ILMATIETEEN LAITOS
METEOROLOGISKA INSTITUTET
FINNISH METEOROLOGICAL INSTITUTE

KONE FOUNDATION

Promising subseasonal forecasting results based on machine learning

EGU General Assembly 2020

AS1.9 S2S prediction
D3205 EGU2020-19133

04 May 2020

Matti Kämäräinen
matti.kamarainen@fmi.fi

@KamarainenMatti

Short summary

KONE FOUNDATION

- LASSO regression and ensembling was used to forecast 2-weekly temperature and precipitation in Tropics and Northern Extratropics
- The method requires minimal amount of tuning and is effective in finding the most relevant predictors
- The achieved skill was high and comparable to the skill of the state-of-the-art dynamical model of ECMWF

Background

- Recently, machine learning methodology has been proposed as an alternative paradigm for making S2S predictions (Cohen et al., 2019) in addition to the traditional dynamical methods
- Kämäräinen et al. (2019) used skillfully LASSO* regression, PCA*, predictor lagging, and bagging* of predictor data to forecast seasonal temperatures in Europe based on reanalyses

Glossary*

LASSO	least absolute shrinkage and selection operator
PCA	principal component analysis
Bagging	bootstrap aggregating ≈ random sampling

References

Cohen, J. et al., 2019: S2S reboot: An argument for greater inclusion of machine learning in subseasonal to seasonal forecasts. Wiley Interdiscip. Rev. Clim. Chang., 10, 1–15, doi:10.1002/wcc.567.

Kämäräinen, M. et al., 2019: Statistical Learning Methods as a Basis for Skillful Seasonal Temperature Forecasts in Europe. J. Clim., 32, 5363–5379, doi:10.1175/JCLI-D-18-0765.1.

Method

KONE FOUNDATION

- Here the earlier method was revised to forecast the subseasonal time scale over the land areas of Tropics and Northern Extratropics
  - Variables (SST, Z, ...) from the 20CRv2c and NCEPv1 reanalyses were decomposed into their leading principal components to be used as predictor variables
  - 2-week means of temperature (T2M) and precipitation rate (PRAT) were the target variables
  - Each season, grid cell, and lead time was predicted using a separate LASSO ensemble with 50 members
    - Predictor selection and weighting in each ensemble member is automated and based on the internal cross-validation of LASSO
  - The output was bias-corrected with ERA-5 reanalysis

Validation metrics

KONE FOUNDATION

- Persistence, climatology, and reforecasts from the ECMWF dynamical model were used as reference forecasts
- ERA-5 was used as observations
- Anomaly correlation coefficient (ACC) and root mean squared error (RMS) were calculated from the LASSO model output, and from the reference forecasts
- ACC and RMS values were transformed to skill scores:

$$ACCS = \frac{ACC_{fcs}+1}{ACC_{ref}+1} - 1$$

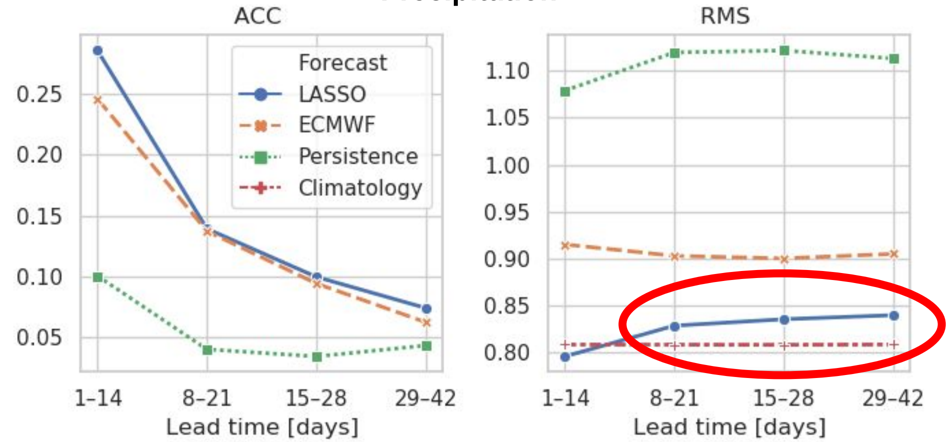
$$RMSS = 1 - \frac{RMS_{fcs}}{RMS_{ref}}$$

