Climate Time Series Analysis:

Paleoclimatology

Manfred Mudelsee

Climate Risk Analysis, Germany, www.mudelsee.com
Alfred Wegener Institute, Germany
Prologue

Climate: knowledge and experience, space–time scales

Time series analysis: statistical concept

Computer: implementation of concept

Mathematical formulas: no sweat!
Chapters

1. Introduction
2. Persistence Models
3. Bootstrap Confidence Intervals
4. Regression I
5.
6.
7.
8.
9.
1. Introduction

1.1 Climate archives, variables and dating
1.2 Noise and statistical distribution
1.3 Persistence
1.4 Spacing
1.5 Aim and structure of this course
1. Introduction

Figure 1.3. Ice core data: deuterium and CO₂ records from the Vostok station

Climate time series
1. Introduction

Climate equation, continuous time

\[ X(T) = X_{\text{trend}}(T) + X_{\text{out}}(T) + S(T) \cdot X_{\text{noise}}(T), \] (1.1)

Climate evolves in time, and a stochastic process (a time-dependent random variable representing a climate variable with not exactly known value) and time series (the observed or sampled process) are central to statistical climate analysis. We shall use a wide definition of trend and decompose a stochastic process, \( X \), as follows:

\[ X(T) = X_{\text{trend}}(T) + X_{\text{out}}(T) + S(T) \cdot X_{\text{noise}}(T), \]

where \( T \) is continuous time, \( X_{\text{trend}}(T) \) is the trend process, \( X_{\text{out}}(T) \) is the outlier process, \( S(T) \) is a variability function scaling \( X_{\text{noise}}(T) \), the noise process. The trend is seen to include all systematic or deterministic, long-term processes such as a linear increase, a step change or a seasonal signal. The trend is described by parameters, for example, the rate of an increase. Outliers are events with an extremely large absolute value and are usually rare. The noise process is assumed to be weakly stationary with zero mean and autocorrelation. Giving \( X_{\text{noise}}(T) \) standard deviation unity enables introduction of \( S(T) \) to honour climate's definition as not only the mean but also the variability of the state of the atmosphere and other compartments (Brückner 1890; Hann 1901; Köppen 1923). A version of Eq. (1.1) is written for discrete time, \( T(i) \), as

\[ X(i) = X_{\text{trend}}(i) + X_{\text{out}}(i) + S(i) \cdot X_{\text{noise}}(i), \] (1.2)

using the abbreviation \( X(i) \equiv X(T(i)) \), etc. However, for unevenly spaced \( T(i) \) this is a problematic step because of a possibly non-unique relation between \( X_{\text{noise}}(T) \) and \( X_{\text{noise}}(i) \), see Section 2.1.2.1. The observed, discrete time series from process \( X(i) \) is the set of size \( n \) of paired values \( t(i) \) and \( x(i) \), compactly written as \( \{t(i), x(i)\}_{i=1}^{n} \).

To restate, the aim of this book is to provide methods for obtaining quantitative estimates of parameters of \( X_{\text{trend}}(T) \), \( X_{\text{out}}(T) \), \( S(T) \) and \( X_{\text{noise}}(T) \) using the observed time series data \( \{t(i), x(i)\}_{i=1}^{n} \).

A problem in climate analysis is that the observation process superimposes on the climatic process. \( X_{\text{noise}}(T) \) may show not only climatic
1. Introduction

Climate equation, discrete time

\[ X(i) = X_{\text{trend}}(i) + X_{\text{out}}(i) + S(i) \cdot X_{\text{noise}}(i), \]

(1.2)

Climate equation, discrete time

\( T(i) \), time points

\( X(i) \equiv X(T(i)), \) etc.
1. Introduction

Stochastic process \( X(i) \)  

“Process level”

Time series \( \{t(i), x(i)\}_{i=1}^{n} \)  

“Sample level”

Time series analysis uses sample to learn about process: trend parameters, probability of extremes, cycles, etc.

Univariate:  \( T(i), X(i) \)

Bivariate:  \( T(i), X(i), Y(i) \)
## 1. Introduction

### 1.1 Climate archives, variables and dating

<table>
<thead>
<tr>
<th>Climate archive</th>
<th>Time range</th>
<th>Geological epoch</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Holocene</td>
<td>Pleistocene</td>
</tr>
<tr>
<td></td>
<td>$10^0$</td>
<td>$10^1$</td>
</tr>
<tr>
<td>Marine sediment core</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ice core</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lake sediment core</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Speleothem</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tree-rings</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Documents</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Direct measurements</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Climate model</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Dating method**

- K/Ar
- Dosimeter
- U/Th
- $^{14}$C
- $^{210}$Pb
- Layer counting
1. Introduction

