CLeaR: An Adaptive Continual Learning Framework for Regression Tasks

Yujiang He (yujiang.he@uni-kassel.de)
Universitat Kassel Fachbereich 16 Elektrotechnik/Informatik https://orcid.org/0000-0003-2159-7419

Bernhard Sick
Universitat Kassel Fachbereich 16 Elektrotechnik/Informatik

Research article

Keywords: Continual learning, Renewable energy forecasts, Regression tasks, Deep neural networks

Posted Date: March 4th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-251554/v1

Version of Record: A version of this preprint was published at AI Perspectives on July 16th, 2021. See the published version at https://doi.org/10.1186/s42467-021-00009-8.

Abstract

Catastrophic forgetting means that a trained neural network model gradually forgets the previously learned tasks when being retrained on new tasks. Overcoming the forgetting problem is a major problem in machine learning. Numerous continual learning algorithms are very successful in incremental learning of classification tasks, where new samples with their labels appear frequently. However, there is currently no research that addresses the catastrophic forgetting problem in regression tasks as far as we know. This problem has emerged as one of the primary constraints in some applications, such as renewable energy forecasts. This article clarifies problem-related definitions and proposes a new methodological framework that can forecast targets and update itself by means of continual learning. The framework consists of forecasting neural networks and buffers, which store newly collected data from a non-stationary data stream in an application. The changed probability distribution of the data stream, which the framework has identified, will be learned sequentially. The framework is called CLeaR (Continual Learning for Regression Tasks), where components can be flexibly customized for a specific application scenario. We design two sets of experiments to evaluate the CLeaR framework concerning fitting error (training), prediction error (test), and forgetting ratio. The first one is based on an artificial time series to explore how hyperparameters affect the CLeaR framework. The second one is designed with data collected from European wind farms to evaluate the CLeaR framework's performance in a real-world application. The experimental results demonstrate that the CLeaR framework can continually acquire knowledge in the data stream and improve the prediction accuracy. The article concludes with further research issues arising from requirements to extend the framework.

Full Text

This preprint is available for download as a PDF.

Figures

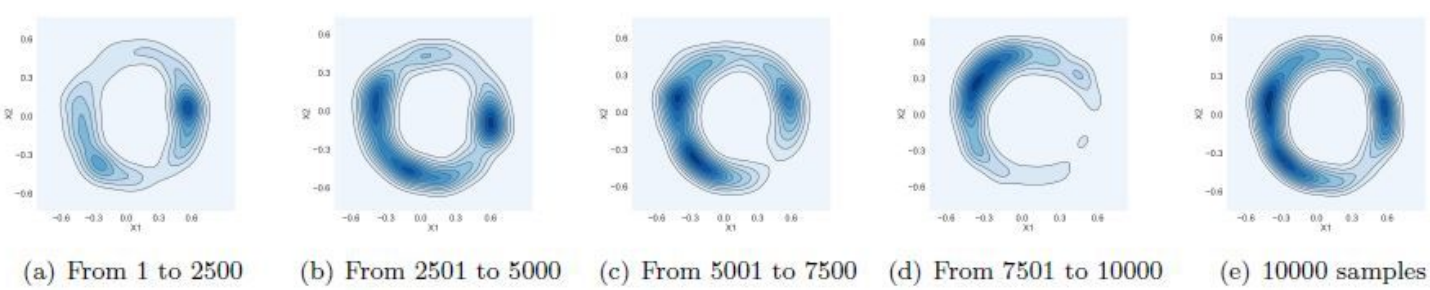


Figure 1

Distributions of the inputs change over time. We split 10000 samples into four sections sequentially over time, each of which has 2500 samples, and visualize their two-dimensional distributions from (a) to (d).

(e) illustrates the distribution of the same 10000 samples. Both axes in the sub-figures, X1 and X2, indicate the first two principal components extracted from seven-dimensional weather features.

