1

Introduction

1.1 Application of Time Series for Forecasting in Engineering

In engineering as well as in other fields such as the natural sciences, environmental science or economics, many processes exist for which ordered sequences of observed values are available. Examples of the latter include the settlement of a bridge measured at specific points in time, traffic loads on roads, snow depths measured over many years, the height of wheat stalks, the diameter of tree trunks or the production output in industry. The observed values, i.e. settlements, traffic loads, snow depths etc., exist for a past observation period. Under certain conditions these constitute a time series.

A time series is a temporally ordered sequence of observed values. Precisely one observed value is assigned to each discrete observation time $\tau \in \mathbf{T}$, where $\mathbf{T}$ represents a set of equidistant points in time. The set of observation time points $\tau = 1, 2, ..., N$ is referred to as the observation period.

In classical time series analysis the observed values are real-valued numbers or natural numbers, i.e. variables to which a precise numerical value is assigned.

Time series comprised of precise observed values are shown in Figs. 1.1, 1.2 and 1.3.

Forecasting of the future progression of time series containing precise observed values and hence forecasting of the process described by the observed values is the subject matter of classical time series analysis [6, 8, 60]. A forecast is possible due to the fact that particular dependencies may be deduced from the significant sequence of the observed values. In order to identify dependencies and laws within a time series, two methods are applied in classical time series analysis: descriptive time series analysis and stochastic models. In descriptive time series analysis descriptive models are used to identify attributes such as trends, seasonal variations or cyclic fluctuations. An important descriptive model is the component model. Stochastic models on the other hand

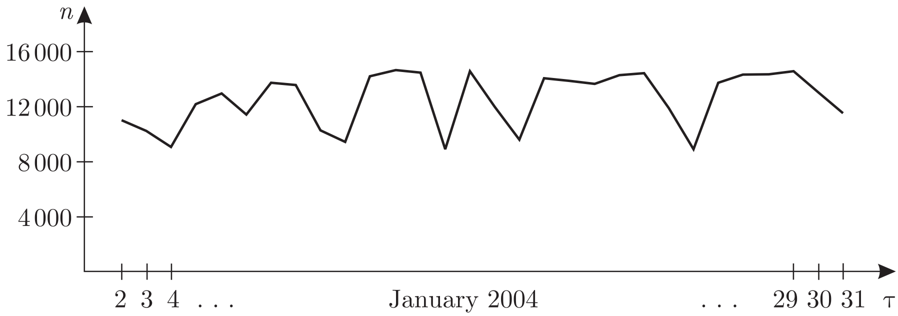

Fig. 1.1. Time series of the number of vehicles crossing the bridge 'Blaues Wunder' in Dresden [Source: Dresden Dept. of Road Construction and Public Works]

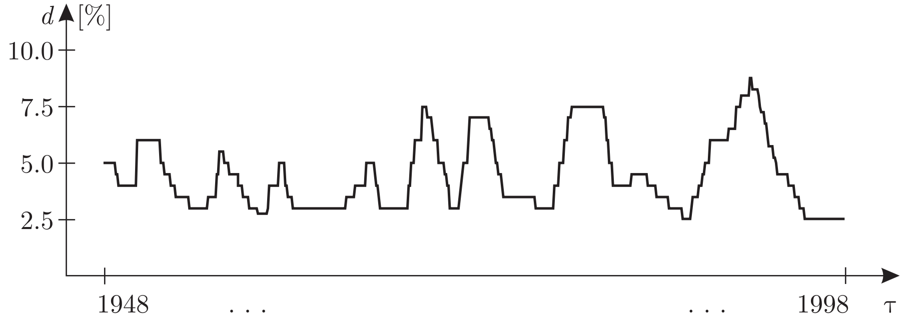

Fig. 1.2. Time series of the discount rate of the 'Deutsche Bundesbank' [Source: 'Deutsche Bundesbank']