1.1 Climate archives, variables and dating

<table>
<thead>
<tr>
<th>Climate archive</th>
<th>Location</th>
<th>Time range (a)</th>
<th>Proxy variable</th>
<th>Resolution (a)</th>
<th>Climate variable</th>
</tr>
</thead>
<tbody>
<tr>
<td>Marine sediment core</td>
<td>Eastern equatorial Pacific</td>
<td>$10^6$</td>
<td>$\delta^{18}$O, benthic foraminifera</td>
<td>$10^3$</td>
<td>Ice volume, bottom water temperature</td>
</tr>
<tr>
<td>Ice core</td>
<td>Antarctica</td>
<td>$10^5$</td>
<td>CO$_2$, air bubbles, $\delta D$, ice</td>
<td>$10^3$</td>
<td>CO$_2$, atmosphere, Air temperature</td>
</tr>
<tr>
<td></td>
<td>Greenland</td>
<td>$10^5$</td>
<td>SO$_4$ content, ice, Ca content, ice, Dust content, ice, Conductivity, $a$ ice</td>
<td>$10^0$</td>
<td>Volcanic activity, Aeolian dust, wind, Soluble material, wind</td>
</tr>
<tr>
<td>Tree-rings</td>
<td>Worldwide</td>
<td>$10^4$</td>
<td>Na content, ice, $\Delta^{14}$C, wood</td>
<td>$10^0$</td>
<td>Seasalt, wind</td>
</tr>
<tr>
<td>Lake sediment core</td>
<td>Boston area</td>
<td>$10^3$</td>
<td>Varve thickness</td>
<td>$10^0$</td>
<td>Wind$^6$</td>
</tr>
<tr>
<td>Speleothem</td>
<td>Southern Oman</td>
<td>$10^4$</td>
<td>$\delta^{18}$O, carbonate</td>
<td>$10^1$</td>
<td>Monsoon rainfall</td>
</tr>
</tbody>
</table>

Book, Table 1.2
1. Introduction

1.1 Climate archives, variables and dating

Figure 1.3. Ice core data: deuterium and CO$_2$ records from the Vostok station (Antarctica) over the past 420,000 years. The core was drilled into the ice (diameter: 12 cm, length: 3623 m) and recovered in segments. The deuterium record (a) was measured on ice material using a mass spectrometer. Values are given in delta notation: $\delta D = [(D/H)_{\text{sample}}/(D/H)_{\text{SMOW}} - 1] \cdot 1000\%$, where (D/H) is the number of D particles over the number of H particles and SMOW is “Standard Mean Ocean Water” standard. Total number of points, $n$, is 3311. The CO$_2$ record (b) was measured on air bubbles enclosed in the ice. Values are given as “parts per million by volume,” $n$ is 283. In b, values (dots) are connected by lines; in a, only lines are shown. The present-day CO$_2$ concentration ($\sim 389$ ppmv) is not recorded in b. The construction of the timescale (named GT4) was achieved using a model of the ice accumulation and flow. Besides glaciological constraints, it further assumed that the points at 110 and 390 ka correspond to dated stages in the marine isotope record. Construction of the CO$_2$ timescale required additional modelling because in the ice core, air bubbles are younger in age than ice at the same depth. One climatological question associated with the data is whether variations in CO$_2$ (the values in air bubbles presenting the atmospheric value accurately) lead over or lag behind those of deuterium (which indicate temperature variations, low $\delta D$ meaning low temperature). (Data from Petit et al. 1999.)

The geologic past increases the proxy error. Spatially extending the range for which a variable is thought to be representative is a further source of error. This is the case, for example, when variations in air temperature in the inversion height above Antarctic station Vostok are used to represent...
1. Introduction

1.2 Noise and statistical distribution

The noise, \( X_{\text{noise}}(T) \), has been written in Eq. (1.1) as a zero-mean and unit-standard deviation process, leaving freedom as regards its other second and higher-order statistical moments, which define its distributional shape and also its spectral and persistence properties (next section). The probability density function (PDF), \( f(x) \), defines

\[
\text{prob} \left( a \leq X_{\text{noise}}(T) \leq a + \delta \right) = \int_{a}^{a+\delta} f(x) \, dx, \quad (1.3)
\]

Probability density function (PDF), \( f(x) \)

Gaussian or normal PDF
1. Introduction

1.2 Noise and statistical distribution

Histogram: inference of PDF

Residuals, “detrended version” of the $x(i)$

Book, Fig. 1.11 (b, Vostok CO$_2$; c, Vostok δD)
1. Introduction

1.3 Persistence

\[ E [X_{\text{noise}}(T_1) \cdot X_{\text{noise}}(T_2)] \quad \text{for} \quad T_1 \neq T_2 \]

Autocovariance

\[ E_X = \text{expectation operator (averaging)} \]

\[ E[X_{\text{noise}}(T)] = 0 \quad \text{(zero mean)} \]

