VSP International
Science Publishers
P.O. Box 346, 3700 AH Zeist
The Netherlands

Lecture Series on Computer and Computational Sciences
Volume 1, 2004, pp. 1-6

Nonlinear Data Fitting for Landslides Modeling

K.E. Parsopoulos*[1], V.A. Kontogianni†[2], S.I. Pytharouli†[3], P.A. Psimoulis†[4], S.C. Stiros†[5] and M.N. Vrahatis*[6]

*Department of Mathematics, University of Patras Artificial Intelligence Research Center (UPAIRC), University of Patras, GR–26110 Patras, Greece

†Department of Civil Engineering, University of Patras, GR–26500 Patras, Greece

Received 31 July, 2004; accepted in revised form 5 September, 2004

Abstract: We consider data fitting schemes that are based on different norms to determine the parameters of curve–models that model landslides in dams. The Particle Swarm Optimization method is employed to minimize the corresponding error norms. The method is applied on real–world data with promising results.

Keywords: Curve Fitting, Optimization, Particle Swarm Optimization, Swarm Intelligence, Landslides Modeling

Mathematics Subject Classification: 65D10, 90C26

1 Introduction

A common problem in physics, earth sciences and engineering is the optimal fitting of a curve to a set of observations of certain parameters versus time (or space). The observations are usually contaminated by various types of errors. The usual procedure that is followed to solve such problems is to test various empirically selected model–curves and estimate the parameters of the curve that minimize the difference between the values obtained through the model and the observed ones.

Traditionally, this task is accomplished using the well–known Least Squares Method (LSQR). More specifically, linear or linearized equations are used and the sum of squares of differences among observations and the corresponding model–curve values is minimized. Therefore, the practitioner has to decide only regarding the most appropriate curve–model (e.g. polynomial, periodic, exponential, mixed, etc.) such that an acceptable fit is obtained.

In some cases, however, the available data are noisy, unevenly distributed versus time, there is no *a priori* knowledge of the variance–covariance matrix or they do not correspond to rather smooth curves (for instance they include offsets, a usual case in tectonic and geotechnical studies [4]). In such cases, the LSQR approach may not be successful, resulting in complex curve–models that lack physical significance and ability to be incorporated to further modeling and analysis. In such cases, the use of different data fitting approaches has been proved very useful [1].

[1] Corresponding author. E-mail: kostasp@math.upatras.gr
[2] E-mail: vkont@civil.upatras.gr
[3] E-mail: spitha@upatras.gr
[4] E-mail: papsimouli@upnet.gr
[5] E-mail: stiros@upatras.gr
[6] E-mail: vrahatis@math.upatras.gr

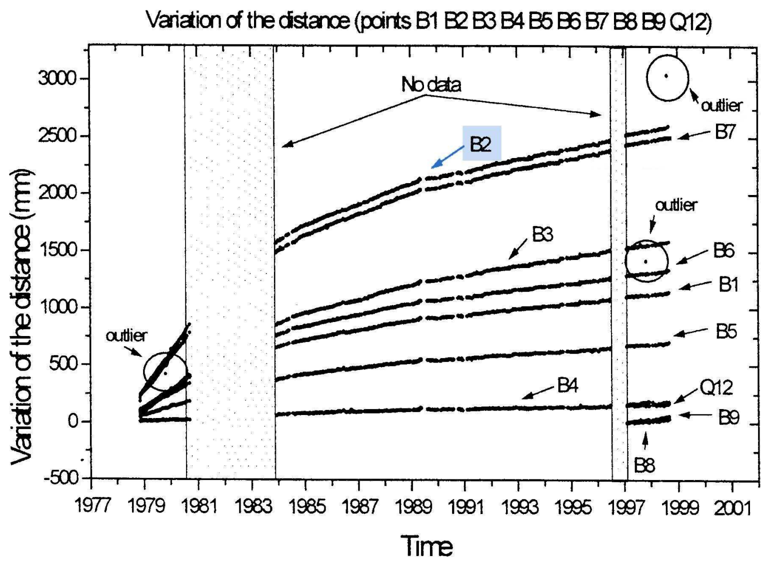

Figure 1: The record of observations for the Polyfyto Dam.

Evolutionary and Swarm Intelligence algorithms have been successfully applied on several data fitting problems [5, 6]. Their ability to work using solely function values even for discontinuous and non–differentiable functions renders them a promising alternative in cases where traditional algorithms, such as LSQR, fail. The aim of this paper is to investigate alternative curve fitting techniques based on the Particle Swarm Optimization (PSO) algorithm and three different norms to cope with a real–life curve fitting problem from the field of Civil Engineering. Results are reported and discussed.