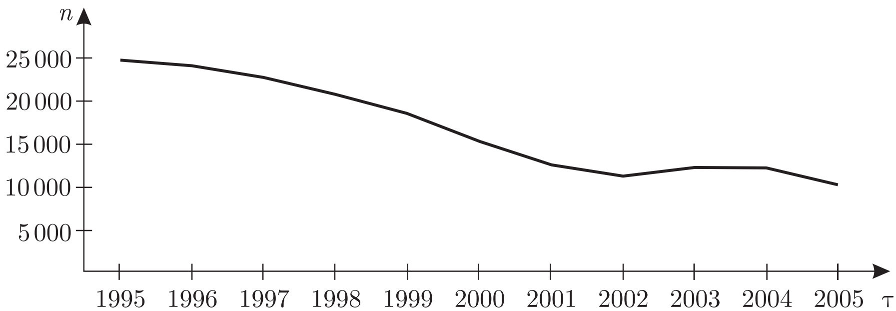

Fig. 1.3. Time series of the number of building approvals in Saxony [Source: Saxony State Office of Statistics]

assume stochastic properties, and treat the time series as the realization of a stochastic process.

If the observed values represent measured values, it is often not possible to assign precise numerical values to the observed data; they then possess data uncertainty. Data uncertainty in engineering practice is mainly due to inaccuracies in measurements, incomplete sets of observations or difficulties in performing measurements, e.g. due to local conditions. The occurrence of data uncertainty also depends on the particular observation scale adopted, i.e. whether a process is described on the microscale, mesoscale or macroscale. For example, although it is theoretically possible to precisely state the commencement of material damage on the microscale, the commencement of damage on the macroscale may be only diagnosed uncertain. Because the observation scale cannot always be chosen arbitrarily, however, the associated uncertainty must be accepted.

Measurement inaccuracy results among other things from the limited precision of a measuring device or from read errors. Geometric data in particular cannot be measured accurately in certain cases. Examples of this include water level measurements on a moving water surface, the thickness of a structural element with a very rough surface or the transport of bed material in a river. The stipulation of some sort of average value, however, means that important information may be lost.

Incomplete sets of observations signalize an information deficit due for example to gaps in a series of measurements resulting from the malfunctioning of measuring devices, irregularities in the reading of measurements or inadequate planning of the measurement regime. The measurement of parameters within a medium or construction is often extremely difficult. The corrosion behavior of steel reinforcement or the position of steel reinforcement in an RC structural element, for example, cannot be measured with absolute certainty. The same applies to crack formation in concrete elements or the quantity of water transported through a flow cross-section.

Sequences of observations may also consist of *linguistic estimates*. Examples of this include a description of concrete flaking on bridges (see Fig. 1.4), a description of the degree of discoloration of a surface or the extent of cloud cover. Linguistic estimates are a priori imprecise, as they express the subjective opinion of an expert. On the other hand, time series of linguistic observations do in fact open up new fields of application in forecasting.

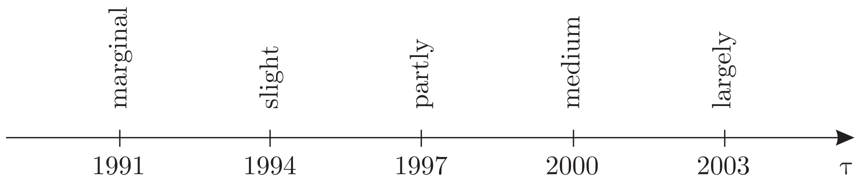

Fig. 1.4. Time series comprised of linguistic estimates of concrete flaking

Fig. 1.5 shows a comparison between a time series comprised of precise observed values and a time series consisting of uncertain observed values. The uncertainty in this case is described by an interval. This is a very simple uncertainty model. In this book the more informative uncertainty models fuzziness and fuzzy randomness are used to describe imprecise data, i.e. uncertain data. An overview of these uncertainty models is given in the following section (Sect. 1.2). New forms of representation of these uncertainty models suitable for time series analysis are derived in Sect. 2.