Positive autocovariance: persistence or "memory"
1. Introduction

1.3 Persistence

Lag-1 scatterplot $x(i-1)$ vs $x(i)$: inference about persistence

Residuals, “detrended version” of the $x(i)$

Book, Fig. 1.12 (b, Vostok CO$_2$; c, Vostok δD)
1. Introduction

1.3 Persistence

\[ X(i) = X_{\text{trend}}(i) + X_{\text{out}}(i) + S(i) \cdot X_{\text{noise}}(i), \]  

\[ \{t(i), x(i)\}_{i=1}^n \]

Residuals:

use sample to determine trend, outliers and variability,
then
subtract trend, remove outliers and divide by variability
(details: Chapters 3 and 4)
1. Introduction

1.4 Spacing

\[ d(i) = t(i) - t(i - 1) \]

Time spacing

\[ \bar{d} \]

Average time spacing
1. Introduction

1.4 Spacing

Figure 1.15. Spacing of selected climatic time series.

- **a** ODP846 δ18O;
- **b** Vostok CO2;
- **c** Vostok δD;
- **d** NGRIP SO4;
- **e** Q5 δ18O. Data are given in Figs. 1.2, 1.3, 1.4 and 1.7.

In **d**, \( d(i) \) is shown for the \( D(i) = 0.5 \) cm data; the time series with \( t(i) = 1a \) (Fig. 1.4) is obtained from the 0.5-cm data using “downsampling.” The ice core data (b–d) reflect to some degree the effects of ice compaction, that means, \( d(i) \) increases with \( t(i) \). The Q5 speleothem spacing time series (e) suggests visually a strong negative correlation with the speleothem δ18O series (Fig. 1.7). This is explained as follows. Low (high) δ18O means strong (weak) Indian monsoonal rainfall, this in turn faster (slower) movement of the rainwater through the soil, weaker (stronger) uptake of soil-CO2, lower (higher) pH of the water, reduced (enhanced) solution of soil-carbonate, less (more) material for calcite precipitation, small (large) annual stalagmite layers and, finally, a higher (lower) temporal spacing because the depth spacing is nearly constant (Fig. 1.7). Note that at places with other soil properties, the relation δ18O–spacing may be different (Burns et al. 2002).

The values of the average spacing, \( \bar{d} \), and the coefficient of variation of spacing, \( CV_d \), which is defined as the standard deviation of the spacing divided by \( \bar{d} \), are as follows.

- **a** \( \bar{d} = 2.40a, CV_d = 0.41; \)
- **b** \( \bar{d} = 1.46a, CV_d = 0.82; \)
- **c** \( \bar{d} = 0.13a, CV_d = 0.85; \)
- **d** \( \bar{d} = 0.32a, CV_d = 0.47; \)
- **e** \( \bar{d} = 5.62a, CV_d = 0.49. \)

If \( n \approx n_{it} \) it is weaker, and only for \( n < n_{it} \) (“downsampling”) it may be absent. That means, interpolation does not allow to go in resolution below the limit set by Eq. (1.4). Second, depending on the type of in-
1.4 Spacing

\[ d(i) = d = \text{const.} \]

Even spacing

Computer model output, observations

Uneven spacing

- Missing values: Observations, documentary data
- Jittered: Observations, proxy data
- Irregular: Proxy data
1. Introduction

1.4 Spacing

Recent \( \tau \) Past

Climate Age, \( T \)

Direct observations, documents, climate model

Archive, sampling
Depth, \( z \)

Sediment core Low resolution

Archive, sampling
Estimated age, \( t \)

Sediment core High resolution

Archive, time series, \( t(i) \)

Ice core

Archive, time series, \( t'(i) \)

"Upsampling", \( t' \)

"Downsampling", \( t' \)

Diffusion

Strong introduced dependence

Weak

No
1. Introduction

1.5 Aim and structure of this course

\[ X(i) = X_{\text{trend}}(i) + X_{\text{out}}(i) + S(i) \cdot X_{\text{noise}}(i), \]  

(1.2)

Climate equation, stochastic process \( X(i) \)

Time series \( \{t(i), x(i)\}_{i=1}^{n} \)

Time series analysis uses sample to learn about process: trend parameters, probability of extremes, cycles, etc.

