



Implementation of the particle filter and the merging particle filter in parallel computing systems

S. Nakano (1,2), G. Ueno (1,2), K. Nakamura (1), T. Higuchi (1,2)

(1) The Institute of Statistical Mathematics, Prediction and Knowledge Discovery Research Center, Minato-ku, Tokyo, Japan
(shiny@ism.ac.jp), (2) Japan Science and Technology Agency

The particle filter (PF) is an ensemble-based algorithm which is applicable to general data assimilation problems with nonlinear dynamical system models. The PF is a rather simple algorithm which provides an approximation of a posterior probability density function by resampling a forecast ensemble. However, the PF requires a large ensemble size in order to achieve sufficient accuracy of the estimation especially for high-dimensional models. This means that we need to run a simulation model extraordinarily many times. Parallel computing offers one potential solution which could enable us to use the sufficiently large ensemble size.

When we implement the PF in a parallel computing system, especially in a distributed computing system consisting of multiple nodes, one major problem is that resampling procedures requires much network traffic between nodes. This network traffic could be crucial especially when we use high-dimensional models. In this study, we assume that internal traffics in each node are much faster than inter-node traffics. Then, we consider a bi-level scheme which reduces inter-node network traffics. In this scheme, resampling procedures are performed locally in each node and inter-node communications are considered separately. An ensemble subset assigned to each node offers an approximation of a probability density function with a smaller ensemble size. If we assign a weight to each node according to an average of likelihoods in members of the ensemble subset, the whole ensemble also offers an approximation of a probability density function as a whole. Inter-node communications are determined by comparing the weights among the multiple nodes. We performed some experiments in which this bi-level scheme was applied to the 40-dimensional Lorenz 96 model. We then discuss the efficiency of this scheme in comparison with the normal scheme.

The network traffic problem would be also crucial when we use the merging particle filter (MPF) which allows us to save computational cost in comparison with the PF. We also consider a similar bi-level approach on the basis of the MPF and discuss the efficiency of it.