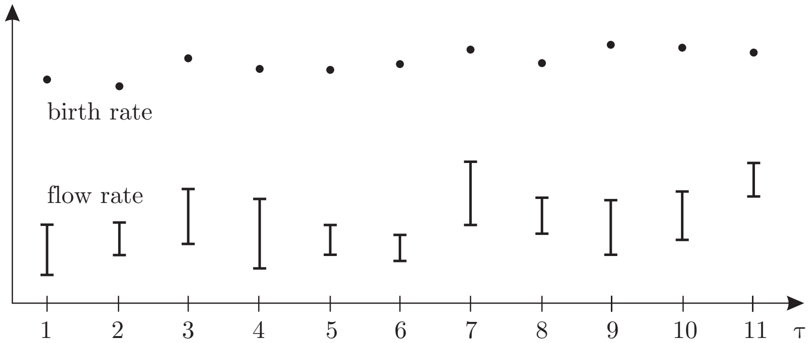

Fig. 1.5. Time series containing precise data versus time series containing uncertain data

By means of the introduced uncertainty models it is possible extend the methods of classical time series analysis in such a way as to permit the forecasting of future uncertain results under due consideration of data uncertainty. By this means it is possible to dispense with the artificial idealization of real data, the forecasting of which may lead to unrealistic results. The decision as to whether the methods of classical time series analysis or the extended methods presented in this book should be applied depends on the particular problem in question and the existing data base.

The subject matter of this book concerns time series comprised of imprecise, i.e. uncertain, observed values. This implies that an individual observed value may be uncertain. By this means it is possible to realistically model the observed values in important practical cases. Because the forecasted values are also uncertain, forecasts are obtained with higher predictive capability.

Three methods are described in the book for forecasting time series comprised of uncertain observed values:

- the fuzzy component model (Sect. 3.2) as an extension of descriptive methods,
- the fuzzy random process as an extension of stochastic models (Sect. 4.2) and

- artificial neural networks for uncertain data as an extension of artificial neural networks for real-valued data (Sect. 4.3)

1.2 Data Uncertainty and Fuzzy Time Series

If it is not possible to assign a precise numerical value to an observed value, the observed value in question possesses uncertainty. How can this uncertainty be described mathematically?

A variety of methods exist for classifying and distinguishing uncertainty. Decisive in this respect are the causes of uncertainty. If the cause is purely random, the uncertainty is referred to as aleatoric uncertainty. This is described with the aid of conventional and highly-developed stochastic models. If the uncertainty is a result of objective and subjective factors, it is then referred as epistemic uncertainty. Models for describing epistemic uncertainty include, among others, fuzziness and intervals. If it is necessary to take account of both aleatoric and epistemic effects, uncertainty is accounted for by the model fuzzy randomness.

In the case of time series the uncertainty of the individual observed values as well as the interpretation of a sequence of uncertain observed values are of interest.

The uncertainty of a single observed value is always epistemic. The uncertain observed value is thus modeled as a fuzzy variable, as illustrated in Fig. 1.6. The major causes of this type of uncertainty have already been outlined in Sect. 1.1.

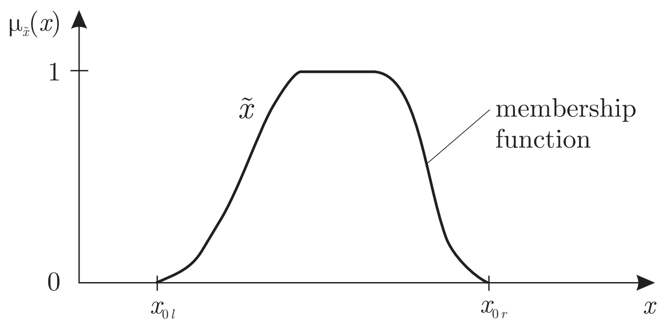

Fig. 1.6. Fuzzy variable $\tilde{x}$

The fuzzy variable $\tilde{x}$ may take on values in the interval $[x_{0\,l}; x_{0\,r}]$. The values are assessed between zero and unity by means of a membership function $\mu_{\tilde{x}}(x)$. This assessment reflects the subjective and objective causes of the existing uncertainty, and may be used to describe the uncertain observed value. Fuzzy variables contain intervals and real numbers as special cases, as illustrated in Fig.1.7.