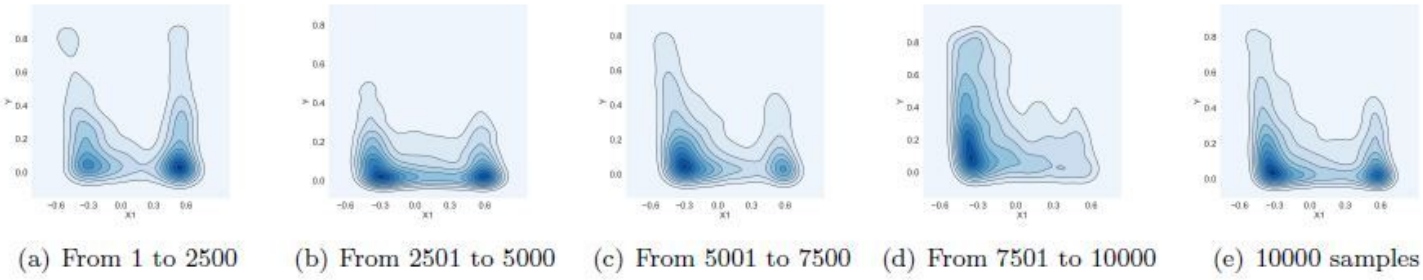


Figure 2

Distributions of the input and the output change over time. We split 10000 samples into four sections sequentially over time, each of which has 2500 samples, and visualize their two-dimensional distributions from (a) to (d). (e) illustrates the distribution of the same 10000 samples. Both axes in the sub-gures, X1 and Y, indicate the rst principal component extracted from seven-dimensional weather features using PCA and the power value, respectively.

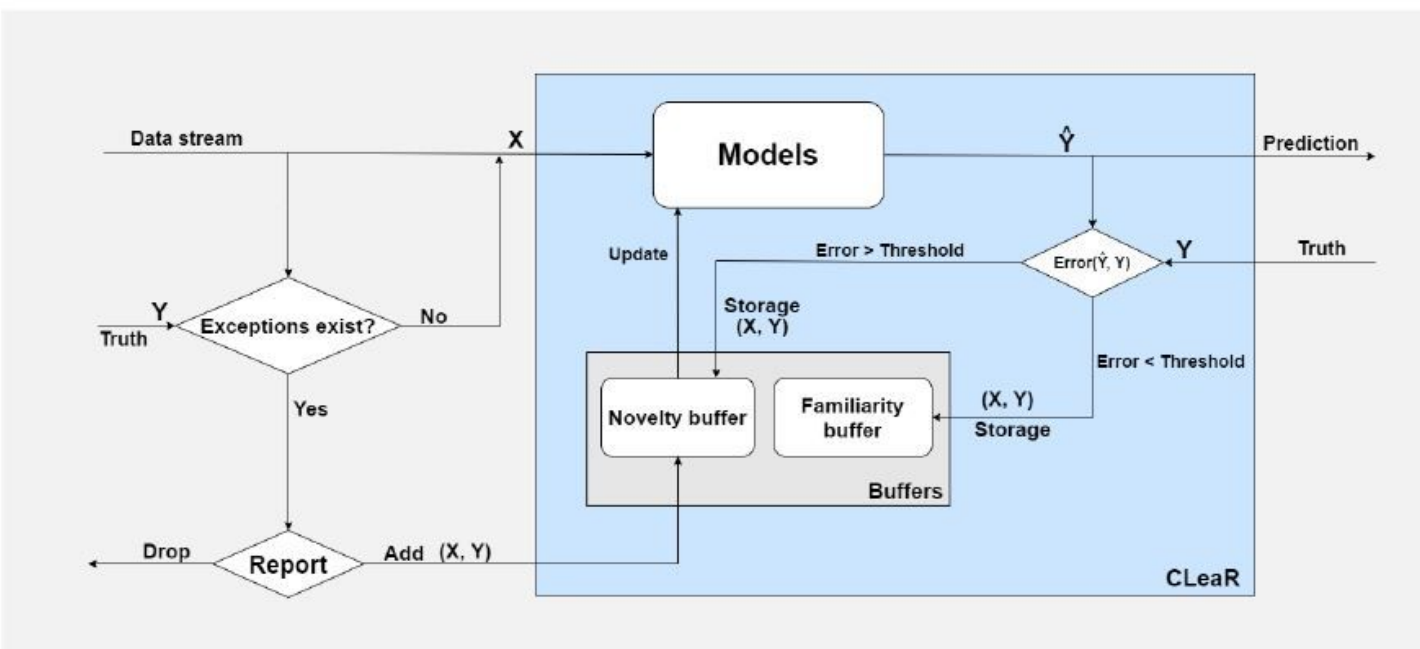


Figure 3

A general work ow of power forecasts comprises four phases: (1) reporting exceptions, (2) predicting targets, (3) storing data, and (4) updating models.