Section 2 is devoted to the description of the problem, while the employed optimization algorithm, PSO, is briefly described in Section 3. Experimental results are reported and discussed in Section 4.

2 Description of the Curve Fitting Problem and Models

The problem investigated here is the monitoring of a landslide of the Polyfyto Dam in the Aliakmonas river in north Greece. A record of observations has been collected in collaboration with the Greek Public Power Corporation s.a.. The record consists of a large number of observations of distance changes obtained by monitoring 7 control points, denoted as $B1 - B7$, on the landslide relative to a stable reference station on stable ground, over a period of 20 years. The record is depicted in Figure 1, along with observations for 3 auxiliary points, $B8$, $B9$ and $Q12$. As we can see, the control point $B2$ exhibited the largest displacement.

The first step in the analysis of the landslide is the determination of a mathematical model, which captures the pattern of the landslide movement and can be used to estimate its future trends [7]. For this purpose, the movement of each control point was individually investigated. In Figure 1 it is clear that almost all points are moving faster in early years, while their movement tends to be stabilized in late years. This effect can be described using different mathematical models, although, just a few models retain the physical meaning of the specific phenomenon. The simplest model that could be used is a polynomial of degree four. However, it exhibits some upward and downward branches that do not fit the observations, and for this purpose, two types of an exponential decay model were adopted,

$$\textbf{Model 1:} \quad f(t) = A\,(1 - \exp(-t/B)) + C, \qquad (1)$$

$$\textbf{Model 2:} \quad f(t) = A\,(1 - \exp(-t/B)) + K\,t + C. \qquad (2)$$

The next step in the analysis is the determination of the unknown parameters A, B, C and K, such that the error among the observations and the corresponding values provided by the model is minimized. For the error measurement, several norms can be used. The most common choices are the ℓ_1, ℓ_2 and ℓ_∞–norms, which are defined as,

$$\|\varepsilon\|_1 = \sum_{i=1}^{m} |\varepsilon_i|, \qquad \|\varepsilon\|_2 = \left(\sum_{i=1}^{m} |\varepsilon_i|^2\right)^{1/2}, \qquad \|\varepsilon\|_\infty = \max_{1 \leqslant i \leqslant m} |\varepsilon_i|,$$

respectively, where m is the number of observations and $\varepsilon_i = M_i - O_i$, $i = 1, \ldots, m$, with O_i being the ith observed value and M_i be the corresponding value implied by the model.

The ℓ_1–norm is the most "fair" norm since it uses the absolute values of the errors. However, it results in non–differentiable minimization problems, therefore, it cannot be used with traditional gradient–based minimizers. On the other hand, the ℓ_2–norm results in differentiable minimization problems but the assumed error values are not always consistent with the actual ones. For example, an absolute error value equal to 10^{-3} becomes 10^{-6}, while an absolute error equal to 10^2 becomes 10^4. The ℓ_∞–norm constitutes the most proper choice in cases where outliers that must be taken seriously into consideration appear in the set of observations, since it minimizes the maximum among all absolute errors.

The performance of LSQR for the determination of the unknown parameters A, B, C and K, is rather poor with the deviation being larger at the edge of the curve where indeed a good fitting is sought. This happens due to the ℓ_2–norm, on which LSQR is based. Thus, alternative fitting techniques that use different norms are of great interest in order to provide more reliable results.

3 Particle Swarm Optimization

Particle Swarm Optimization (PSO) is a swarm intelligence optimization algorithm developed by Eberhart and Kennedy [3]. It employs a population, called a *swarm*, $\mathbb{S} = \{x_1, \ldots, x_N\}$, of search points, called *particles*, $x_i = (x_{i1}, x_{i2}, \ldots, x_{in})^\top$, $i = 1, \ldots, N$, which probe the search space, $S \subset \mathbb{R}^n$, simultaneously. The algorithm works iteratively. Each particle is initialized to a random position in the search space. Then, at each iteration, each particle moves with an adaptable *velocity*, $v_i = (v_{i1}, v_{i2}, \ldots, v_{in})^\top$, while retaining in a memory the best position, $p_i = (p_{i1}, p_{i2}, \ldots, p_{in})^\top \in S$, it has ever visited in the search space. In minimization problems, best positions have lower function values. The particle's movement is also influenced by the experience of the rest particles, i.e., by their best positions. This is performed through the concept of neighborhood. More specifically, each particle is assigned a neighborhood which consists of some prespecified particles. Then, the particles that comprise the neighborhood share their experience by exchanging information. There are two main variants of PSO with respect to the number of particles that comprise the neighborhoods. In the *global* variant, the whole swarm is considered as the neighborhood of each particle, while, in the *local* variant, smaller neighborhoods are used. Neighboring particles are determined based rather on their indices than their actual distance in the search space [6].