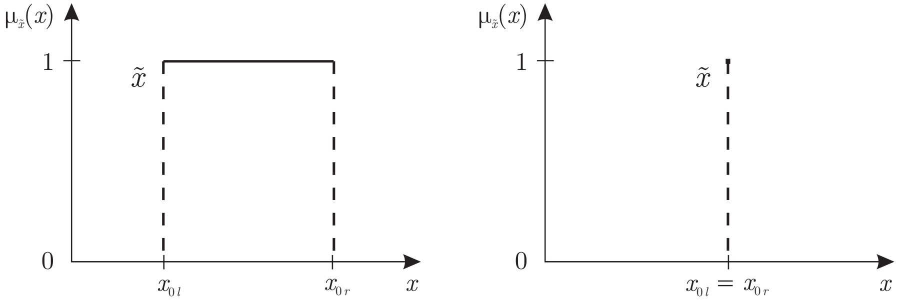

Fig. 1.7. Interval and real-valued number as special cases of a fuzzy variable $\tilde{x}$

Modeling of the individual observed values as fuzzy variables results in so-called fuzzy time series, as shown by way of example in Fig. 1.8. Starting from the uncertain observed values, the aim is to forecast future uncertain values. For this purpose the dependencies existing in the sequence of uncertain observed values are analyzed and modeled.

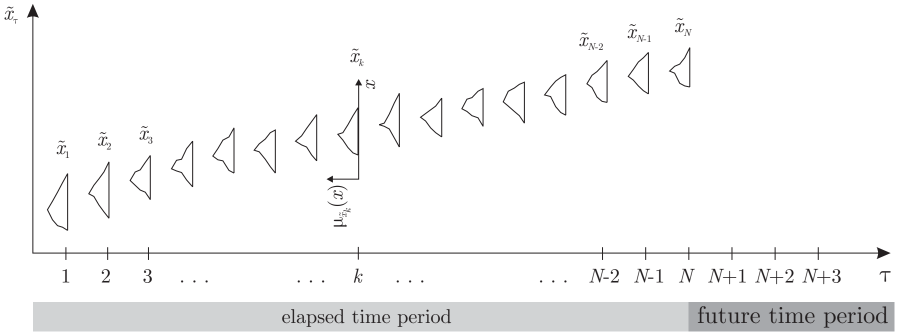

Fig. 1.8. Time series containing fuzzy variables

Modeling as a fuzzy random process. Forecasts are possible if it may be assumed that the fuzzy time series may be modeled with the aid of a fuzzy random process. A fuzzy random process is defined as a family of fuzzy random variables $\tilde{X}_\tau$.

Fuzzy random variables, as introduced in Sect. 2.2, belong to the uncertainty model fuzzy randomness. A time series of fuzzy data may be viewed as a random realization of a fuzzy random process. The realizations of this process are uncertain and thus referred to as fuzzy variables.

Only one sequence of uncertain observed values is available for determining the underlying fuzzy random process. Methods for specifying the fuzzy random process in any given case are developed in Sect. 3.5. A knowledge of this process is a precondition for the forecast. The required forecasting methods are formulated in Sect. 4.2. By means of a new incremental discretization of the fuzzy variables and fuzzy random variables the uncertainty is fully retained in the forecast. The uncertainty is also not artificially increased. This incremental representation is absolutely necessary for a *direct* description as well as for modeling and forecasting purposes. 'direct' implies that the sequence of the fuzzy variables is retained during the description, modeling and forecasting phases. No form of defuzzification or refuzzification is undertaken.

