Uncertainty analysis of a spatiotemporal model for submerged vegetation colonization

Ligia Flávia Antunes Batista[1,2], *Nilton Nobuhiro Imai*[2] ,*Luiz Henrique da Silva Rotta*[2] , *Fernanda Sayuri Yoshino Watanabe* [2] *and Edivaldo Domingues Velini*[3]

[1] Federal University of Technology - Parana (UTFPR), Londrina, PR, 86036-370, Brazil
ligia@utfpr.edu.br
[2] Sao Paulo State University, Presidente Prudente, SP, 19060-900, Brazil
nnimai@fct.unesp.br, luizhrotta@yahoo.com.br, fernandasyw@yahoo.com.br
[3] Sao Paulo State University, Botucatu, SP, 18603-970, Brazil
velini@fct.unesp.br

Abstract

This work presents an uncertainty analysis applied to the results of an ecological model. This model describes the development of submerged macrophytes colonization in a brazilian reservoir, between Sao Paulo and Parana states. To build the model we map the submerged vegetation with hydroacoustic technique to estimate submerged canopy height. Data about the light penetration into the water were also collected in some points. The dynamic model was elaborated with two variables: depth and attenuation coefficient (kt). Monte Carlo technique was used to evaluate how the existing uncertainty in the data acquisition process and measurement tools, propagated to the kriging interpolation, affects the model results. It was possible to evaluate the model output histograms, and the Root Mean Square Error (RMSE) of each simulated point in relation to the observed one. The confidence intervals were also calculated with the 5th and 95th percentiles. With this uncertainty analysis, the interval time and the points with the lowest uncertainty could be identified.

Keywords: Monte Carlo, kriging, mapping, ecology, macrophytes.

1. Introduction

Models are simplified representations of the reality. Because of this, they have a limited scope, due to unknown features or natural changes which cannot be completely evaluated. So, there are sources of uncertainty that must be analyzed in order to determine their influences.

Among the sources of uncertainty, it is possible to enumerate: measurement errors, errors in mathematical formulation, parameters or due to unpredictable changes (Mulligan and Wainwright, 2004). In spatial explicit models, there are also interpolation and positioning errors (Heuvelink *et al*., 2010).

Uncertainty analysis can be performed analytically, using Taylor series or numerical methods, as Monte Carlo technique, an approach that consists in repeatedly running the model, using as inputs the variable values sampled from their distribution probability functions. After simulation, it is possible to analyze the percentiles of the results to estimate the confidence intervals that quantify uncertainty (van Nes and Scheffer 2003).

Many related works have been investigating issues of uncertainty in the modelling process (Refsgaard *et al*., 2007; Heuvelink *et al.,* 2007; Yeh and Li, 2003).

The application domain of this modelling is the colonization in reservoir by a kind of submerged vegetation called macrophytes. This vegetation type has developed excessively in tropical water bodies, causing ecological and economic damages (Michelan *et al*., 2010; Sousa *et al*., 2010).

The infestation of macrophytes hinders navigation, and recreation. In economic terms, hydroelectric power plants are seriously affected, since it is necessary to stop the production many times during the year to remove vegetation retained in turbines. Because of these issues, a way of controlling the proliferation must be applied in order to manage further economic and ecological concerns.

The aim of this study is to analyze uncertainty of the model results. This model describes the spatiotemporal dynamics of submerged macrophytes in reservoirs.

The paper is structured as follows. In Section 2 the applied methodology is presented. In Section 3 the preliminary results are discussed. In Section 4 we draw some conclusions of the work.

2. Methodology

The study area is located in Paranapanema river, near Santo Inácio city, state of Paraná, in Brazil, at coordinates UTM 22S 424900E, 7497750N and 425100E, 7497900N, as illustrated in Figure 1. It is a portion of about 20000m^2 of Taquaruçu reservoir, where the depth ranges from 0.7m to 5m. Four field surveys had been made, from April to August, 2010.

In these surveys, the submerged vegetation was measured with an echosounder which, based on hydroacoustic technique (sonar), estimates bathymetry and the canopy height (Sabol *et al.*, 2002; Valley *et al.,* 2005). Biosonics DT-X model was used, which operates in frequencies between 38 kHz and 1000 kHz and measures depths below 1000 meters. In order to map it is necessary to ride on boat, with the transducer coupled to a GPS (Global Positioning System). Each point provides information about coordinates, depth, and canopy height. Figure 1 shows the points acquired in a survey.

