An Introduction to FLARECAST

FLARECAST Second Users Workshop

Manolis K. Georgoulis

Project Coordinator, FLARECAST

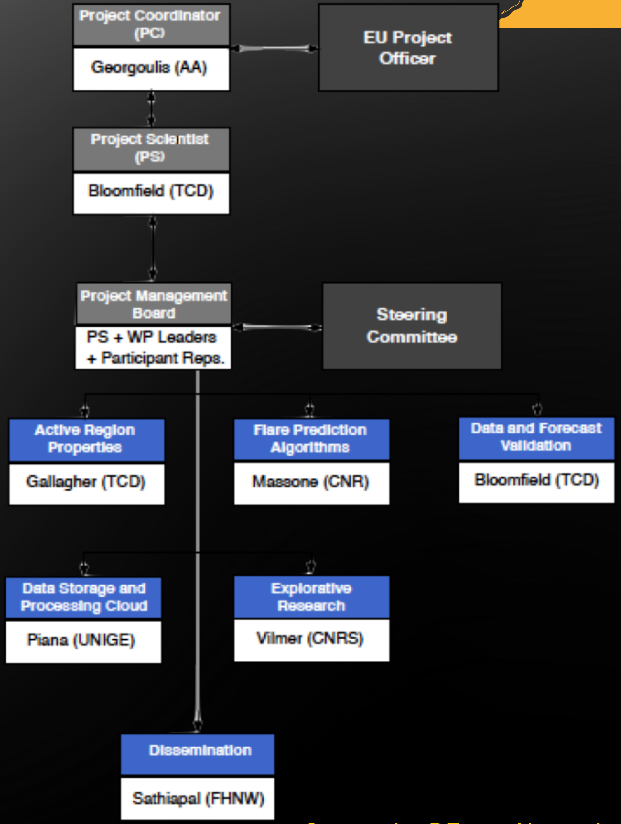
RCAAM of the Academy of Athens

29 November 2017

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

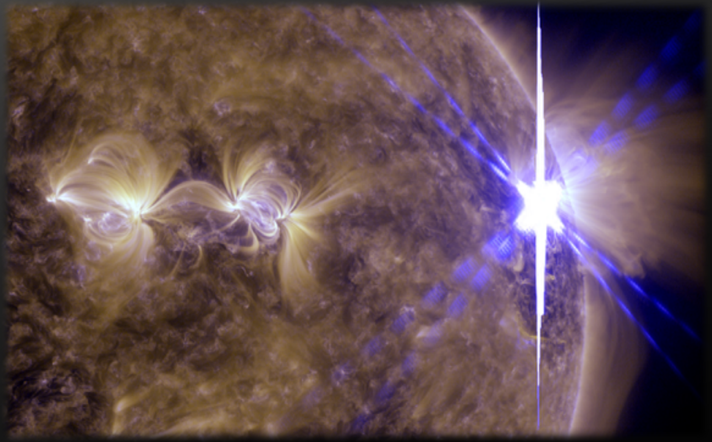
FLARECAST trivia

- ❑ Funded by Horizon 2020 PROTEC-1-2014: Space Weather (Grant Agreement No. 640216)
- ❑ Period of performance: January 2015 – December 2017 (3 years)
- ❑ Max budget: 2.416 MEUR
- ❑ A Consortium of nine (9) institutes in six countries: AA (GR), TCD (IE), UNIGE (IT), CNR (IT), CNRS (FR), UPSud (FR), FHNW (CH), Met Office (UK), UNN (UK)
- ❑ Five-member Steering Committee: N. Hurlburt (US, Chair), D. Biesecker (US), G. Barnes (US), S. Villa (IT), P. Russo (NL)



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.

