UNIVERSITÀ DEGLI STUDI DI BARI
"A. Moro"

Dipartimento Interateneo di Fisica "M. Merlin"

SCUOLA DI DOTTORATO DI RICERCA IN FISICA

CICLO XXXI

Settore Scientifico Disciplinare FIS/07

Tecniche modellistiche e statistiche a supporto delle previsioni dei *Wind Days*

PhD Thesis by:

Dott. Andrea Tateo

Director of PhD Programme:

Prof. Giuseppe Iaselli

Tutor:

Prof. Roberto Bellotti

Co-Tutors:

Dott.ssa. Micaela Menegotto

Dott. Mario Marcello Miglietta

Introduction

Evaluation of the role played by physical parameterizations, by grid resolution and by uncertainty on the initial condition in meteorological medium term forecasting (predictions of wind, PBL height and Wind Days).

NUMERICAL WEATHER PREDICTION MODELS

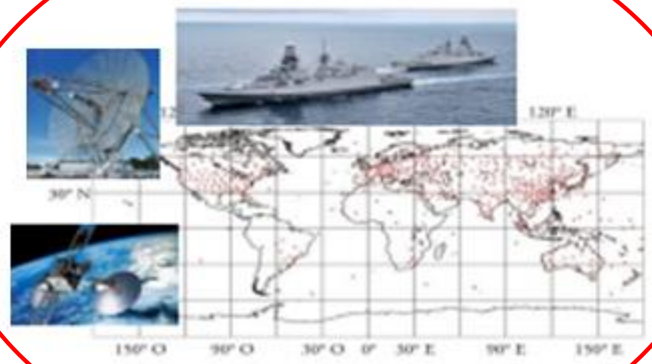
WHAT ARE THEY?

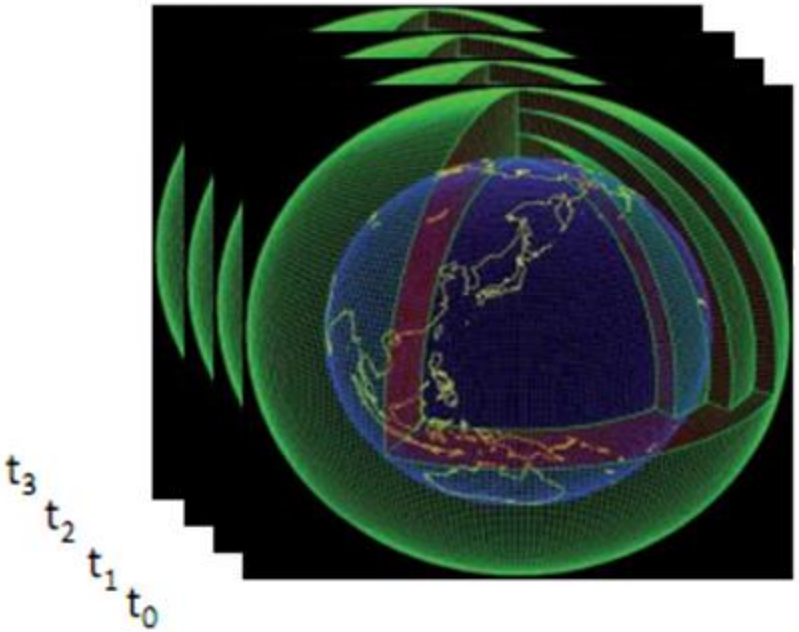
- NWP models are algorithms that solve systems of differential equations that represent the Atmosphere
- They allow **short** (*nowcasting - to 6h*), **medium** (to 3 or 5 days) and **long-term** (seasonal) forecasting
- Depending on the spatial and time resolution NWP models are meant to be Global (as GFS, ECMWF) and Regional (as WRF, COSMO)

