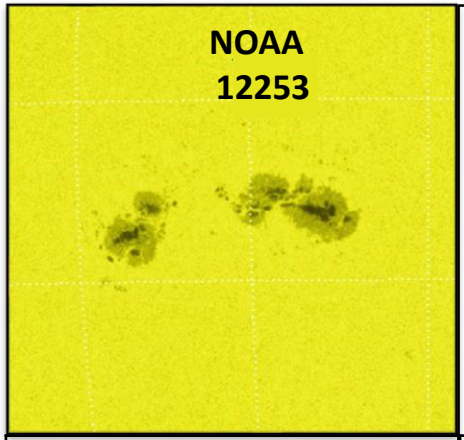
Sunspot Classifications & Solar Flare Prediction

Does machine learning improve upon Poisson-based prediction models?

Aoife E. McCloskey[1], Peter T. Gallagher[2], D. Shaun Bloomfield[3]

DLR Institute for Solar-Terrestrial Physics[1] | Dublin Institute for Advanced Studies[2] | Northumbria University[3]

Contact: aoife.mccloskey@dlr.de



AR Database

Sources

NOAA Solar Region Summary
- Zpc, MtWil, Location, Area, ..

NOAA Solar Event Reports
- AR Flares (C-, M- & X-Class)

Time Range

1988-2019 (SC22, 23 &24)

McIntosh Class (Zpc)			NOAA Active Regions		
Day 1	Day 2	Day 3	SC22	SC23	SC24
Dai	Eac	Ekc	2,708	3,063	1,763

Poisson-based Probability Forecast

$$P(f) = 1 - e^{-\lambda}$$

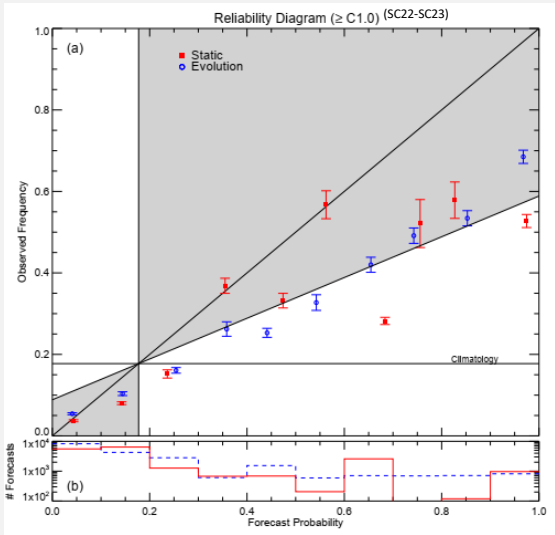
- **Poisson-Static** "snapshot" *Gallagher et al. 2002*
- **Poisson- Evolution** 24-hr change *McCloskey et al. 2018*

Machine Learning

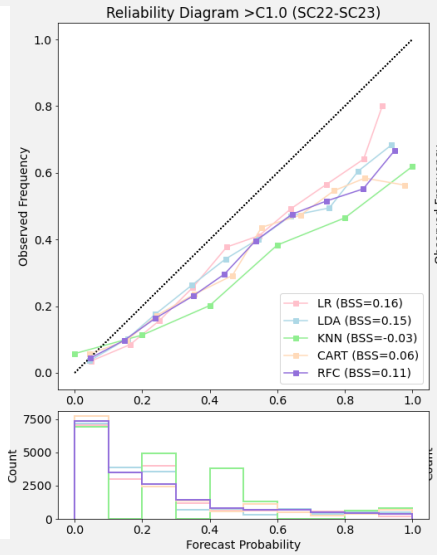
Logistic Regression (LR)
Lin. Disc. Analysis (LDA)
K-Nearest Neighbour (KNN)
Regression Tree (CART)
Random Forest (RFC)

→ 24-hr Forecast Probabilities [0 -> 1]

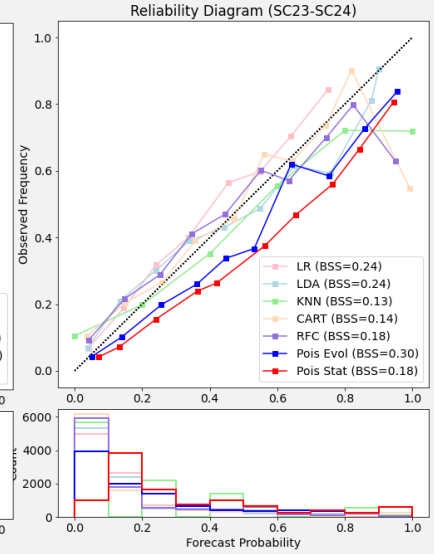
→ 24-hr Classification [Yes/No]



