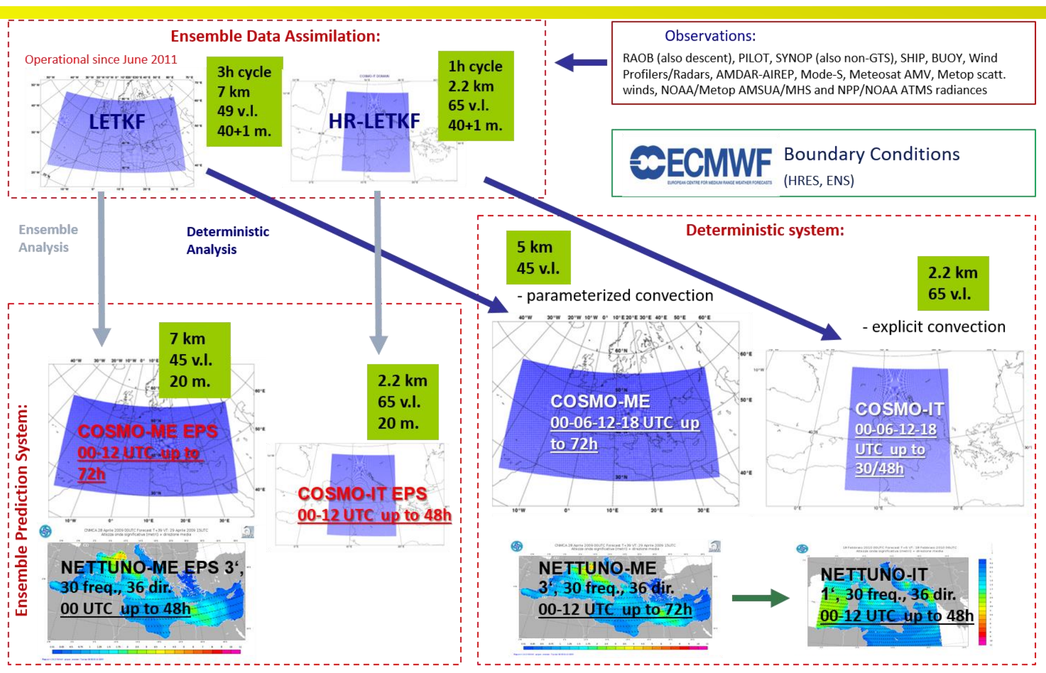
Post-Processing of NWP Model Output by Machine Learning Algorithms for Severe Weather Forecasting

Antonio Vocino[1], Francesca Marcucci[1], Raffaele Golino[2], Fabio Del Frate[3]

ABSTRACT: The potential of modern Machine Learning (ML) algorithms for the diagnosis of atmospheric instability and the early detection of meteorological strong convective systems is investigated in this study, recently conceived and hosted at the Italian Air-Force Meteorological Centre (COMET) and co-funded in the framework of EUMETNET-SRNWP-EPS Project (2019-23 Phase). The research focuses on the post-processing of NWP model output to provide the forecasters with improved Decision Support Systems, specifically designed for aviation hazards and severe weather phenomena, such as thunderstorms, fog, icing, turbulence. Preliminary results are shown, based on a set of significant case studies over Italy comparing new ML prognostic tool against classical multi-variate numerical model output. Further activities on this topic are planned, including more systematic objective verification campaigns and optimization of the ML engine algorithms.

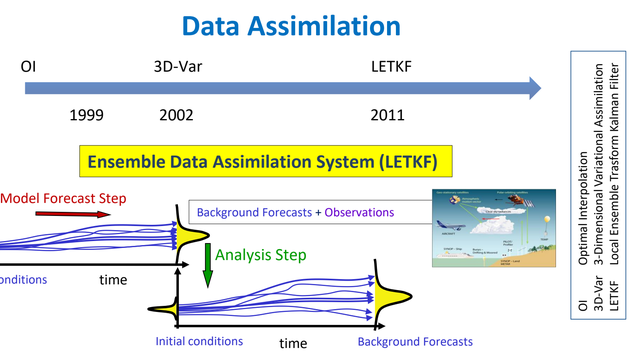
KEYWORDS: Numerical Weather Prediction, atmospheric models, Machine Learning, GPU accelerators, High Performance Computing.

