Chapter 13
How to Describe and Propagate Uncertainty When Processing Time Series: Metrological and Computational Challenges, with Potential Applications to Environmental Studies

Christian Servin, Martine Ceberio, Aline Jaimes,
Craig Tweedie, and Vladik Kreinovich

Abstract. Time series comes from measurements, and often, measurement inaccuracy needs to be taken into account, especially in such volatile application areas as meteorology and economics. Traditionally, when we deal with an individual measurement or with a sample of measurement results, we subdivide a measurement error into random and systematic components: systematic error does not change from measurement to measurement while random errors corresponding to different measurements are independent. In time series, when we measure the same quantity at different times, we can also have correlation between measurement errors corresponding to nearby moments of time. To capture this correlation, environmental science researchers proposed to consider the third type of measurement errors: periodic. This extended classification of measurement error may seem ad hoc at first glance, but it leads to a good description of the actual errors. In this paper, we provide a theoretical explanation for this semi-empirical classification, and we show how to efficiently propagate all types of uncertainty via computations.

1 Formulation of the Problem

In many applications areas – e.g., in meteorology, in financial analysis – the value of the important variable (temperature, stock price, etc.) changes with time. In order to adequately predict the corresponding value, we need to analyze the observed time series and to make a prediction based on this analysis; see, e.g., [3, 20].

All the values that form the time series come from measurements or from expert estimates. Neither measurements nor expert estimates are 100% accurate, especially in such volatile application areas as meteorology and economics. Thus, the

Christian Servin · Martine Ceberio · Aline Jaimes · Craig Tweedie · Vladik Kreinovich
Cyber-ShARE Center,
University of Texas at El Paso,
El Paso, TX 79968, USA
e-mail: vladik@utep.edu

W. Pedrycz & S.-M. Chen (Eds.): Time Series Analysis, Model. & Applications, ISRL 47, pp. 279–299.
DOI: 10.1007/978-3-642-33439-9_13

actual values of the corresponding variables are, in general, slightly different from the observed values x_t. These measurement uncertainties affects the result of data processing.

For example, in meteorological and environmental applications, we measure, at different locations, temperature, humidity, wind speed and direction, flows of carbon dioxide and water between the soil and atmosphere, intensity of the sunlight, reflectivity of the plants, plant surface, etc. Based on these *local* measurement results, we estimate the *regional* characteristics such as the carbon fluxes describing the region as a whole – and then use these estimates for predictions. These predictions range from short-term meteorological predictions of weather to short-term environmental predictions of the distribution and survival of different ecosystems and species to long-term predictions of climate change; see, e.g., [1, 12]. Many of these quantities are difficult to measure accurately: for example, the random effects of turbulence and the resulting rapidly changing wind speeds and directions strongly affect our ability to accurately measure carbon dioxide and water flows; see, e.g., [18]. The resulting measurement inaccuracy is one of the main reasons why it is difficult to forecast meteorological, ecological, and climatological phenomena.

It is therefore desirable to describe how the corresponding measurement uncertainty affects the result of data processing. In this paper, we analyze this problem, describe the related challenges, and show how these challenges can be overcome.

2 Traditional Approach to Measurement Errors

When we are interested in the value x of some quantity that we can measure directly, we apply an appropriate measuring instrument and get the measurement result $\widetilde{x}$. In the ideal world, the measurement result $\widetilde{x}$ is exactly equal to the desired value x. In practice, however, there is noise, there are imperfection, there are other factors which influence the measurement result. As a consequence, the measurement result $\widetilde{x}$ is, in general, different from the actual (unknown) value x of the quantity of interest, and the *measurement error* $\Delta x \stackrel{\text{def}}{=} \widetilde{x} - x$ is different from 0.

Because of this, if we repeatedly measure the same quantity by the same measuring instrument, we get, in general, slightly different results. Some of these results are more frequent, some less frequent. For each interval of possible values, we can find the frequency with which the measurement result gets into this interval; at first, some of these frequencies change a lot with each new measurement, but eventually, once we have a large number of measurements, these frequencies stabilize – and become *probabilities* of different values of $\widetilde{x}$ and, correspondingly, probabilities of different values of measurement error Δx. In other words, the measurement error becomes a *random variable*.

Usually, it is assumed that random variables corresponding to different measurement errors are statistically independent from each other. In statistics, independence of two events A and B means that the probability of A does not depend on B, i.e., that the conditional probability $P(A\,|\,B)$ of A under condition B is equal to the unconditional probability $P(A)$ of the event A.