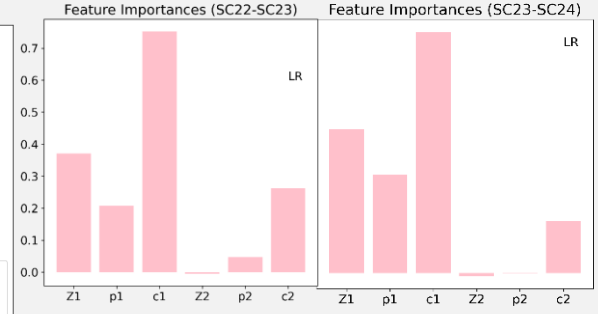
BSS_{stat} = -0.09 | BSS_{evol} = 0.09



BSS_{LR} = 0.16



BSS_{evol} = 0.30 | BSS_{LR} = 0.24

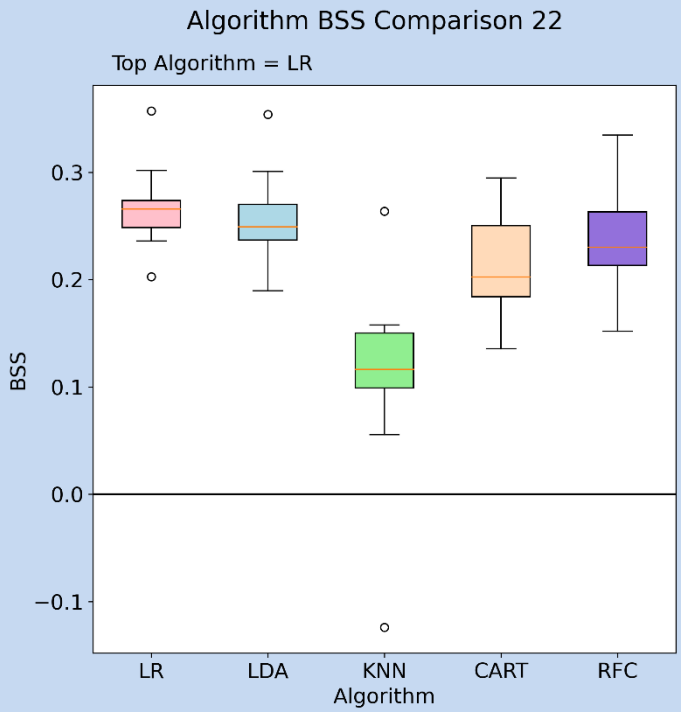


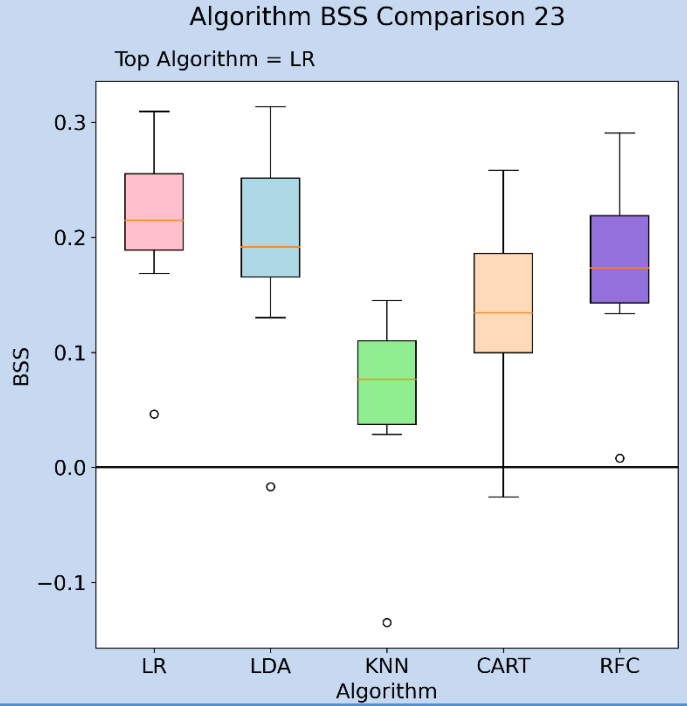
Key Results:

- SC22 -> SC23 : all **over-forecast**
- SC23->SC24
  Poisson models **over-forecast**
  ML models more **"reliable"**
- Feature Importance – **Compactness** "Mag. flux filling"

