A computationally efficient procedure for combining ecological datasets by means of sequential consensus inference

David Conesa[1], Mario Figueira[1], Antonio López Quílez[1], Iosu Paradinas[2]

[1]Universitat de València/Valencia Bayesian Research Group; [2]AZTI - Centro de Investigación Marina y Alimentaria

In ecology and environmental sciences, combining diverse datasets has become an essential tool for managing the increasing complexity and volume of ecological data. However, as data complexity and volume grow, the computational demands of previously proposed models for data integration escalate, creating significant challenges for practical implementation. This study introduces a sequential consensus Bayesian inference procedure designed to offer the flexibility of integrated models while significantly reducing computational costs.

The method is based on sequentially updating some model parameters and hyperparameters, and combining information about random effects after the sequential procedure is complete. The implementation of the approach is provided through two different algorithms. The strengths, limitations, and practical use of the method are explained and discussed throughout the methodology and examples.

Finally, we demonstrate the method's performance using three different examples—one simulated and two with real ecological data—highlighting its strengths and limitations in practical ecological and environmental applications.

Keywords: Geostatistics, INLA, Preferential sampling, sequential inference, SPDE.