Univariate: \( T(i), X(i) \)

Bivariate: \( T(i), X(i), Y(i) \)
1. Introduction

1.5 Aim and structure of this course

Interpolation: a step away from the original data

Book, Fig. 1.3 (modified)
2. Persistence Models

2.1 First-order autoregressive model

2.2

2.3

2.4

2.5
2. Persistence Models

2.1 First-order autoregressive model

\[
X_{\text{noise}}(1) = \mathcal{E}_{N(0, 1)}(1),
\]
\[
X_{\text{noise}}(i) = a \cdot X_{\text{noise}}(i - 1) + \mathcal{E}_{N(0, 1 - a^2)}(i), \quad i = 2, \ldots, n. \tag{2.1}
\]

AR(1) process, even spacing

\(-1 < a < 1\)

Innovation term

\[
\mathcal{E}_{N(\mu, \sigma^2)}(\cdot)
\]

Mean = \(\mu = 0\)

Variance = \(\sigma^2 = 1\)

Gaussian shape

\[
f(x) = (2\pi)^{-1/2} \sigma^{-1} \exp\left[-(x-\mu)^2/(2\sigma^2)\right]
\]

\[
f(x) = (2\pi)^{-1/2} \exp\left(-x^2/2\right)
\]
2. Persistence Models

2.1 First-order autoregressive model

\[ X_{\text{noise}}(1) = \mathcal{E}_N(0, 1)(1), \]
\[ X_{\text{noise}}(i) = a \cdot X_{\text{noise}}(i - 1) + \mathcal{E}_N(0, 1-a^2)(i), \quad i = 2, \ldots, n. \] (2.1)

AR(1) process, even spacing

Mean: \( E[X_{\text{noise}}(i)] = 0 \) for all \( i \)

Variance: \( \text{VAR}[X_{\text{noise}}(i)] = 1 \) for all \( i \)

Strict stationarity
2. Persistence Models

2.1 First-order autoregressive model

$$X_{\text{noise}}(1) = \mathcal{E}_{N(0, 1)}(1),$$

$$X_{\text{noise}}(i) = \exp \left\{ - \frac{[T(i) - T(i - 1)]}{\tau} \right\} \cdot X_{\text{noise}}(i - 1)$$

$$+ \mathcal{E}_{N(0, 1 - \exp\{-2[T(i)-T(i-1)]/\tau\})}(i), \quad i = 2, \ldots, n. \quad (2.9)$$

AR(1) process, uneven spacing

Persistence time $\tau$

$\tau = 0$: no memory

$\tau > 0$: memory

Heteroscedastic innovation term

If $T(i) - T(i-1) = d = \text{const.}$:

even spacing with $a = \exp(-d/\tau)$
2. Persistence Models

2.1 First-order autoregressive model

\[ X_{\text{noise}}(1) = \mathcal{E}_{N(0, 1)}(1), \]
\[ X_{\text{noise}}(i) = \exp \left\{ - \frac{T(i) - T(i - 1)}{\tau} \right\} \cdot X_{\text{noise}}(i - 1) \]
\[ + \mathcal{E}_{N(0, 1 - \exp\{-2[T(i) - T(i-1)]/\tau\})}(i), \quad i = 2, \ldots, n. \quad (2.9) \]

AR(1) process, uneven spacing

Mean: \( E[X_{\text{noise}}(i)] = 0 \) for all \( i \)

Variance: \( \text{VAR}[X_{\text{noise}}(i)] = 1 \) for all \( i \)

Strict stationarity
2. Persistence Models

2.1 First-order autoregressive model

\[
E [X_{\text{noise}}(i + h) \cdot X_{\text{noise}}(i)] = \exp \left[ -|T(i + h) - T(i)|/\tau \right].
\]

Autocorrelation

Figure 2.2. Autocorrelation function of the AR(1) process, \(a > 0\). In the case of even
2. Persistence Models

2.1 First-order autoregressive model

\[ X_{\text{noise}}(1) = \mathcal{E}_{N(0, 1)}(1), \]
\[ X_{\text{noise}}(i) = \exp \left\{ - \frac{T(i) - T(i - 1)}{\tau} \right\} \cdot X_{\text{noise}}(i - 1) \]
\[ + \mathcal{E}_{N(0, 1 - \exp\{-2\frac{T(i) - T(i-1)}{\tau}\}}(i), \quad i = 2, \ldots, n. \]  

AR(1) process, uneven spacing

Least-squares estimation

Residuals \( x_{\text{noise}}(i) \), “detrended version” of the \( x(i) \)
3. Bootstrap Confidence Intervals

3.1 Error bars and confidence intervals
3.2 Bootstrap principle
3.3 Bootstrap resampling
3.4 Bootstrap confidence intervals
3.5 Examples
3.6
3. Bootstrap Confidence Intervals

3.1 Error bars and confidence intervals

\[ X(i) = X_{\text{trend}}(i) + X_{\text{out}}(i) + S(i) \cdot X_{\text{noise}}(i), \]  \hspace{1cm} (1.2)

Climate equation, stochastic process \( X(i) \)

Time series \( \{t(i), x(i)\}_{i=1}^{n} \)

Time series analysis uses sample to learn about process: trend parameters, probability of extremes, cycles, etc.
3. Bootstrap Confidence Intervals

3.1 Error bars and confidence intervals

Estimator

A recipe how to calculate parameter(s) of interest

Estimation

We apply recipe to data and get result.