Brier Skill Scores

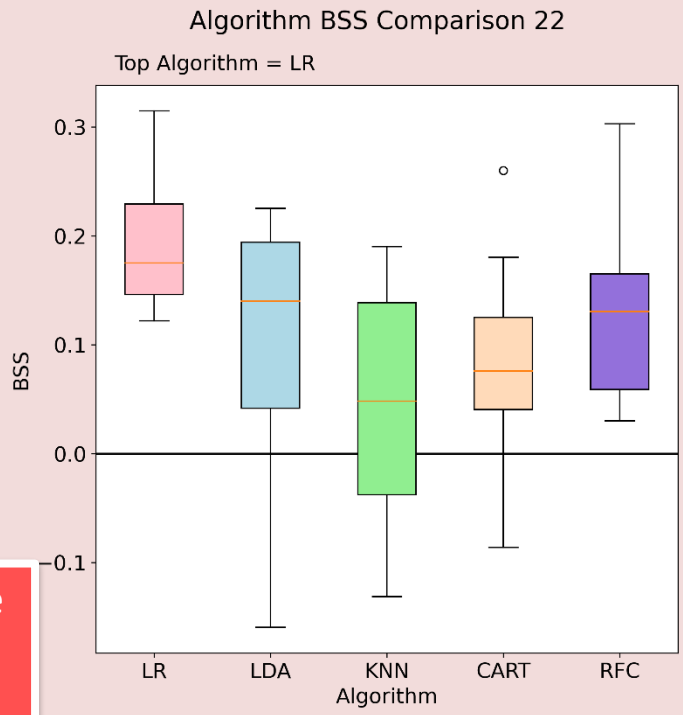
>C1.0 Flares

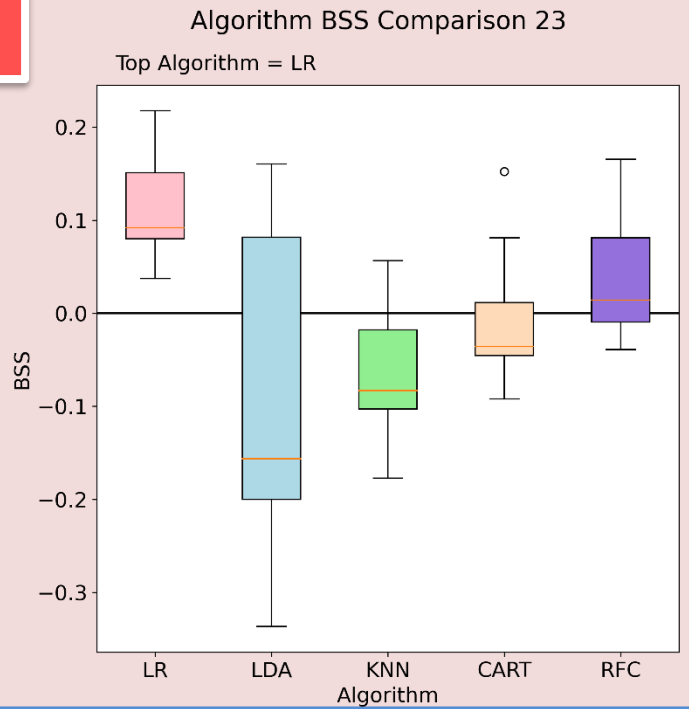




How well do the ML models perform probabilistically?

