Application of Machine learning techniques to simulate the evaporative fraction in corn crops

Terenzio Zenone[1], Luca Vitale[1], Daniela Famulari[2], and Vincenzo Magliulo[1]
[1]National Research Council. Institute for Agricultural and Forestry Systems in the Mediterranean.
P.le Enrico Fermi 1, 80055, Portici (Italy) [2]National Research Council CNR. Institute for Bioeconomy, Via P. Gobetti, 101 - 40129 Bologna; Italy,

Background

- The evaporative fraction (EF) represents an important biophysical parameter reflecting the distribution of surface available energy.
- In this study, we have evaluated the performance of five Machine Learning (ML) classes of algorithms:
  the Linear regression (LR), Regression Tree (RT), Support vector Machine (SVM) Ensembles of Tree (ETs) and Gaussian process regression (GPR) to predict the EF.
- The adopted methodology consisted of three main steps that include: (i) selection of the EF predictors; (ii) comparison of the different classes of ML; (iii) application, cross-validation of the selected ML algorithms and comparison with the observed data.

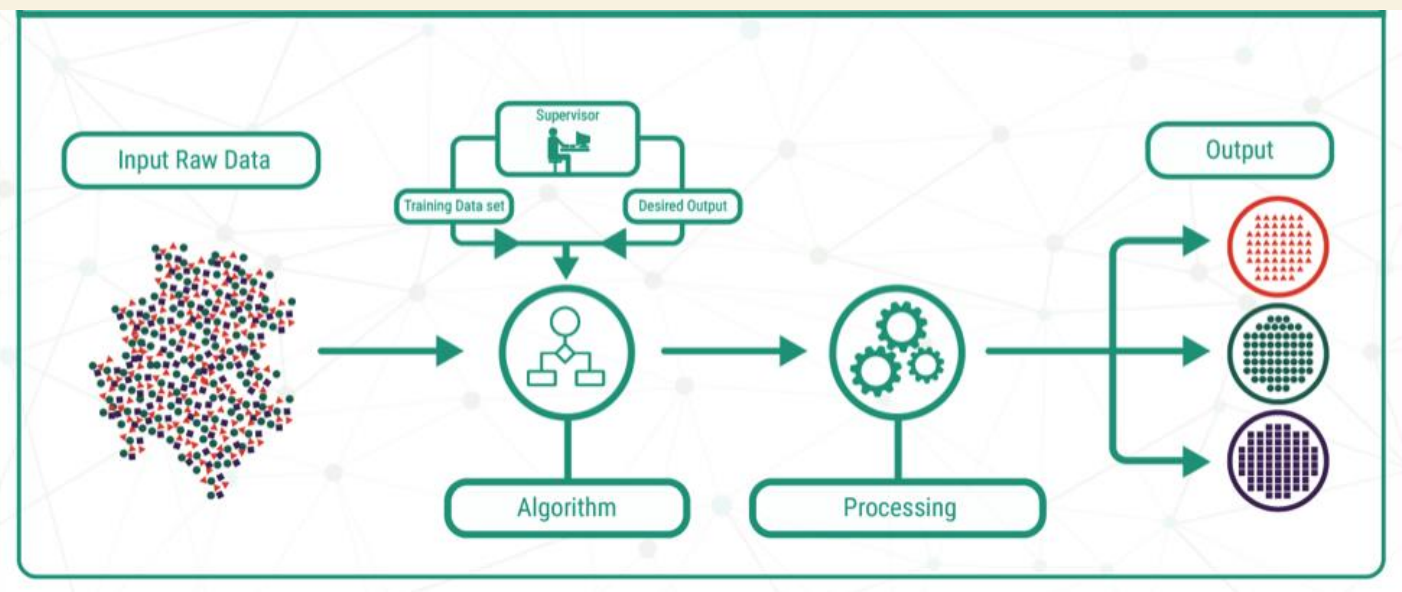

Figure 1 Schematic approach to a machine learning model

Souce: a Beginners Guide to Machine Learning Swason 2018 modified

