FLARECAST

The Next Level in Automated Solar Flare Forecasting

The EU FLARECAST Project

Manolis K. Georgoulis* & the FLARECAST Team

*RCAAM of the Academy of Athens

AGU FALL MEETING
New Orleans
11-15 Dec. 2017

H2020-PROTEC-2014 RIA; Project No.: 640216

What is FLARECAST?

FLARECAST is an EC H2020 project aiming to develop an advanced solar flare prediction system based on automatically extracted physical properties of solar active regions, coupled with state-of-the-art solar flare prediction methods and validated using the most appropriate forecast verification measures.

Top-level objectives:

- ❑ **Science:** Understand the drivers of solar flare activity and improve flare prediction
- ❑ **R2O:** Provide a globally accessible flare prediction service that facilitates expansion
- ❑ **Communication:** Engage with SWx end users and inform policy makers and the public

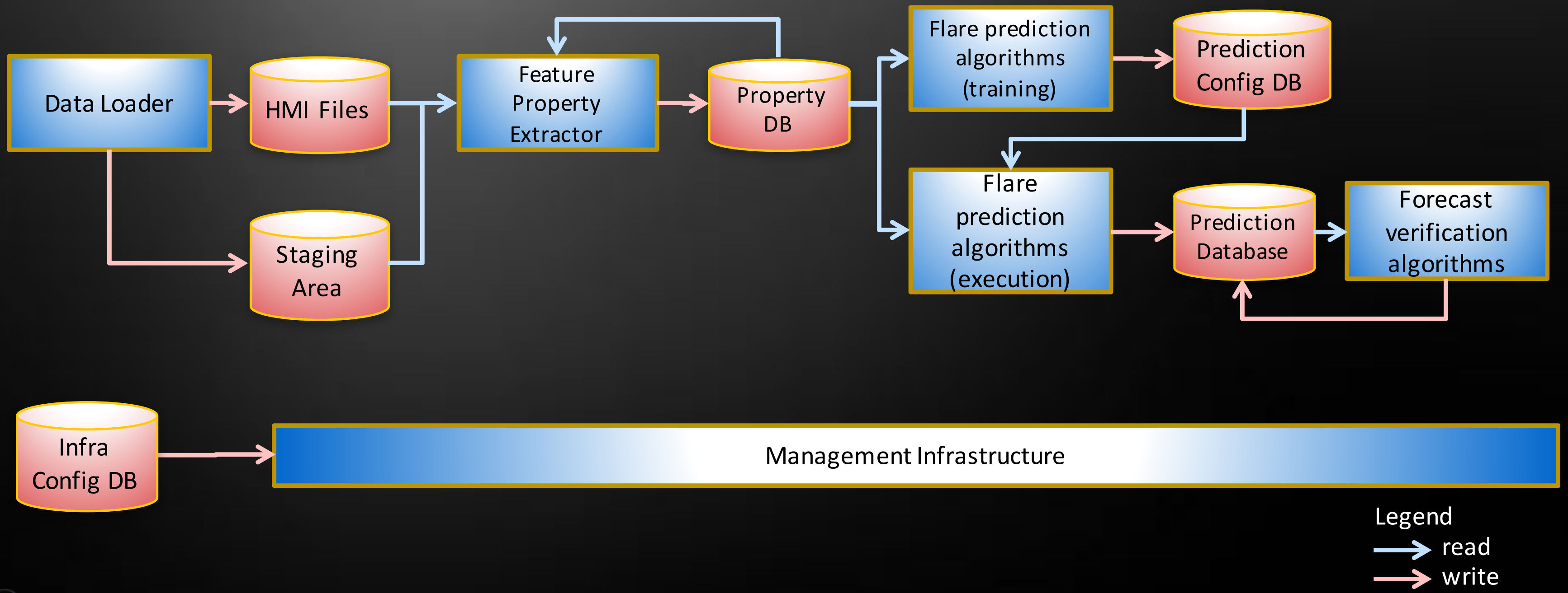
How do we do it - FLARECAST Architecture

Step 1: Data acquisition

Step 2: Feature property extraction

Step 3: Prediction training / execution

Step 4: Data verification



FLARECAST steps and data types in a nutshell

FLARECAST

- Four steps; three data types:
  - Step 1: Data acquisition
  - Step 2: Feature property extraction
  - Step 3: Prediction training / execution
  - Step 4: Forecast verification

External data:

- SDO / HMI NRT SHARPs
- NOAA / SWPC SRS data
  - Active region numbers
  - AR locations
  - Flare occurrences