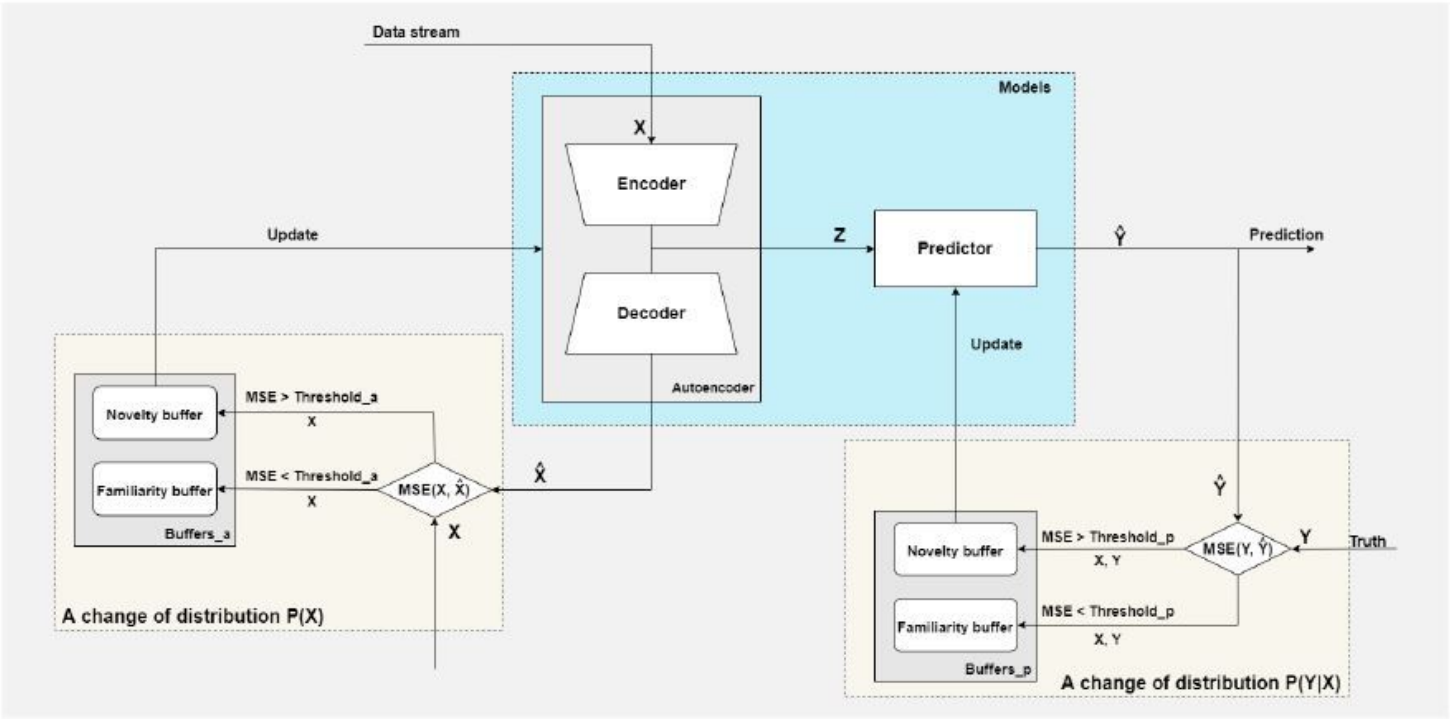


Figure 4

please see the manuscript file for the full caption

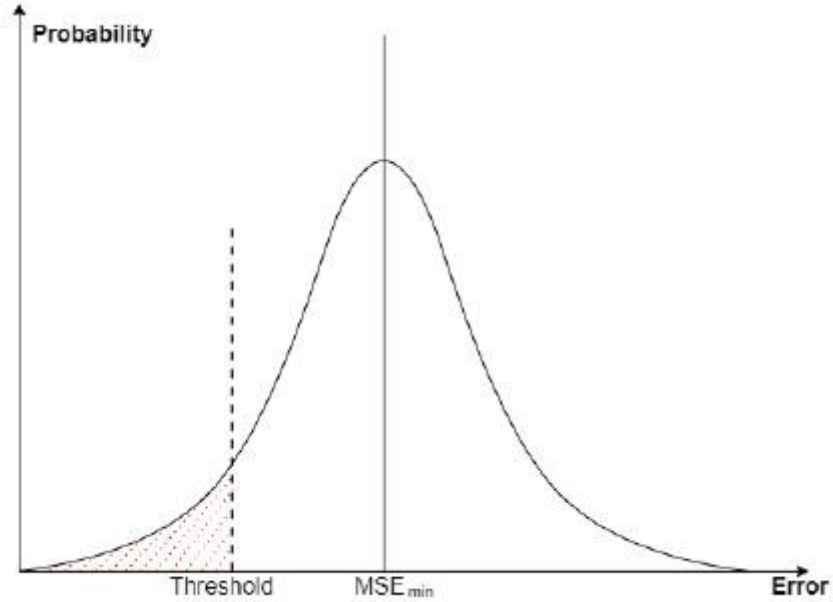
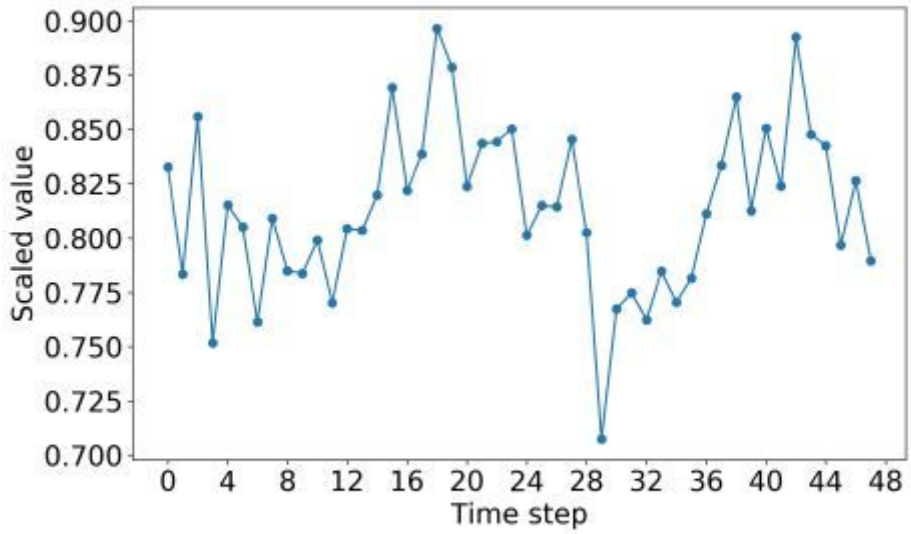
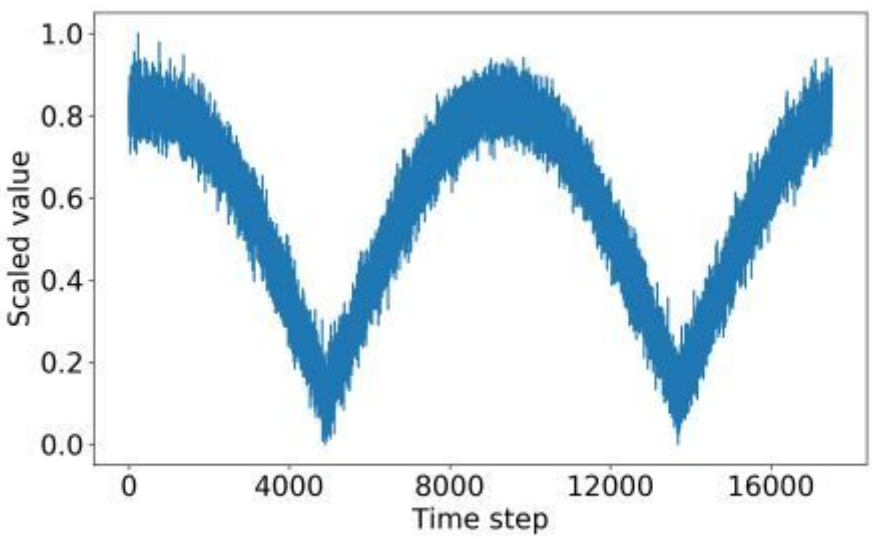


Figure 5

please see the manuscript file for the full caption



(a) $x_i^d(t)$ over 48 hours (2 days)



(b) $x_i^y(t)$ over 17520 hours (2 years)

Figure 6

The two figures illustrate the fluctuation of fifith dimension sequence over 48 hours (2 days) and 17520 hours (2 years), respectively.