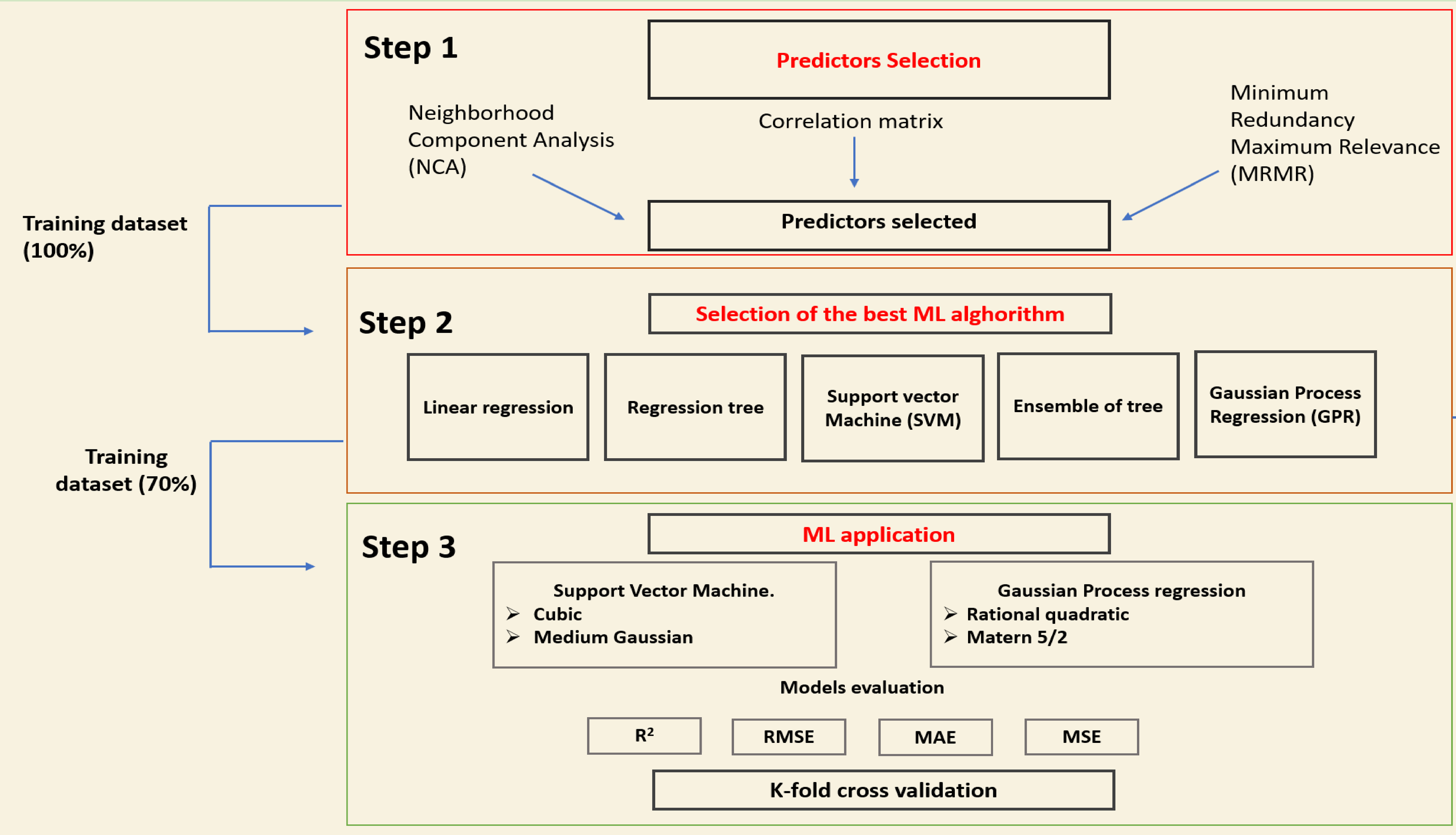

Figure 2. Flowchart of the methodology adopted to select the predictor variables and select the ML algorithms

Results

- The predictor selected were: CO_2 Net Ecosystem exchange, soil water content, Vapour pressure deficit, Air temperature, and soil heat flux (Fig 3).
- Support Vector Machine (SVM) and the Gaussian Processes Regression (GPR) were the best classes of ML at predicting the EF.
- Four different algorithms: Cubic SVM, Medium Gaussian SVM, the Matern 5/2 GPR, and the Rational quadratic GPR were used.
- The comparison between observed and predicted EF in all 4 algorithms, during the training phase (Fig 4) were within the 95% confidence interval: the R^2 value between observed and predicted EF was 0.76 (RMSE 0.05) for the Medium Gaussian SVM, 0.99 (RMSE 0.01) for the Rational Quadratic GPR, 0.94 (RMSE 0.02) for the Mater 5/2 GPR, and 0.83 (RMSE 0.05) for the Cubic SVM algorithms.
- Similar results were obtained during the testing phase Fig 5). The results of the cross-validation analysis indicate that the R^2 values obtained between all iterations for each of the 4 adopted ML algorithms were basically constant, confirming the ability of ML as a tool to predict EF.