NWP System



Operational NWP System

The Italian Air Force Meteorological Centre operates a complete NWP system, including an ensemble based data assimilation system and a set of nested, limited area atmospheric and wave models, in both deterministic and ensemble configurations, providing the high-resolution forecasting fields feeding the generation of timely and accurate meteorological products for the end users.



Quality control of incoming observations is performed on-line through ECMWF pre-processing tool (SAPP) and background QC in LETKF assimilation system

LIMITED AREA DETERMINISTIC MODEL – COSMO-IT	
MODEL	COSMO
Domain size	576x 701
Grid spacing	0.02° (2.2 km)
Number of layers / top	65 / ~22 Km
Time step and integration scheme	25 sec Runge-Kutta HE-VI time splitting
Forecast range/step	48 hrs/1 h
Initial time of model run	00/06/12/18 UTC
Lateral boundary conditions	IFS (ECMWF)
L.B.C. update frequency	1 hrs
Initial state	HR-LETKF deterministic analysis
Initialization	None
External analysis	snow cover, SST
Status / Hardware	Operational / HP Cluster Linux (COMET), CRAY (ECMWF)
ADDITIONAL FEATURES	Provides atmosferic forcing to the NETTUNO sea state model (WAM) 1' resolution

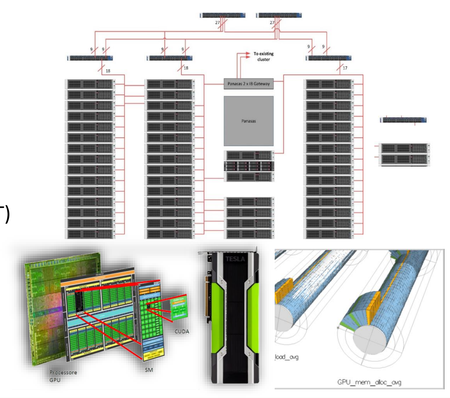
Mediterranean Sea Forecasting (NETTUNO)

in collaboration with ECMWF **and** ISMAR

Surface winds from **COSMO-ME** and **COSMO-IT** are used as atmospheric forcing in **WAM 4.0** model *(Komen et al, 1994)*

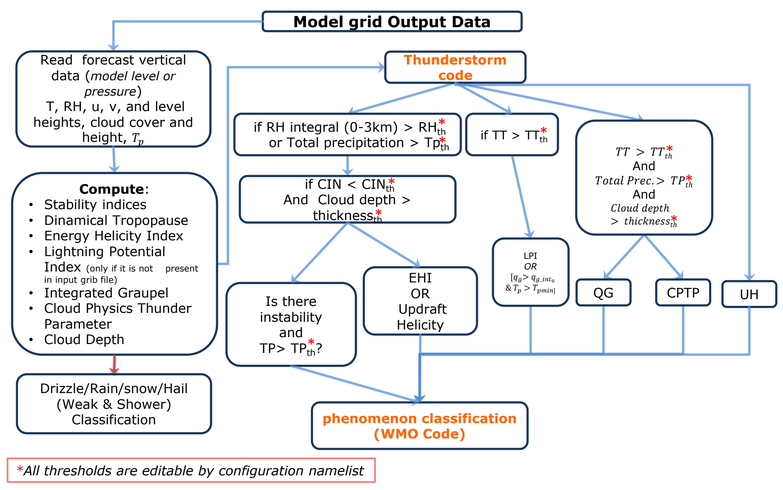
CONFIGURATION:
Lat-Lon regular grid, mesh size 3' (NETTUNO-ME) / 1' (NETTUNO-IT)
Spectral discretization with 30 frequencies and 36 directions
Initial state from previous run (warm start)
Initial time of model run 00/12 UTC
Forecast range to 72 h (NETTUNO-ME) / 48 h (NETTUNO-IT)
OUTPUT FIELDS:
Significant wave height, Mean wave direction, mean wave period