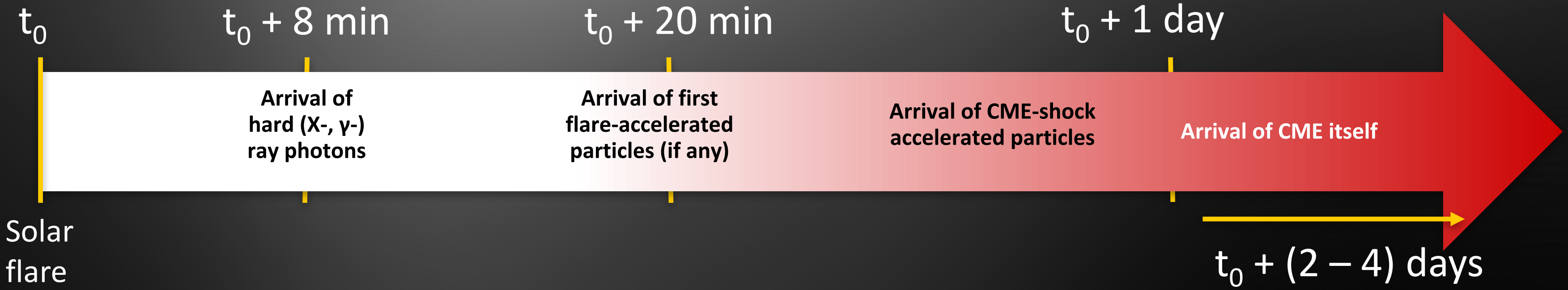


FLARECAST top-level objectives:

- To understand the drivers of solar flare activity and improve flare prediction
- To provide a globally accessible flare prediction service that facilitates expansion
- To engage with space weather end users and inform policy makers and the public

So... Why do we need advance flare prediction?

FLARECAST



- There is no early warning for flare X- and γ-ray photons
- There is a slim (few min) early-warning window for possible flare-only particulate
- From the flare class, one can effectively proceed to CME prediction for major flares
- Flares are the primary agents for solar radio bursts
- Dot-connecting exercises (from predicted flare location, surroundings, orientation) can be made to assess possible eruption impact and combine with other SWx prediction efforts (CMEs, SEPs)

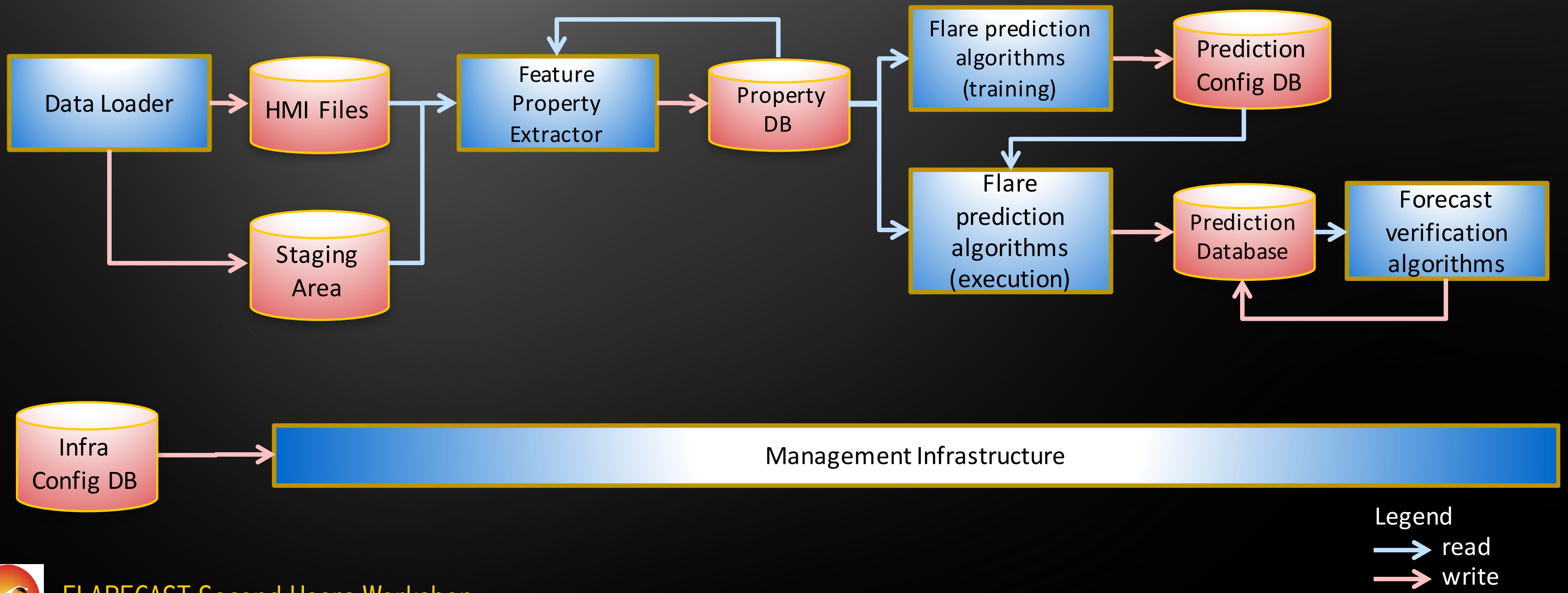
How do we do it - FLARECAST Architecture