NEED OF

INITIAL and **BOUNDARY** conditions + **OROGRAPHY**

GENERAL SCHEME …





Forecasts on regular spatial and temporal grid, with resolutions of:
Δx: 0,5km : 1 km
Δt: 1 minute : 1h

Complex interpolation algorithms
+
Global Model

Forecasts on regular spatial and temporal grid, with resolutions of:
Δx: 16km : 50 km
Δt: 3h :6h

Regional Model

ISSUES

1. Initial and boundary conditions are not exactly known

   Measurements are:

   - Inhomogeneous (different instruments for the same variable)
   - Asynchronous (not at the same time)
   - Not uniformly distributed
   - Measurements are not available for each point of the spatial domain
   - Instrumental errors

2. Moreover, the real domain is continuous but the computational domain is discrete;

3. In order to solve numerical equations, physical approximations are needed (parameterizations).

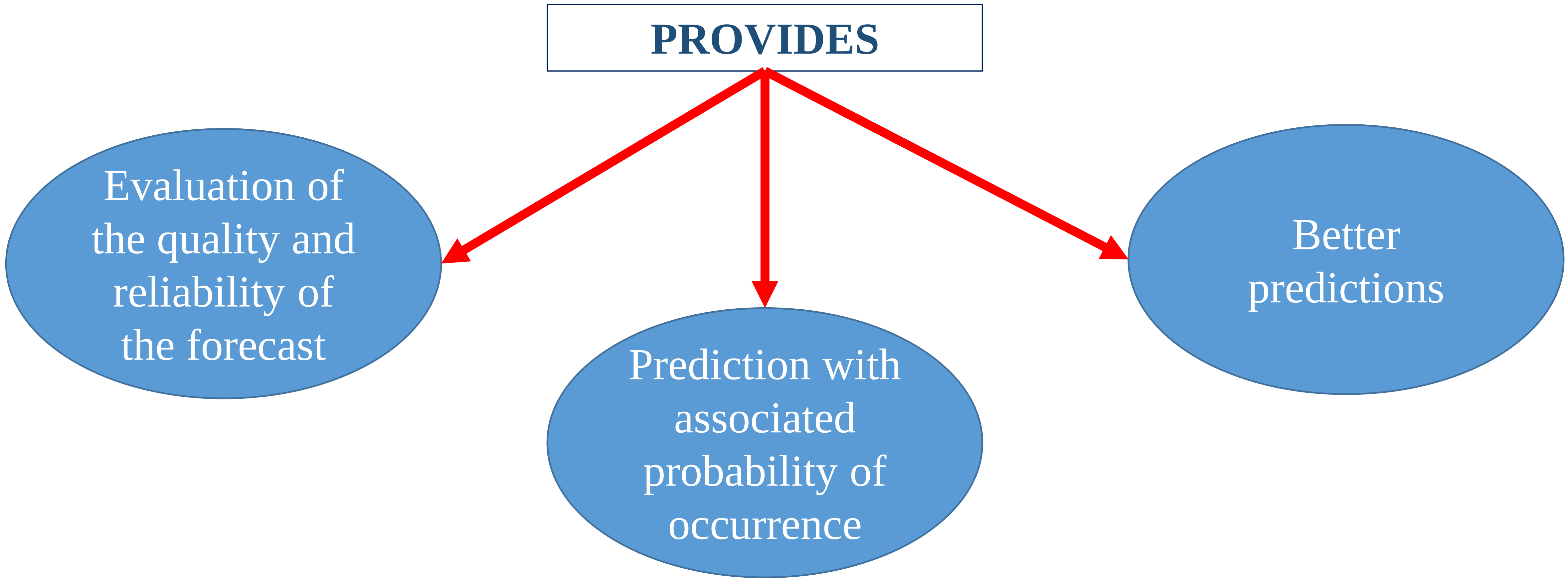
STATISTICAL APPROACH: the ensembles

In the “Chaos theory” of E. Lorenz[1],
the Atmosphere is defined as “chaotic system”,
i.e.
very small changes in the initial state could greatly affect the evolution of system.