Modeling using Artificial Neural Networks. As an alternative to fuzzy random processes, methods for modeling and forecasting fuzzy time series using Artificial Neural Networks are developed in Sect. 3.6 and Sect. 4.3, respectively. The conventional multilayer perceptrons associated with the latter are extended in such a way that they may be applied to time series for fuzzy variables. A precondition for this extension is again the new incremental discretization mentioned in the foregoing. An Artificial Neural Network is first trained in an optimization process. Training is carried out on the basis of the particular fuzzy time series concerned. Different forecasting strategies are developed for forecasting purposes.

The use of Artificial Neural Networks does not require an explicit determination of the underlying fuzzy random process.

1.3 Examples of Fuzzy Time Series

The practical relevance of fuzzy time series is demonstrated by the following two examples. Further examples are given in Sect. 5.

The total global ozone change between 1965 and 2000 is shown in Fig. 1.9. The time series from 1980 onwards reflects the uncertainty of the measured data. The reason for this uncertainty is due to inaccuracies in measurements. The uncertainty is hence epistemic in nature and may be modeled by fuzziness.

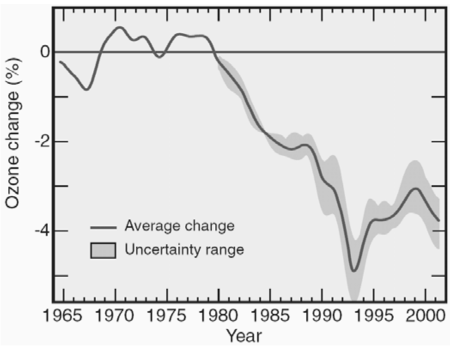

Fig. 1.9. Uncertain time series of total global ozone change [69]

The second example (Fig. 1.10) concerns measurements of the earth pressure acting on a wall. Several closely arranged pressure transducers are installed on the wall. The measured values differ from one pressure transducer to the next. The different observed values signalize uncertainty. Instead of computing an average value, this uncertainty is taken into consideration. Fuzzy variables are constructed for the measured values at each point in time. Fig. 1.10 shows a cut-out segment of the obtained fuzzy time series. The complete time series begins with measurements made in 1999.

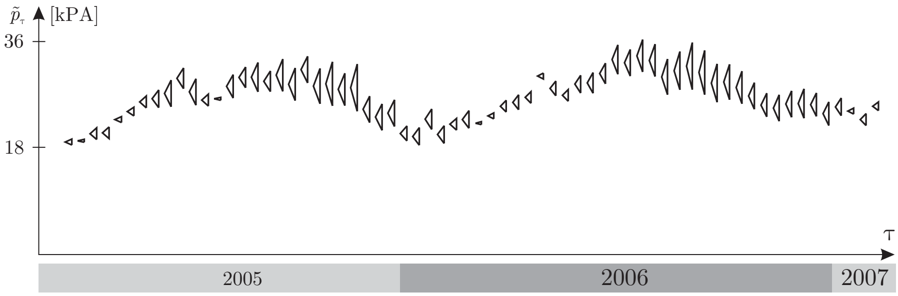

Fig. 1.10. Uncertain time series of earth pressure measurements (cut-out segment) [14]

Fuzzy time series analogous to those presented in the above examples are frequently encountered in engineering and environmental science. These share the common feature of measurable observed values. Forecasting of the latter is possible using the forecasting strategies developed in Sect. 4. If the forecasts of measurable observed values are combined with a computational model, it is also possible to forecast non-measurable observed values such as the damage state of a structure. Model-based forecasting strategies for this purpose are developed and demonstrated by way of examples in Sect. 5.

http://www.springer.com/978-3-540-37173-1

Uncertainty Forecasting in Engineering
Möller, B.; Reuter, U.
2007, XIV, 202 p., Hardcover
ISBN: 978-3-540-37173-1