1 Introduction

It has been proclaimed that 'data are the resource of the 21st century' [1]. Looking back in history, most people can agree on that oil was the most important natural resource in the 20th century. However, if we take a closer look at oil, we see that it is not very useful on its own. Surely, oil is flammable, but things like cars, planes, or chemical industry plants are needed to make oil the useful resource we know and therefore, to make it an important resource. Hence, tools and equipments are needed to create a benefit using the natural resource. The same situation applies to data. Without adequate tools data is nothing more than ones and zeroes. However, with the right methods data can be analyzed to extract useful information to gain new insights. Like diesel needs a different consumption engine than gasoline some type of data requires different methods than other types. This work focuses on time series data. In almost all everyday life situations time series data can occur. For instance, sensors monitoring the weather and collecting the air and ground temperature or the amount of rainfall, sensors monitoring vital parameters like the heart rate of a person, or smartphones counting a person's steps, all results in time series data. Time series data is also given by financial markets, e.g. stock market prices, or by social-economical data, e.g. unemployment rates or gross national products.

An important characteristic of time series data is that different data points are usually not independent from each other. That is why such a type of data requires special treatment. An example for this situation is the bootstrap method, see section 1.2 for an introduction. If this method is applied in the same manner as it is applied for data which consists of independent data points, the bootstrap method would give in general not the answer it should. Or one may even say, it could give a wrong answer. This problem is tackled in chapter 2.

An important question in the time series setting is the question of forecast; given data up to today what can be predicted for the future. For a specific class of time series chapter 3 gives insights how a good prediction can be achieved.

1.1 Time Series Fundamentals

A stochastic process is a family of random variables $\{X_t, t \in T\}$, where T is some index set, defined on a probability space $(\Omega, \mathcal{F}, P)$, c.f. (Brockwell and Davis, 1991, Chapter 1). In time series analysis

[1] Angela Merkel (Chancellor of Germany), Hanover, 2016: http://www.cebit.de/de/news-trends/news/bundeskanzlerin-merkel-daten-sind-die-rohstoffe-des-21-jahrhunderts-1190

the index set is usually given by $\mathbb{Z}$, hence $\{X_t, t \in \mathbb{Z}\}$. We distinguish here between univariate time series, where the stochastic process is $\mathbb{C}$-valued, and multivariate time series, where a vector-valued process is considered, $\mathbb{C}^d, d \geq 1$. chapter 2 deals mainly with univariate time series, whereas multivariate time series are in focus of chapter 3. Throughout this work we concentrate on stationary time series. A time series $X = \{X_t, t \in \mathbb{Z}\}$ is said to be stationary if $E|X_t^2| < \infty, EX_t = \mu$ for all $t \in \mathbb{Z}$, and for all $t, h \in \mathbb{Z}$ we have

$$\mathrm{Cov}(X_{t+h}, X_t) = E[(X_{t+h} - EX_{t+h})(X_t - E(X_t))^\top] = E[(X_h - EX_h)(X_0 - E(X_0))^\top] =: \gamma(h). \tag{1.1.1}$$

The function $\gamma(h), h \in \mathbb{Z}$, defined by (1.1.1) is called the autocovariance function of time series X. Furthermore, a time series $\{X_t, t \in \mathbb{Z}\}$ is said to be strictly stationary if $P^{X_{t_1},\ldots,X_{t_k}} = P^{X_{t_1+h},\ldots,X_{t_k+h}}$ for all $t_1, \ldots, t_k \in \mathbb{Z}$ and $h \in \mathbb{Z}$, where $P^{X_{t_1},\ldots,X_{t_k}}$ denotes the joint distribution of $X_{t_1}, \ldots, X_{t_k}$.