In order to take into account the uncertainty in the initial conditions and to improve the deterministic prediction
in the last few years “probabilistic” forecasting systems have been developed:
EPS
(Ensemble Prediction System)

[1]Lorenz, E. N. (1995). *The essence of chaos*. University of Washington Press.

STATISTICAL APPROACH: the ensembles



Analysis

Analysis 1: Ensemble of parameterizations

[Predicted atmospheric variables: WS10, WD10, height of PBL]

a) Performance evaluations of 5 different parameterization used in WRF model.
b) Post-processing technique to improve deterministic predictions.

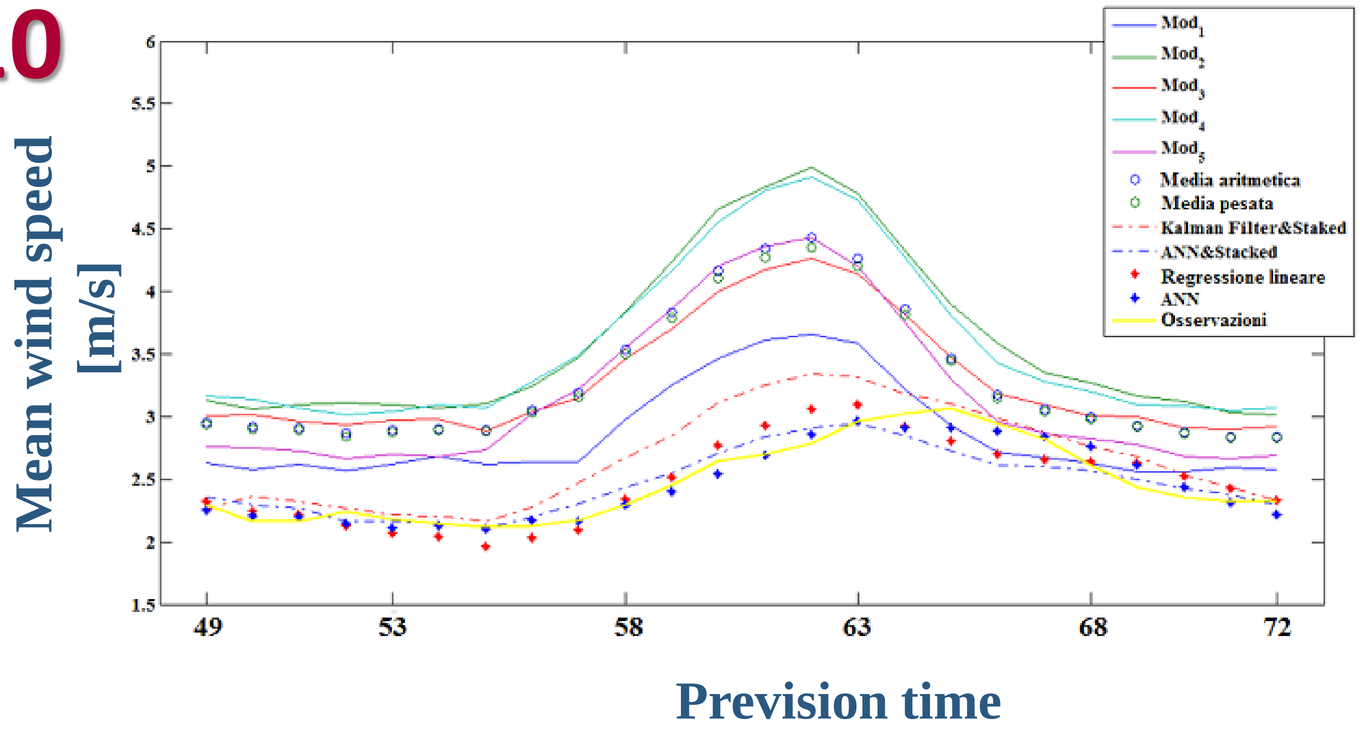
Analysis 2: Ensemble of perturbations

[Wind Day prediction]

a) Role evaluation of the grid resolution
b) Comparison of deterministic and probabilistic forecasts
c) Comparison of categorical and probabilistic approach
d) Evaluation and validation of a new method to estimate the probability of occurrence

