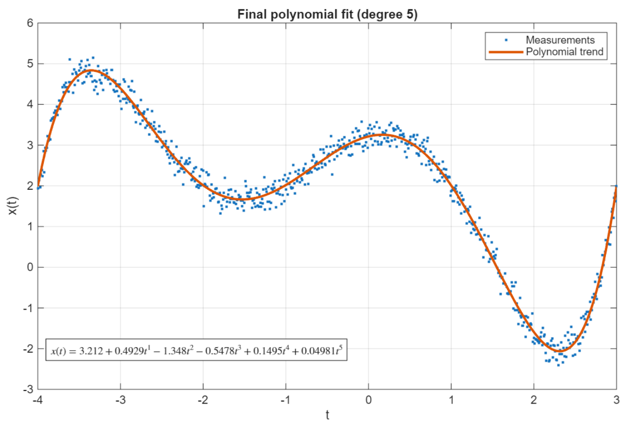

Figure 15: Final polynomial fit of degree $k = 5$ to the time series.

Interpretation

The gradual increase of the polynomial degree provides a clear picture of how the deterministic structure of the time series is absorbed by the trend function. For $k = 1, 2, 3$, and 4, the highest-order coefficients are all statistically significant, yet the corresponding residuals retain recognisable systematic behaviour. This indicates that these lower-order models do not fully capture the curvature present in the data.

At degree $k = 5$, the situation changes noticeably. All coefficients are highly significant, the empirical standard deviation decreases sharply, and the residuals fluctuate around zero in a manner consistent with an aperiodic, non-systematic signal. When the degree is increased to $k = 6$, the new coefficient β_6 is no longer significant, and neither $\hat{s}_0$ nor the residual structure shows any relevant improvement. According to the criteria of the assignment, the polynomial of degree five is therefore the appropriate choice for modelling the deterministic component of the series.

Removing this fifth-degree trend yields the detrended signal shown in Fig. 16. The resulting series oscillates around zero with no remaining long-term drift or curvature, indicating that the essential trend has been successfully separated from the stochastic fluctuations. This provides a suitable basis for subsequent analysis steps, where only the random component of the process is of interest.

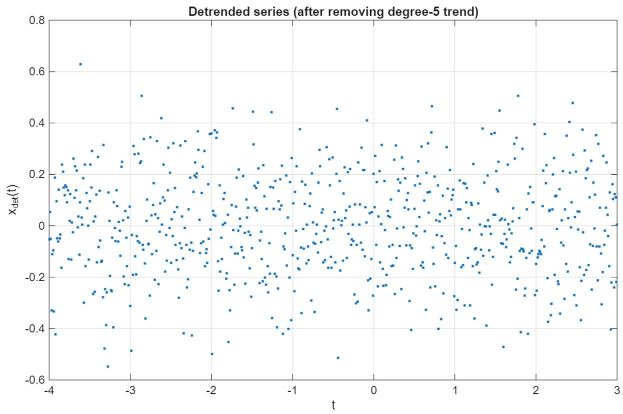

Figure 16: Detrended series after removing the degree-5 polynomial trend.

5. Task 3: Nonlinear trend models

The files `Exercise3-4.txt`, `Exercise3-5.txt` and `Exercise3-6.txt` contain three time series $x(t)$, where the first column is the time t and the second column the measured values $x(t)$. All measurements are treated as equally weighted and uncorrelated, and the time axis is regarded as error–free.

5.1. Visual inspection of the time series

Figure 17 shows the three raw series. Series 4 exhibits a rapidly increasing, convex curve; the growth rate appears to be proportional to the current value, which is typical for an exponential law. Series 5 and 6 both show a monotone growth from a low (initially negative for Series 5) value towards a horizontal asymptote, characteristic of logistic-type saturation behaviour.