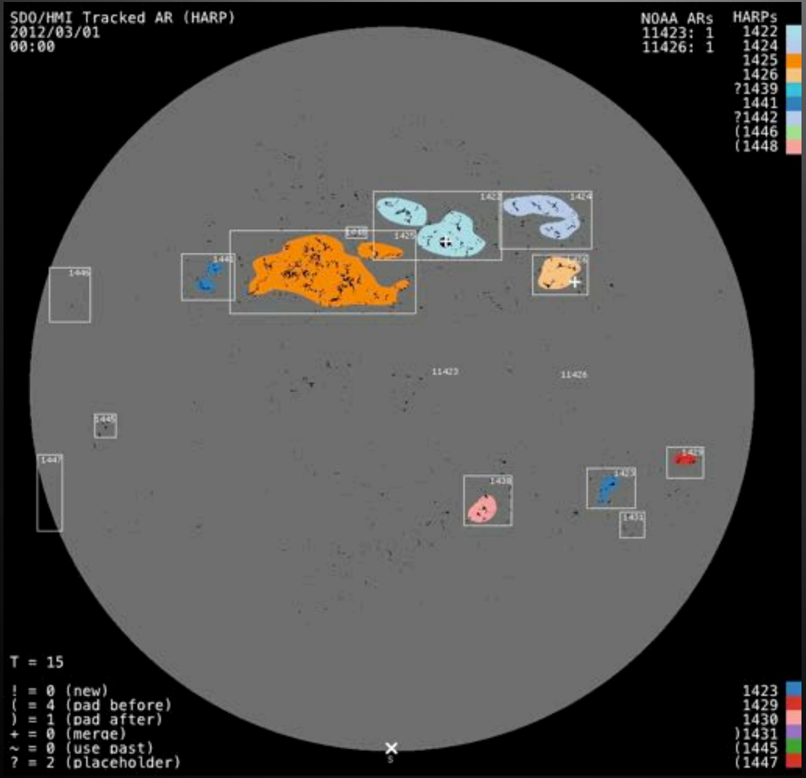
Science data:

- Extracted properties
- Prediction algorithm config.
- Predictions
- Validation

Infrastructure data:

- Algorithm management
- Workflow management

Step 1: Data acquisition



- SDO / HMI data
  - SHARP vector magnetograms - NRT (hmi.sharp_720s_nrt)
  - LOS magnetograms (hmi.M_720s)
  - SHARP vector magnetograms - definitive (hmi.sharp_720s)
- SRS active region (SWPC)

  (YYYY_events.tar.gz)
- Flare association (GOES)

Step 2: Feature property extraction

FLARECAST

- ❑ Pretty much everything proposed as promising for flare prediction over the past 25 years

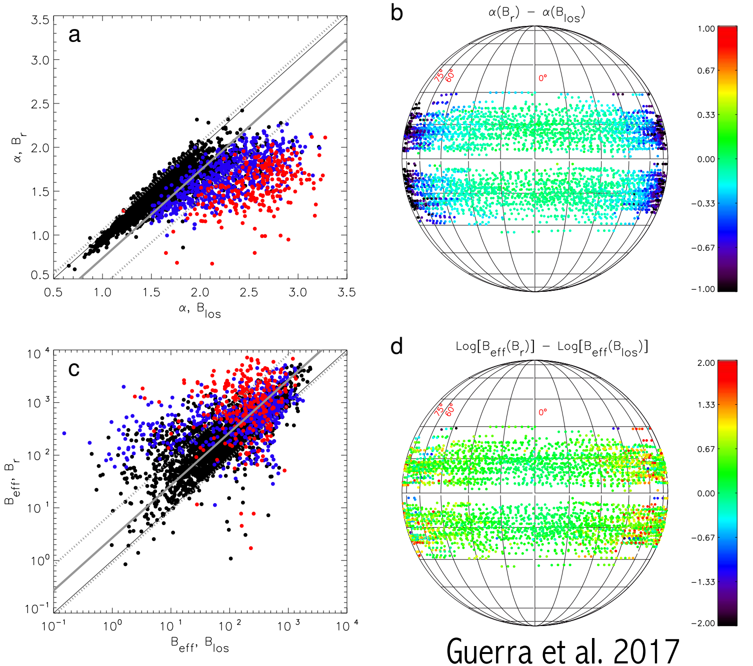
SWPC catalogues			(To do / In progress / Under testing / Delivered)
Details	Solar Region Summary properties	TCD	Delivered
Details	GOES X-ray events	TCD	Delivered
Vector magnetograms			
Details	SHARP properties (Bobra et al. 2014)	TCD	Delivered
Details	Magnetic helicity injection rate (Berger & Field 1984)	TCD	Delivered
Details	Magnetic energy injection rate (Kusano et al. 2002)	TCD	Delivered
Details	Non-neutralized currents (Georgoulis et al., 2012)	AA	Delivered
Details	Flow field characteristics (Deng et al. 2006; Wang et al. 2014)	TCD	Delivered
	Magnetic bipolar feature characteristics	TCD	Under testing (further investigated in WP6)
Intensity images			
	Flow field characteristics	TCD	Under testing (further investigated in WP6)