Step 1: Data acquisition

Step 2: Feature property extraction

Step 3: Prediction training / execution

Step 4: Data verification



FLARECAST steps and data types in more detail

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

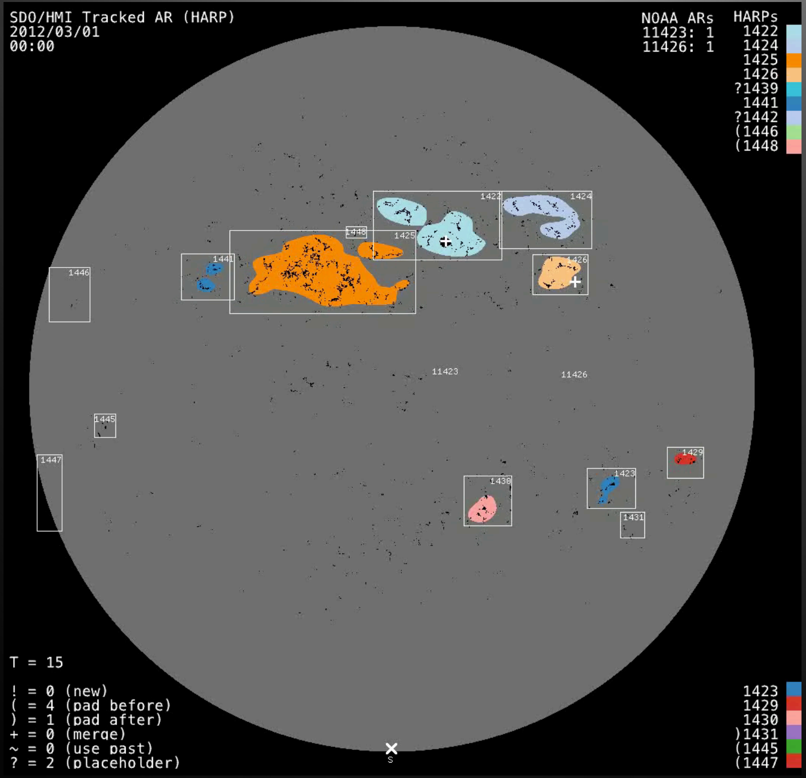
cf. Shaun's talk

Infrastructure data:

- Algorithm management
- Workflow management

cf. André's talk

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

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

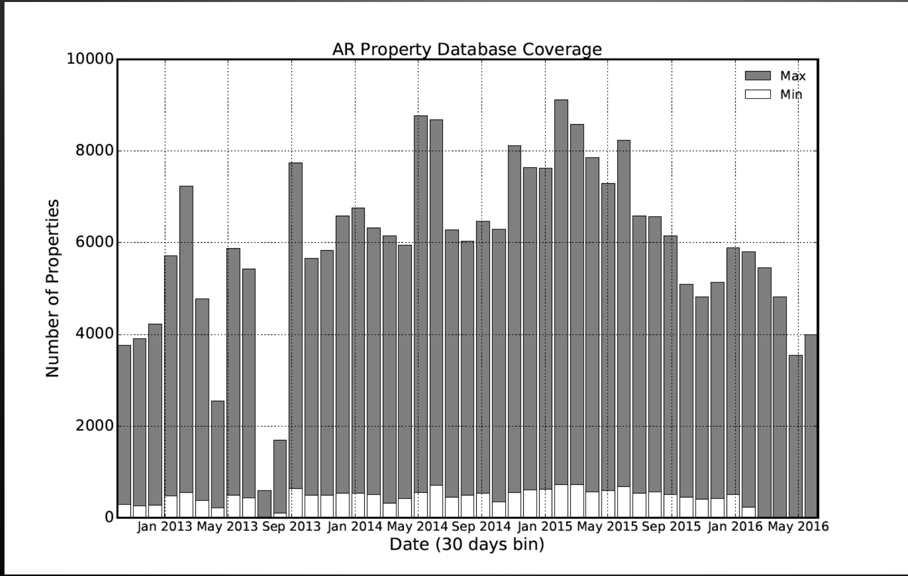
SWPC catalogues			(To do / In progress / Under testing / Delivered)
Details	Solar Region Summary properties	TCD	Delivered
Details	GOES X-ray events	TCD	Delivered

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)

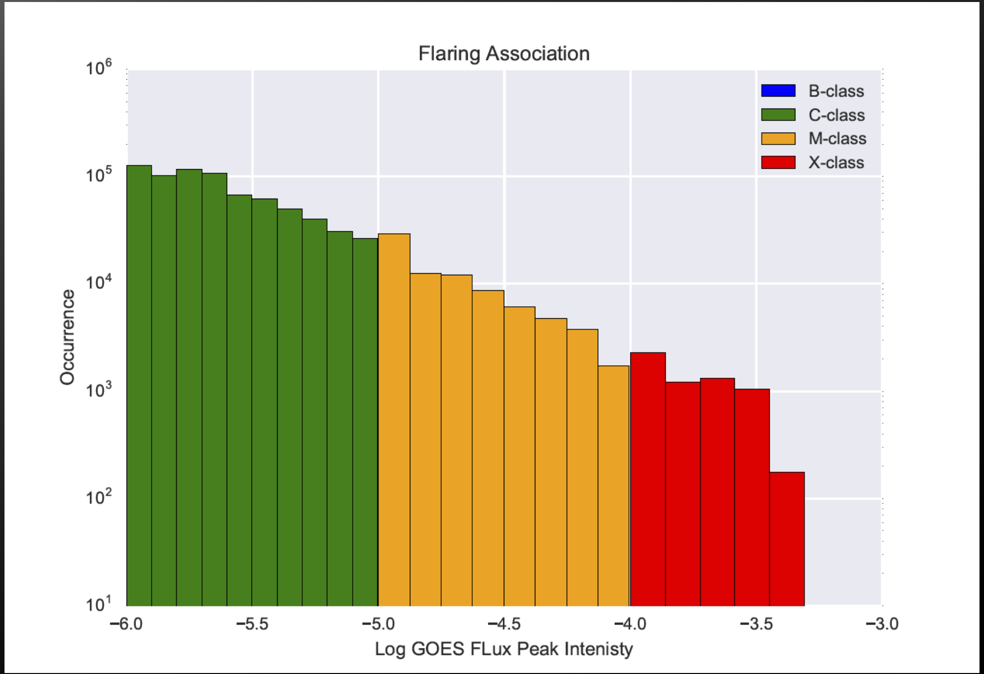
More than 100 features (predictors) for each magnetogram!

