EMS Annual Meeting Abstracts
Vol. 20, EMS2023-495, 2023, updated on 21 Jul 2024
https://doi.org/10.5194/ems2023-495
EMS Annual Meeting 2023

Comparison of Daily Urban Temperature Forecast Performance by Traditional and Machine Learning-based Approaches

Yu-ting Kwok[1], Wan-hin Mok[2], Ming-chun Lam[1], and Wai-kin Wong[1]
[1]Hong Kong Observatory, Hong Kong, China
[2]Department of Mathematics, The Hong Kong University of Science and Technology, Hong Kong, China

Skillful location-specific weather forecasts are invaluable for the public and concerned communities in planning daily activities and preparedness actions for weather-related impacts. Despite continuous enhancements in numerical weather prediction (NWP) models, it remains challenging to capture the intricate variations in meteorological conditions within a city with highly heterogeneous landscape and diverse urban environment such as Hong Kong. In support of regional weather forecasting services, the Hong Kong Observatory (HKO) has developed the Objective Consensus Forecast (OCF) system for various weather elements, including a 9-day automatic forecast in daily maximum and minimum air temperatures (T_{max}, T_{min}). OCF is a past performance-weighted multi-model consensus forecast which employs Kalman Filter (KF) as an adaptive post-processing method for NWP forecasts at different weather stations. Though the performance of OCF has been largely satisfactory in the past decade, recent trends of machine learning (ML) methods in weather forecasting applications have driven efforts to improve the post-processing of NWP forecasts.

In this study, two ML-based approaches to improve location-specific daily T_{max} and T_{min} are presented. The first (OCF-CB) makes use of CatBoost, a ML technique based on gradient boosting on decision trees, to adjust the current OCF outputs. Besides the time-varied temperature prediction series and past errors against observations of an ensemble of global NWP models, other predictors that affect the diurnal variation of temperature (i.e. relative humidity, cloud cover, amount of rainfall, wind speed and direction) are used as model training data. The second (OCF-RF) aims to consider the effects of surface cover, urban geometry, and demographics to the screen-level temperature at each weather station. It uses a random forest model trained with different station environmental parameters, alongside other forecasted weather elements, to predict and correct the forecast errors of each NWP model. Biases of the corrected outputs are further reduced by a KF and multi-model consensus approach, similar to the current OCF system.

The daily T_{max} and T_{min} forecasted by both ML-based approaches are generally found to outperform the OCF for the verification period (2020-2022). OCF-CB shows more appreciable improvement in forecast performance in spring, as it can quickly readjust based on recent forecast errors and capture the changeable weather due to competence between the northeast monsoon

and southerly airstream. On the other hand, using urban parameters allows OCF-RF to better represent the higher urban temperatures in spring and summer, and the lower rural T_{min} in winter due to radiation cooling. Further work is underway to integrate the proposed ML-based temperature forecast approaches into the operational system for improving weather forecasting services. Applications in urban climate studies and gridded forecasts for urban heat risk assessments at refined spatial resolutions would also be explored.