Results: grid aggregations

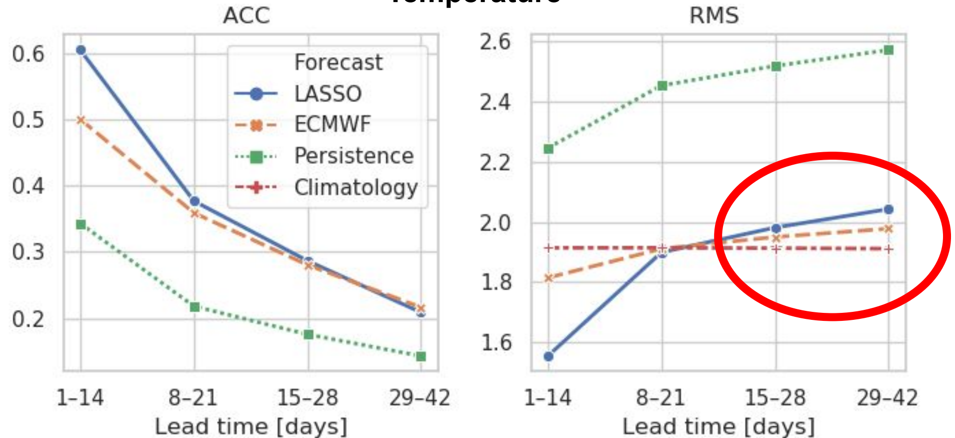
KONE FOUNDATION

- Out of all references, only climatology is slightly better than the mean of LASSO ensemble after 2–4 weeks in the RMS sense
- Otherwise LASSO ensemble performs better or similarly than the reference forecasts
- Note: ACC is not defined for climatology

Precipitation



Temperature



ILMATIETEEN LAITOS
METEOROLOGISKA INSTITUTET
FINNISH METEOROLOGICAL INSTITUTE

Results: T2M spatial

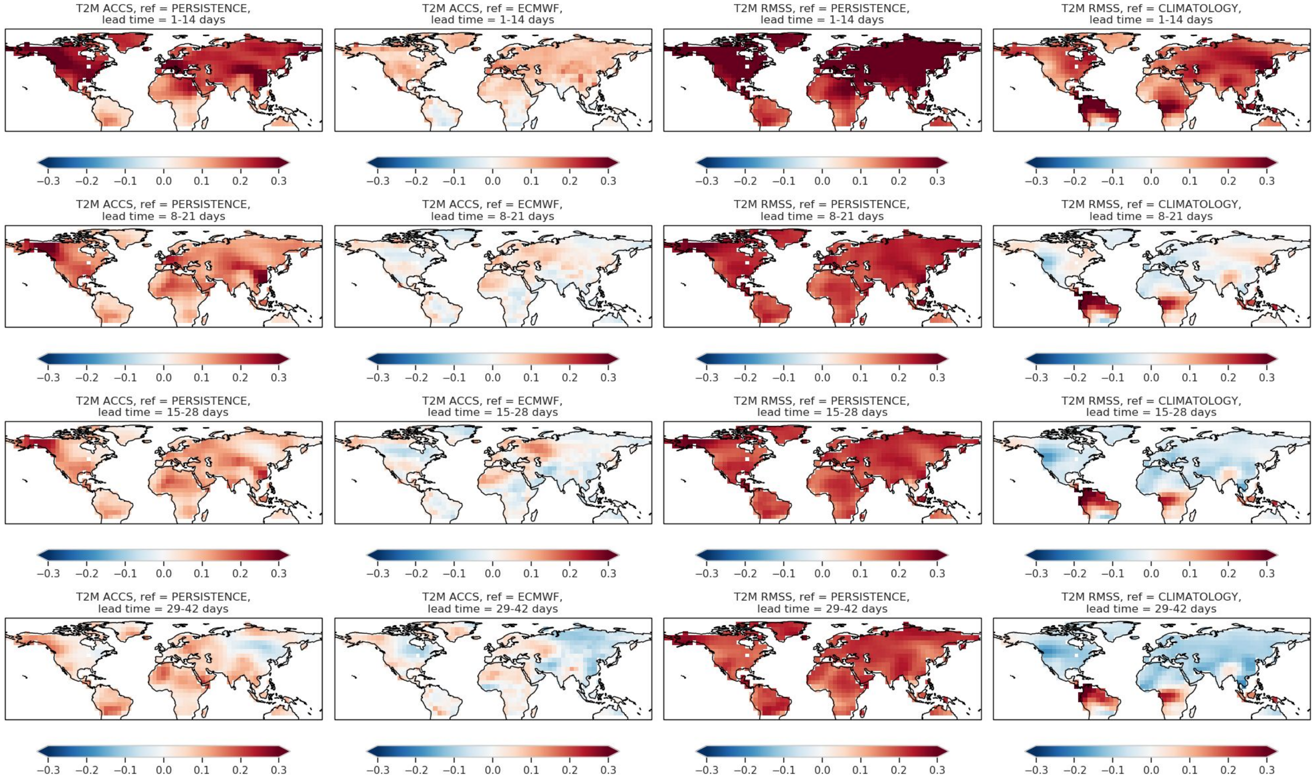
KONE FOUNDATION

Red shades = LASSO ensemble outperforms the reference
Blue shades = Reference outperforms the LASSO ensemble
Near-white shades = equally good performance