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	$^{L}WL_{SG}$ (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

Overview of the FLARECAST Property Database



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)

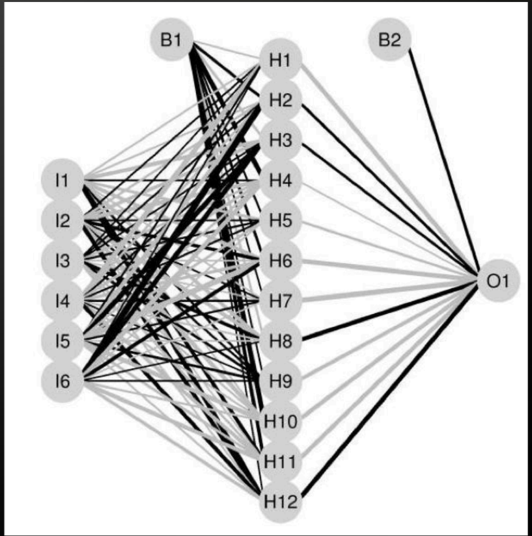
Time series preprocessing	
	Discrete Wavelet Transform
	Principal Component Analysis

Categories of FLARECAST prediction algorithms

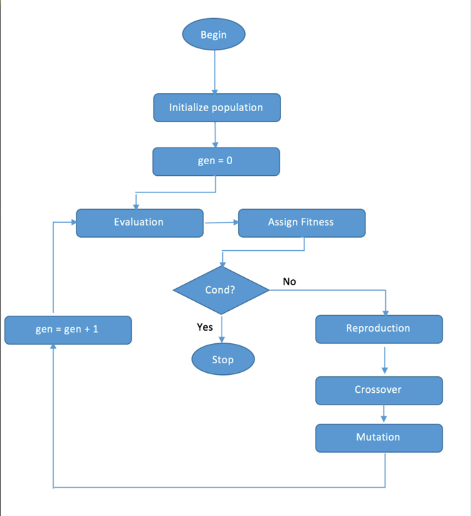
- ❑ Statistical
- ❑ Supervised learning
- ❑ Unsupervised learning
- ❑ Timeseries analysis

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

(white)	Non machine-learning
(yellow)	Machine-learning
(purple)	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:

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

Different skill scores for different purposes:

- Heidke (HSS - ref. random prediction)

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

- Appleman (HSS - ref. climatology)

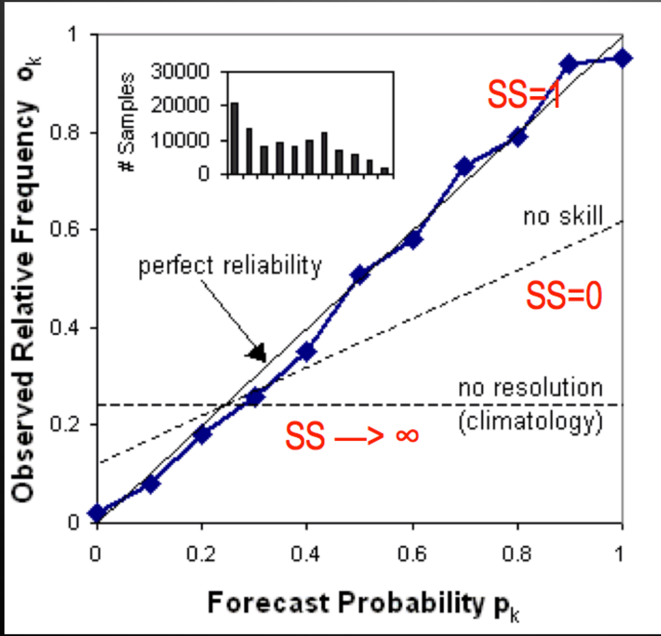
$$ApSS = \frac{\mathrm{TP} - \mathrm{FP}}{N}$$

- 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: $$SS = 1 - \frac{\mathrm{MSE}_{forecast}}{\mathrm{MSE}_{ref}}$$