LOS magnetograms			
Details	Effective connected magnetic field strength (B_{eff}) (Georgoulis & Rust, 2007)	AA	Delivered
Details	Fractal dimension (Georgoulis, 2012)	AA	Delivered
Details	Multi-fractal structure function s(q) inertial range index k (Georgoulis, 2012)	AA	Delivered
Details	Fourier power spectral index (Guerra et. al., 2015)	TCD	Delivered
Details	CWT power spectral index (Hewett et. al., 2008)	TCD	Delivered
Details	Generalised correlation dimension (Georgoulis, 2012)	AA	Delivered
Details	Holder exponent h (Conlon et al., 2010)	AA	Delivered
Details	Hausdorff dimension D(h) (Conlon et al., 2010)	AA	Delivered
	WTMM (Conlon et al., 2010)	TCD	Under testing (further investigated in WP6)
Details	Decay index (Zuccarello et al. 2014)	TCD	Delivered
Details	Magnetic polarity inversion line characteristics (Mason & Hoeksema 2010)	TCD	Delivered
Details	3D magnetic null point (Greene 1992)	TCD	Delivered
Details	R (Schrijver 2007)	TCD	Delivered
Details	LWL_{SG} (Falconer et al. 2008) *	TCD	Delivered
Details	Ising energy (Ahmed et al. 2010)	AA	Delivered
Details	WG_M and S_{l-f} (Korsos et al. 2015)	AA	Delivered
Details	Magnetic helicity injection rate proxy (Park et al. 2013)	TCD	Delivered

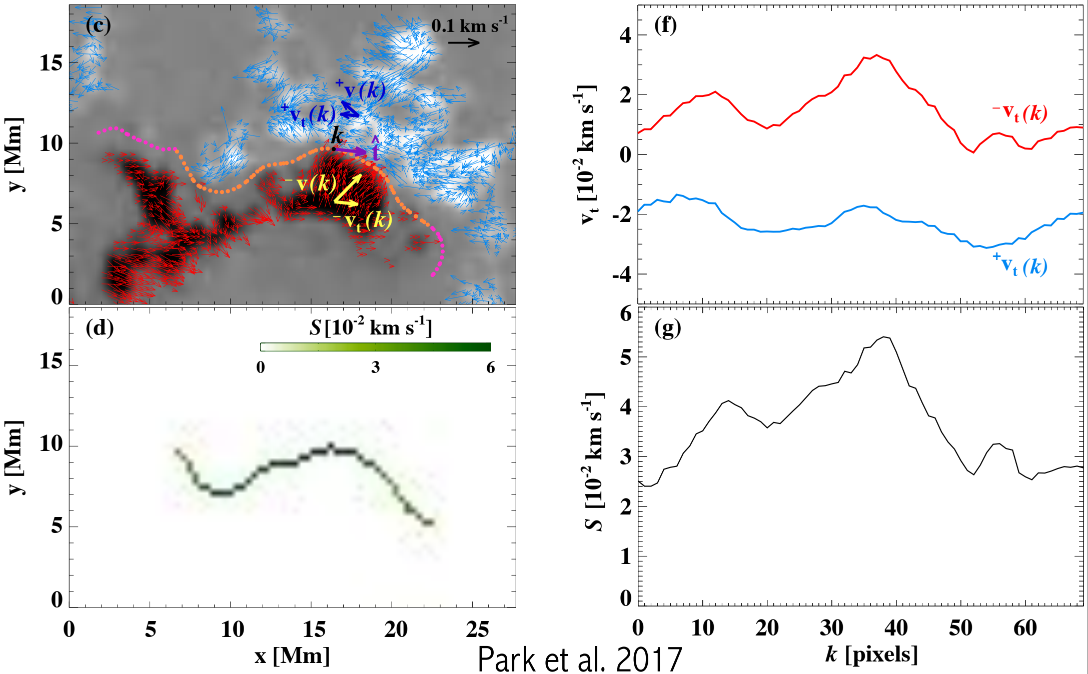
More than 100 features (predictors) for each magnetogram!