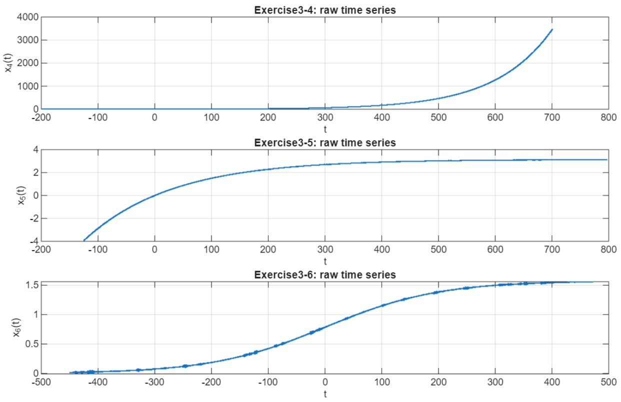

Figure 17: Raw time series for `Exercise3-4`, `Exercise3-5`, and `Exercise3-6`.

5.2. Choice of nonlinear trend models

Based on the qualitative shapes in Fig. 17, the following nonlinear trend functions are adopted.

Series 4: Exponential model

For `Exercise3-4` the trend is modelled as

$$x_4(t) = a_4\, \mathrm{e}^{b_4 t}, \tag{21}$$

which becomes linear in the parameters after the logarithmic transform $\ln x_4(t) = \ln a_4 + b_4 t$. A linear least–squares adjustment of $(t, \ln x_4)$ yields

$$\hat{a}_4 = 3.1416, \qquad \hat{b}_4 = 0.0100, \tag{22}$$

so that the estimated exponential trend is

$$x_4(t) \approx 3.1416\, \mathrm{e}^{0.01t}. \tag{23}$$

The corresponding fit is shown in Fig. 18.

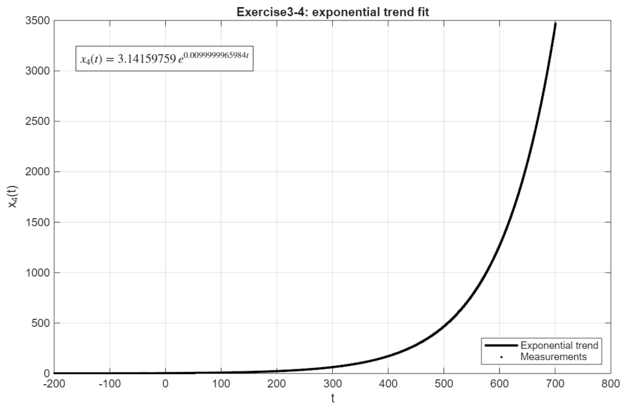

Figure 18: Exponential trend fit for `Exercise3-4`.

Series 5: Logistic trend with offset

The second series is described by a shifted logistic curve

$$x_5(t) = c_5 + \frac{L_5}{1 + \exp\bigl(-k_5(t - t_{0,5})\bigr)}, \tag{24}$$

where c_5 represents a vertical offset, L_5 the amplitude between lower and upper asymptote, k_5 the growth rate and $t_{0,5}$ the inflection time. This model is nonlinear in the parameters and is therefore adjusted iteratively using `lsqcurvefit`. The estimated parameters are

$$\begin{aligned} \hat{c}_5 &= -4284.41, & \hat{L}_5 &= 4287.55, \\ \hat{k}_5 &= 0.006505, & \hat{t}_{0,5} &= -1109.52, \end{aligned} \tag{25}$$

so that

$$x_5(t) \approx -4284.41 + \frac{4287.55}{1 + \exp\bigl(-0.006505\,(t + 1109.52)\bigr)}. \tag{26}$$

The fitted logistic curve closely follows the observed saturation behaviour (Fig. 19).

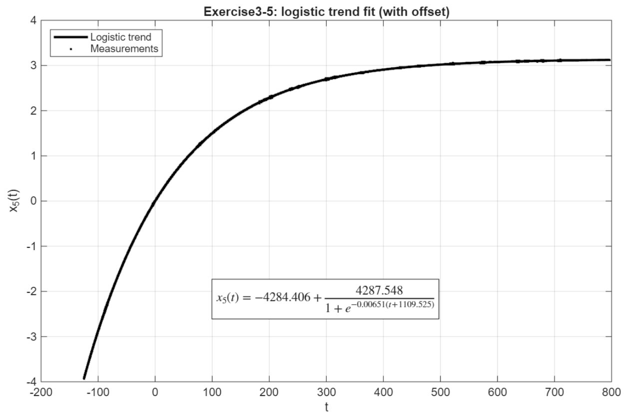

Figure 19: Logistic trend with offset for `Exercise3-5`.