Analysis 1: Ensemble of Parameterization

WS10



Time of analysis
08/2015-03/2016

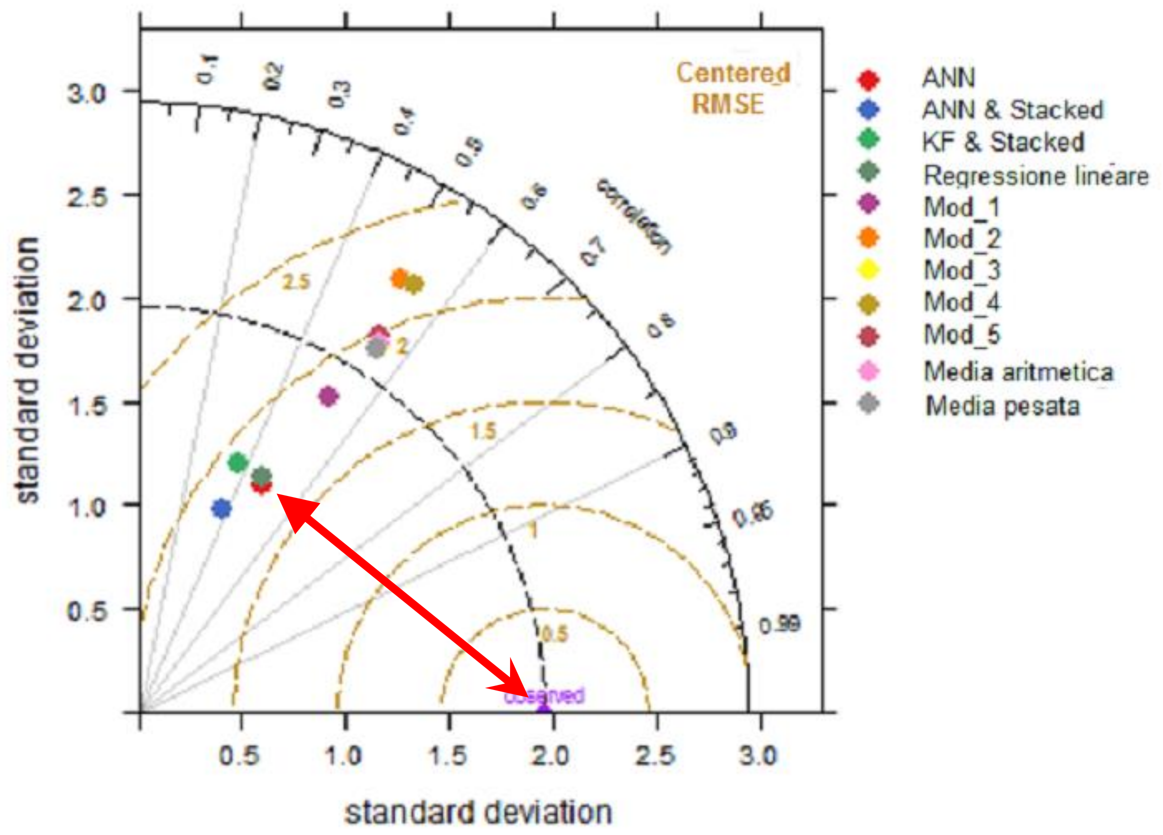
Data
WRF a 4 Km con input GFS
+72h (49^{th}-72^{th})