Property extraction: examples

- ❑ A number of papers published; more follow



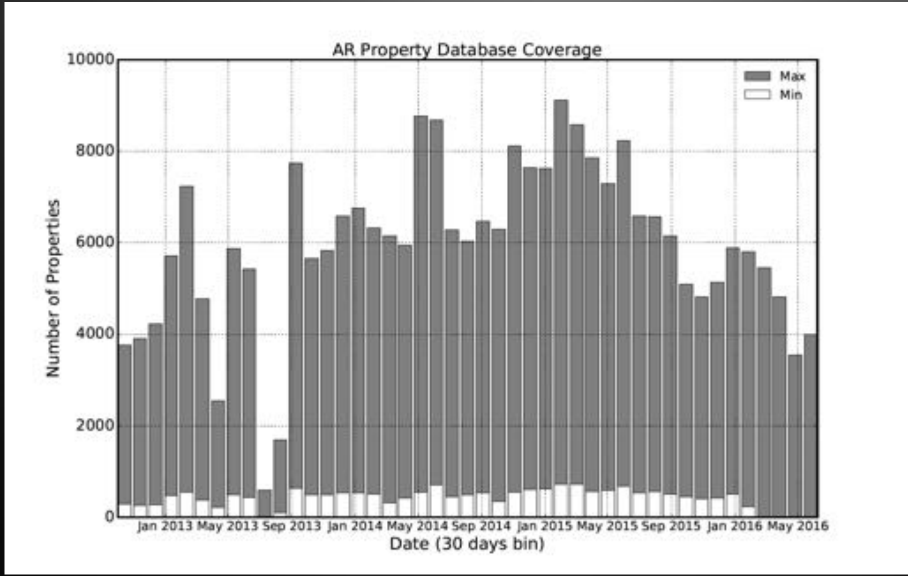
Guerra et al. 2017



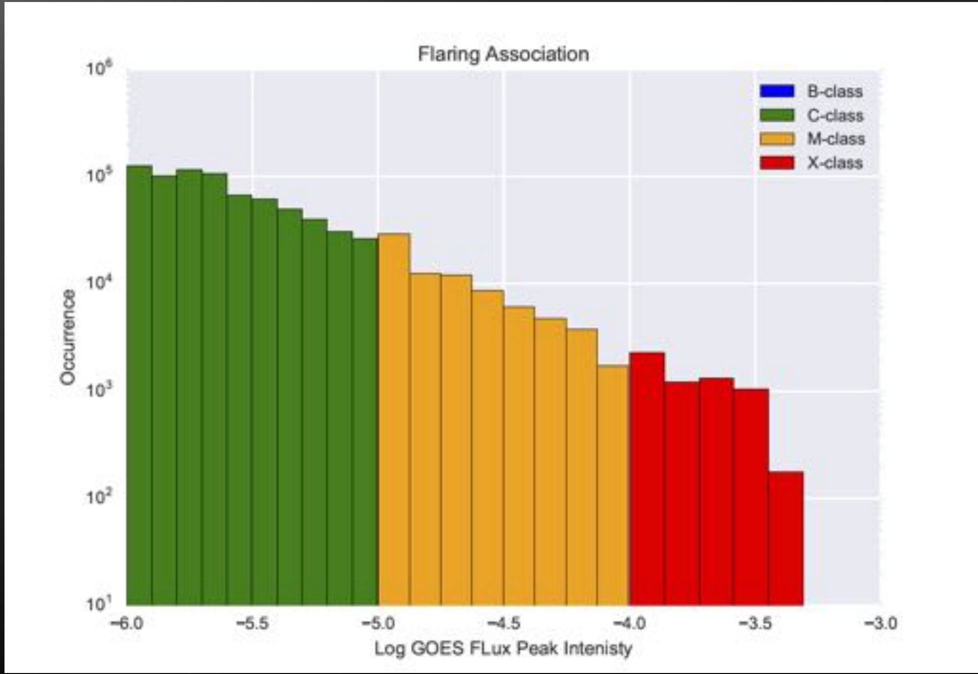
Park et al. 2017

Overview of the FLARECAST Property Database

FLARECAST



Number of properties per month



Flare association

Step 3: Prediction training / execution

FLARECAST

- A total of 22 prediction algorithms tested, most of them in points in time and some in timeseries

Statistical	
	Poisson (McIntosh point-in-time)
	Poisson (McIntosh 24-hour evolution)
	Probit regression
	Logit regression
	Linear regression
	Bayesian binary quantile regression with lasso
Unsupervised learning	
	k-means clustering
	Probabilistic k-means (Fuzzy C-means) clustering
	Possibilistic k-means clustering

Supervised learning	
	Discriminant analysis
	Multi-layer perceptron with back-propagation (point-in-time)
	Multi-layer perceptron with back-propagation (time series)
	Radial basis function networks (point-in-time)
	Radial basis function networks (time series)
	Recurrent neural network with evolutionary algorithm (point-in-time)
	Recurrent neural network with evolutionary algorithm (time series)
	Support vector machine (point-in-time)
	Support vector machine (time series)

Time series preprocessing	
	Discrete Wavelet Transform
	Principal Component Analysis

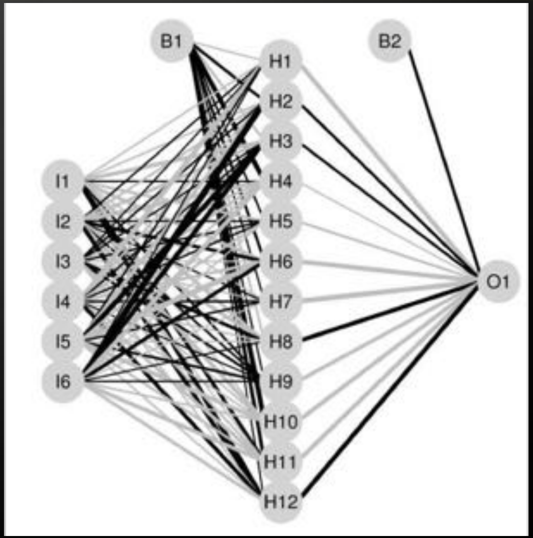
Categories of FLARECAST prediction algorithms

FLARECAST

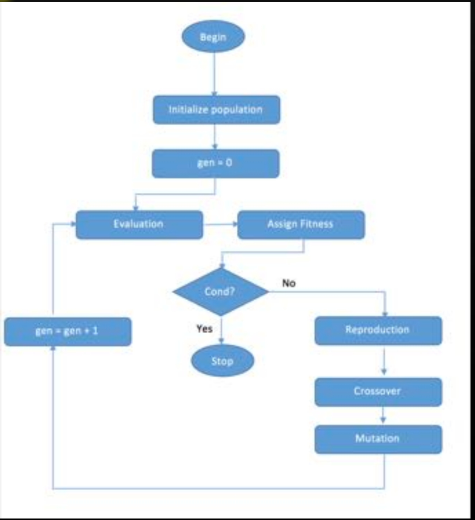
- ❑ Statistical
- ❑ Supervised learning
- ❑ Unsupervised learning
- ❑ Timeseries analysis