A simple example for a stationary time series is white noise. It is given by an uncorrelated time series $\{\varepsilon_t, t \in \mathbb{Z}\}$ with $E\varepsilon_t = 0$ and $\mathrm{Var}\varepsilon_t = \Sigma_\varepsilon^2 < \infty$ for all $t \in \mathbb{Z}$. Furthermore, we introduce here two important time series models: moving average (MA) models and autoregressive (AR) models. Based on some white noise $\{\varepsilon_t, t \in \mathbb{Z}\}$ a moving average process $\{X_t, t \in \mathbb{Z}\}$ of order q is defined by

$$X_t = \sum_{j=0}^{q} B_j \varepsilon_{t-j}, t \in \mathbb{Z}, \tag{1.1.2}$$

where $B_0, \ldots, B_q \in \mathbb{C}^{d\times d}$, B_0 is usually normalized to the identity matrix and $B_q \neq 0$. An autoregressive process $\{X_t, t \in \mathbb{Z}\}$ of order p is defined by

$$X_t = \sum_{j=1}^{p} A_j X_{t-j} + \varepsilon_t, t \in \mathbb{Z}, \tag{1.1.3}$$

where $A_1, \ldots, A_p \in \mathbb{C}^{d\times d}$ and $A_p \neq 0$. Both models are special cases of autoregressive moving average (ARMA) models of order (p, q) which are given by

$$X_t - \sum_{j=1}^{p} A_j X_{t-j} = \sum_{j=1}^{q} B_j \varepsilon_{t-j} + \varepsilon_t, t \in \mathbb{Z}, \tag{1.1.4}$$

where $A_1, \ldots, A_p, B_1, \ldots, B_q \in \mathbb{C}^{d\times d}$ and $A_p, B_q \neq 0$.

A stationary time series and its properties can be expressed either in the time domain or in the frequency domain. The frequency domain is only used in chapter 2, hence in the univariate case. In order to simplify notation it is described here for the univariate case. However, the results given here can be transfered to the multivariate case. The autocovariance function describes the second-

order properties of a time series in time domain. Herglotz's Theorem, c.f. (Brockwell and Davis, 1991, Theorem 4.3.1) gives the corresponding representation in the frequency domain: A function $\gamma : \mathbb{Z} \to \mathbb{C}$ is non-negative definite (hence, an autocovariance function) if and only if

$$\gamma(h) = \int_{(-\pi,\pi]} \exp(ih\nu) dF(\nu), \text{ for all } h \in \mathbb{Z},$$

where $F(\cdot)$ is a right-continuous, non-decreasing, bounded function on $[-\pi, \pi]$ and $F(-\pi) = 0$. The function F is called the spectral distribution function of γ and if $F(\lambda) = \int_{-\pi}^{\lambda} f(\nu) d\nu, -\pi \leq \lambda \leq \pi$, then f is called a spectral density of γ. Furthermore, we have, c.f. (Brockwell and Davis, 1991, Theorem 4.3.2), that if $\sum_{h\in\mathbb{Z}} |\gamma(h)| < \infty$ then

$$\gamma(h) = \int_{-\pi}^{\pi} \exp(ih\nu) f(\nu) d\nu, h \in \mathbb{Z},$$

where

$$f(\lambda) = \frac{1}{2\pi} \sum_{h\in\mathbb{Z}} \exp(-ih\lambda)\gamma(h).$$

Throughout this work, we denote by second-order properties of a time series the properties defined by the entire autocovariance function in time domain or by the spectral density in frequency domain, respectively.

For a stationary ARMA model given by (1.1.4), the spectral density can be directly derived by using the corresponding AR and MA polynomials. The following theorem, c.f. (Brockwell and Davis, 1991, Theorem 4.4.2), gives insight: Let $X = \{X_t, t \in Z\}$ be an ARMA(p, q) process satisfying $A(L)X_t = B(L)\varepsilon_t, \{\varepsilon_t, t \in \mathbb{Z}\}$ is some white noise with variance σ^2, L is the lag-operator, and $A(z) = 1 - \sum_{j=1}^{p} a_j z^j, B(z) = 1 + \sum_{j=1}^{q} b_j$. If the polynomials $A(z)$ and $B(z)$ have no common zeroes and $A(z) \neq 0$ for $|z| = 1$, then X has spectral density

$$f(\lambda) = \frac{\sigma^2}{2\pi} \frac{A\left(\exp(-i\lambda)\right)}{B\left(\exp(-i\lambda)\right)}, -\pi \leq \lambda \leq \pi.$$