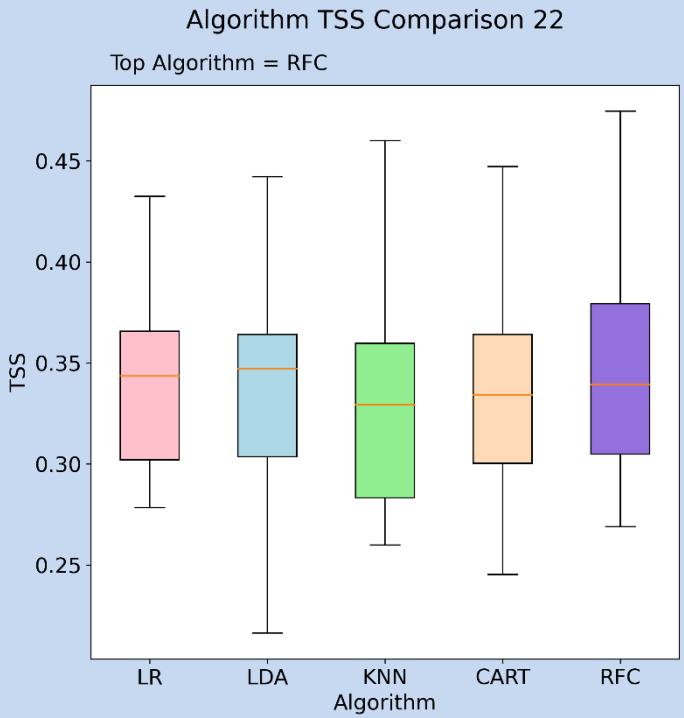
>M1.0 Flares

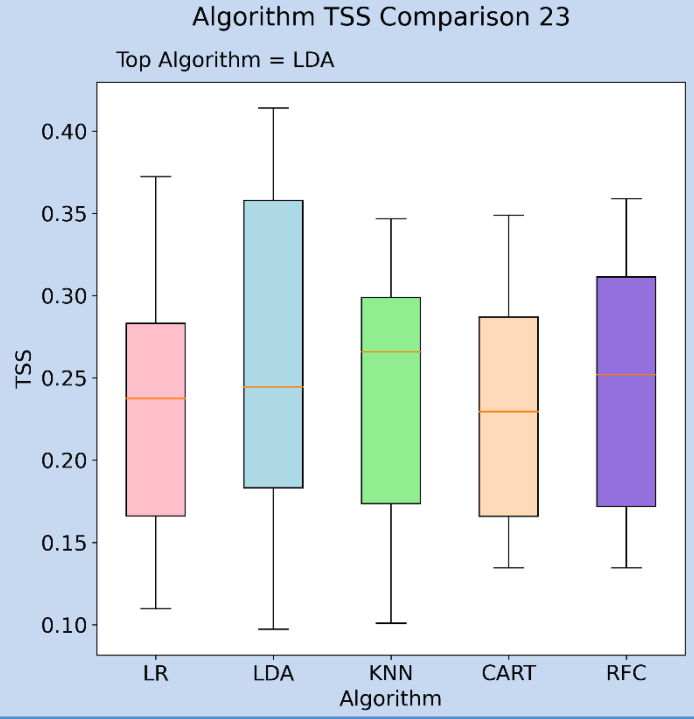




True Skill Statistic

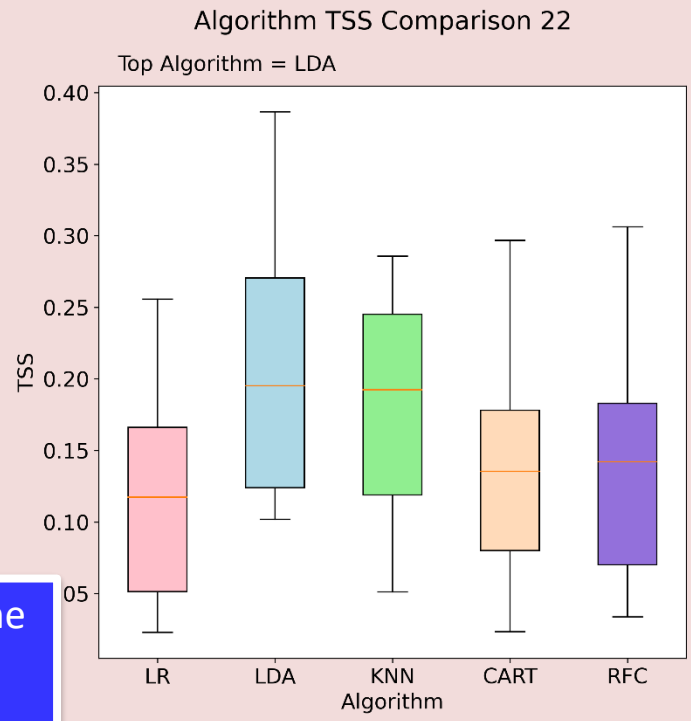
>C1.0 Flares

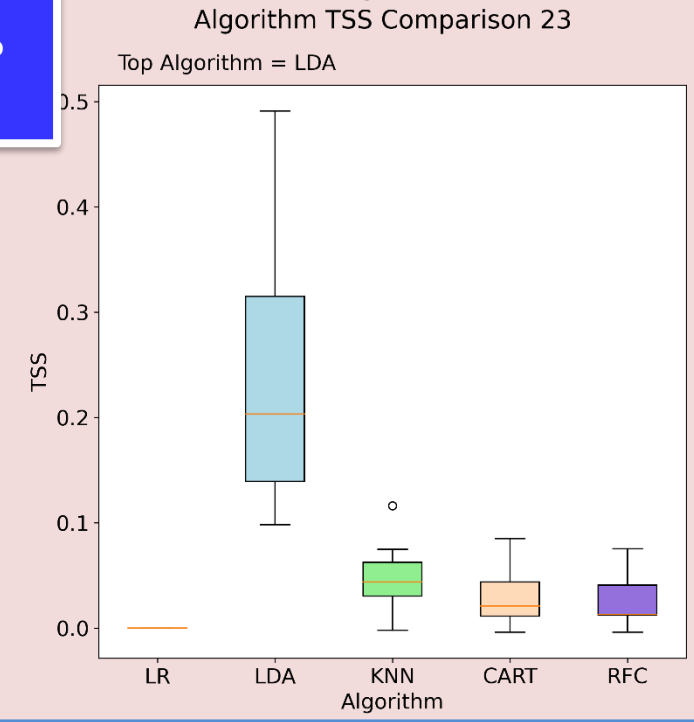




How well do the ML models perform classification? [Yes/No]