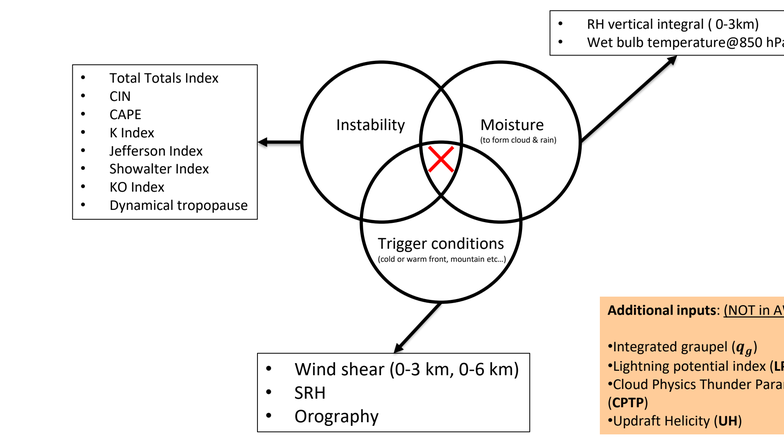
COMET HPC cluster

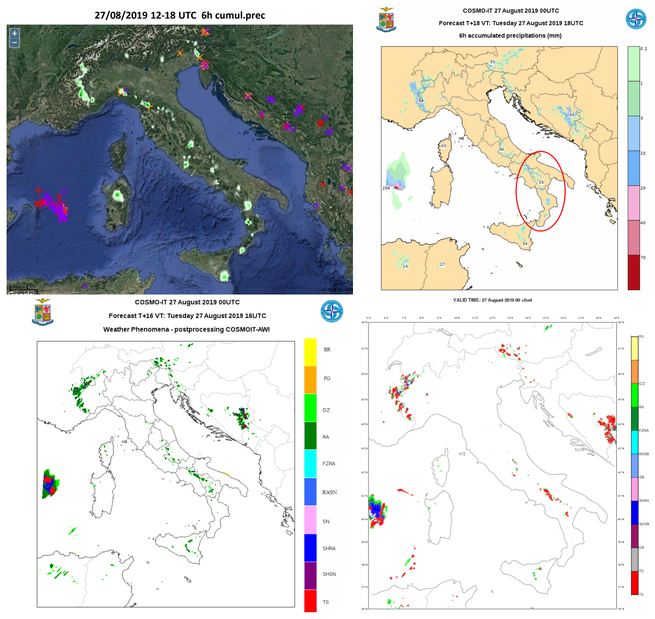


- 66 x DL380 G9 Computing Nodes (2 x 12 Haswell cores - 256 GB) + 132 x K80 GPUs
- 10 x DL380 G9 Computing Nodes (2 x 14 Haswell cores - 256 GB) + 20 x K80 GPUs
- 7 x DL380 G10 Computing Nodes (2 x 12 Haswell cores - 256 GB) + 14 x P100 GPUs
- 2 x DL380 G9 Management Nodes
- 4 x PANASAS AS12 (1 Director Blade each)
- 8 x Infiniband 36p FDR switches
- ~300 TFLOPS peak performance

Post-Processing







Decision-Tree algorithm for Thunderstorm

As major component of the NWP cascading system, a set of decision-tree based applications running at COMET compose the "Automatic Weather Interpreter" (AWI), providing the forecasters with an expert system able to integrate the direct model output fields and to predict the associated, most probable significant weather phenomenon. Such multi-class classifier relies on a very complex and optimized algorithm with customized thresholds and conditions. In particular, the Thunderstorm class is detailed here, as the baseline for the investigation of a new approach demonstrating how Machine Learning could enhance the exploitation of the information content of NWP model state.