Besides the autocovariance, the time series itself can be expressed in frequency domain by using an orthogonal increment process. Since it is not used in this work, we are not going into detail here. It is more important that both domains contain the same amount of information. The only difference is the way this information is given. This different point of view can be enlightening for some applications, see section 1.5 in Brillinger (2001) for applications of the frequency domain. The autocovariance as well as the spectral density can be estimated with some observations $X_1, \ldots, X_n$. An estimator for the autocovariance is the sample autocovariance given by

$$\hat{\gamma}_n(h) = \frac{1}{n} \sum_{t=1}^{n-h} \left(X_{t+h} - \frac{1}{n} \sum_{s=1}^{n} X_s \right) \left(X_t - \frac{1}{n} \sum_{s=1}^{n} X_s \right), 0 \leq h \leq n-1, \tag{1.1.5}$$

$\hat{\gamma}_n(h) = 0, h \geq n, \hat{\gamma}_n(h) = \hat{\gamma}_n(-h)$. The corresponding spectral density is given by

$$I_n(\lambda) = \frac{1}{2\pi} \sum_{h \in \mathbb{Z}} \hat{\gamma}_n(h) \exp(-ih\lambda) = \frac{1}{2\pi n} \left| \sum_{t=1}^{n} \left(X_t - \frac{1}{n} \sum_{s=1}^{n} X_s \right) \exp(-it\lambda) \right|^2, -\pi \leq \lambda \leq \pi,$$

and I_n is called the Peridogram. Point-wise consistency of $\hat{\gamma}_n$, hence, for a given h, can be established, c.f. (Brockwell and Davis, 1991, Section 7.2). However, the periodogram I_n is inconsistent, c.f. (Kreiss and Neuhaus, 2006, Satz 12.7), which also implies that the absolute error for all sample autocovariance do not vanish, hence $\sum_{h=0}^{\infty} |\gamma(h) - \hat{\gamma}_n(h)| \neq o_P(1)$. This also implies that $\hat{\Sigma}_n = [\hat{\gamma}_n(i-j)]_{i,j=1,\dots,n}$ is not a consistent estimator of the autocovariance matrix $\Sigma_n = [\gamma(i-j)]_{i,j=1,\dots,n}$, c.f. McMurry and Politis (2010). Additional smoothing is required to get consistent estimators of the second-order properties. This can be achieved by using a truncated autocovariance estimator such as $\tilde{\gamma}_n(h) = k(h/M(n))\hat{\gamma}_n(h)$, where k is some kernel with support $[-1,1]$ and $M(n) < n$ such that $\tilde{\gamma}_n(h) = 0$ for $h > M(n)$. The resulting spectral density estimators $\hat{f}(\lambda) = 1/(2\pi) \sum_{h \in \mathbb{Z}} \tilde{\gamma}_n(h) \exp(-ih\lambda)$ are denoted as lag-window estimators and give consistent results, see Jentsch and Subba Rao (2015) as well as section 2.2.3 for details. Since the spectral density and the autocovariance describe the same information, only in different domains, such a truncation leads also to consistent estimators $\tilde{\Sigma}_n = [\tilde{\gamma}_n(i-j)]_{i,j=1,\dots,n}$ for the autocovariance matrix Σ_n, see Wu and Pourahmadi (2009) and McMurry and Politis (2010) for details. The spectral density plays a major role in chapter 2.