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

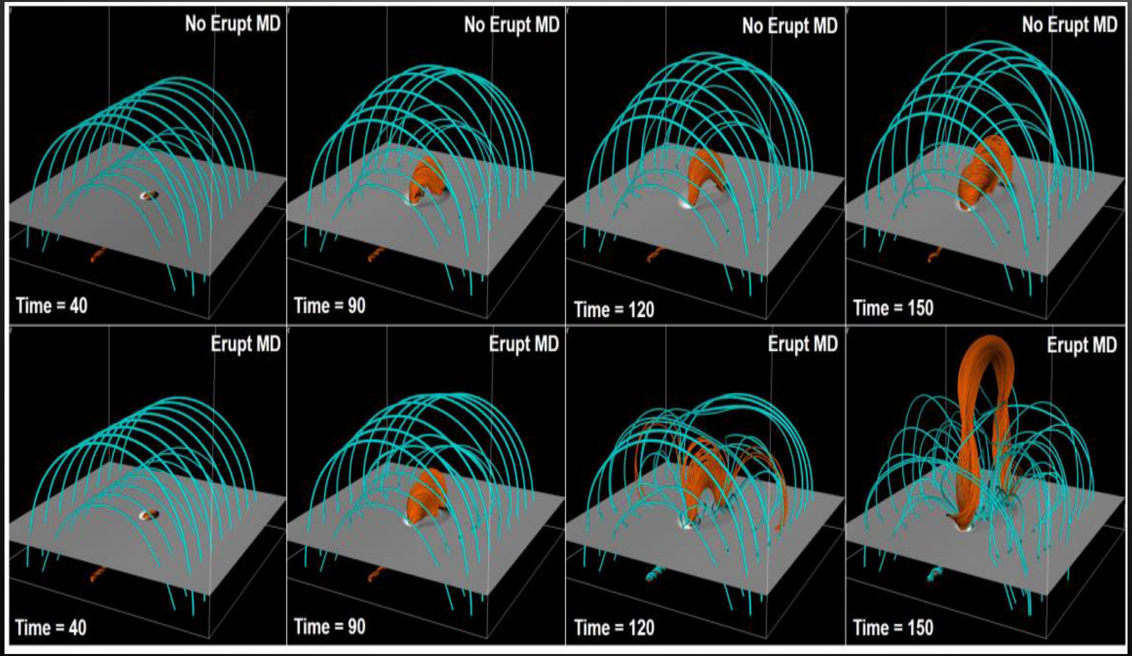
- Brier skill score:

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

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

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

FLARECAST Science: explorative research

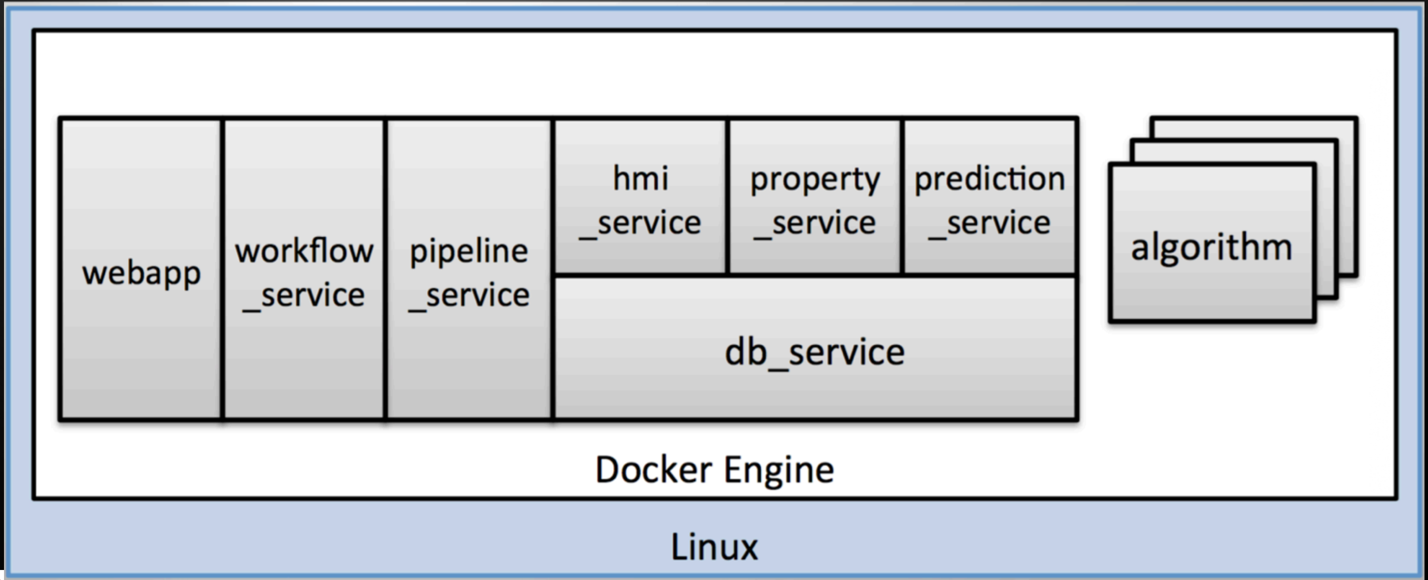


Study of eruptive flares in synthetic MHD configurations

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

FLARECAST top-level objectives: Technology

- ❑ API accessible databases
- ❑ Open-source Architecture based on Docker engine and containers
- ❑ Pick-and-mix installation



FLARECAST top-level objectives: communication

FLARECAST

- ❑ Communicating with the scientific community

http://flarecast.eu/research/publications

Pages / Management Edit Favourite Watching Share •••

/ Publications and Conferences

FLARECAST Publication Plan

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

At least seventeen (17) envisioned refereed papers, of which:

- ❑ Six (6) are already published
- ❑ Three (3) are under review
- ❑ Eight (8) are in preparation

Al , either in open-access journals or in ArXiv

FLARECAST top-level objectives: communication

- ❑ Communicating with industry and government

- ❑ 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/industry/first-stakeholder-workshop http://flarecast.eu/outreach-activities

FLARECAST top-level objectives: communication

@FLARECAST_EU

❑ Communicating with the public

Il Secolo XIX, Genova, 13 September 2016


20 MARTEDÌ 13 SETTEMBRE 2016 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

FRANCESCO MARGIOCCO

PER NOI CHE SIAMO sulla terraferma non rappresentano un problema, anzi ci incantano quando creano quei giochi di luce che sono le aurore polari. Le cose cambiano se invece di starcene coi piedi per terra voliamo su un aereo a 10 mila metri d'altezza. È lì che le tempeste solari, o brillamenti, possono farci male, ed è per questo che gli scienziati sono alla ricerca del modo migliore per prevederle. Su mandato della Commissione europea, un gruppo dell'Università di Genova e del Cnr sta creando insieme ad una squadra internazionale di tecnici e scienziati un

Michele Piana

Anna Massone

probabilità che una tempesta solare si verifichi nell'immediato futuro». Flarecast, questo il nome del sistema per prevedere le tempeste solari, è finanziato dalla Commissione europea, nell'ambito del programma Horizon 2020, con 3 milioni di euro, e oltre all'Italia coinvolge altri cinque Paesi coordinati da Manolis Georgoulis, fisico dell'Academy of Athens.

Tra un anno e mezzo Flarecast sarà pronto. A gennaio Georgoulis e colleghi lo presenteranno a Londra ad un simposio di aziende produttrici di satelliti, interessate a limitare i danni ai loro apparecchi. Negli Stati Uniti un si-

COSA SONO?

Tempeste solari
Improvvisi rilasci di energia – pari a qualche milione di bombe di Hiroshima – e di materia da parte della superficie del sole

PERCHÈ PREVEDERLE?

Le tempeste solari non danneggiano la terra che è protetta dall'atmosfera. Lo scontro tra le tempeste e l'atmosfera produce le aurore boreali.
Le tempeste solari possono però danneggiare gli aerei in volo e i loro passeggeri, esposti ai raggi x, possono colpire e danneggiare i satelliti e colpiscono frequentemente anche la Stazione spaziale internazionale il cui equipaggio in caso di tempesta si ripara