After acquiring the samples, the data were imported to an Geographical Information Systems (GIS), interpolated with kriging method, and the growth of the vegetation was calculated by the difference among surfaces of two consecutive surveys. Because of this, growth is refered by time intervals, called in this text by s1-s2, s2-s3 and s3-s4, which are related to the differences observed from survey 1 to 2, 2 to 3 and 3 to 4, respectively.

The model that describes the vertical plant development can be called hybrid, because is based on the logistic theoretical model (Renshaw, 1991), by replacing the growth rate by a linear combination of the variables kt and depth. Besides, it has an empirical characteristic, due to the calibration process of the weights, carried out using the interpolated data.

After elaborating a conceptual model, an equation which describes canopy height variation rate (*dh*) was formulated (Equation (1)), where *depth* is measured by echosounder, *kt* represents attenuation coefficient of the light penetration into the water, k_1 and k_2 are the weights of the linear combination, *h* is the canopy height and the *0.7* constant is the carrying capacity, obtained of the maximum value of canopy height found during the surveys.

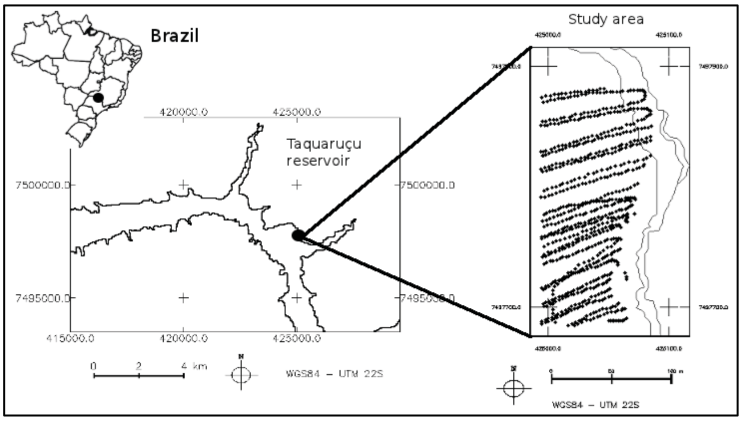

Figure 1: Study area

$$\frac{dh}{dt} = (depth * k_1 + kt * k_2) * \left(\frac{1-h}{0.7} \right) \tag{1}$$

To calibrate the coefficients k_1 and k_2 of Equation (1), the evolutionary approach of genetic algorithms has been used (Madsen, 2003). To validate the results, root mean square error (RMSE) metric was computed and the value obtained was 0.02067m, about 18% of the maximum growth value of the validation data.

In order to verify the uncertainty of the model, the two input parameters (*depth* and *kt*) and the initial condition of the dependent variable were represented by a normal distribution. The mean parameter of this distribution was specified using the measured value for the *kt* variable. The interpolated value for *depth* and the initial condition of the differential equation, which describes the canopy development, were used as mean of their distributions. The standard deviation of *kt* was calculated from the collected data in each date. The variance map generated by the kriging interpolation (3m resolution) produced the standard deviation of both, depth and canopy height.

Twelve sample elements used in validation phase were submitted to one hundred thousand simulations. These sample elements represent four georeferenced points changing in three time intervals. changing in three time intervals.

3. Preliminary results

The histograms calculated with simulated data for each validation element are in Figure 2. All points of interval s1-s2 and the point 2 of interval s3-s4 present less flattening, and it resulted in a lower uncertainty.

Figure 3 shows the RMSE boxplots of simulated data, by validation point. The median of simulated data is smaller than the RMSE found in validation model (lower horizontal line at 0.02m), in 8 of 12 points. The upper horizontal line represent the maximum residual found in validation phase (0.04 m). It shows that only the points p1, in s3-s4 and p4, in s1-s2, had the third quartile above 0.04 m.