Average	KF & ANN	ANN
Weighted Average	ANNs & ANN	Linear Regression

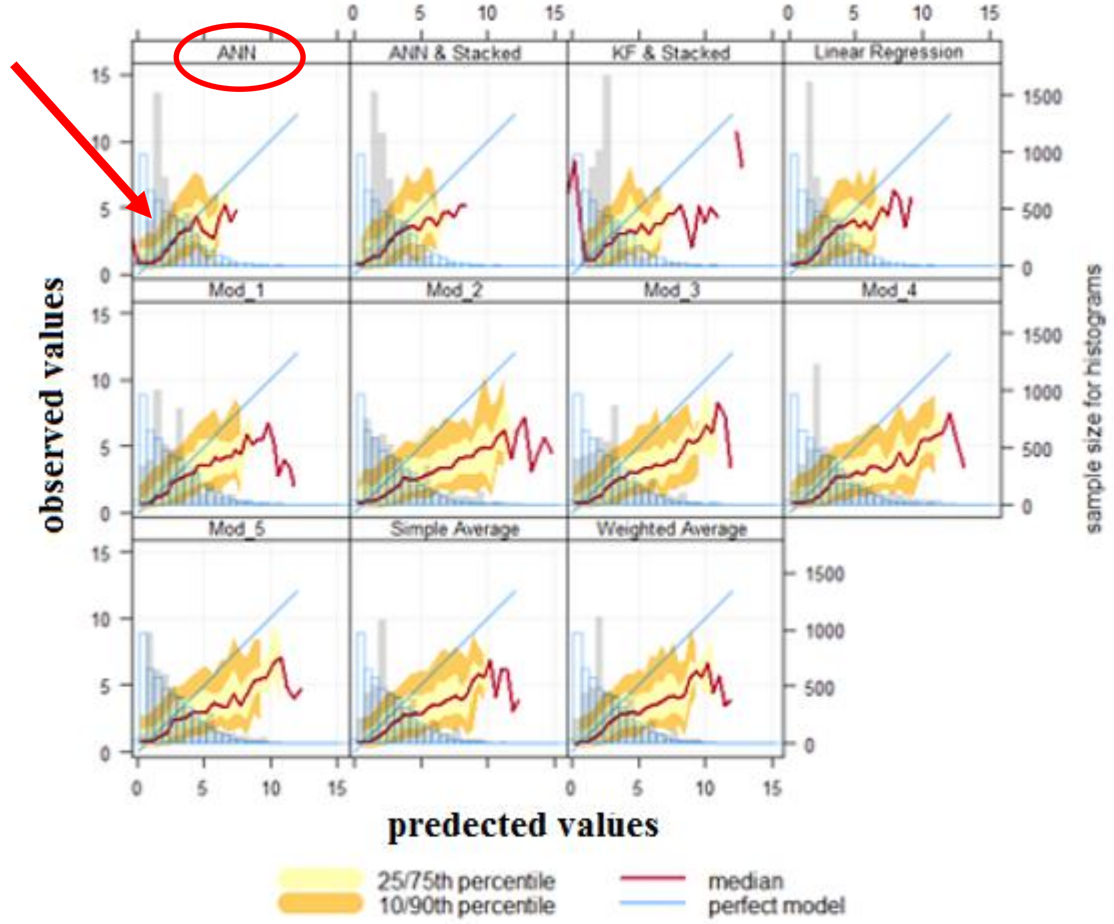
ANALYSIS 1: ENSEMBLE OF PARAMETERIZATION

WS10

Taylor Diagram



Conditional Quintile Plot



ANALYSIS 1: ENSEMBLE OF PARAMETERIZATION

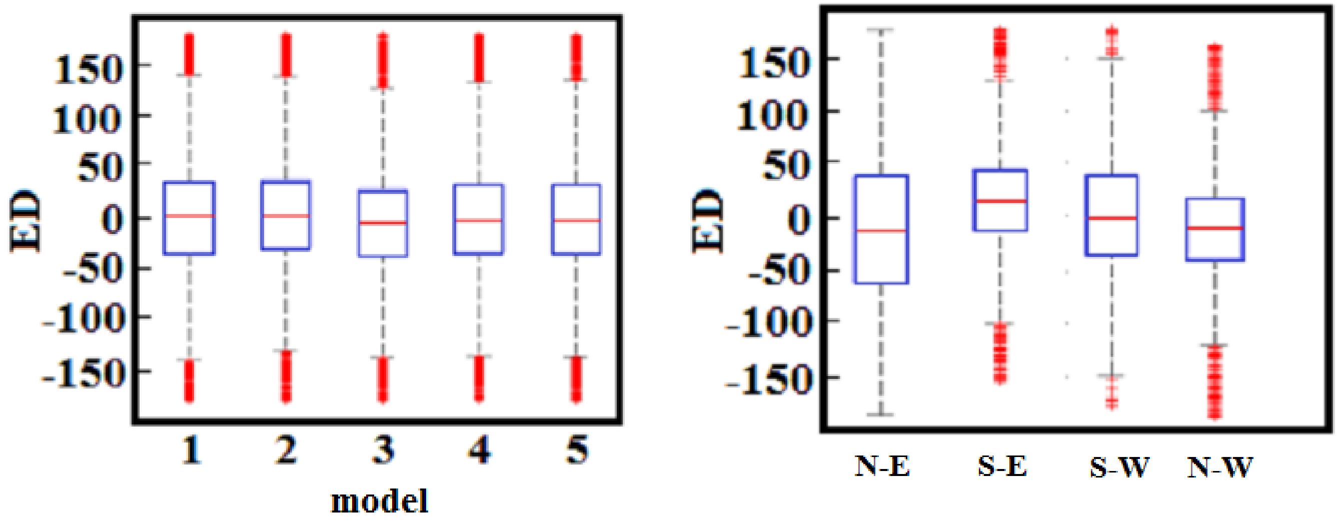
WD10

TIME OF ANALYSIS
08/2015-03/2016

DATA