EU Researchers Night,
TCD, Dublin, 30.09.2016

EU Researchers Night,
Athens, 30.09.2017

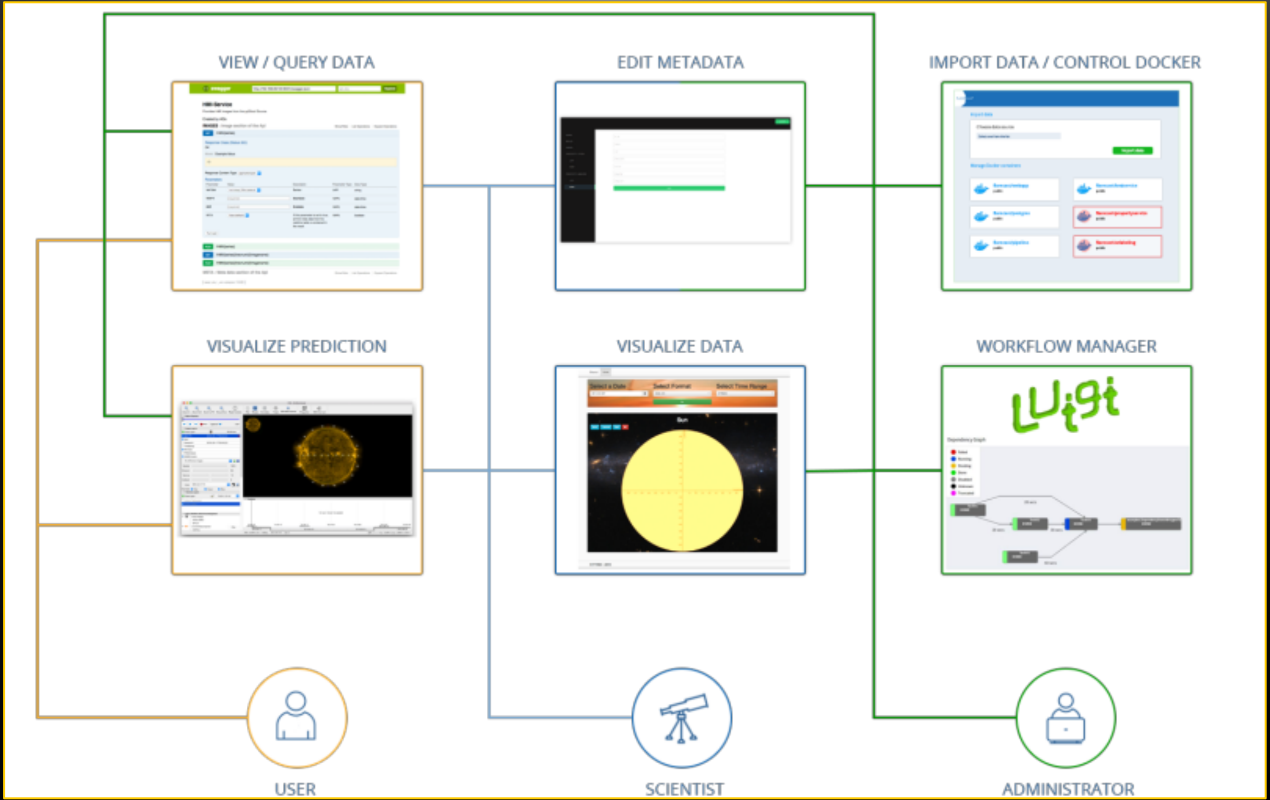
AA, Athens, 13.11.2017

Science Café,
Zurich, 11.11.2016

Fet de la Science, Paris, 16.10.2016

http://flarecast.eu/outreach-activities

FLARECAST service: how it will work



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

Summary

- FLARECAST is arguably the most systematic, cost- and effort-intensive solar flare prediction project worldwide at this time.

- The project has multi-faceted objectives that aim to understand and forecast the flaring phenomenon, at the same time raising awareness in the scientific community, stakeholders and the public about the challenge and benefit of the effort.

- FLARECAST data, codes and infrastructure are fully and openly accessible worldwide and can be used to avoid effort duplication in future SWx forecasting efforts.

- FLARECAST will be finalized in the next few months. Any idea, comment or suggestion that can be taken into account at this point is welcome.