Estimate

Numerical value of result

Hat notation

Denotes estimator/estimate.
3. Bootstrap Confidence Intervals

3.1 Error bars and confidence intervals

<table>
<thead>
<tr>
<th>True parameter</th>
<th>Sample</th>
<th>Estimator, estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \theta )</td>
<td>Time series</td>
<td>( \hat{\theta} )</td>
</tr>
<tr>
<td>(“Theta”)</td>
<td></td>
<td>(“Theta hat”)</td>
</tr>
</tbody>
</table>

\[
\theta \neq \theta
\]

\( n < \infty \),

Observation errors
(measurement, proxy, calibration, climate model, etc.)

\[
\hat{\theta} \neq \theta
\]
3. Bootstrap Confidence Intervals

3.1 Error bars and confidence intervals

Estimate does not equal true (but unknown) value, it deviates.

However, we can say how large this deviation typically is: error bars, confidence intervals, etc.

We need statistical language, we make a statistical inference.
3. Bootstrap Confidence Intervals

3.2 Bootstrap principle

\[
\{t(i), x(i)\}_{i=1}^{n} \rightarrow \hat{\theta} \rightarrow \{t^*(i), x^*(i)\}_{i=1}^{\theta^{*1}} \rightarrow \text{Confidence interval } \text{CI}_{\hat{\theta}, 1-2\alpha}
\]

\[
\{t^*(i), x^*(i)\}_{i=1}^{\theta^{*2}} \rightarrow \hat{\theta}^{*2}
\]

\[
\{t^*(i), x^*(i)\}_{i=1}^{\theta^{*B}} \rightarrow \hat{\theta}^{*B}
\]

\[
\{t(i), x(i)\}_{i=1}^{n} \rightarrow \hat{\theta}
\]

\[
\{t(i), x(i)\}_{i=1}^{n} \rightarrow \text{Sample, estimate }
\]

\[
\{t^*(i), x^*(i)\}_{i=1}^{\theta^{*1}} \rightarrow \text{Replications}
\]

\[
\{t^*(i), x^*(i)\}_{i=1}^{\theta^{*2}} \rightarrow \text{Replications}
\]

\[
\{t^*(i), x^*(i)\}_{i=1}^{\theta^{*B}} \rightarrow \text{Replications}
\]

35

Book, Fig. 3.3
3. Bootstrap Confidence Intervals

3.2 Bootstrap principle

Bootstrap

Resampling

Moving block bootstrap (MBB)
Autoregressive bootstrap (ARB)

Other

Confidence interval construction

Normal
Student’s $t$
Percentile
BCa
3. Bootstrap Confidence Intervals

3.3 Bootstrap resampling

Moving block bootstrap (MBB)

**Algorithm 3.1. Moving block bootstrap algorithm (MBB).**

1. **Step 1** Data
   \[ \{ t(i), x(i) \}_{i=1}^n \]

2. **Step 2** Resampled times unchanged
   \[ \{ t^* (i) \}_{i=1}^n \] = \[ \{ t(i) \}_{i=1}^n \]

3. **Step 3** Blocks
   \[ j \] (see above)
   \[ \{ x(i) \}_{j-1}^{j+l-1} \] \[ i = j, j = 1, \ldots, n-l+1 \]

4. **Step 4** Set counter \( c = 1 \)
   Start resampling

5. **Step 5** Draw random block \( j^* \) \[ j^* \in \{ 1, \ldots, n-l+1 \} \]

6. **Step 6** Insert block data
   \[ \{ x^* (i) \}_{c+l-1}^{c+l-1} \] \[ i = c, c = \{ x(i) \}_{j^*+l-1}^{j^*+l-1} \] \[ i = j^*, j^* = 1, \ldots, n-l+1 \]

   If \( x^* (n) \) has been inserted
   Stop inserting and exit

7. **Step 7** Increase counter \( c \rightarrow c + l \)

8. **Step 8** Go to Step 5
   End resampling

---

Means of Monte Carlo simulations of real-world conditions, as is done in subsequent parts of this book.

Bühlmann and Künsch (1999) presented a fully data-driven block length selector (Algorithm 3.2). They showed the equivalence of \( l_{opt} \) selection and smoothing in spectral estimation (Chapter 5).

