Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 119 (2017) 325–333

Procedia Computer Science

www.elsevier.com/locate/procedia

6th International Young Scientists Conference in HPC and Simulation, YSC 2017, 1 – 3 November 2017, Kotka, Finland

Spatially adaptive ensemble optimal interpolation of in-situ observations into numerical vector field models

Anton Gusarov[a,*], Anna Kalyuzhnaya[a], and Alexander Boukhanovsky[a]

[a]*ITMO University, St.Petersburg, Russia*

Abstract

Ensemble optimal interpolation (EnOI) is a well-known, having a relatively low computational cost yet powerful method for correcting outputs of numerical models in accordance with *in-situ* observations. Although, more advanced methods exist, e.g. variational analysis, this technique is widely used among different areas including meteorology and oceanography. Meteorological fields possess spatial inhomogeneity so as the quality of available measurements can vary between locations. This affects efficiency of the correction scheme and consequently motivates the need for adaptive choice of the correction parameters. In this paper we study how the ridge regularization influences the EnOI outcomes regarding statistical measures of fit between corrected and measured time series. Our numerical experiments for the wind field in southwestern Arctic region show that the optimal values of regularization parameter change from one group of observation points to another. We found also that these groups can be identified by clustering analysis based on estimated mutual covariances between time series of the observation points. As a result, we can adapt the EnOI scheme to each geographic sub-region and therefore to achieve more accurate correction results.

Peer-review under responsibility of the scientific committee of the 6th International Young Scientist conference in HPC and Simulation

Keywords: Ensemble optimal interpolation; ridge regularization; agglomerative clustering; wind field; Arctic.

1. Introduction

Even well-tuned numerical models of complex atmospheric phenomena are unable to reproduce the real process completely at all the scales in both time and space. Discrepancies between observations and model arise, indicating limits of the used numerical model. This can be justified by the multi-scale nature of the complex physical phenomena. The computer model can only simulate behavior of certain scale components while measurements contain all the constituent scales per moment as well as the noise.

Conceptually, with a sufficient amount of these model-observation differences, an additive correction might be inferred to actualize information provided by the modeling outcomes. This approach is generally known as Data

* Corresponding author. Tel.: +7-962-682-2415.
E-mail address: gusarovas@niuitmo.ru

1877-0509
Peer-review under responsibility of the scientific committee of the 6th International Young Scientist conference in HPC and Simulation
10.1016/j.procs.2017.11.191

Assimilation (DA), and refers to the family of methods that provide a fusion between modeled and observed data both in time and space also giving an estimated uncertainty of the derived result [9]. The goal of DA is to obtain more accurate estimate of the studied physical system state. This approach can be also considered as a problem of the field reconstruction based on a set of irregularly located corrections, i.e., differences between observations and modeling. Next, this reconstructed and mapped onto the original model regular grid field is to be added to the original field obtained from modeling eventually giving the result termed as the analysis. Further, in this paper we will understand the term DA as the model result post-correction, rather than the use of DA for refinement of numerical models initial conditions.

Statistical or optimal interpolation (OI) is a numerically efficient and widely used in meteorological operational practice method of DA. It produces the robust minimum variance estimate of the true atmospheric process state [8]. In the OI the relation between data in different locations is formulated essentially as a function of distance (including mean and spatially varying components) and than used in estimation of the gain matrix which is the key element of the whole correction scheme.

The OI applied for meteorology and oceanography has been widely reported in the literature. An example of buoy data interpolation into the wind fields model integrated with the wave model WAM Cycle-4 at the Arabian Sea is described in [12]. The root mean-square error (RMSE) reduction of 30 – 50% for the wave heights has been claimed. Satellite sea surface temperature (SST) became the subject of the OI in [6] and modification of OI hybridized with the simplified Kalman filter (KF) data assimilation in [10]. Climatological field (satellite SST) has been used as the background field. Results has been validated at North Sea – Baltic Sea region for the year 2001 so that RMSE is reported to be reduced from 1.13°C to 0.70°C. In [1] performance of OI for the case of wave data assimilation is reported in application to the South China Sea. The background error covariance matrix is constructed in this research using an isotropic error covariance function. Validation experiments are performed by the satellite altimeter data from Jason-1(2) and Envisat as well as using buoy observations for June – August 2010. The RMSE reduction is reported in this study of up to 45%.

To our knowledge, the previously published works do not describe a disciplined way of how their numerical implementations of OI handle spatial inhomogeneity of the fields and measurements to control quality of the assimilation product.

To examine this question, in this paper we demonstrate the adaptive modification of the ensemble optimal interpolation (EnOI) method used to post-correct outputs of the WRF (Weather Research and Forecasting) numerical model [15] of the wind field in southwestern part of the Arctic region having real observations at weather stations. We have formulated the correction scheme in a form of the ridge regression where the regularization parameter controls the estimated innovation weights. Next, we have explored how the optimal ridge parameter varies in space. Revealed spatial variability of the ridge parameter indicates that the spatially adaptive implementation of EnOI scheme based on the observation points clustering can be developed to obtain more accurate assimilation results.

The outline of this paper is as follows. In Section 2 we describe the multivariate EnOI method formulated as the ridge regression problem. Section 3 outlines the wind model and observations defining the numerical experiment setup. Then we continue this section by comparing two cases of numerical OI realizations with and without considering spatial variation of EnOI efficiency. Section 4 provides a summary of the obtained results.

2. Method for numerical model correction

Our goal is to apply the method of EnOI [3] to post-process the numerically simulated wind field in the way that it maximally fits measurements at weather stations. The basic principle behind this technique is to interpolate discrepancies between measurements and simulation over the numerical model grid points following the spatiotemporal consistency of the modeled field. However, its statistical properties are in general non-stationary, spatially inhomogeneous as well as the quality of measurements vary from station to station. Thus, such a correction can introduce undesirable artifacts (outliers) into the output.

We will show that such a harmful effect of interpolation might be controlled by introducing the direct ridge regularization with the tunable weights assigned to observation anomalies.