- ❑ Machine-learning methods
  - Standard
  - Advanced
  - Innovative

	Non machine-learning
	Machine-learning
	Timeseries (not impleme-nted in this release of FLARECAST



Typical example of multi-layer perceptron



Typical flowchart of a genetic algorithm

Step 4: Forecast verification – binary

FLARECAST

Binary validation: Flare (YES) or No Flare (NO)

	Forecast Flare	Forecast No-flare
Observed Flare	TP	FN
Observed No-flare	FP	TN

2 x 2 contingency table

- TP : true positives
- FN : false negatives
- FP : false positives
- TN : true negatives

Generalized skill score:

$$SS = \frac{\text{score} - \text{score}_{ref}}{\text{score}_{perfect} - \text{score}_{ref}}$$

Different skill scores for different purposes:

- Heidke (HSS - ref. random prediction)

$$HSS = \frac{2(\mathrm{TP} + \mathrm{TN}) - N}{N}$$

- Appleman (HSS - ref. climatology)

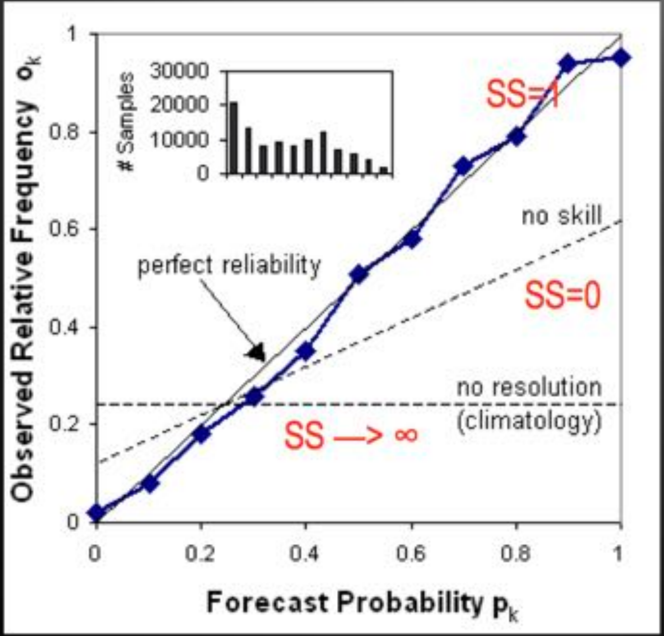
$$ApSS = \frac{\mathrm{TP} - \mathrm{FP}}{N}$$

- True skill statistic (TSS)

$$TSS = \mathrm{POD} - \mathrm{POFD}$$

Step 4: Forecast verification – probabilistic

A probability 0 < p < 1 is assigned to each prediction



Reliabillity diagram

- Correlates forecast probability with observed frequency
- Reliability, skill, resolution
- Generalized skill score: $$SS = 1 - \frac{\mathrm{MSE}_{forecast}}{\mathrm{MSE}_{ref}}$$

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(o_i - p_i)^2$$

- Brier skill score:

$$BSS = 1 - \frac{\mathrm{MSE}(\bar{o}, p)}{\mathrm{MSE}(\bar{o}, \tilde{o})}$$

$$\bar{o} \equiv \{0, 1\}$$

$$\tilde{o} : \text{climatology}$$

Validation: preliminary results

FLARECAST

- Properties relying on line-of-sight field used (full earthward solar disk)