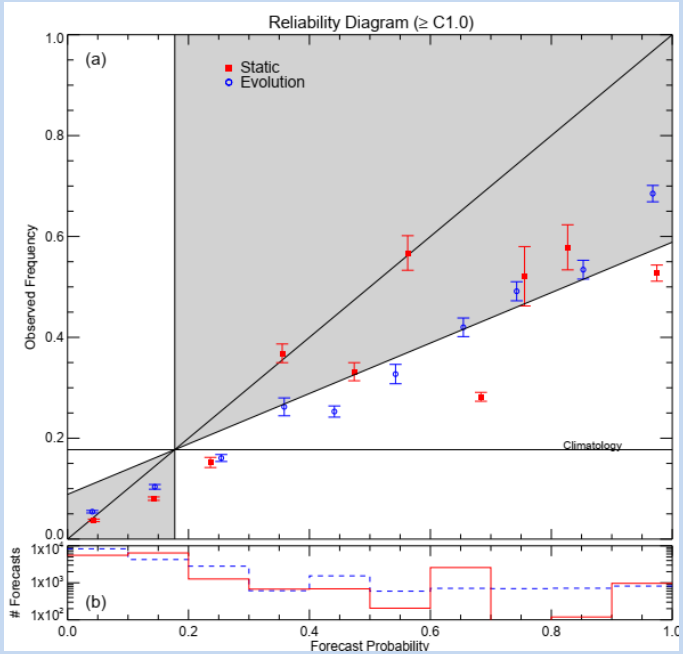
>M1.0 Flares

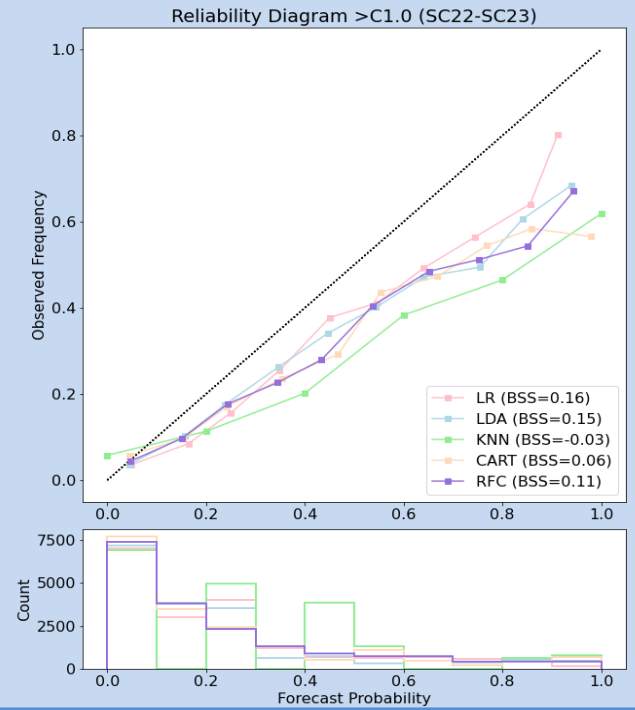




Reliability Diagrams
Solar Cycle 22

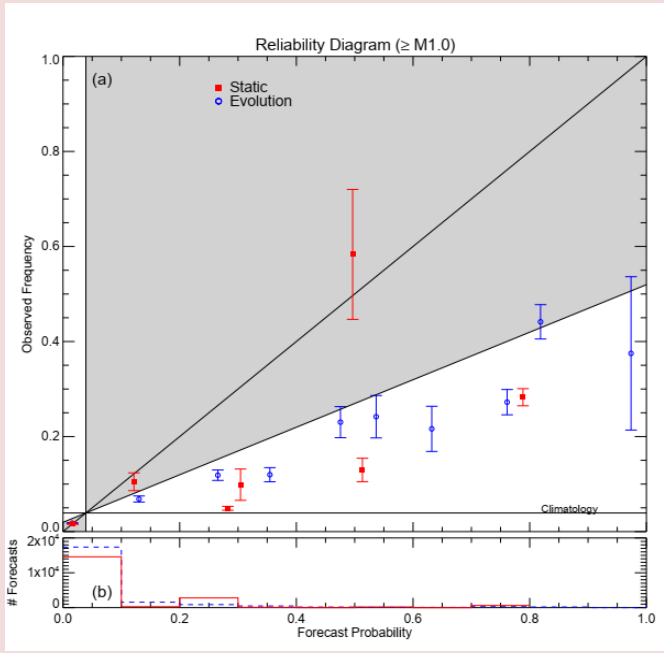
>C1.0 Flares

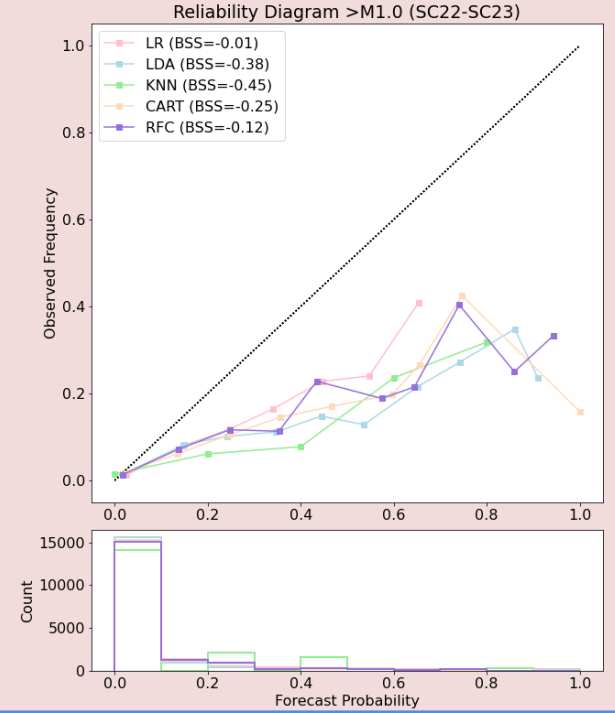




How reliable is the forecast compared to what we observe?