Series 6: Logistic trend without offset

The third series exhibits a similar sigmoidal shape, but with a lower asymptote close to zero. Hence a standard logistic model without offset is sufficient:

$$x_6(t) = \frac{L_6}{1 + \exp\bigl(-k_6(t - t_{0,6})\bigr)}. \tag{27}$$

Nonlinear least–squares adjustment gives

$$\hat{L}_6 = 1.57082, \qquad \hat{k}_6 = 0.0099992, \qquad \hat{t}_{0,6} \approx 0, \tag{28}$$

so that

$$x_6(t) \approx \frac{1.57082}{1 + \exp\bigl(-0.0099992\, t\bigr)}. \tag{29}$$

The resulting fit is displayed in Fig. 20.

5.3. Detrending

For each series, the fitted nonlinear trend $\hat{x}_i(t)$ is subtracted from the observations,

$$x_{i,\mathrm{det}}(t) = x_i(t) - \hat{x}_i(t), \tag{30}$$

to obtain a detrended signal with mean approximately equal to zero.

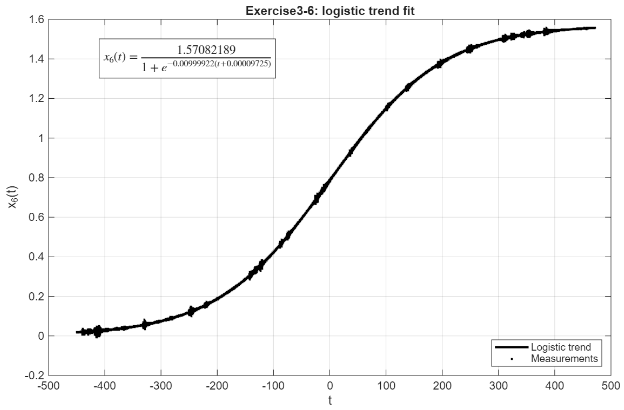

Figure 20: Logistic trend without offset for `Exercise3-6`.

Figure 21 shows the detrended series for `Exercise3-4`. The amplitudes of $x_{4,\mathrm{det}}(t)$ are of order 10^{-3}, i.e. several orders of magnitude smaller than the original exponential trend, which confirms that the chosen model captures the dominant growth behaviour very well.

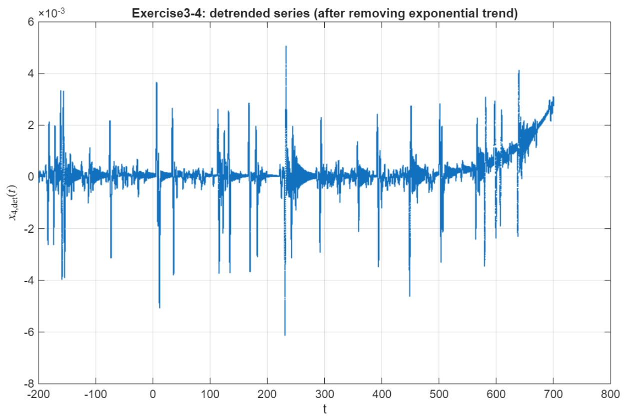

Figure 21: Detrended series for `Exercise3-4` after removing the exponential trend.

The remaining fluctuations are not perfectly homogeneous in time: occasional spikes and slight changes in variance are visible, in particular towards the end of the record. This suggests that, beyond the smooth exponential trend, the process contains local disturbances or time–dependent noise levels. Nevertheless, no systematic bias or long-term curvature remains, so the exponential model provides an adequate nonlinear trend function for this series, and the detrended signal is suitable for further stochastic analysis (e.g. correlation or spectral methods).

