Seasonal prediction of land surface air temperature in mid-latitudes: The value of machine learning to improve prediction skills

*Pascal Oettli[1], Ingo Richter[1], Masami Nonaka[1], Hiroyuki Koshiba[2], Yosuke Tokiya[2], Itsumi Hoshino[2], Swadhin Behera[1]

1. Japan Agency for Marine-Earth Science and Technology, 2. JERA

In the Kantō region, a strong correlation exists between summer and winter temperatures and the electric power demand. Identification of the sources of variability in this region is crucial for power demand forecasting few months ahead.

Due to the ocean memory effect, sea-surface temperature (SST) anomalies are considered as a source of predictability of surface air temperature (SAT) anomalies, with different time lag. In this way, conditions in the equatorial Pacific are known to influence SAT anomalies in Japan, both in summer and winter. Thus, predicting tropical SST conditions few months ahead helps to estimate the sign of the SAT anomalies in the Kantō region.

Over the years, SINTEX-F2 seasonal prediction system proved its ability to accurately predict SST anomalies few months in advance. Nevertheless, skills drastically drops when it comes to predict SAT anomalies in the mid-latitudes, particularly because the teleconnection patterns are not captured well by the dynamical system.

We propose a hybrid approach which utilizes SINTEX-F2 to provide predictors (i.e., SST) of the SAT in the Kantō region to a statistical modelling system, consisting of a set of three different machine learning algorithms. The statistical component is aimed to restore teleconnections between SST and SAT, particularly in mid-latitudes.

Results show that the hybrid model outperforms both the SINTEX-F2 prediction of SAT and the 2-month lead persistence. This is also true when prediction skill is assessed for each calendar month separately.

Despite the model's strong performance, there are also some limitations, such as the limited number of samples available to calibrate the statistical modelling and to infer reliable statistics to evaluate the output. Although the hybrid system outperforms the dynamical prediction system, it still only reproduces around 40% of the variability of the SAT in the Kantō region.

Keywords: Seasonal prediction, Statistical modelling, Hybrid prediction system