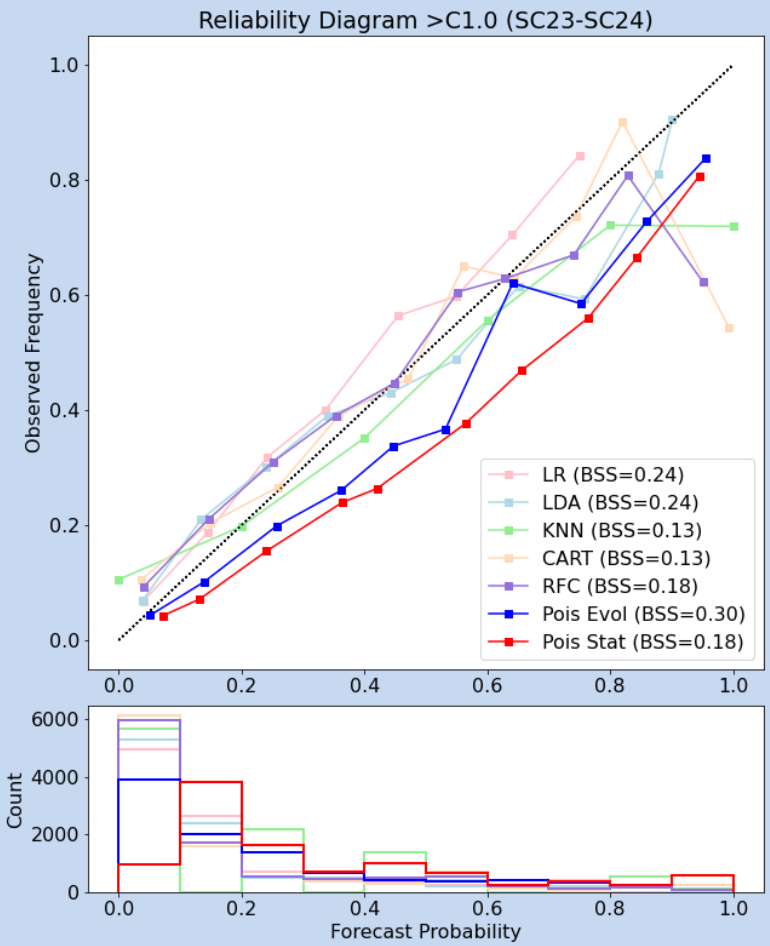
>M1.0 Flares





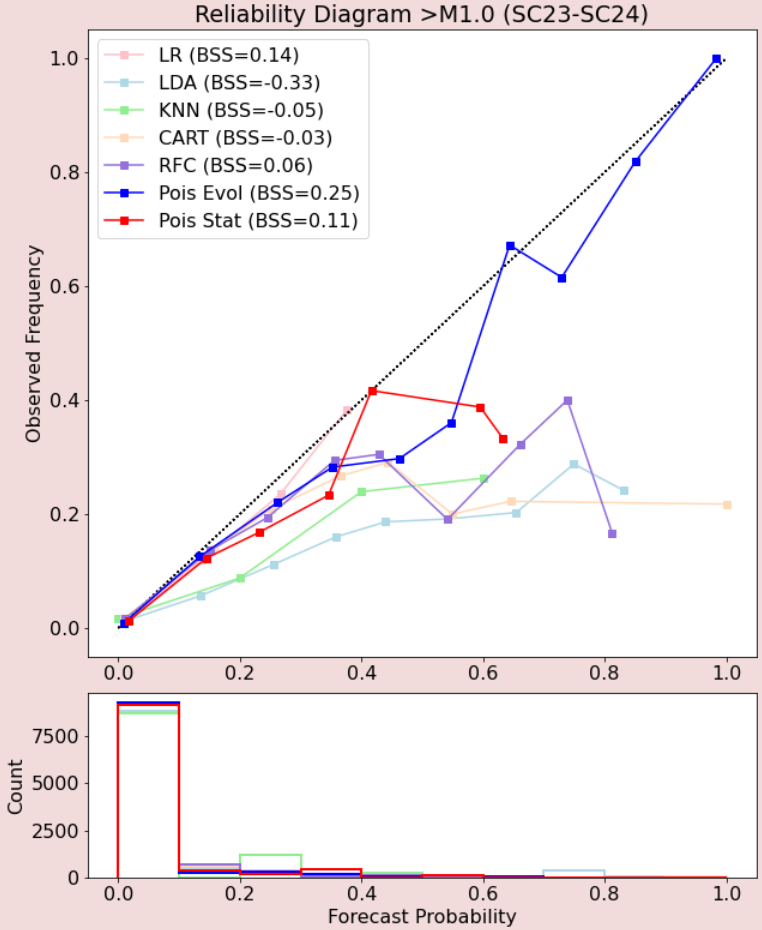
Reliability Diagrams
Solar Cycle 23

>C1.0 Flares



How reliable is the forecast compared to what we observe?