6. Task 4: Comparison of detrended time series

The detrended signals obtained from `Exercise3-5` and `Exercise3-6` are now analysed and compared. Although these two series were fitted using different nonlinear trend functions (a logistic curve with offset for Series 5 and a standard logistic curve for Series 6), the underlying raw data in both files are identical. Hence, once the nonlinear trends are removed, the detrended signals should coincide.

6.1. Detrended series

Figures 22 and 23 show the detrended time series for `Exercise3-5` and `Exercise3-6`, respectively. Both signals exhibit a similar structure: high-frequency fluctuations with intermittent bursts of increased amplitude, occurring at the same temporal locations. The overall amplitude is small (on the order of 10^{-2}), which indicates that the nonlinear trend models successfully capture the dominant behaviour of the original series.

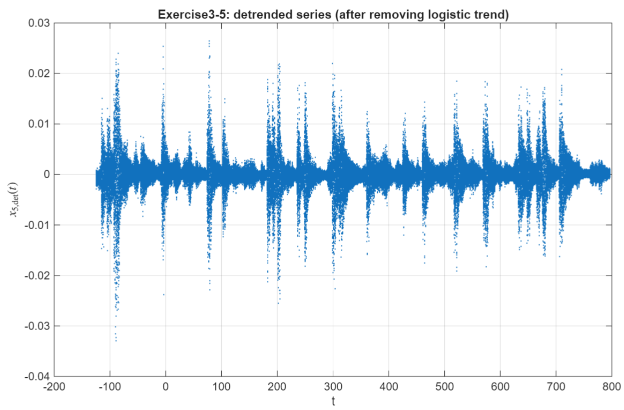

Figure 22: Detrended series for `Exercise3-5` after removing the logistic trend with offset.

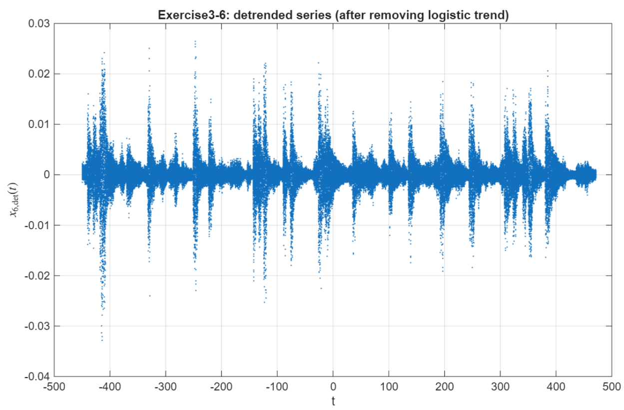

Figure 23: Detrended series for `Exercise3-6` after removing the standard logistic trend.

6.2. Graphical comparison

To assess the agreement between the two detrended signals, Fig. 24 displays $x_{5,\mathrm{det}}(t)$ and $x_{6,\mathrm{det}}(t)$ together. The two curves overlap almost perfectly. Their fluctuations share the same amplitude, frequency content and burst structure. No persistent offset or systematic deviation is visible.

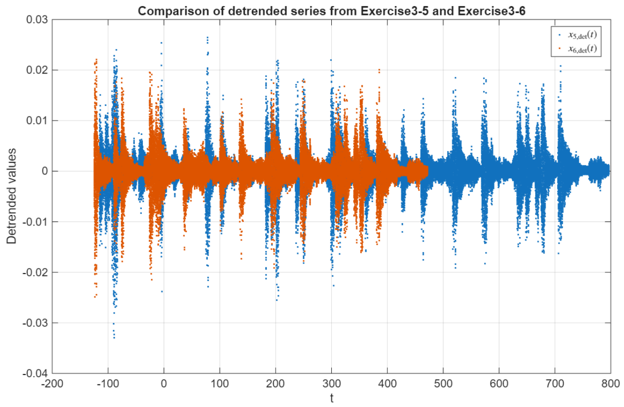

Figure 24: Comparison of the detrended series from `Exercise3-5` and `Exercise3-6`. Both signals follow the same pattern.

