CORE
Metadata, citation and similar papers at core.ac.uk
Provided by Okayama University Scientific Achievement Repository

Application of the Particle Filter to COHERENS V2 model of Lake Säkylän Pyhäjärvi

Akiko Mano* and Takayuki Shuku†

*Ochanomizu University, 2-1-1, Ohtsuka, Bunkyo-ku, Tokyo 112-8610, Japan,
mano.akiko@ocha.ac.jp
† Graduate School of Environmental and Life Science, Okayama University, Japan
shuku@cc.okayama-u.ac.jp

Abstract:

The particle filter (PF, Gordon et al, 1993; Kitagawa, 1996) was applied to determine the optimal parameters for the sediment transport for the lake Säkylän Pyhäjärvi model with COHERENS Version 2. After having obtained the optimal parameters, time series of the TSM (Total Suspended Matters) concentrations at the automatic station were compared among the model result of (1) before- (2) after applying PF and (3) the monitored observation data. The model result with using the optimal parameters after applying PF was well-reproduced the observation data. The application of PF can be considered as a quite usable way to improve the model reproducibility.

1. Introduction

Our research was focused on suspended solid simulation in Lake Säkylän Pyhäjärvi (Fig. 1), the largest lake in southwestern Finland, with the COHERENS V2 model (COHERENS version 2(Luyten, 2011)). The aim of this study is to apply PF to the model to determine the optimal parameters related to sediment transport to improve the model reproducibility.
The lake has two inlets from Ylaneenjoki and Pyhäjoki rivers and one outlet from Eurajoki river. These inflow and outflow river water and meteorological conditions such as wind are considered to be the main driving forces of water circulation in the lake.

2. Data and materials

To put it plainly, the model itself and all data such as the boundary conditions, the initial conditions and the observation data were based on ones used in satellite data assimilation (Mano, 2013) as described below.

The boundary conditions at river mouths were used the following data. Time series of river discharge and TSM loading at each river mouth were extracted from Hertta-database of Finnish

Environmental Administration. Those of temperature were done from VEMALA, the water quality component of the Watershed Simulation and Forecasting System (Vehviläinen B, et al., 2005) of the Finnish Environment Institute. Meteorological data was obtained from Finnish Meteorological Institute which contains wind speed and direction, air temperature, humidity, cloud coverage and air pressure. The initial condition for the TSM concentration on the lake surface was converted from Turbidity of the satellite data on June 26 in 2009 constructed in Finnish Environment Institute. The observation data was monitored on the surface of the lake at the automatic station at the main basin of the lake by Finnish Environment Institute.
The parameters related to the sediment transport were used as shown in Table 1. PF was applied for these four parameters.

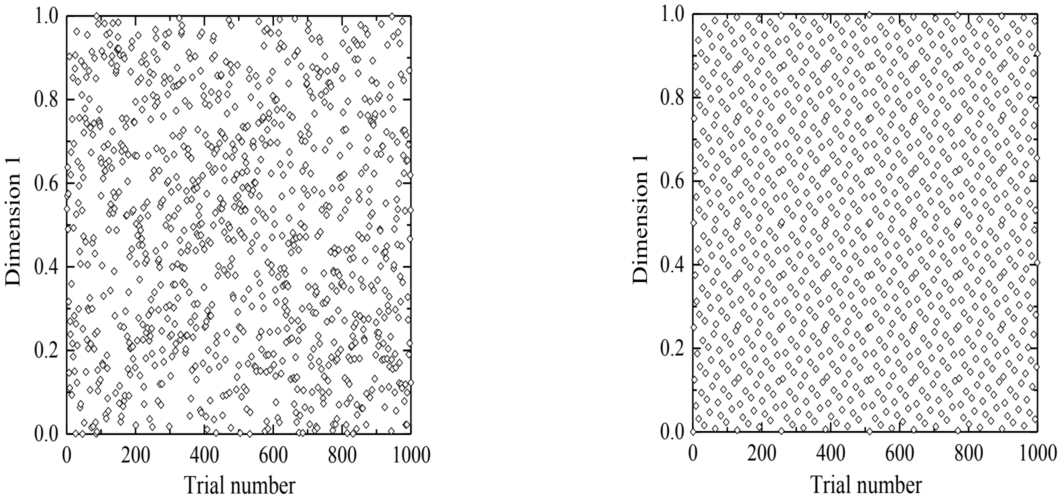
Table 1 Four parameters related to the sediment transport. Original value means one used in the model for satellite data assimilation.