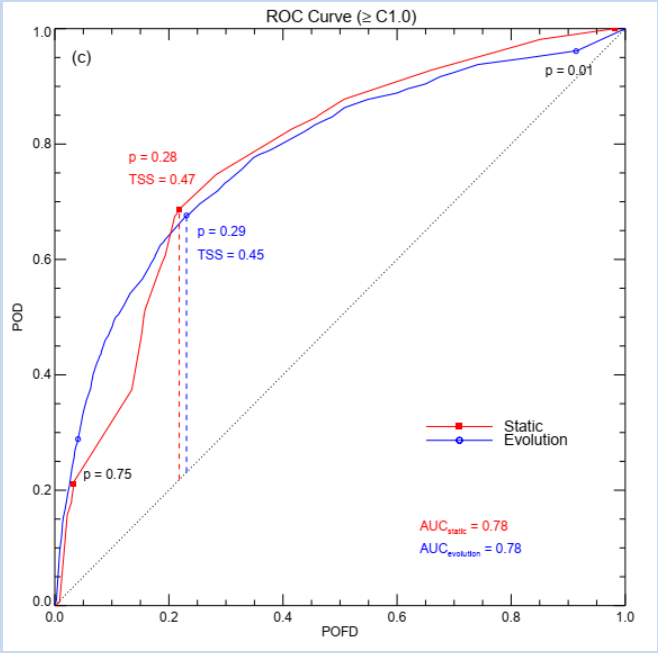
>M1.0 Flares

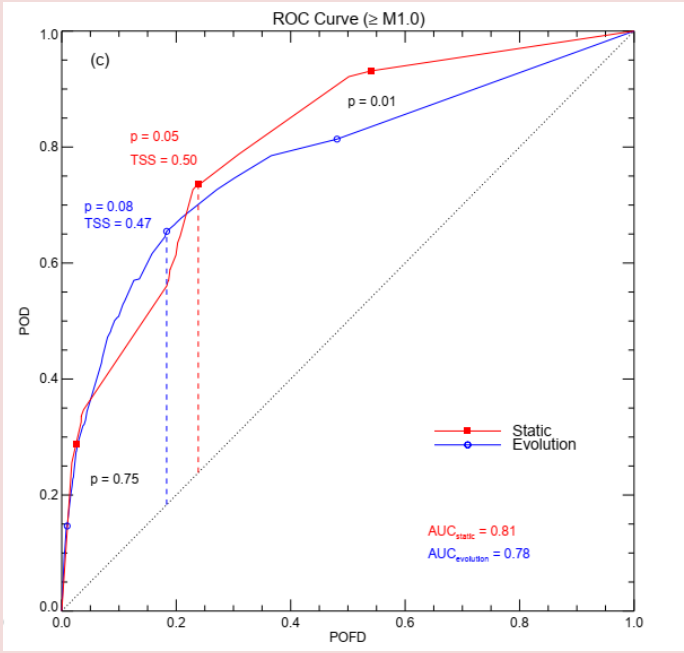


ROC Curves
Solar Cycle 22

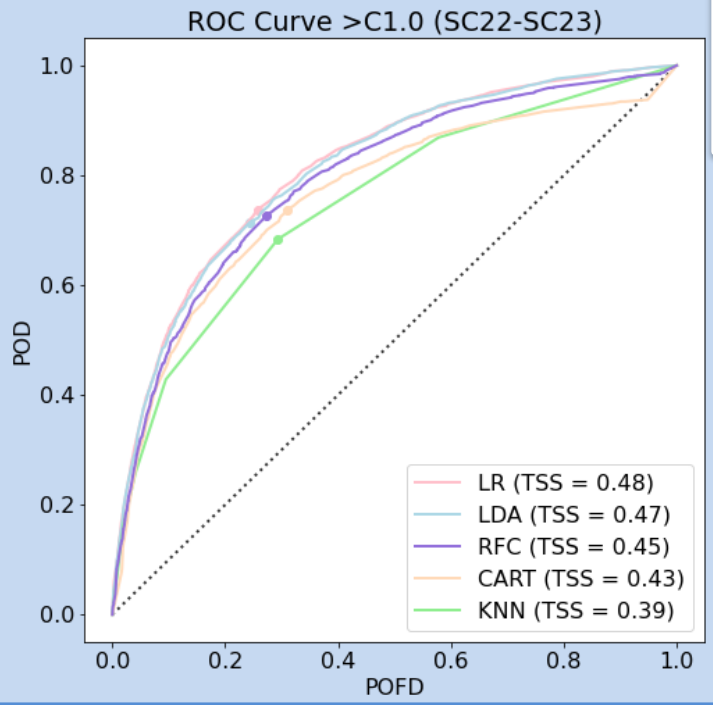
>C1.0 Flares

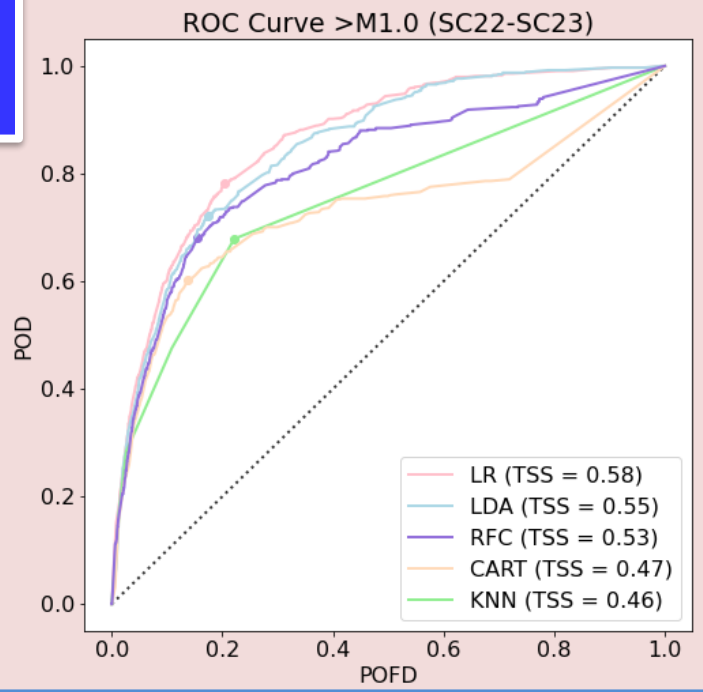
>M1.0 Flares