Machine Learning

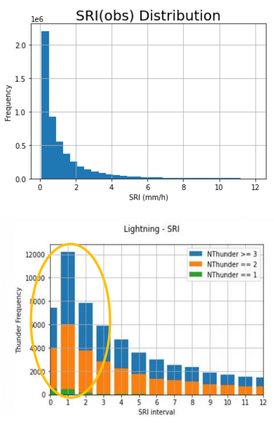
Datasets

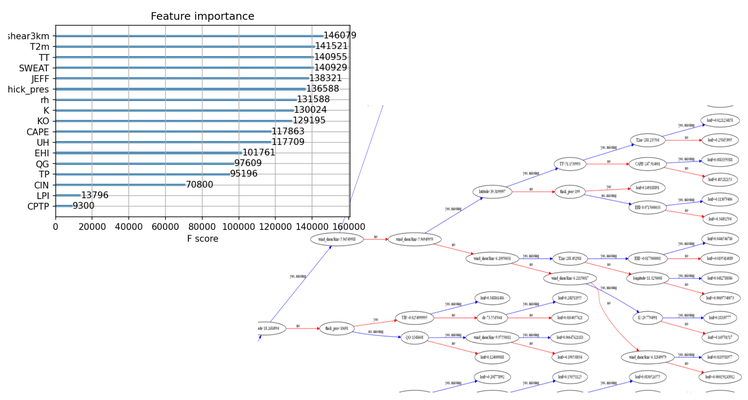
Predictors (called «features») used as INPUT for ML model to forecast the event (TS/SH):
✓ *full set of AWI/FOGTH predictors*

Observations used as TARGET for ML to train/validate and test the algorithm :
✓ *SRI and lightnings*

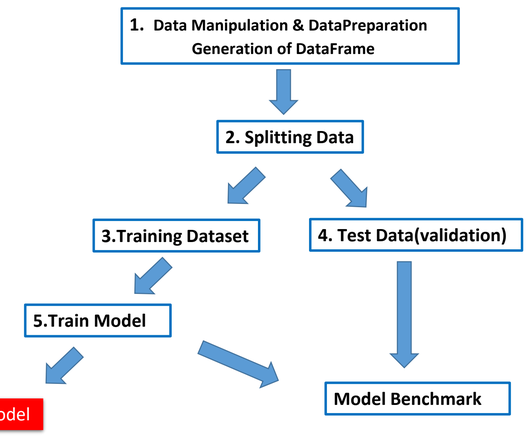
«Features» list:

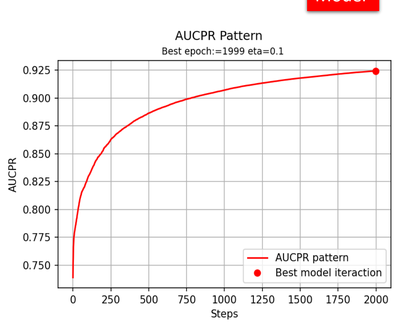
3km shear	Cloud dept	QG
T2m	K index	Total precipitation
Total Total index	KO index	CIN
SWAET index	CAPE	LPI
Jefferson Index	Updraft Helicity	CPTP index
RH	Energy Helicity Index	

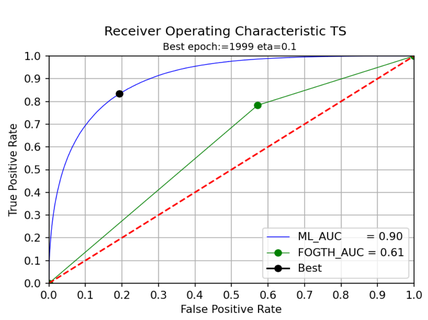




Steps







Results

Preliminary results show that the ML tool based on gradient boosting adaptive multi-index consensus algorithm set up as binary classifier for TS events outperforms traditional, static decision-tree post-processing driven by direct model output.

Libraries Starter Pack (free tools)

NumPy · python · scikit learn · pandas · jupyterhub · RAPIDS CUDF – CUML – CUGRAPH · matplotlib · DASK · dmlc XGBoost

Future work

Temporal extension and potential of new datasets of observations (e.g. METAR reports for aerodrome sites) will be evaluated to improve the target classification.
Furtherly, alternative methods to better balance the classes and select the input features will be implemented.
Finally, the application of the same methodology for other weather hazards (fog, icing, turbulence) is planned.

References
Official web site: http://www.meteoam.it
WMO Progress Report on the GDPFS and NWP research activities of the Italian Air Force Meteorological Service, available on-line at: https://community.wmo.int/wmo-technical-progress-report-global-data-processing-and-forecasting-system-gdpfs-and-numerical-weather-prediction-nwp-research-2020

Affiliations
1) COMET – Italian Air Force Meteorological Centre - Via Pratica di Mare 45, 00071 Pomezia (Rome), Italy
2) Geo-K - Via Del Politecnico 1, 00133 Rome, Italy
3) University of Rome Tor Vergata - Via Del Politecnico 1, 00133 Rome, Italy

Corresponding author: *antonio.vocino@aeronautica.difesa.it*