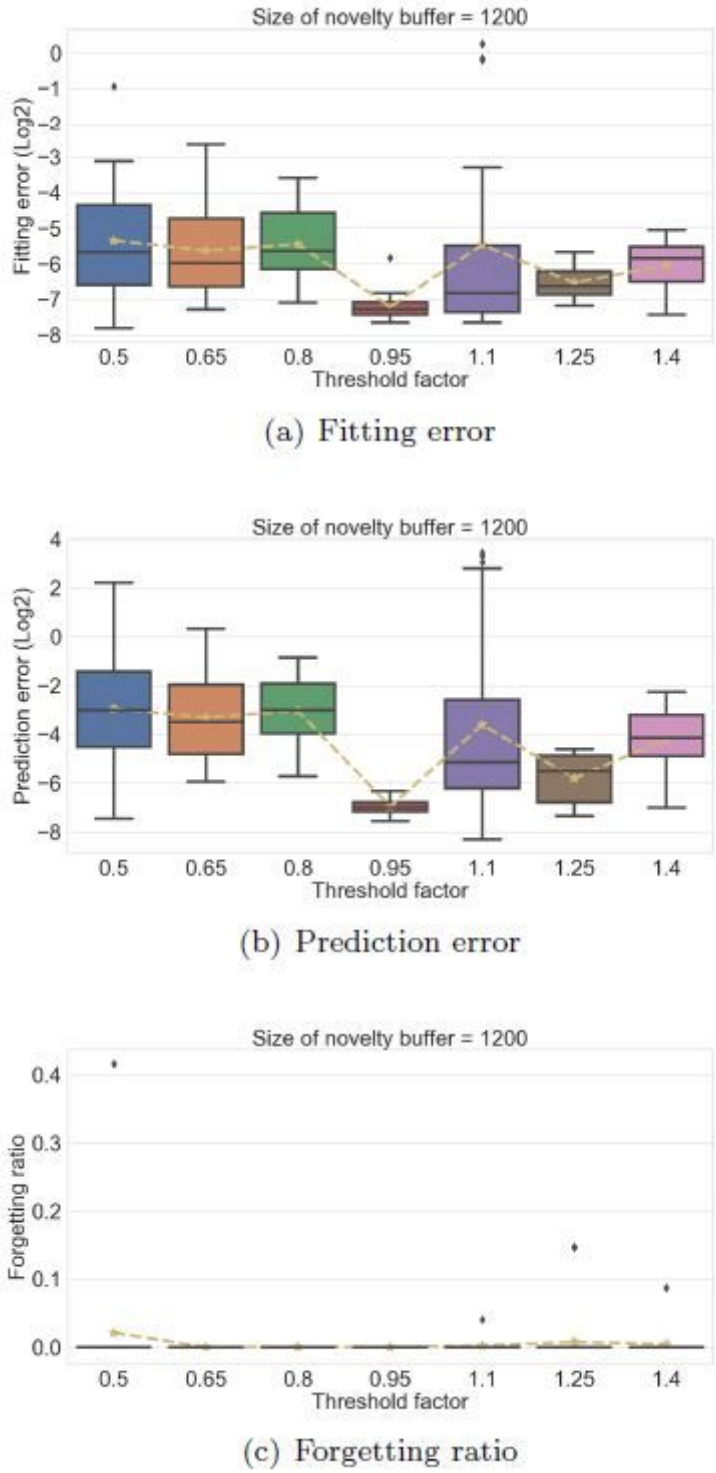


Figure 7

The influence of the threshold factor with the novelty buffer size fixed to 1200. Experiments are repeated 20 times with different initializations under the same setting. Black diamonds above boxes denote outliers, and the black lines in boxes denote the medians. Yellow stars connected by the dashed line indicate the mean values of 20 repetitions.

