Environmental Modelling & Software 40 (2013) 171–180

Contents lists available at SciVerse ScienceDirect

Environmental Modelling & Software

journal homepage: www.elsevier.com/locate/envsoft

Set-membership approach for identification of parameter and prediction uncertainty in power-law relationships: The case of sediment yield

Karel J. Keesman [a,b,*], John P. Norton [b,c], Barry F.W. Croke [b], Lachlan T.H. Newham [b], Anthony J. Jakeman [b]

[a] Systems and Control Group, Wageningen University, Bornse Weilanden 9, 6708 WG Wageningen, The Netherlands
[b] Fenner School of Environment and Society, National Centre for Groundwater Research and Training, The Australian National University, Canberra, ACT 0200, Australia
[c] School of Electronic, Electrical & Computer Engineering, University of Birmingham, Birmingham B15 2TT, UK

ARTICLE INFO

Article history:
Received 21 June 2012
Received in revised form
5 September 2012
Accepted 14 September 2012
Available online 25 October 2012

Keywords:
Set-membership estimation
Least-squares
Uncertainty
Power laws
Erosion

ABSTRACT

Power laws are used to describe a large variety of natural and industrial phenomena. Consequently, they are used in a wide range of scientific research and management applications. This paper focuses on the identification of bounds on the parameter and prediction uncertainty in a power-law relation from experimental data, assuming known bounds on the error between model output and observations. The prediction uncertainty bounds can subsequently be used as constraints, for example in optimisation and scenario studies. The set-membership approach involves identification and removal of outliers, estimation of the feasible parameter set, evaluation of the feasible model-output set and tuning of the specified bounds on model-output error. As an example the procedure is applied to data of scattered sediment yield versus catchment area (Wasson, 1994). The key result is an un-falsified relationship between sediment yield and catchment area with uncertainty bounds on its parameters. The set-membership results are compared with the results from a conventional least-squares approach with first-order variance propagation, assuming a zero-mean, symmetrical error distribution.

1. Introduction

In recent decades uncertainty analysis has become an essential part of environmental model building (see e.g. Beck, 1987; Beven and Binley, 1992; Refsgaard et al., 2007). Given a well-defined method of characterizing uncertainty, the practical aims are to quantify the propagation of error in the data on which a model is based (observation error) to uncertainty in the model-parameter estimates, or propagation of model-parameter uncertainty to prediction uncertainty. This paper is a contribution to conventional and relatively unknown methods for quantifying these uncertainties.

In a wide range of scientific and other applications, power laws are used to describe natural and industrial phenomena quantitatively (e.g. Suki, 2002; Visser and Yunes, 2003; Glazier, 2005; Newman, 2005; Nacher and Akutsu, 2007; Ames et al., 2009; Deng and Jung, 2009; Lima-Mendez and Van Helden, 2009; Martinez, 2009; Millington et al., 2009). In particular, power-law probability densities are widely applied in earth sciences, linguistics, biology, economics and social sciences (e.g. Plerou et al., 2004; Doyne Farmer and Lillo, 2004; Gupta and Campanha, 2005; Arnold and Bauer, 2006; Schlicht and Iwasa, 2007) to relate sizes to frequency of occurrence. These laws embody the phenomenon that large is rare and small is common. In power-law densities, the exponent in the power law is always negative (since the density must integrate to unity). But the application of power laws is not limited to negative exponents. For instance, an exponent of 0.5 gives a square root relation between the free outflow rate from a tank and depth, as is derived from Bernoulli's law. An exponent greater than 1 gives an increasing input–output relation, as is frequently seen in biology. In fact, many well-known laws in physics are expressed in terms of a power law function, for instance, the Stefan–Boltzmann law, the inverse-square laws of Newtonian gravity and electrostatics, van der Waals' force model, Kepler's third law, the square–cube law (ratio of surface area to volume) and the twentieth century's best-known equation $E = mc^2$; also Pareto's principle follows a power law. In the pre-computer era, scientists plotted all kinds of phenomena on log–log scales to arrive at a linear relationship between variables related by a power law. To summarise, power laws are a widely useful way of describing relationships.

Apart from some derived relationships in physics, power laws are most frequently obtained directly from experimental data (Clauset et al., 2009). Experimental data always contain

* Corresponding author. Systems and Control Group, Wageningen University, Bornse Weilanden 9, 6708 WG Wageningen, The Netherlands.
E-mail address: karel.keesman@wur.nl (K.J. Keesman).

1364-8152/$ – see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.envsoft.2012.09.007

measurement and sampling errors, usually characterized statistically. Consequently, the estimates of the power-law parameters are statistical. However, the presumed statistical properties are not always valid, for instance in case of very limited data or after non-linear transformation of the data. As an alternative, a bounded-error characterization of errors has often been employed in recent decades, sometimes under the name of the set-membership approach.

The objective of the paper is to present a set-membership approach to the identification of error bounds on the parameters and hence on the output of a power law, and to compare these results with those obtained from a conventional least-squares approach with first-order variance propagation, assuming a zero-mean, symmetrical error distribution. As an example, the approach is applied to data on river sediment yield and catchment area (Wasson, 1994).