1.2 Overview of Bootstrap Methods for Time Series

In statistics when a certain quantity is estimated with a given statistic often the questions occurs how precise the estimation is and what deviation can be expected in x out of 100 cases. To answer such questions it is helpful to derive the distribution of the statistic. However, it is usually the case that it is not possible to derive the exact distribution. Instead, a consistent approximation is used. Bootstrap methods can be used to estimate the distribution of a given statistic. In its basic form the bootstrap method was introduced by Efron (1979). For a given statistic T the idea is as follows; Based on a sample $X = (X_1, \dots, X_n)$ new samples $(X_1^{*,j}, \dots, X_n^{*,j}), j = 1, \dots, N$ are created by using the empirical distribution function given by the sample X. Then the statistic is evaluated for each new sample, hence, we obtain $T_1^* = T(X_1^{*,1}, \dots, X_n^{*,1}), \dots, T_N^*$. The empirical distribution function of $T_1^*, \dots, T_N^*$ is then used as an approximation of the distribution function of T. N is the number of bootstrap samples and is similar to the number of trials in a Monte Carlo simulation. However, nothing is said about the performance of this approximation. We say that a bootstrap method is valid if $c_n(T_n - ET_n)$ and $c_n(T_n^* - E^*T_n^*)$ have the same limiting distribution, where c_n is such that $c_n(T_n - ET_n)$ converges to a non-degenerate distribution. Or more precisely, (Kreiss and Paparodi-

tis, 2017, Definition 1.41), let $(\Omega_n, \mathcal{A}_n, P_n), n \in \mathbb{N}$, be a sequence of statistical experiments and L_n a sequence of random variables on Ω_n. Given $x_n \in \Omega_n$, let $(\Omega_n^*, \mathcal{A}_n^*, P_n^*), n \in \mathbb{N}$, be a corresponding bootstrap statistical experiment and L_n^* bootstrap random variables. Denote by $\mathcal{L}_n$ the distribution of L_n and by $\mathcal{L}_n^*(x_n)$ the distribution of L_n^* given x_n. We denote the bootstrap proposal L_n^* as valid (or consistent, respectively) for L_n if and only if $\lim_{n\to\infty} d(\mathcal{L}_n, \mathcal{L}_n^*(x_n)) = 0$, in P_n-probability, where d is some distance measure between distributions. For features in probability see section 1.3.1 and especially Definition 1.7 in Kreiss and Paparoditis (2017). Possible distance measures are the Kolmogorov's distance, c.f. section 1.4.2 in Kreiss and Paparoditis (2017), and the Mallow's distance, c.f. section 1.4.3 in Kreiss and Paparoditis (2017). In this work the Mallow's distance is mainly considered. If the data consists of independent and identically distributed data points the bootstrap proposal of Efron (1979) is valid for most statistics and settings. However, time series data is considered here, hence the data points are dependent. In this case, the classical bootstrap proposal is not even valid for the sample mean $1/n \sum_{t=1}^{n} X_t$. That is why several new bootstrap ideas have been proposed to overcome this shortcoming of the classical bootstrap proposal. These ideas can be grouped and in the following only the basic concepts of the three most important groups are presented. The review paper by Kreiss and Paparoditis (2011) is recommended for a more exhaustive overview of the several bootstrap ideas. Further details can be found in Kreiss and Paparoditis (2017) and Lahiri (2003).

An intuitive extension of the classical proposal is the block bootstrap. In the classical proposal new samples are generated by drawing with replacement from the original sample. However, this destroys the dependent structure. In order to retain the dependent structure, the idea is to generate new samples by drawing with replacement from blocks of data points. Hence, within such a block a fraction of the dependence structure of the data is kept. In order to fully capture the dependence structure of the underlying process it is necessary that the block length increases to infinity as the sample size increases to infinity. For a valid approximation it is also necessary that the number of blocks increases as well. Many authors have adapted this idea. Some work with non-overlapping blocks has been done by Carlstein (1986) or Hall (1985), with overlapping blocks by Künsch (1989) or even overlapping blocks with random block length by Politis and Romano (1994). Furthermore, it is possible to taper the block-ends to get a smoother transition between blocks, c.f. Paparoditis and Politis (2001). The block bootstrap idea does not require that the underlying process follows some parametric structure. However, all block bootstrap variations have in common that they are in general very sensitive regarding the choice of the block length.

The setting of the residual bootstrap is that the underlying process $X = \{X_t, t \in \mathbb{Z}\}$ possesses some structure which can be expressed by $X_t = f(\varepsilon_t, \dots)$, where f is some unknown function and $\{\varepsilon_t, t \in \mathbb{Z}\}$ is a process which is less dependent than X. The ε's are denoted as the residuals.