- Skill scores for temperature indicate that the LASSO ensemble works well in many regions of Tropics
- In Extratropics LASSO ensemble skill surpasses the ECMWF model skill eg. in Europe

References/scores

Lead times



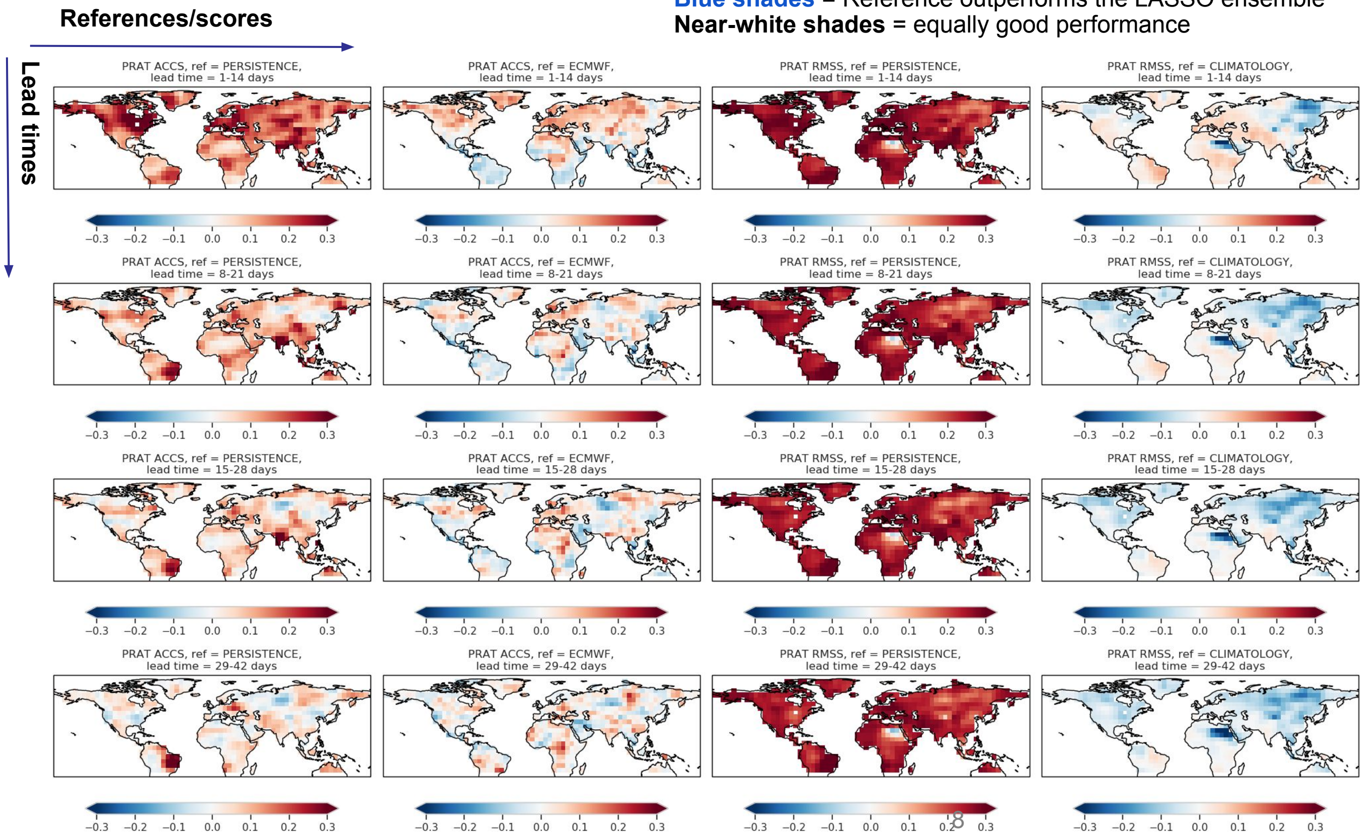
ILMATIETEEN LAITOS
METEOROLOGISKA INSTITUTET
FINNISH METEOROLOGICAL INSTITUTE

Results: PRAT spatial

KONE FOUNDATION

- Qualitatively similar results for precipitation

Red shades = LASSO ensemble outperforms the reference
Blue shades = Reference outperforms the LASSO ensemble
Near-white shades = equally good performance



ILMATIETEEN LAITOS
METEOROLOGISKA INSTITUTET
FINNISH METEOROLOGICAL INSTITUTE

Conclusion

- Machine learning should be used in subseasonal forecasting