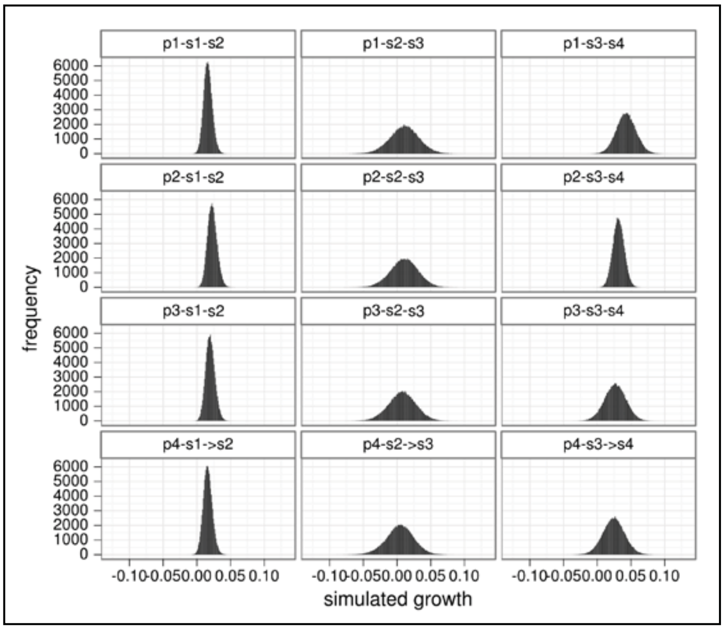

Figure 2: Simulated growth histograms

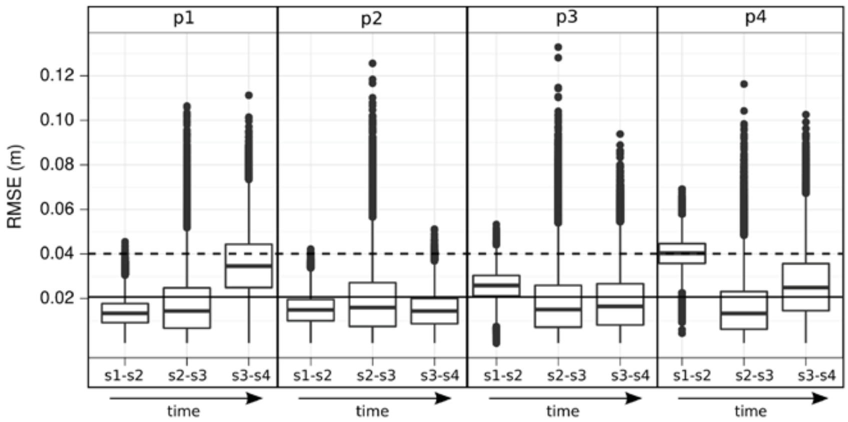

Figure 3: RMSE boxplot by point. Lower horizontal line points the RMSE and the upper horizontal line (dashed) marks the maximum residual.

The confidence interval of each point is presented in Figure 4. It was determined with the 5th and 95th percentiles of simulated data. In this graphic it is possible to compare the confidence intervals of stochastic results, for each point, in relation to the deterministic results and the observed values.

By analyzing the graphic, it is noted that 3 of 12 observed points of this plot are outside of the confidence interval: point p1, from s3-s4 and the points p3 and p4, from s1-s2 time interval.

We also noticed in Figure 4 that, concerned with interval confidence width, the lowest uncertainty is in the interval s1-s2 for all points. The point p2, of interval s3-s4, also has low uncertainty compared to the others of the same interval. In spatial terms, the point p2 presents the lowest uncertainty, given the confidence interval width in each time period. Besides, this is the point that presents the best result compared with observed data, since the temporal trend is well represented.