The residuals are often uncorrelated, in some cases they are even independent. Based on a sample $X_1, \dots, X_n$ the idea of the residual bootstrap is to estimate f and the residuals. Afterwards the classical bootstrap approach on the residuals is being used. Hence, a new bootstrap observation of X_t is given by $X_t^* = \hat{f}(\varepsilon_t^*, \dots)$, where (ε_t^*) is sampled by the empirical distribution function given by $\hat{\varepsilon}_t, t = 1, \dots, n$. A classical example here is the case when X_t is an AR(p) process. Hence, $X_t = \sum_{j=1}^{p} a_j X_{t-j} + \varepsilon_t$. However, this bootstrap idea is not restricted to finite models. The AR-sieve bootstrap has the idea of approximating the dependence structure with AR-models of increasing order, see Kreiss (1992), Bühlmann (1997), Paparoditis and Streitberg (1991), and Kreiss et al. (2011). The linear process bootstrap by McMurry and Politis (2010) is another bootstrap proposal which does not require a specific finite model. This method is described in more detail in section 2.8.

A special form of the residual bootstrap is the frequency domain bootstrap, c.f. Franke and Hardle (1992), Hurvich and Zeger (1987) or Dahlhaus et al. (1996). For a time series $X_t = \sum_{j\in\mathbb{Z}} \phi_j \varepsilon_{t-j}, t \in \mathbb{Z}$, such bootstrap methods use the following approximation of the periodogram for linear processes at Fourier frequencies λ_j, $I_n(\lambda_j) \approx f(\lambda_j) I_{n,\varepsilon}(\lambda_j)$, where f is the spectral density of X and $I_{n,\varepsilon} = (2\pi n)^{-1} |\sum_{t=1}^{n} \varepsilon_t \exp(it\lambda_j)|^2$ is the periodogram of the residuals ε_t. Furthermore, we have under some conditions that the periodogram is asymptotically independent for different Fourier frequencies, c.f. (Brillinger, 2001, Theorem 5.2.6) or (Brockwell and Davis, 1991, Theorem 10.3.2). Hence, given some spectral density estimator $\hat{f}_n$ residuals $\tilde{\varepsilon}_k$ can be obtained by $\tilde{\varepsilon}_k = I_n(\lambda_k)/\hat{f}_n(\lambda_k)$. After normalization, those residuals can be resampled i.i.d. to obtain bootstrap values for the periodogram. Statistics as the sample autocovariance, sample autocorrelation, or spectral density estimators can be expressed by the integrated periodogram given by $\int_0^{2\pi} W(\lambda) I_n(\lambda) d\lambda)$, for some function $W : [0, 2\pi] \to \mathbb{R}$, see section 12.7 in Kreiss and Neuhaus (2006) for details. Therefore, the frequency domain bootstrap can be applied to those statistics. This bootstrap scheme creates new samples in the frequency domain. Some authors, c.f. Jentsch and Kreiss (2010) or Kirch et al. (2011), extended the idea of the frequency domain bootstrap to create also samples in the time domain.

The residuals used within such a residual bootstrap procedure can be bootstrapped wild. Hence, instead of using the estimated residuals some predefined distribution is used to sample residuals. Usually the residuals are sampled i.i.d., however it is possible to give these residuals also a predefined dependent structure.

1.3 Network Fundamentals

In its most general form a network denotes simply a collection of interconnected things, see (Kolaczyk, 2009, Chapter 1). Network data occur in many different fields such as social sciences, biology, physics or logistics. For instance, a social network of friendships between 34 members of a karate club, Zachary (1977), a network representing the topology of the western states power grid of the

United States, Watts and Strogatz (1998), or a network of human contact which could help to understand epidemics, Rocha et al. (2011).