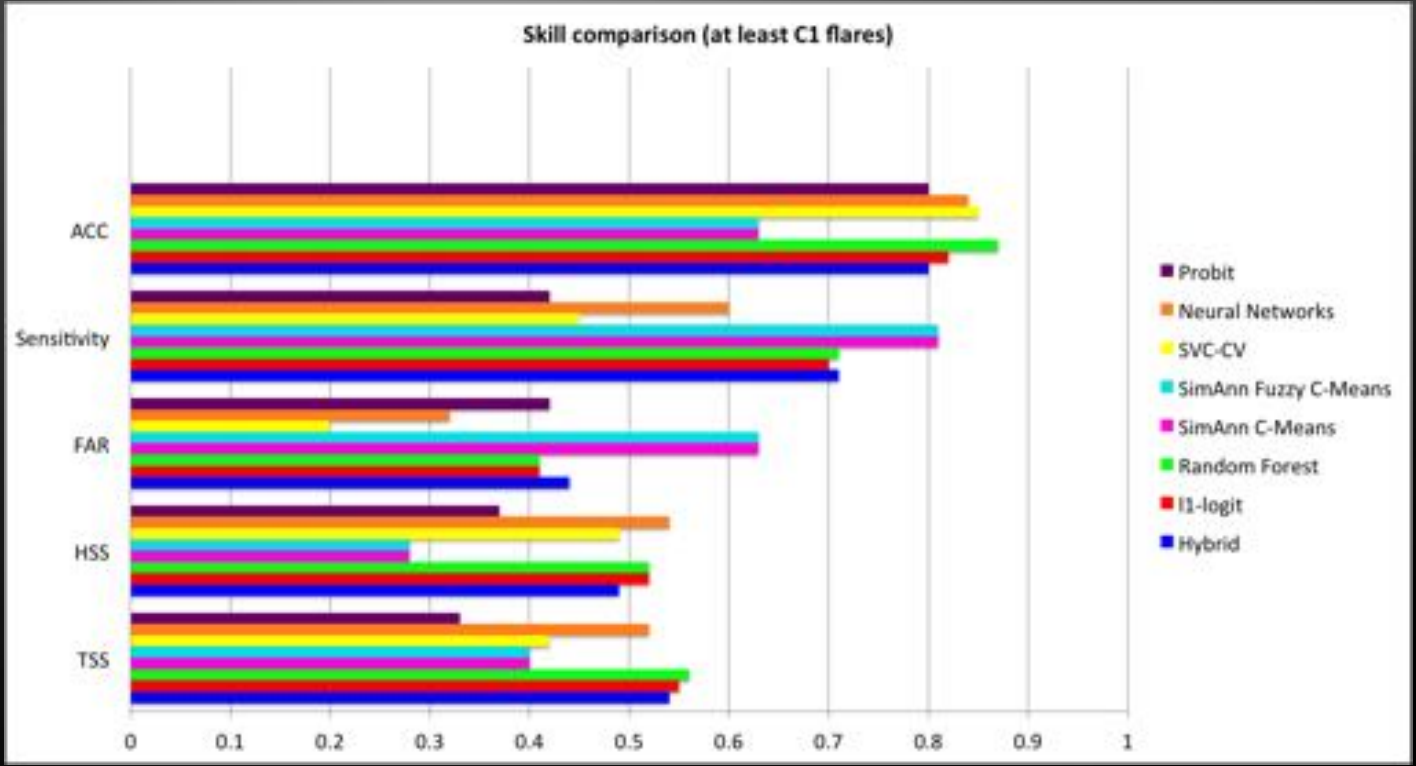
- Event definition:
  - ≥ M1.0 flares within 24 hours
- Trained on 14-Sep-2012 to 31-Dec-2014
- Tested on 1-Jan-2015 to 31-Mar-2016
- Only showing verification for flare yes/no classifying algorithms

Prediction Algorithm	Probability of Detection POD	Probability of False Detection POFD	True Skill Statistic TSS
Hybrid Lasso	0.94	0.20	0.74
Hybrid Logit	0.90	0.20	0.70
Random Forest	0.71	0.07	0.65
Probabilistic K-means	0.65	0.40	0.25
Support Vector Classifier	0.14	0.02	0.12
K-means	0.02	0.01	0.01
Sim. Ann. K-means	0.00	0.32	-0.32
Fuzzy K-means	0.08	0.66	-0.57

Courtesy: Shaun Bloomfield

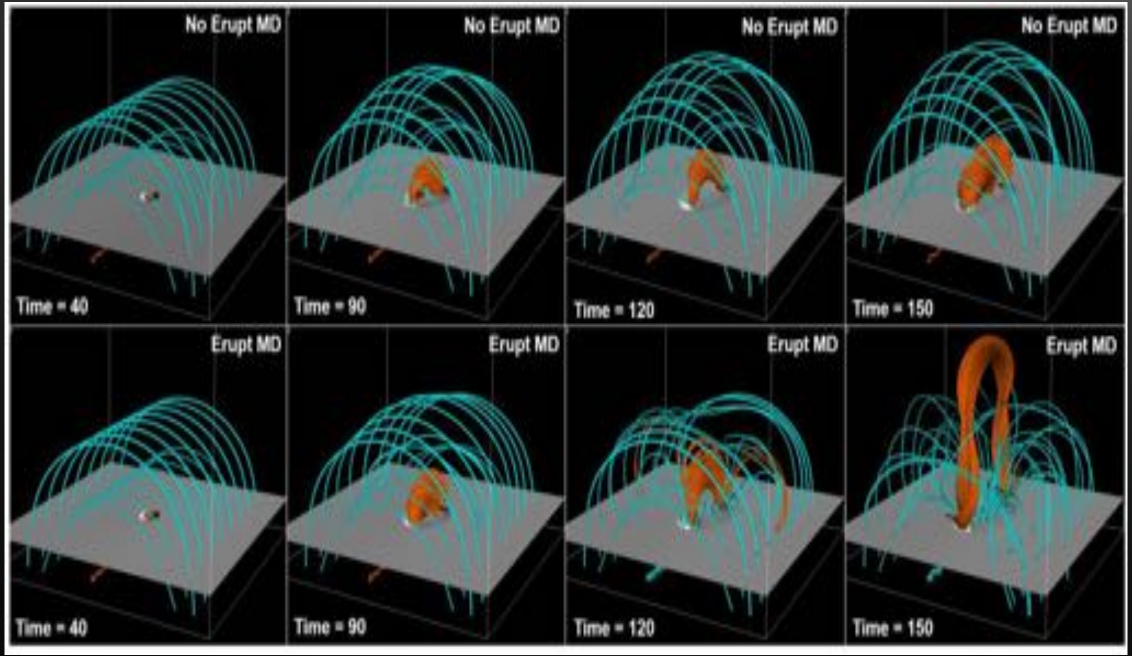
Validation: preliminary results

FLARECAST



Massone et al., (2018)