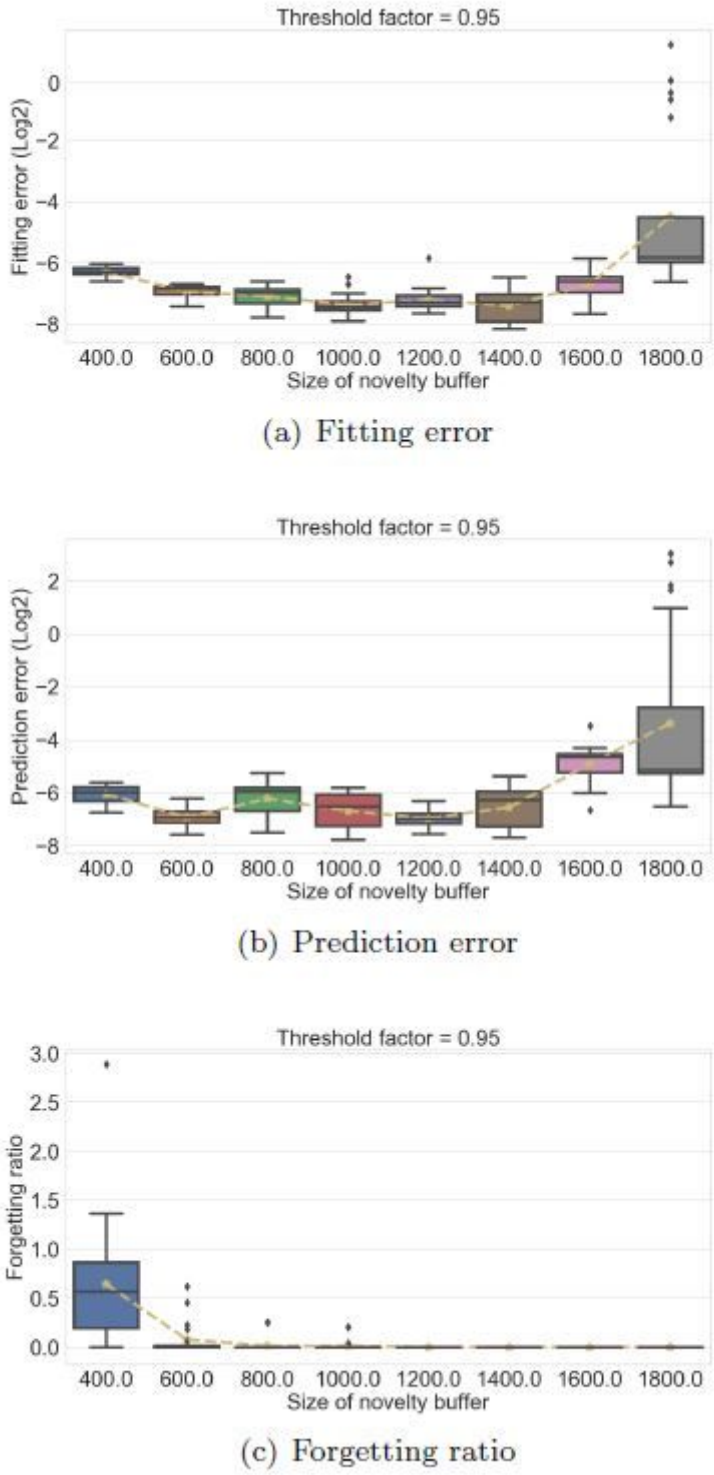


Figure 8

The influence of novelty buffer size with the threshold factor fixed to 0.95. Experiments are repeated 20 times with different initializations under the same setting. Black diamonds above boxes denote outliers, and the black lines in boxes denote the medians. Yellow stars connected by the dashed line indicate the mean values of 20 repetitions.

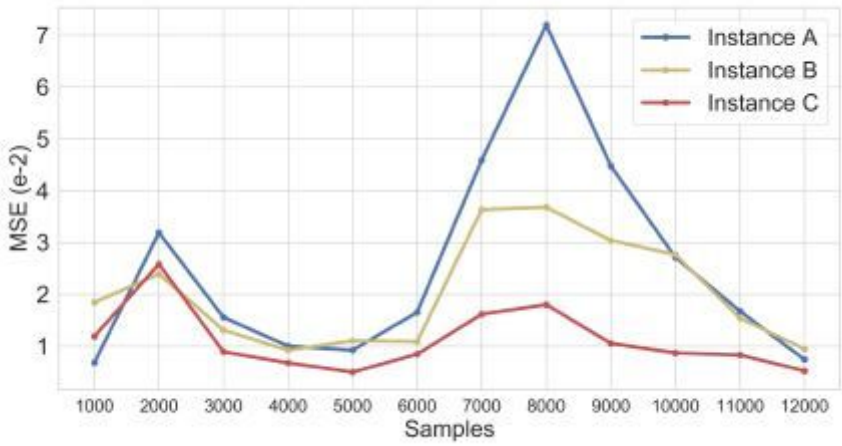


Figure 9

Errors of the three instances about one wind farm. Each point refers to the mean error over 1000 samples.