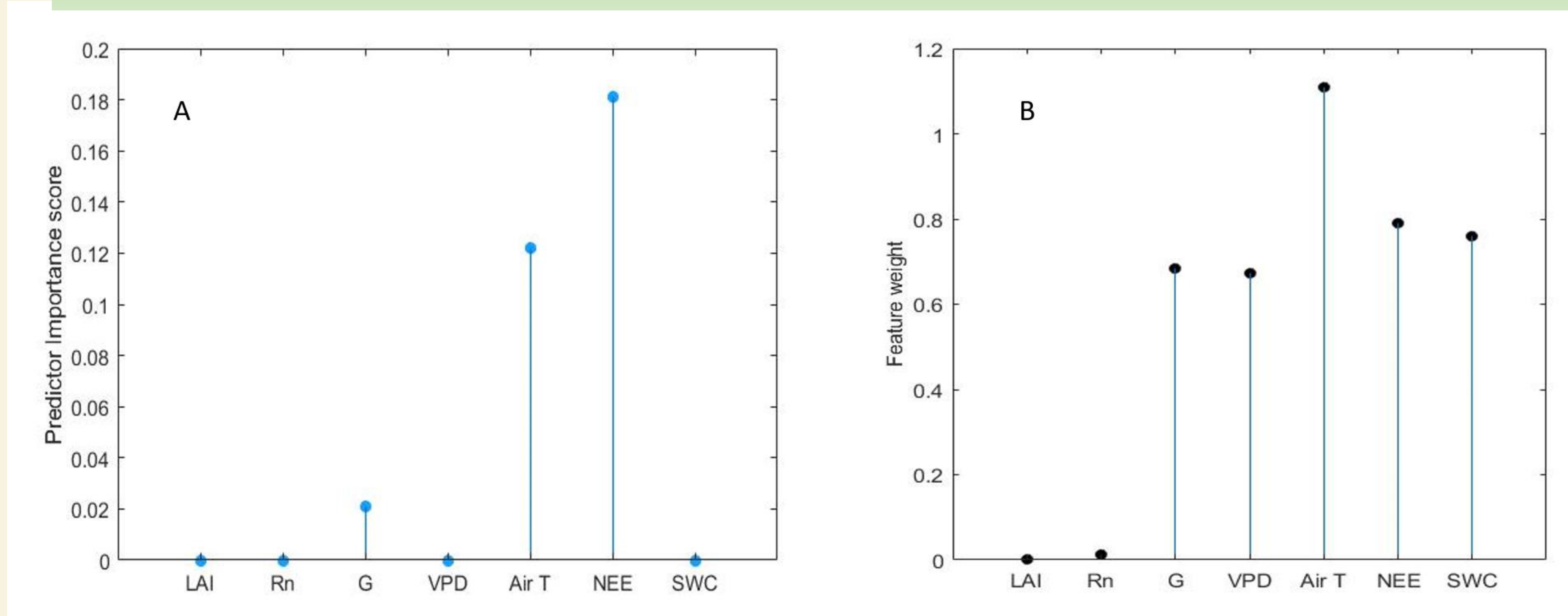

Figure 3 Results of the feature selection analysis by NCA and MRMR algorithms.