Let g_i be the index of the best particle in the neighborhood of x_i, i.e., the index of the particle that attained the best position among all the particles of the neighborhood. The indices of the particles are considered in a cyclic order, i.e., 1 is the index that follows after N. At each iteration,

Table 1: Computed solutions for the two models.

Model	Norm	A	B	C	K
Model 1	ℓ_1	2377.10	2447.27	236.68	
	ℓ_2	2391.74	2468.11	246.08	
	ℓ_∞	2397.33	2588.62	283.42	
Model 2	ℓ_1	1626.50	1431.98	0.106	178.76
	ℓ_2	1624.57	1464.39	0.106	187.87
	ℓ_∞	1686.59	1582.49	0.096	202.47

the swarm is updated according to the equations [2, 8],

$$v_i^{(k+1)} = \chi\left[v_i^{(k)} + c_1 r_1\left(p_i^{(k)} - x_i^{(k)}\right) + c_2 r_2\left(p_{g_i}^{(k)} - x_i^{(k)}\right)\right], \quad (3)$$

$$x_i^{(k+1)} = x_i^{(k)} + v_i^{(k+1)}, \quad (4)$$

where $i = 1, \ldots, N$; k is the iterations' counter; χ is a parameter called *constriction factor* that controls the velocity's magnitude; c_1 and c_2 are positive acceleration parameters, called *cognitive* and *social* parameter, respectively; and r_1, r_2 are random vectors that consist of random values uniformly distributed in $[0, 1]$. All vector operations in Eqs. (3) and (4) are performed componentwise. A stability analysis of PSO, as well as recommendations regarding the selection of its parameters are provided in [2, 8].

PSO has been applied on ℓ_1–norm errors–in–variables data fitting problems with very promising results, exhibiting superior performance even than the well–known Trust Region methods [5]. Therefore it was selected for the error minimization in our problem using the ℓ_1, ℓ_2 and ℓ_∞–norms.

4 Results and Discussion

The PSO algorithm was used for the determination of parameters of the two models defined in Eqs. (1) and (2), minimizing the error defined through the ℓ_1, ℓ_2 and ℓ_∞–norms, which will be denoted as $L1$, $L2$ and $L3$, respectively. We concentrated on the case of the control point $B2$, which had the largest displacement in our set of observations. The data set for $B2$ consisted of 404 observations. For the PSO, the default parameters, $\chi = 0.729$ and $c_1 = c_2 = 2.05$ were used. The swarm size was equal to 60 for Model 1 and 80 for Model 2. The algorithm was let to run for 5000 iterations. We conducted 100 independent experiments for each model and norm.

In all experiments, the same solutions (model parameters) were computed and they are reported in Table 1. The absolute error for each observation was also recorded for the detected model parameters. The mean value and the standard deviation of these absolute error values as well as the typical error for a single observation,

$$S0 = \sqrt{\frac{\varepsilon^2}{m-n}},$$

where m is the number of observations and n is the dimension of the problem were computed for the three norms. For Model 1, the plot of the actual data along with the corresponding model values for each norm, a boxplot with the distribution of the absolute error for the 404 observations

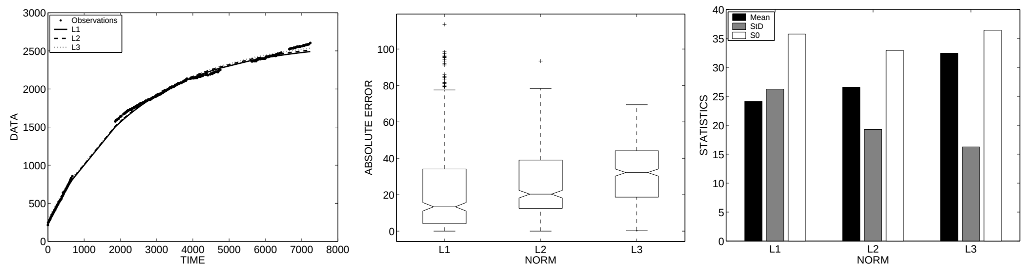

Figure 2: Plot of the actual data along with the corresponding model values for each norm (left), boxplot with the distribution of the absolute error for all observations for the computed model parameters (center), and statistics of absolute error (right) for Model 1. Labels $L1$, $L2$ and $L3$ correspond to the norms ℓ_1, ℓ_2 and ℓ_∞, respectively.