Berkowitz and Kilian (2000) presented a brute-force block length selector:

1. Approximatethedatageneratingprocessbyaparametricmodel(e.g., ARMA).
2. Generate Monte Carlo samples from this fitted model.
3. Bootstrap Confidence Intervals

3.3 Bootstrap resampling

Step 1 Data
\[\{t(i), x(i)\}_{i=1}^{n}\]

Step 2 Resampled times
\[\{t^*(i)\}_{i=1}^{n} = \{t(i)\}_{i=1}^{n}\]
unchanged

Step 3 Residuals (Eq. 3.29)
\[r(i) = \left[x(i) - \hat{x}_{\text{trend}}(i) - \hat{x}_{\text{out}}(i)\right] / \hat{S}(i)\]

Step 4 Apply MBB
(Algorithm 3.1)
to residuals
\[\{r(i)\}_{i=1}^{n}\]

Step 5 Resampled residuals
\[\{r^*(i)\}_{i=1}^{n}\]

Step 6 Use resampled residuals
to produce resamples
\[x^*(i) = \hat{x}_{\text{trend}}(i) + \hat{x}_{\text{out}}(i) + \hat{S}(i) \cdot r^*(i)\]

Algorithm 3.3. MBB for realistic climate processes, which comprise trend, outlier and variability components.
Bootstrap confidence intervals

Estimation of \( \hat{\theta} \) is repeated for the resamples, \( \{t^*_{b(i)}, x^*_{b(i)}\}_{i=1}^n \), \( b = 1, \ldots, B \). This yields the bootstrap replications, \( \{\hat{\theta}^*_b\}_{b=1}^B \). The replications are used to construct equi-tailed \((1 - 2\alpha)\) confidence intervals, \( \text{CI} \hat{\theta}, 1 - 2\alpha \), see Fig. 3.3.

Two approaches, standard error based and percentile based, dominate theory and practice of bootstrap CI construction. The estimated bootstrap standard error is the sample standard error of the replications,

\[
\hat{\text{se}}_{\hat{\theta}^*} = \left\{ \sum_{b=1}^B \left[ \hat{\theta}^*_b - \left\langle \hat{\theta}^*_b \right\rangle \right]^2 \right\}^{1/2} (B - 1),
\]

(3.30)

where

\[
\left\langle \hat{\theta}^*_b \right\rangle = \frac{1}{B} \sum_{b=1}^B \hat{\theta}^*_b.
\]

Bootstrap standard error
3. Bootstrap Confidence Intervals

3.4 Bootstrap confidence intervals

Student’s $t$ confidence interval

$$\text{CI}_{\hat{\theta},1-2\alpha} = \left[ \hat{\theta} + t_{\nu}(\alpha) \cdot \hat{\text{se}}_{\hat{\theta}^*}; \hat{\theta} - t_{\nu}(\alpha) \cdot \hat{\text{se}}_{\hat{\theta}^*} \right], \quad (3.32)$$

where $t_{\nu}(\alpha)$ is the percentage point of the $t$ distribution function with $\nu$ degrees of freedom (Section 3.9). It is in practice presumably always more accurate to prefer, as this book does, Student’s $t$ CIs over normal CIs because they recognize the reduction of degrees of freedom. (For data sizes above, say, 30, the difference becomes negligible.)

- $t_{\nu}(\alpha)$: Percentage point of Student’s $t$-distribution with $\nu$ degrees of freedom ($\nu = n - \text{number of parameters}$)
- Takes into account that not only standard error, but also fit value estimated.
- Negligible difference to normal CI if $\nu$ above ~ 30
3. Bootstrap Confidence Intervals

3.5 Examples

Vostok, CO₂

Glacial, [140 ka; 177 ka], \( n = 13 \)
Interglacial, [115 ka; 130 ka], \( n = 24 \)

Bootstrap resampling

Block bootstrap (not MBB, but SB)

Block length (average) \( NINT \left( 4 \cdot \frac{\tau}{d} \right) \)

\( B = 2000 \)

Confidence interval

95% Student’s \( t \)
4. Regression I

4.1

4.2

4.3 Nonparametric regression or smoothing
4. Regression I

\[ X(i) = X_{\text{trend}}(i) + X_{\text{out}}(i) + S(i) \cdot X_{\text{noise}}(i), \]  \hspace{1cm} (1.2)

Climate equation, stochastic process \( X(i) \)

Time series \( \{t(i), x(i)\}_{i=1}^{n} \)

Time series analysis uses sample to learn about process: trend parameters, probability of extremes, cycles, etc.