A graph structure is used to describe this mathematically. A graph $G = (V, E)$ is a mathematical structure consisting of a set V of vertices and a set E of edges. In this work the vertices are labeled by $1, \dots, n$ such that $V = \{1, \dots, n\}$. Here, we consider directed edges. Consequently, E consists of ordered pairs $\{u, v\}, u, v \in V$. In the undirected case there is no distinction between $\{u, v\}$ and $\{v, u\}$. An edge $\{u, u\}$ is denoted as a loop and it is also possible that an edge $\{u, v\}$ is contained multiple times in E. Such edges are denoted as multi-edges. Graphs with directed edges and multi-edges are also denoted as multi-digraphs, see (Kolaczyk, 2009, Chapter 2). The connectivity of a graph G can be captured in an $n \times n$ matrix A with entries $A_{ij} = |\{e \in E : e = (i, j)\}|$. The matrix A is called the adjacency matrix and entry i, j gives the number of edges from vertex i to vertex j. The row sum $d_i^{\text{out}} = A_{i+} = \sum_{j=1}^{n} A_{ij}$ gives the number of edges which are going out from vertex i and d_i^{out} is denoted as the out-degree. The number of edges going into vertex i is given by the column sum $d_i^{\text{in}} = A_{+i} = \sum_{j=1}^{n} A_{ji}$ and is denoted as the in-degree. A graph with no multi-edges can contain at most n^2 edges. Hence, the density of a graph with no multiple edges can be defined by $\text{den}(G) = |E|/(n^2)$. We denoted a network as sparse if $|E| = O(n)$ and dense if $|E| = O(n^2)$.

In the example of the karate club, Zachary (1977), a vertex represents a person and an edge between two vertices represents friendship between the corresponding persons.

In this work a dynamic network is given by a family of graphs $\{G_t = (V_t, E_t), t \in \mathbb{Z}\}$ and a static network is given by a single graph G. That is why often the terms 'graph' and 'network' are used inter-changeably. If a static number of vertices is considered, then a dynamic network can be described by a time-dependent adjacency matrix $\mathbf{Ad} = \{Ad_t, t \in \mathbb{Z}\}$.

Several statistical models have been developed to describe such network data. An important model class is the exponential random graph model (ERGM), see section 6.5 in Kolaczyk (2009). We denote that a random vector Z belongs to an exponential family if its probability function can be expressed in the form $P_\Theta(Z = z) = \exp(\Theta^\top g(z) - \phi(\Theta))$, where $\Theta \in \mathbb{R}^p$ is a vector of parameters, g is a p-dimensional function of z, and $\phi(\Theta)$ is a normalization term, c.f. equation (6.23) in Kolaczyk (2009) or section 4.4 in Mood (1970). Let $Y_{ij}, i, j = 1, \dots, n$ be a binary random variable indicating the presence or absence of an edge from vertex i to vertex j. Then, an exponential random graph model is a model for which the joint distribution of elements in Y is specified in exponential family form. A special case of the ERGMs is the Bernoulli random graph model. For these models, it is considered that the edges are independent to each other, hence, $Y_{i,j}$ is independent to $Y_{s,k}$ for any $i, j \neq s, k$. Furthermore, $Y_{i,j}, i, j = 1 \dots, n$ is Bernoulli distributed and often it is further simplified that all edges share one common parameter. An ERGM describes a static network, however, Hanneke and Xing (2007) have extended these models to dynamic networks. In the dynamic setting, a common

assumption is that the network possesses some form of Markov property, c.f. Crane (2015). Hence, for a dynamic network with a static number of vertices this means that $\mathbf{Ad} = \{Ad_t, t \in \mathbb{Z}\}$ is a Markov process.

Bibliography

Brillinger, D. (2001). *Time Series: Data Analysis and Theory*. Classics in Applied Mathematics. Society for Industrial and Applied Mathematics.

Brockwell, P. and Davis, R. A. (1991). *Time Series: Theory and Methods (2nd edition)*. Springer, New York.

Bühlmann, P. (1997). Sieve bootstrap for time series. *Bernoulli*, 3(2):123–148.

Carlstein, E. (1986). The use of subseries values for estimating the variance of a general statistic from a stationary sequence. *The Annals of Statistics*, pages 1171–1179.

Crane, H. (2015). Time-varying network models. *Bernoulli*, 21(3):1670–1696.

