

Machine learning techniques for spatial interpolation of the IASI Land Surface Temperature, Dew Point Temperature and water deficit index

F. Della Rocca^(1,2), I. De Feis⁽¹⁾, G. Liuzzi⁽³⁾, R. Giosa⁽³⁾, P. Pasquariello⁽³⁾, G. Masiello⁽³⁾, C. Serio⁽³⁾
(1) IAC, National Council of Research, (2) University of Naples Federico II, (3) University of Basilicata, Department of Engineering.

Context and motivation

Climate change has increased drought frequency, making their assessment crucial. In this context, satellite data can provide significant assistance due to its large spatial coverage and continuous data supply. To improve the knowledge of drought events, we studied two important climate variables (Land Surface Temperature and Dew Point Temperature) retrieved from the Infrared Atmospheric Sounder Interferometer (IASI), and the derived index Water Deficit Index (wdi)[1] that we have already proven to be useful in detecting drought events[2]. The IASI profiling capability for surface parameters, atmospheric temperature, and water vapor is leveraged in this new index. Unfortunately, infrared sensors such as IASI cannot penetrate thick cloud layers, so observations are blinded to surface emissions under cloudiness bringing sparse and not homogeneous distributed data over a given spatial region. For this reason, we exploited the capability of machine learning algorithms (Boosting, Random Forest and Neural Network) to convert IASI L2 data to a regular grid L3. Specifically, we trained a model that can predict the variables of interest over a 0.05° regular grid, using data from other sensors as a proxy together with vegetational products, soil indexes, and territorial and geographic information as covariates. The goodness of the proposed approaches has been tested over the Po Valley region, which experienced an intense drought in the last three years causing high vegetation and soil water stress, considering 9 years of IASI data (2015-2023). Overall, we found that these methods can yield good results, allowing a simultaneous retrieval of missing information over a regular grid and downscaling.

Material and Methods

LST, Tw are retrieved from IASI radiances (L1) with the ϕ -IASI package[3], an inverse radiative transfer code for the optimal estimation of the thermodynamic state of the atmosphere.

Spatio-temporal colocalization is used to build monthly training datasets. For each variable we used the following formulas:

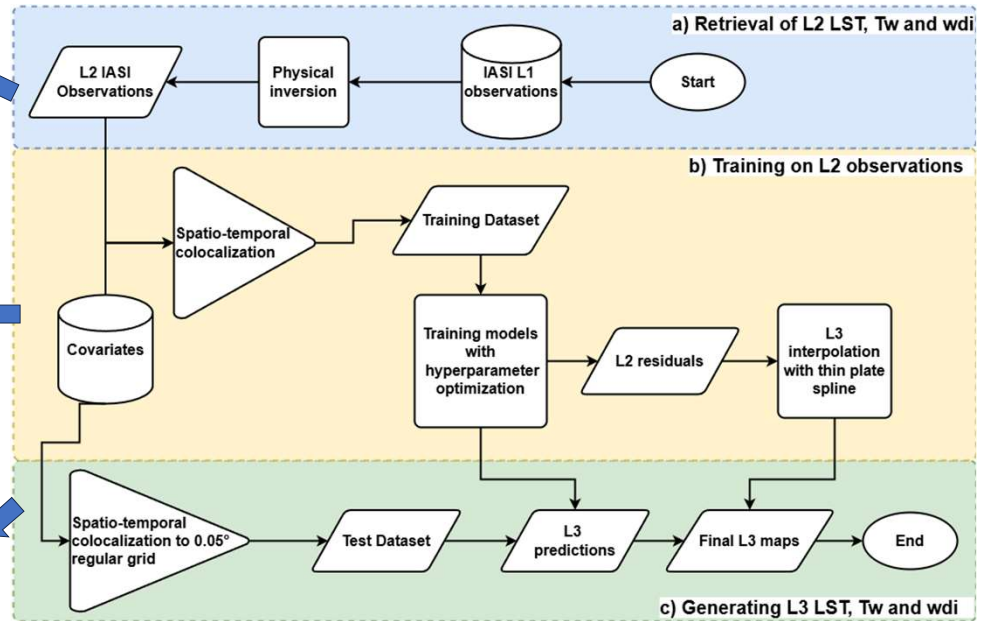
$$LST_{IASI} = f\left(\begin{matrix} lat, lon, dem, minutes, DOY, year, \\ CORINE, LST_{SEVIRI}, LST_{AVHRR}, LAI, FVC, NDVI \end{matrix}\right)$$