In Section 2, background material on power laws with the conventional statistical error approaches and the set-membership approach is presented. Section 3 gives results of applying the set-membership approach to data of Wasson and results from a statistical error approach using conventional least-squares with first-order variance propagation. The paper concludes with a discussion on both approaches and possible extensions of the set-membership approach to more general cases, and some concluding remarks.

2. Background

2.1. Statistical error approach

Power laws express the scalar output y as a single-term polynomial in a scalar input x:

$$y = ax^k + o\left(x^k\right) \equiv f(x) + e \tag{1}$$

where a, k are real constants and $o(x^k)$ is asymptotically small and treated as part of the model-output error. The parameter k is called the scaling exponent, since a property of a power law is scaling invariance: if x is multiplied by a constant c then $f(cx) = a(cx)^k = c^k f(x) \propto f(x)$, merely scaling the function. Allometric scaling laws, for instance, frequently used to describe the relation between biological variables, are some of the best-known power laws in nature. Notice from (1) that for $k \in \{-1, 0, 0.5, 1, 2\}$ some typical power law functions, as the hyperbolic, constant, square root, linear and parabolic function, result. In what follows, and from the viewpoint of parameter estimation, (1) is called a non-linear regression, with unknown parameters a and k.

Notice, however, from (1) that when k is given, a can be simply estimated from the resulting linear regression using ordinary least-squares (LS) estimation. On the contrary, the estimation of the exponent k is not so easy. There are many ways of estimating the scaling exponent in a power law from data. However, not all of them yield unbiased and consistent estimates. A commonly applied technique is to apply a (natural) logarithm transformation to the deterministic part $y = ax^k$ of (1), which results in the linear regression

$$\ln y = \ln a + k \ln x \tag{2}$$

It is evident that logarithmic transformation distorts the error e (see e.g. Bartlett, 1947; Box and Cox, 1964). In this specific case, using a Taylor series expansion,

$$\ln y = \ln\left(ax^k + e\right) \cong \ln a + k \ln x + \frac{e}{ax^k} - \frac{e^2}{2ax^{2k}} \tag{3}$$

Hence

$$E[\ln y] \cong E[\ln a + k \ln x] - \frac{\sigma_e^2}{2ax^{2k}} \tag{4}$$

with σ_e^2 the variance of e and assuming that $E[e]$ is zero. Consequently, the error term in the transformed equation is not zero-mean, and ordinary least squares (LS) would give biased estimates for $\ln a$ and k and thus biased predictions. An alternative is to assume zero-mean, additive error e in the log-transformed model, avoiding the approximation in (3) and yielding unbiased ordinary LS estimates of $\ln a$ and k. This model would be more appropriate if the error in the power law output were multiplicative and the zero-mean assumption more plausibly applied to log-transformed error. If the error distribution is to be considered, e.g. to allow maximum-likelihood (ML) estimation, logarithmic transformation of the data alters the distribution of the original error, typically making it less convenient for ML estimation. For example, if e is assumed to be Gaussian, the transformed error is log-normally distributed. ML estimation of a and k from (1), assuming e to be Gaussian, is algebraically but not computationally straightforward, as iterative solution for k is necessary. In any case, with limited data the assumption of known error distribution, as in an ML estimation procedure, is often questionable and cannot be adequately tested.

Non-linear least-squares estimation of the parameters a and k in the original non-linear regression $y = ax^k$ avoids all these statistical considerations, though at the cost of requiring an iterative procedure to find a solution. The fitting can be performed using ordinary non-linear least squares techniques if the variance of the dependent variable is constant over the range of the independent variable. If this is not the case, then a weighted version of a non-linear least-squares method should be used. Generally, the weights should be equal to the reciprocal of the variance of each observation (e.g. Croke, 2007).

2.2. Set-membership (bounded-error) approach

Given a small data set, as is common in practice, an alternative route for exploring the values of, and uncertainty in, the model parameters and subsequently the model predictions is set-membership estimation (Walter, 1990; Norton, 1994, 1995; Milanese et al., 1996). Consider the non-linear regression model in vector form

$$\mathbf{y} = \mathbf{F}(\vartheta) + \mathbf{e} \tag{5}$$

where $\mathbf{y} \in \mathbb{R}^N$ contains the N observed output values and $\mathbf{F}(\vartheta)$ is a generally non-linear vector function mapping the unknown parameter vector $\vartheta \in \mathbb{R}^m$ into the model output. $\mathbf{F}$ incorporates the input (regressors) data as in (1), of course, but here the focus is on its dependence on the parameters constituting ϑ. In set-membership estimation, the error vector $\mathbf{e}$ is assumed to be bounded in a given norm, but no other assumptions are made about its distribution or statistical properties. In what follows, we assume that

$$\|\mathbf{e}\|_\infty \le \varepsilon \tag{6}$$

where ε is a fixed positive number. In other words, the largest individual error is assumed to be bounded by $\pm\varepsilon$, so $-\varepsilon \le e_i \le \varepsilon$ for $i = 1,\ldots, N$. A measurement uncertainty set (MUS), containing all