Dahlhaus, R., Janas, D., et al. (1996). A frequency domain bootstrap for ratio statistics in time series analysis. *The Annals of Statistics*, 24(5):1934–1963.

Efron, B. (1979). Bootstrap methods: Another look at the jackknife. *Annals of Statistics*, 7:1–26.

Franke, J. and Hardle, W. (1992). On bootstrapping kernel spectral estimates. *The Annals of Statistics*, pages 121–145.

Hall, P. (1985). Resampling a coverage pattern. *Stochastic processes and their applications*, 20(2):231–246.

Hanneke, S. and Xing, E. (2007). Discrete temporal models of social networks. *Statistical Network Analysis: Models, Issues, and New Directions Edited by:Edoardo M. Airoldi, David M. Blei, Stephen E. Fienberg, Anna Goldenberg, Eric P. Xing, Alice X. Zheng*.

Hurvich, C. M. and Zeger, S. (1987). *Frequency domain bootstrap methods for time series*. New York University, Graduate School of Business Administration.

Jentsch, C. and Kreiss, J.-P. (2010). The multiple hybrid bootstrap—resampling multivariate linear processes. *Journal of Multivariate Analysis*, 101(10):2320–2345.

Jentsch, C. and Subba Rao, S. (2015). A test for second order stationarity of a multivariate time series. *Journal of Econometrics*, 185(1):124–161.

Kirch, C., Politis, D. N., et al. (2011). Tft-bootstrap: Resampling time series in the frequency domain to obtain replicates in the time domain. *The Annals of Statistics*, 39(3):1427–1470.

Kolaczyk, E. D. (2009). *Statistical Analysis of Network Data: Methods and Models*. Springer Publishing Company, Incorporated, 1st edition.

Kreiss, J.-P. (1992). *Bootstrap procedures for AR(∞)-processes*, volume 376. Springer, In Bootstrapping and Related Techniques of *Lecture Notes in Economics and Mathematical Systems*. 107–113.

Kreiss, J.-P. and Neuhaus, G. (2006). *Einführung in die Zeitreihenanalyse*. Statistik und ihre Anwendungen. Springer.

Kreiss, J.-P. and Paparoditis, E. (2011). Bootstrap methods for dependent data: A review. *Journal of the Korean Statistical Society*, 40(4):357–378.

Kreiss, J.-P. and Paparoditis, E. (2017). *Bootstrap for Time Series: Theory and Applications*. to appear.

Kreiss, J.-P., Paparoditis, E., and Politis, D. N. (2011). On the range of validity of the autoregressive sieve bootstrap. *Ann. Statist.*, 39(4):2103–2130.

Künsch, H. R. (1989). The jackknife and the bootstrap for general stationary observations. *Ann. Statist.*, 17(3):1217–1241.

Lahiri, S. (2003). *Resampling Methods for Dependent Data*. Springer Series in Statistics. Springer.

McMurry, T. L. and Politis, D. N. (2010). Banded and tapered estimates for autocovariance matrices and the linear process bootstrap. *Journal of Time Series Analysis*, 31(6):471–482.

Mood, A. M. (1970). Introduction to the theory of statistics. Third edition.

Paparoditis, E. and Politis, D. N. (2001). Tapered block bootstrap. *Biometrika*, 88(4):1105–1119.

Paparoditis, E. and Streitberg, B. (1991). Order identification statistics in stationary autoregressive moving-average models: Vector autocorrelations and the bootstrap. *Journal of Time Series Analysis*, 13(5):415–434.

Politis, D. N. and Romano, J. P. (1994). The stationary bootstrap. *Journal of the American Statistical Association*, 89(428):1303–1313.

Rocha, L. E., Liljeros, F., and Holme, P. (2011). Simulated epidemics in an empirical spatiotemporal network of 50,185 sexual contacts. *PLoS computational biology*, 7(3):e1001109.

Shao, X. (2010). The dependent wild bootstrap. *Journal of the American Statistical Association*, 105(489):218–235.

Watts, D. J. and Strogatz, S. H. (1998). Collective dynamics of small-worldnetworks. *nature*, 393(6684):440.