Parameter name	Notes	Original value
bio_sinkrate	settling velocity [m/s]	-0.000 000 03
sedimentationrate	sedimentation rate (ex. 0.1 means that 10 % of sedimentation occurs par day.)	0.1
alphas	resuspension rate parameter [g/m2/s]	0.0025
rnsed	resuspension exponential factor	3.0

3. Methods

PF is a method for estimating probability density functions (PDFs) of state variables (Gordon, 1993; Kitagawa, 1996), and it can be applied to parameter identification of input parameters in numerical simulation. The considered sets of parameters called "particles" in PF. The simulation code of the PF was developed in (Shuku et al., 2012). Here, four parameters shown in Table 1 were regarded as "particles", and the observation data was regarded as the reference data. Preparing a number of sets of four parameters and performing a number of model calculations from COHERENS with using the parameters ("particles") were required for PF.

The PF usually requires about 1000 sets of parameters (particles) by a random function to find out the optimal ones. Nevertheless, by using not a random function but LDS (Low-discrepancy sequence (Tezuka, 1995)), the required sets will be reduced to 200 sets. This is because LDS generates more efficient numbers for PF to converge the results than a random function does. The images of the numbers using random function and LDS are shown in Fig. 1. Because of the time limitation, LDS was used for generating 200 particles. An application developed by one of the authors (Shuku) for generating numbers with LDS was used in this study, and the mean and the deviation for it are shown in Table 2.



(a) Random (b) LDS

Fig. 1 Image of the numbers using a random function(a) and LDS(b).

Table 2 Mean and deviation for an application with LDS for four parameters for generating 200 sets of parameters.

Parameter name	Mean	Deviation	Considered range
bio_sinkrate	-0.000 002 2	-0.000 001	-0.000 003 2 ~ -0.000 001 2
sedimentationrate	0.5	0.5	0.0 ~ 1.0
alphas	0.001 501	0.001 499	2.00E-06 ~ 0.003
rnsed	11. 5	10.505	1 ~ 22

Procedures were as follows. Some test calculations were performed with the parameters to guess the optimal parameters approximately in advance in order to narrow the range of values to be considered as particles. Here, only "bio_sinkrate" was focused because of the time limitation and that its sensitivity was the largest of the four as mentioned in the report (Mano, 2013). The tested values were in the hundredfold to hundredth range of the original value, -0.000 000 03. As the results of the test calculations, -0.000 002 20 was assumed to be close in value to the optimal one according to time series of the TSM concentration at the lake surface at the automatic station with the monitored observation data. An application for generating sets of parameters with LDS in an Excel file format was made by one of the authors (Shuku) for this study. 200 particles were generated with the setting shown in Table 1. The generated particles are shown in Fig. A-1 and listed in Table A-1. The parameters for each case were read from the file named "param.txt "(See Table A-1) automatically. This is because a directory's name of a case was made to be equivalent to a line number in "param.txt". Extracted time series of TSM concentration at the automatic station were gathered from 200 cases. With the data and observation data provide the optimal parameters with PF.

The calculation term was 13:00, June 26 ~ 12:00, July 6, the period when the CHL concentrations seem to be low according to satellite image and monitored the CHL concentration at the automatic station. For the PF application, calculation term is June 26, 13:00(close to the time when satellite data was obtained, 12h59m14s) ~ 12:00, June 29. Only the results from June 28 12:00 ~ June 29 12:00 were used for running the PF application. For comparing the before-after results, the ending time was extended until 12:00, July 6. The calculation term started from June 26 for all cases, the satellite data of the day was used as the initial value of the TSM concentration on the lake surface.

4. Results

Fig. 2 shows time series of the TSM concentration extracted from 200 cases and the observation data for 24 hours. The observation data is TSM concentration with no biomass converted from Turbidity monitored at the automatic station.