Ignore \( X_{\text{out}}(i) \),
focus on \( X_{\text{trend}}(i) \) and also \( S(i) \).
4. Regression I

\[ X_{\text{trend}}(i) \]

Parametric trend model
  - Linear
  - Nonlinear

Nonparametric trend model
4. Regression I

4.3 Nonparametric regression or smoothing

\( X_{\text{trend}}(T) \) estimated via averaging points \( x(i) \) in neighbourhood around \( T \)

Short-term noise components smoothed away

No parametric trend model
4. Regression I

4.3 Nonparametric regression or smoothing

Running mean

\[
\hat{X}_{\text{trend}}^\text{PC}(T) = h^{-1} \sum_{i=1}^{n} [T(i) - T(i - 1)] K \left[ \frac{T - T(i)}{h} \right] X(i), \quad (4.52)
\]

Kernel smoothing

- \( K \), kernel function (\( K = 0 \) outside, \( K = \text{const.} > 0 \) inside)
- \( h \), bandwidth

---

Table 4.14. Monte Carlo experiment, break regression with timescale errors and AR(1) noise of normal shape: CI coverage performance.

<table>
<thead>
<tr>
<th>n sim</th>
<th>CI coverage performance</th>
</tr>
</thead>
<tbody>
<tr>
<td>475</td>
<td></td>
</tr>
</tbody>
</table>

\( n \) sim = 475 random samples were generated from \( X(i) = X_{\text{break}}(i) + X_{\text{noise}}(i) \), where the prescribed break model parameters (Eq. 4.47) are \( x_1 = 2.0, t_2 = 0.5, n, x_2 = 1.0 \) and \( x_3 = 4.0 \), the prescribed times are \( T_{\text{true}}(i) = i, i = 1, \ldots, n \) and the noise is a Gaussian AR(1) process (Eq. 2.1) with \( a = 1/e \approx 0.37 \). Timescale error simulations were performed as in the experiment on ramp regression (Table 4.13). Bootstrap BCa CIs used timescale-ARB resampling (Algorithm 4.5), \( B = 1999 \) and \( \alpha = 0.025 \).

The Monte Carlo experiment (Table 4.14) reveals that for data sizes above 100–200, the time parameter (\( t_2 \)), the level parameters (\( x_1, x_2, x_3 \)) and the slopes (\( \beta_1, \beta_2 \)) have excellent coverage performance also for heteroscedastic timescale errors.

---

Priestley and Chao (1972)
4. Regression I

4.3 Nonparametric regression or smoothing

\[ K(y) = (2\pi)^{-1/2} \exp(-y^2/2) \]

Gaussian kernel

\[ K(y) = 1 \text{ for } |y| \leq 1/2 \]

Uniform kernel

\[ K(y) = 0.75 \ (1 - y^2) \text{ for } |y| \leq 1 \]

Epanechnikov kernel
4. Regression I
4.3 Nonparametric regression or smoothing

Smoothing problem

selection of bandwidth

<table>
<thead>
<tr>
<th>Bandwidth</th>
<th>Standard error</th>
<th>Bias</th>
</tr>
</thead>
<tbody>
<tr>
<td>Large</td>
<td>Small (many points)</td>
<td>Large (loose details)</td>
</tr>
<tr>
<td>Small</td>
<td>Large</td>
<td>Small</td>
</tr>
</tbody>
</table>

(choice of \( K \) more of “cosmetic” interest)

Analogy: density estimation
4. Regression I

4.3 Nonparametric regression or smoothing—Stalagmite example

<table>
<thead>
<tr>
<th>Age</th>
<th>Depth</th>
<th>δ¹⁸O</th>
</tr>
</thead>
<tbody>
<tr>
<td>8162 ± 83 a</td>
<td>0 mm</td>
<td>-6.23 ± 0.08 %</td>
</tr>
<tr>
<td>6525 ± 55 a</td>
<td>100 mm</td>
<td>-5.25 ± 0.08 %</td>
</tr>
<tr>
<td>3278 ± 77 a</td>
<td></td>
<td></td>
</tr>
<tr>
<td>910 ± 61 a</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

© Climate Risk Analysis

49
4. Regression I

4.3 Nonparametric regression or smoothing—Stalagmite example

Fohlmeister et al. (in prep.) [ISCAM age model]
4. Regression I

4.3 Nonparametric regression or smoothing—Stalagmite example

Oxygen isotope proxy
winter temperature
winter precipitation

Fohlmeister et al. (in prep.) [ISCAM age model]
4. Regression I

4.3 Nonparametric regression or smoothing—Stalagmite example

Mudelsee et al. (in prep.) [ISCAM age model]
4. Regression I

4.3 Nonparametric regression or smoothing—Stalagmite example

Mudelsee et al. (in prep.) [ISCAM age model]
4. Regression I

4.3 Nonparametric regression or smoothing

Bootstrap error band (pointwise)
Take fit curve
Add blocks of residuals†
Resample: \{t(i), x^*(i)\}
Re-estimate: replication
Repeat many times
For each T: error bar (replications)
Band: connect error bars over T

Book, Algorithm 3.1, Fig. 3.3, Section 4.2.2.3 [† used here actually not MBB but ARB (Algorithm 3.5)]
4. Regression I

4.3 Nonparametric regression or smoothing—Stalagmite example

Mudelsee et al. (in prep.) [ISCAM age model]
4. Regression I
4.3 Nonparametric regression or smoothing

Bootstrap error band (pointwise)

Take fit curve
Add blocks of residuals^†
Resample: \{t^*(i), x^*(i)\}
Re-estimate: replication
Repeat many times
For each T: error bar (replications)
Band: connect error bars over T

Algorithm 3.1.
Moving block bootstrap algorithm (MBB). Note: An equation like
\{t^*(i)\}_{i=1}^n = \{t(i)\}_{i=1}^n is used to denote
t^*(i) = t(i), i = 1, ..., n.