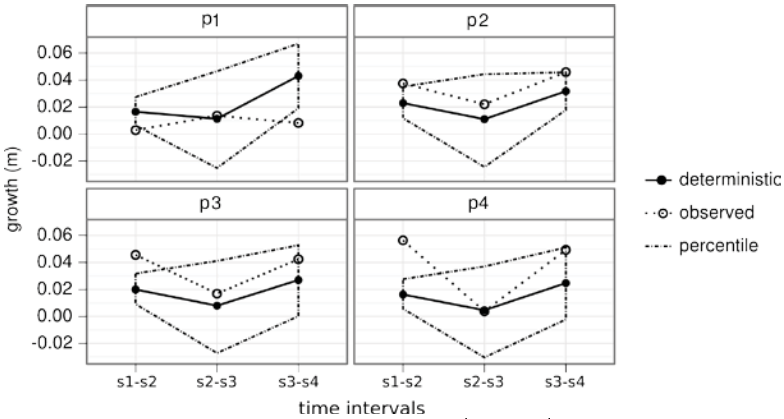

Figure 4: Confidence intervals of 5th and 95th percentiles

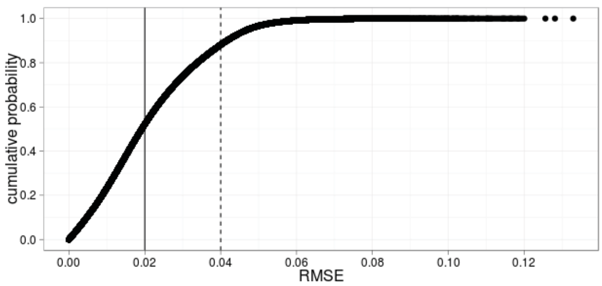

Figure 5: Cumulative distribution of RMSE

In Figure 5, the cumulative probability for RMSE of simulated data is showed. The vertical line in this plot marks the cumulative probability for 0.02m, about the same RMSE obtained in validation phase. The probability calculated based on simulations to this value was 0.52, indicating that about half of simulated data were above RMSE. When the the maximum residual was considered (0.04m), 88.1% of simulated data were below this maximum value.

3. Conclusion

With this analysis, we could realize that the uncertainty is distributed in a heterogeneous way, both in space and time. The s1-s2 interval presented the smallest confidence interval and the point p2 had the lowest uncertainty of the twelve points considered. The 88.1% percentage of points below maximum residual is considered reasonable.

Future works include the evaluation of a bigger amount of points in order to generate surfaces to represent uncertainty.

Acknowledgments

We are thankful to FCT/UNESP and to UTFPR by the support provided.

References

Heuvelink, G. B. M., Brown, J. D., Loon, E. E. Van. (2007), "A probabilistic framework for representing and simulating uncertain environmental variables." *International Journal of Geographical Information Science,* Vol. 21(5): 497-513.

Heuvelink, G.B.M., Burgers, S.L.G.E., Tiktak, A., Van Den Berg, F. (2010), "Uncertainty and stochastic sensitivity analysis of the GeoPEARL pesticide leaching model." *Geoderma,* Vol. 155 (3-4) : 186-192.

Madsen, H. 2003. "Parameter estimation in distributed hydrological catchment modelling using automatic calibration with multiple objectives." *Advances in Water Resources* 26 (2): 205-216.

Michelan, T. S., Thomaz, S. M., Mormul, R. P., Carvalho, P. (2010), "Effects of an exotic invasive macrophyte (tropical signalgrass) on native plant community composition, species richness and functional diversity." *Freshwater Biology* 55 (6):1315-1326.

Mulligan, M.; WainWright, J. (2004), *Environmental modelling: finding simplicity in complexity*, Wiley, London, U.K., 408p.

Refsgaard, J. C., van der Sluijs, J. P., Højberg, A. L., Vanrolleghem, P. A. (2007), "Uncertainty in the environmental modelling process – A framework and guidance." *Environmental Modelling & Software* 22 (11):1543-1556.

Renshaw, E. (1991), *Modelling Biological Populations in Space and Time*, Cambridge University Press, Cambridge,U.K., 402p.

Sabol, B. M., Melton, Jr., R. E., Chamberlain, R., Doering, P, Haunert, K. (2002), "Evaluation of a digital echo sounder system for detection of submersed aquatic vegetation." *Estuaries* 25 (1):133-141.

Sousa, W. T. Z., Thomaz, S. M., Murphy, K. J. (2010). "Response of native Egeria najas Planch. and invasive Hydrilla verticillata (L.f.) Royle to altered hydroecological regime in a subtropical river." *Aquatic Botany* 92 (1):40-48.

Valley, R., Drake, M., Anderson, C. (2005), "Evaluation of alternative interpolation techniques for the mapping of remotely-sensed submersed vegetation abundance." *Aquatic Botany* 81 (1): 13-25.

van Nes, E. H., Scheffer, M. (2003), "Alternative attractors may boost uncertainty and sensitivity in ecological models." *Ecological Modelling,* Vol. *159(2-3):*117-124.

Yeh, Anthony Gar-on, and Xia Li. 2003. Error Propagation and Model Uncertainties of Cellular Automata in Urban Simulation with GIS. In *7th International Conference on GeoComputation*, 1-16. United Kingdom: University of Southampton.