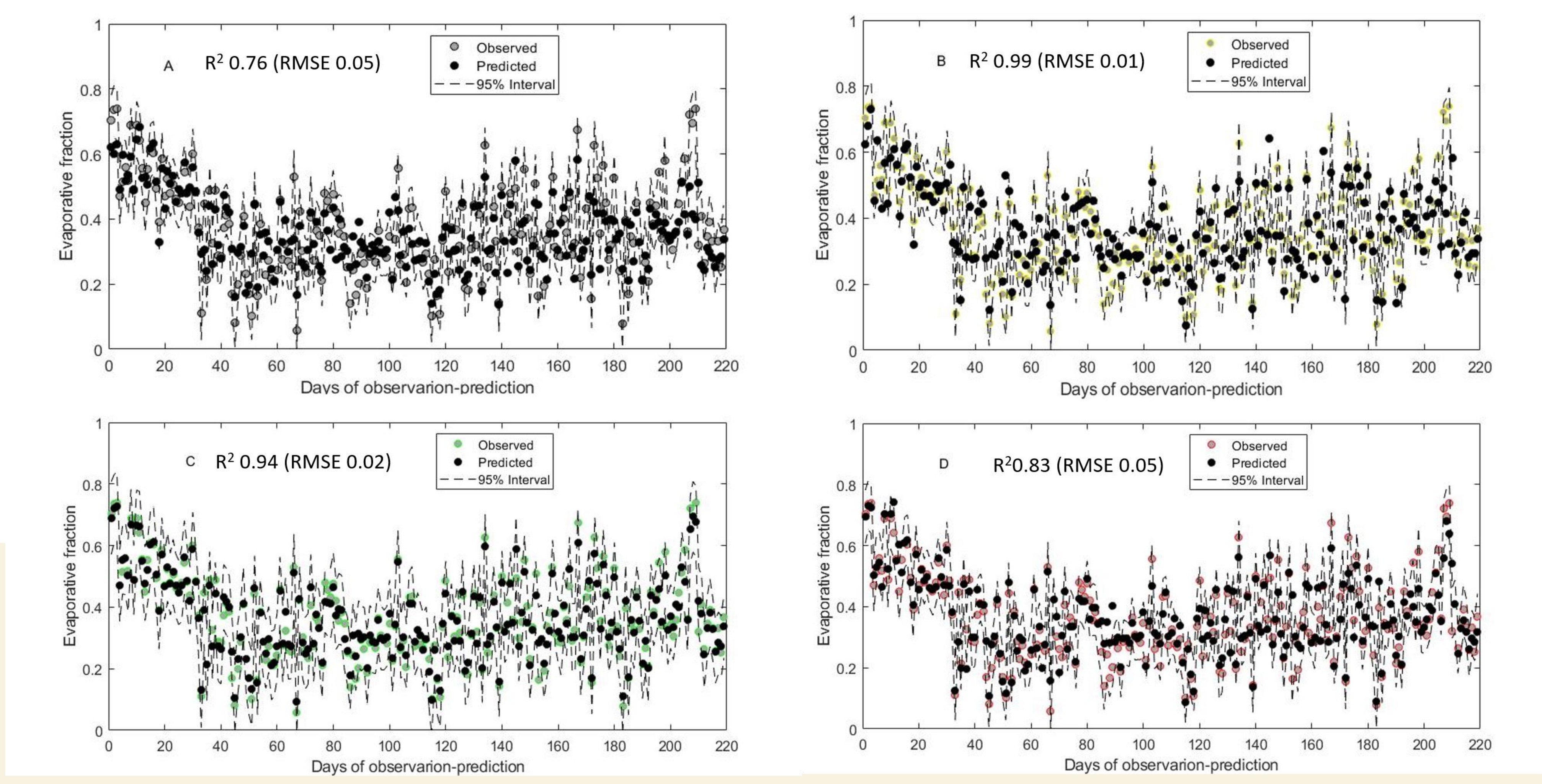

Figure 4. Comparison between observed and predicted daily EF, during the training phase, of the four different approaches: SVM polynomial model (A), GRP rational quadratic (B), GRP Matern 5/2 (C) and SVM cubic Gaussian (D). Coloured circles represent the observed values, black circles the predicted values, and the dashed lines represent the 95 % confidence interval.