Bühlmann and Künsch (1999) presented a fully data-driven block
length selector (Algorithm 3.2). They showed the equivalence of
selection and smoothing in spectral estimation (Chapter 5).

Berkowitz and Kilian (2000) presented a brute-force block length se-
lector:
1 Approximate the data generating process by a parametric model (e.g.,
ARMA).
2 Generate Monte Carlo samples from this fitted model.

Sample, estimate
\{t(i), x(i)\}_{i=1}^n

Confidence interval
\hat{\theta}^\pm
Replications
\{\hat{\theta}_1, \hat{\theta}_2, \hat{\theta}_n\}
Resamples
\{t^{*(i)}, x^{*(i)}\}_{i=1}^n

Book, Algorithm 3.1, Fig. 3.3, Section 4.2.2.3 [^† used here actually not MBB but ARB (Algorithm 3.5)]
4. Regression I

4.3 Nonparametric regression or smoothing—Stalagmite example

Mudelsee et al. (in prep.) [ISCAM age model]
4. Regression I

4.3 Nonparametric regression or smoothing—Stalagmite example

Mudelsee et al. (in prep.) [ISCAM age model]
4. Regression I

4.3 Nonparametric regression or smoothing—Stalagmite example

SMudelsee et al. (in prep.) [† STALAGE age model]
4. Regression I

4.3 Nonparametric regression or smoothing—Stalagmite example

Mudelsee et al. (in prep.) [ISCAM age model]
4. Regression I

4.3 Nonparametric regression or smoothing—Stalagmite example

Next steps

Discuss age models
ISCAM, STALAGE
agreement/deviations
accuracy
Discuss bandwidth selection
Physical arguments (e.g., solar cycles)
Monte Carlo experiments
Coverage performance of error bands
1. Introduction
2. Persistence Models
3. Bootstrap Confidence Intervals
4. Regression I
5. Spectral Analysis
6. Extreme Value Time Series
7. Correlation
8. Regression II
9. Future Directions

\[ X(T) = X_{\text{trend}}(T) + X_{\text{out}}(T) + S(T) \cdot X_{\text{noise}}(T), \quad (1.1) \]
Climate Time Series Analysis: Conclusions

(1) $\hat{\theta} \neq \theta$

and importance of computational techniques (bootstrap) to obtain realistic error bars.

(2) Paleo challenges: uneven spacing uncertain timescales
Climate Time Series Analysis: Conclusions

(1) \( \hat{\theta} \neq \theta \) and importance of computational techniques (bootstrap) to obtain realistic error bars.

(2) Paleo challenges: uneven spacing uncertain timescales

Thanks!
Postdoc job: www.climate-risk-analysis.com
4. Regression I

4.3 Nonparametric regression or smoothing

\[ \hat{X}_{\text{trend}}^{\text{GM}}(T) = h^{-1} \sum_{i=1}^{n} \left[ \int_{s(i-1)}^{s(i)} K \left( \frac{T - y}{h} \right) dy \right] X(i), \quad (4.53) \]

where \( T(i - 1) \leq s(i - 1) \leq T(i) \) (e.g., \( s(i - 1) = [T(i - 1) + T(i)]/2 \) with \( s(0) \) and \( s(n) \) being the upper and lower bounds of the \( T \) interval, respectively).

\[ \hat{X}_{\text{trend}}^{\text{PC}}(T) = h^{-1} \sum_{i=1}^{n} [T(i) - T(i - 1)] K \left[ \frac{T - T(i)}{h} \right] X(i), \quad (4.52) \]

Kernel smoothing

- \( K \), kernel function (\( K = 0 \) outside, \( K = \text{const.} > 0 \) inside)
- \( h \), bandwidth
Bootstrap resampling

Autoregressive bootstrap (ARB)

- Take noise residuals $r(i)$
- Fit AR(1) model
- Take white-noise residuals
- Draw randomly from $e(i)$ with replacement (ordinary bootstrap)
- Plug in resampled $e(i)$
- Plug in resampled $r(i)$

$$
e(i) = [r(i) - \hat{a}' \cdot r(i - 1)]$$