FLARECAST Science: explorative research

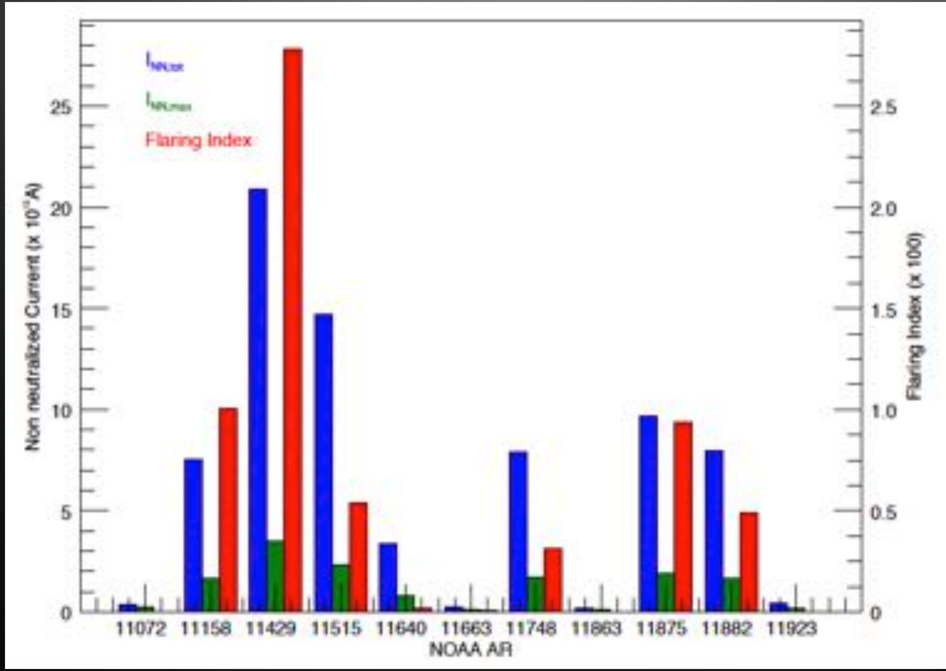


Pariat et al. 2017

- ❑ Understand solar magnetic eruptions
- ❑ Improve future flare prediction, involving use of timeseries
- ❑ Investigate suitability of forecast window and latency
- ❑ Advance CME prediction

Study of eruptive flares in synthetic MHD configurations

FLARECAST Science: explorative research



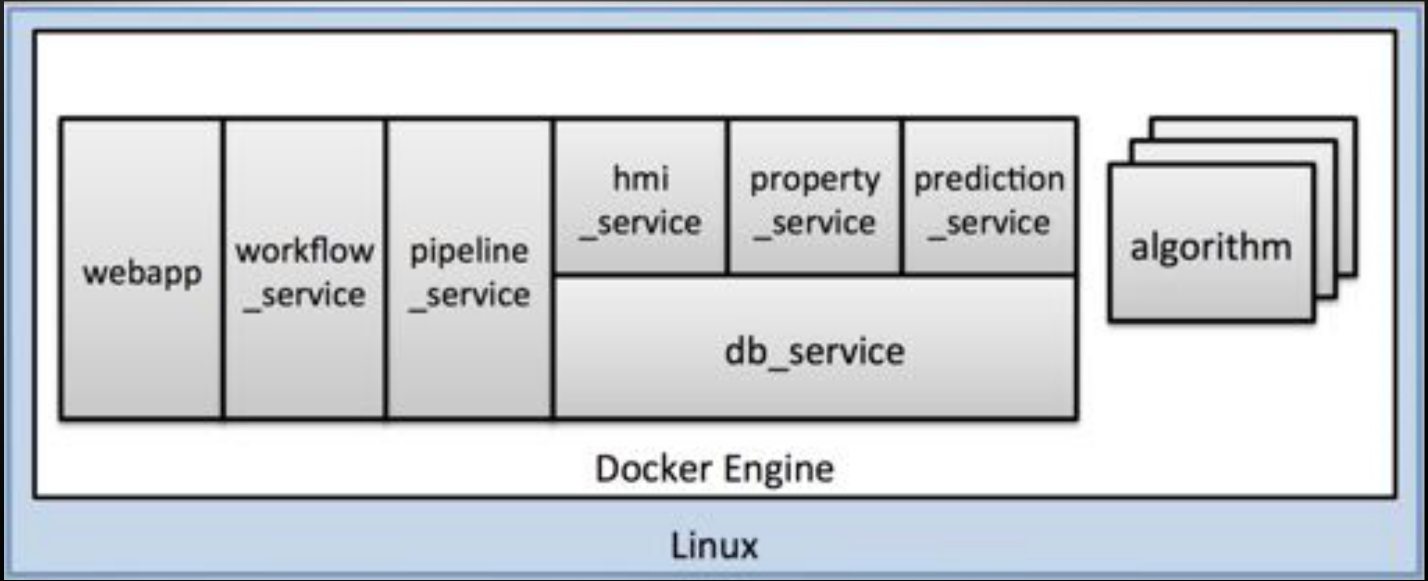
Kontogiannis et al. 2017

- ❑ Understand solar magnetic eruptions
- ❑ Improve future flare prediction, involving use of timeseries
- ❑ Investigate suitability of forecast window and latency
- ❑ Advance CME prediction