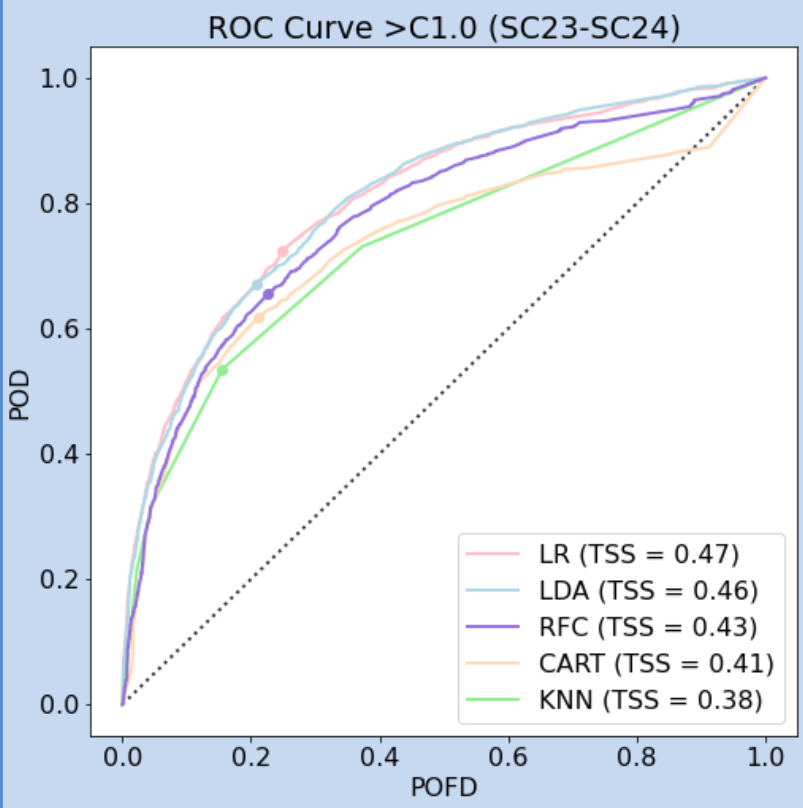
How well do the ML models perform classification? [Yes/No]

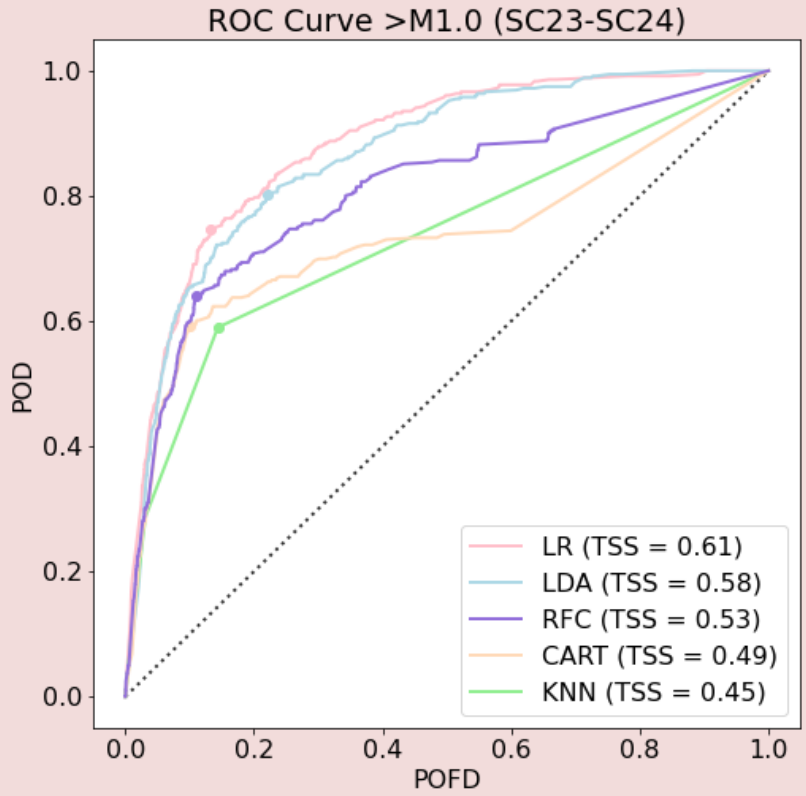




ROC Curves
Solar Cycle 23

>C1.0 Flares



How well do the ML models perform classification? [Yes/No]

>M1.0 Flares