$$Tw_{IASI} = f\left(\begin{matrix} lat, lon, dem, minutes, DOY, year, Tw_{ERA5}, Wind_{ERA5}, \\ Evapo_{ERA5}, Tai_{ERA5}, Dowl_shortwave_{ERA5} \end{matrix}\right)$$

$$WDI_{IASI} = LST_{IASI} - Tw_{IASI}$$

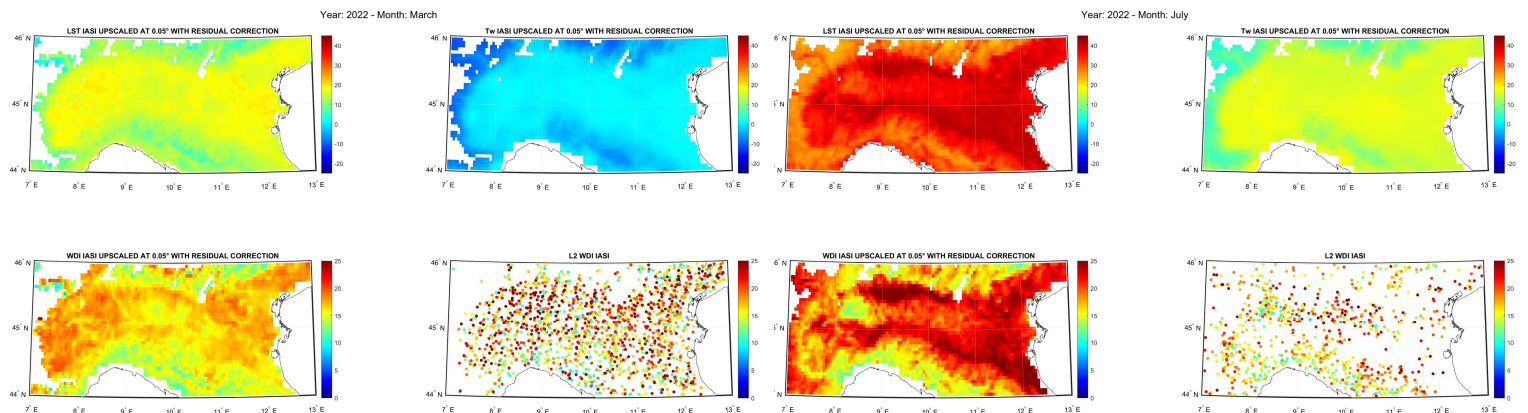
We tested three ML algorithms: Random Forest (RF), Gradient Boosting (GBT) and Neural Network (NN).

Using the same covariates, we built a test dataset by colocalizing them onto a 0.05° regular grid. This dataset is then used to predict the IASI variable. Finally, residual correction is performed by interpolating the residuals using thin plate spline.



Results

The three ML methods are compared using MAE and RMSE selecting the best-performing algorithm for each variable. Specifically, we used GBT for LST (MAE: 0.19, RMSE: 0.26) and RF for TW (MAE: 1.27, RMSE: 1.64). Here, we present two examples for the year 2022, when northern Italy experienced a severe drought, with the worst precipitation deficit since 1961.



Future Developments

- Testing other ML algorithms such as Gaussian Process Regression, Support Vector Machine, Graph Neural Network, Stacking Regression.
- Using higher resolution covariates to further improve the downscaling (such as the Era5 @2.2Km).
- Validate the results with in-situ data.