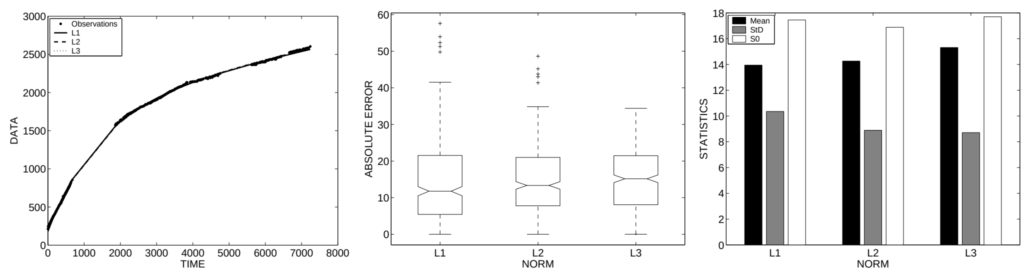

Figure 3: Plot of the actual data along with the corresponding model values for each norm (left), boxplot with the distribution of the absolute error for all observations for the computed model parameters (center), and statistics of absolute error (right) for Model 2. Labels $L1$, $L2$ and $L3$ correspond to the norms ℓ_1, ℓ_2 and ℓ_∞, respectively.

for the computed model parameters, as well as a bar plot with the mean value and standard deviation of absolute error and the quantity $S0$, are depicted in Figure 2. Figure 3 reports the corresponding graphs for Model 2. The boxplot produces a box and whisker plot for the sample of 404 absolute error values. The box has lines at the lower quartile, median, and upper quartile values. The whiskers are lines extending from each end of the box to show the extent of the rest of the data. Outliers are data with values beyond the ends of the whiskers. Notches represent a robust estimate of the uncertainty about the medians for box to box comparison. All displacement units in figures are in millimeters (mm).

In the case of Model 1, the ℓ_1–norm exhibited the smallest mean absolute error, followed by ℓ_2 and ℓ_∞. The latter norm had the smallest standard deviation of absolute error, which implies its robustness. Finally, the best value of $S0$ was obtained using the ℓ_2–norm. Model 2 provided a far better fit, although the corresponding model is more complex and harder to be incorporated in further analysis. The same comments with Model 1 can be made for the mean value, standard deviation and $S0$, although the differences between the different norms are smaller than in the case of Model 1. The ℓ_∞–norm is much better, especially at the edges of the intervals covered by observations.

Concluding, the three different approaches through PSO using ℓ_1, ℓ_2 and ℓ_∞–norms resulted in an efficient scheme that optimizes the exponential decay models considered for the curve fitting problem of the Polyfyto Dam, providing further intuition on tackling similar problems. Further work is needed toward the direction of estimating the future trends of the landslide.

Acknowledgment

This work was partially supported by the PENED 2001 Project awarded by the Greek Secretariat of Research and Technology.

References

[1] Amiri–Simkooei, A., Formulation of ℓ_1 norm minimization in gauss–markov models, *J. Surveying Engineering, ASCE*, **129**(1), pp. 37–43, 2003.

[2] Clerc, M. and Kennedy, J., The particle swarm–explosion, stability, and convergence in a multidimensional complex space, *IEEE Transactions on Evolutionary Computation*, **6**(1), pp. 58–73, 2002.

[3] Kennedy, J. and Eberhart, R.C., *Swarm Intelligence*, Morgan Kaufmann Publishers, 2001.

[4] Liu, Q.W. and Chen, Y.Q., Combining the geodetic models of vertical crustal deformation, *J. Geodesy*, **72**, pp. 673–683, 1998.

[5] Parsopoulos, K.E., Laskari, E.C. and Vrahatis, M.N., Solving ℓ_1 norm errors-in-variables problems using particle swarm optimizer, In M.H. Hamza, editor, *Artificial Intelligence and Applications*, pp. 185–190. IASTED/ACTA Press, 2001.

[6] Parsopoulos, K.E. and Vrahatis, M.N., On the computation of all global minimizers through particle swarm optimization, *IEEE Transactions on Evolutionary Computation*, **8**(3), pp. 211–224, 2004.

[7] Stiros, S., Vichas, C. and Skourtis, C., Landslide monitoring based on geodetically–derived distance changes, *J. Surveying Engineering, ASCE*, 2004. in press.

[8] Trelea, I.C., The particle swarm optimization algorithm: Convergence analysis and parameter selection, *Information Processing Letters*, **85**, pp. 317–325, 2003.