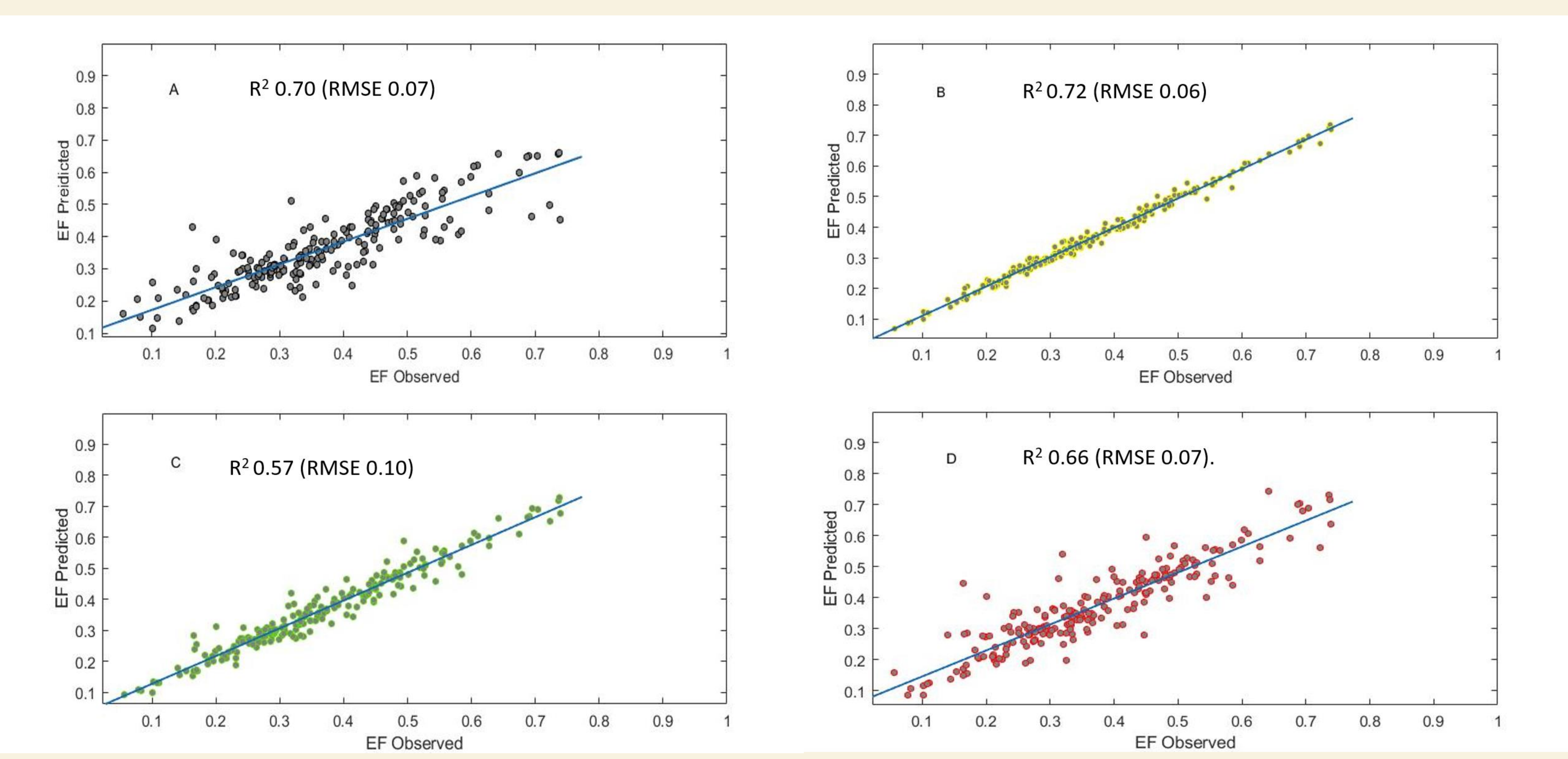

Figure 5. Scatter plot between observed and predicted EF, during the testing phase of the four different approaches: the SVM polynomial model (A), GRP rational quadratic (B), GRP Matern 5/2 (C), and SVM cubic Gaussian (D).

Conclusions

From this study, daily EF predictions have been reliably derived for a corn crop in a Mediterranean region using ML algorithms.

- The application to other geographical areas and crops requires further improvements, applying model training based on diverse data sources from different soils, climate, cropping systems and agronomic managements.
- Our results show that Support Vector Machine and Gaussian Process Regression algorithms are able, with limited input measurement data, to explain a substantial portion of the EF variability.
- The performance of the tested ML algorithms has proven to be comparable to the conventional remote sensing-based models and can be used when the sky conditions are not suitable for remote sensing observations. In addition, ML algorithms facilitate the interpretation of interactions between the predictors and the EF. Our results suggest also that in principle the integration of ML algorithms with remote sensing-based models could be an opportunity to improve the predictability of EF.

ICOS INTEGRATED CARBON OBSERVATION SYSTEM

Italian National Research Council

CNR ISAFOM

Contact information: Terenzio Zenone CNR ISaFom. Terenzio.zenone@isafom.cnr.it
Office Phone 06.4993 27525 Mobile +39 335 8369630

Source: Zenone et al 2022. Application of machine learning techniques to simulate the evaporative fraction and its relationship with environmental variables in corn crops. Ecological process (in press).
https://ecologicalprocesses.springeropen.com/