Feasibility of non-neutralized currents in active regions as flare predictors

FLARECAST top-level objectives: Technology

- API accessible databases (api.flarecast.eu)
- Open-source Architecture based on Docker engine and containers
- Pick-and-mix installation



FLARECAST top-level objectives: communication

- ❑ Communicating with the scientific community

http://flarecast.eu/research/publications

Pages / Management / Publications and Conferences

Edit · Favourite · Watching · Share

FLARECAST Publication Plan

Created by D. Shaun Bloomfield, last modified by Etienne Pariat on Mar 31, 2017

At least nineteen (19) envisioned refereed papers, of which:

- ❑ Six (6) are already published; two (2) are in press
- ❑ Three (3) are under review
- ❑ At least eight (8) are in preparation

Al , either in open-access journals or in ArXiv

FLARECAST top-level objectives: communication

- Communicating with industry and government

http://flarecast.eu/industry/first-stakeholder-workshop

- First Stakeholders Workshop, Met Office 12-13 January 2017

- Second Users Workshop, ESWW14, 29 November 2017

http://flarecast.eu/second-stakeholder-workshop http://flarecast.eu/outreach-activities

FLARECAST top-level objectives: communication

FLARECAST

- Communicating with the public

Il Secolo XIX, Genova, 13 September 2016


20 genova IL SECOLO XIX

Il futuro è in città

UN SOFTWARE SARÀ CAPACE DI PREVEDERE I BRILLAMENTI, FENOMENI PERICOLOSI NELLO SPAZIO

Ricercatori a caccia di tempeste solari

Università e Cnr stanno creando una squadra di tecnici e scienziati internazionali

COSA SONO?

PERCHÉ PREVEDERLE?

EU Researchers Night, TCD, Dublin, 30.09.2016

EU Researchers Night, Athens, 30.09.2017

AA, Athens, 13.11.2017

Science Café, Zurich, 11.11.2016

Fet de la Science, Paris, 16.10.2016

FLARECAST user interface: how it will work

- Three different levels of service exploitation:
  - **Administrator:** control, workflow manager
  - **Scientist:** edit metadata; visualize data
  - **End user:** view / query data; visualize prediction

Summary

FLARECAST

- ❑ FLARECAST is arguably the most systematic, cost- and effort-intensive solar flare prediction project worldwide at this time.
- ❑ The project has diverse objectives, comprising Science, R2O and Communication. Impressively diverse expertise has been used
- ❑ FLARECAST data, codes and infrastructure are <u>fully and openly accessible worldwide</u> and can be used to avoid effort duplication in future SWx forecasting efforts.
- ❑ FLARECAST architecture is modular and expandable. Integrated SWx forecasting platforms might conceivably use and expand it

FLARECAST / FLARECAST-citing papers (so far)

FLARECAST

- A. McCloskey, et al.: Flaring Rates and the Evolution of Sunspot Group McIntosh Classifications, *Solar Physics*, **291**, 1711, 2016, DOI: 10.1007/s11207-016-0933-y
- G. Barnes, et al.: A Comparison of Flare-Forecasting Methods. I. Results from the "All-Clear" Workshop, *Astrophysical Journal*, **829**, article.id 89, 2016, DOI: 10.3847/0004-637X/829/2/89
- S. Murray, et al.: Flare Forecasting at the Met Office Space Weather Prediction Center, *Space Weather*, **15**, 577, 2017, DOI: 10.1002/2016SW001579
- E. Pariat, et al.: Relative Magnetic Helicity as a Diagnostic of Solar Eruptity, *Astronomy & Astrophysics*, **601**, A125, 2017, DOI: 10.1051/0004-6361/201630043
- C. Guennou, et al.: Testing Predictors of Eruptivity Using Parametric Flux Emergence Simulations, *J. Space Weather & Space Climate*, **7**, A17, 2017, DOI: 10.1051/swsc/2017015
- I. Kontogiannis, et al.: Non-Neutralized Electric Currents in Solar Active Regions and Flare Productivity, *Solar Physics*, **292**, 159, 2017, DOI: 10.1007/s11207-017-1185-1
- K. Florios, et al.: Forecasting Solar Flares Using Magnetogram-Based Predictors and Machine Learning, *Solar Physics*, 2017, in press
- A. M. Massone, et al.: Machine Learning for Solar Flare Forecasting, In Machine-Learning Techniques for Space Weather (E. Camporeale, S. Wing, J. Johnson, Editors), Elsevier, 2018, in press

Three more have been submitted; around ten are in preparation

BACKUP SLIDES