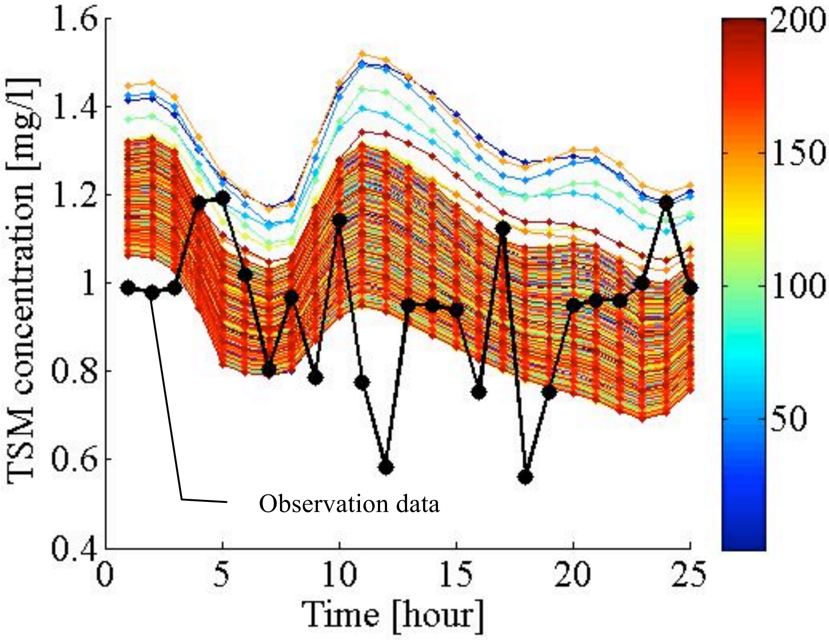


Fig. 2 Time series of TSM concentration on the lake surface at the automatic station. 200 cases of model result and an observation data.

PF was applied to the data of 200 cases and observation data. Fig. 3 shows the PF results for four parameters. The converged values show the optimal parameters. Thus the optimal values of parameters are shown in Table 3.

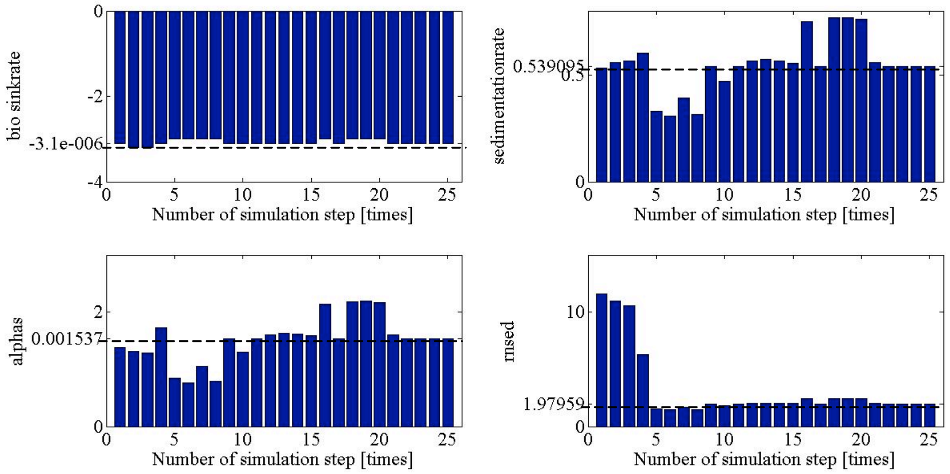


Fig. 3 PF result for four parameters

Table 3 Original values and optimal values for four parameters

Parameter name	Original value	Optimal value
bio_sinkrate	-0.000 000 03	-0.000 003 1
sedimentationrate	0.1	0.539 095
alphas	0.0025	0.001 537
rnsed	3. 0	1.979 59

5. Discussions and conclusions

Comparisons with the results before-after applying the optimal parameters were performed. Fig. 4 shows the time series of TSM concentration at the automatic station. “Before” and “After” are the results from calculations with using the original value and the optimal value in Table 3, respectively. As PF was applied to the results of “Before” from day 3 to 4 in Fig. 4 to obtain the optimal values for the four parameters, the results after day 4 shows the pure predictions. According to Fig. 4, the results of “After” are dramatically well-reproduced the observations. Thus it can be regarded that applying PF improved the model prediction in this case.

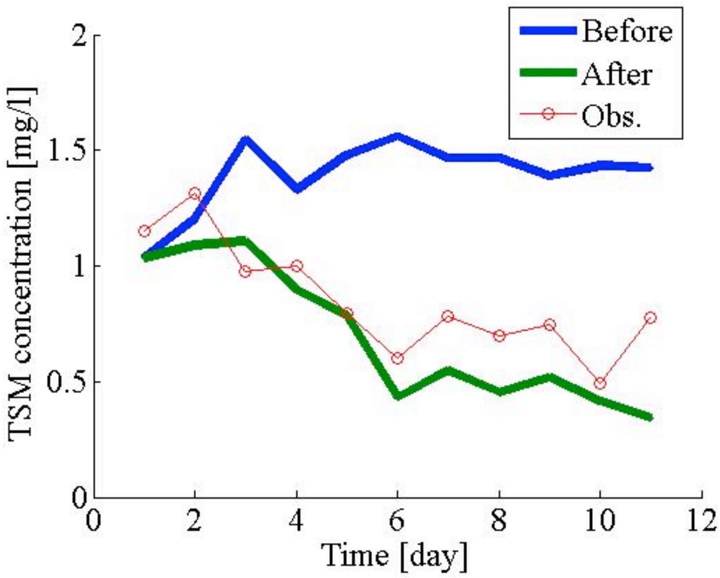


Fig. 4 Time series of TSM concentration at the automatic station.