6.3. Difference between detrended signals

A quantitative evaluation is obtained from the pointwise difference

$$d(t) = x_{5,\mathrm{det}}(t) - x_{6,\mathrm{det}}(t). \tag{31}$$

As shown in Fig. 25, the difference fluctuates around zero without any visible trend. The magnitude of $d(t)$ is comparable to the noise level of the individual detrended series and does not exhibit any systematic temporal structure.

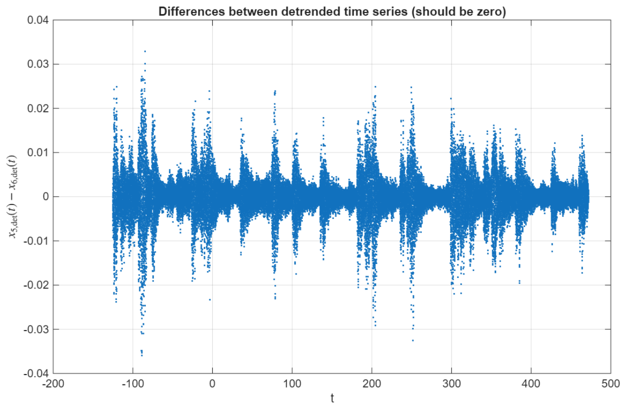

Figure 25: Differences $d(t)$ between the detrended series. The values oscillate around zero, consistent with the expectation that both files contain the same underlying measurements.

6.4. Interpretation

The detrended signals from `Exercise3-5` and `Exercise3-6` are effectively identical. Three observations support this conclusion:

- The detrended series (Figs. 22–23) display the same fluctuation pattern and noise characteristics.
- Their direct superposition (Fig. 24) shows no visual discrepancy.
- The pointwise differences (Fig. 25) form a zero-mean, structureless noise sequence.

This behaviour confirms the exercise background: both datasets are based on the same underlying time series, and the two nonlinear trend models remove only the smooth trend component without introducing artefacts. The remaining signal can therefore be regarded as the true stochastic component of the process, suitable for further analysis.

7. Conclusions

The assignment illustrates how deterministic trends of different types can be identified, modelled and removed from time series using regression techniques within the Gauss–Markov framework.

In Task 1, simple linear and quadratic models were sufficient to describe the dominant behaviour of two short data sets. After detrending, both series revealed almost identical harmonic residuals, demonstrating that apparently different curves can share the same stochastic component once their trends are properly removed.

Task 2 showed that the choice of polynomial degree should not rely solely on a visual impression of the fit. A stepwise increase of the degree, combined with t-tests on the regression coefficients and inspection of the residuals, indicated that lower-order models ($k \leq 4$) did not fully absorb the systematic structure. A fifth-degree polynomial provided a statistically consistent description: all coefficients were highly significant, the empirical variance factor decreased sharply, and the residuals behaved like trend–free noise. Adding a sixth-degree term did not yield any meaningful improvement, underscoring the importance of significance testing for model selection and for avoiding over-parameterisation.

Task 3 extended the analysis to genuinely nonlinear trends. An exponential model captured the strong convex growth in `Exercise3-4`, while logistic functions—with and without offset—described the sigmoidal saturation behaviour in `Exercise3-5` and `Exercise3-6`. In all cases the nonlinear regression left residuals whose amplitude was several orders of magnitude smaller than the original trend, confirming that the chosen functional forms were appropriate for the underlying physics represented by the synthetic data.

Finally, Task 4 compared the detrended series from `Exercise3-5` and `Exercise3-6`, which are based on the same original measurements but were fitted with slightly different logistic models. The detrended signals coincided both visually and in their pointwise differences, which fluctuated around zero without systematic patterns. This confirms that, provided the trend is modelled adequately, different but equivalent parametrisations of the same underlying law lead to the same residual process.

Overall, the exercises demonstrate that careful trend modelling—guided by visual inspection, statistical testing and residual analysis—is essential for separating deterministic and stochastic components in time series. Once an appropriate trend has been removed, the remaining signal is well suited for subsequent stochastic investigations such as correlation or spectral analysis, which then reflect the intrinsic dynamics of the process rather than artefacts of its deterministic evolution.