Fig. 5 shows the comparison of vertical profile at the automatic station on 11th day of the calculation term. In the case of after applying optimal parameters, the trend, the deeper part has the larger TSM concentration, was reproduced as monitored at the automatic station (Mano, 2013).

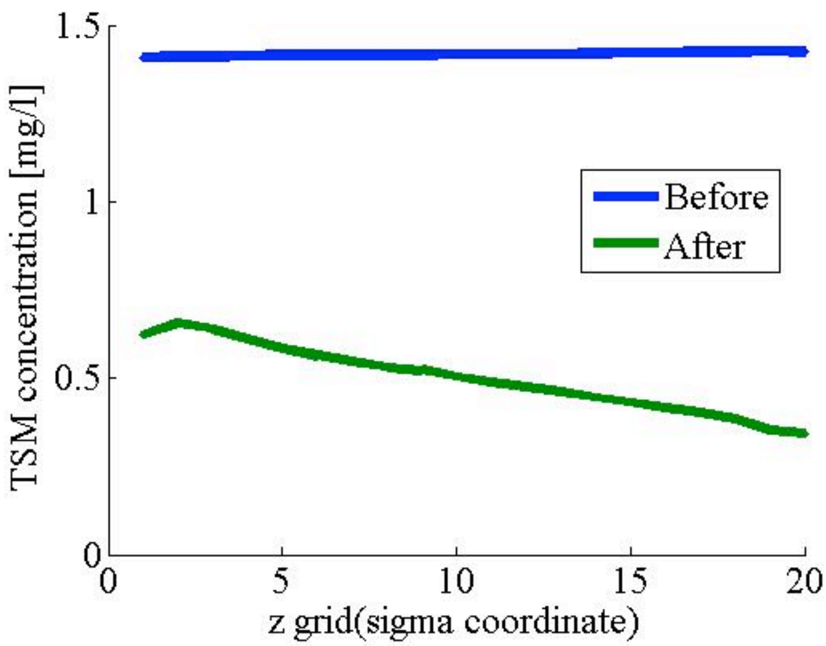


Fig. 5 Comparison of vertical profile of before-after applying the optimal parameters determined from PF on 11th day of the calculation term at the automatic station. z=20 and z=1 mean the lake surface and the bottom, respectively.

According to these results, to determine the optimal parameters with using PF seems to be very usable. This is because, not only time series at the surface but also the vertical profile, to apply PF increased the reproducibility of the model dramatically. However, to make sure the usability of PF, more tests would be required.

Two ideas for future works to make sure the usability of PF would be considered. One would be to apply PF to other calculation term, for example, not for 24 hours from June 28 12:00 but for 24 hours from on June 27 12:00. The other would be to apply PF from June 26 13:00 ~ July 6 12:00 (11 days) with using a coarser grid model. To apply PF for longer term could provide us more reliable clues for consider the usability of PF.

Acknowledgment

The authors would like to express sincere gratitude to Shin-ichi Nishimura at Okayama University and Janne Ropponen at SYKE for their guidance and kind help throughout this work.

References

[1] N. J. Gordon, D. J. Salmond, A. F. M. Smith, "Novel approach to nonlinear/non-Gaussian Bayesian state estimation", *IEE Proceedings-F*, **140**, 2, pp. 107-113, 1993.

[2] G. Kitagawa, "Monte Carlo filter and smoother for non-Gaussian nonlinear state space models", *J. Compt. Graph. Statist.,* **5**, 1, pp.1-25, 1996.

[3] P. Luyten, "COHERENS — A Coupled Hydrodynamical-Ecological Model for Regional and Shelf Seas", User Documentation. Version 2.0. *RBINS-MUMM Report, Royal Belgian Institute of Natural Sciences*, **1202** pp., 2011.

[4] A. Mano, "Report for river and lake modelling group: Assimilation of remote sensing data to COHERENS V2 model of lake Sakylan Pyhajarvi", **20** pp., 2013.

[5] T. Shuku, A. Murakami, S. Nishimura, K. Fujisawa and K. Nakamura, "Parameter identification for Cam-Clay model in partial loading model tests using the particle filter", *Soils and Foundations*, **52** 2, pp. 279-298, 2012.

[6] S. Tezuka, "Uniform random numbers: Theory and Practice", *Kluwer Academic Publishers*: Boston, 1995.