WRF A 4 KM CON INPUT GFS
+72H (49TH-72TH)

$$ED = \begin{cases} WD_{pred} - WD_{obs} & \text{if } |WD_{pred} - WD_{obs}| \leq 180^o \\ -|360 - |WD_{pred} - WD_{obs}|| & \text{if } (WD_{pred} - WD_{obs}) > 180^o \\ +|360 - |WD_{pred} - WD_{obs}|| & \text{if } (WD_{pred} - WD_{obs}) < -180^o \end{cases}$$

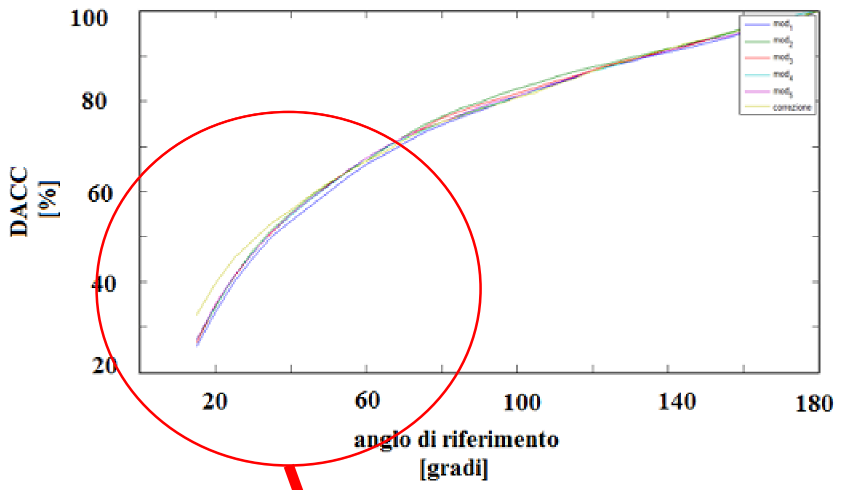
MODELL 1

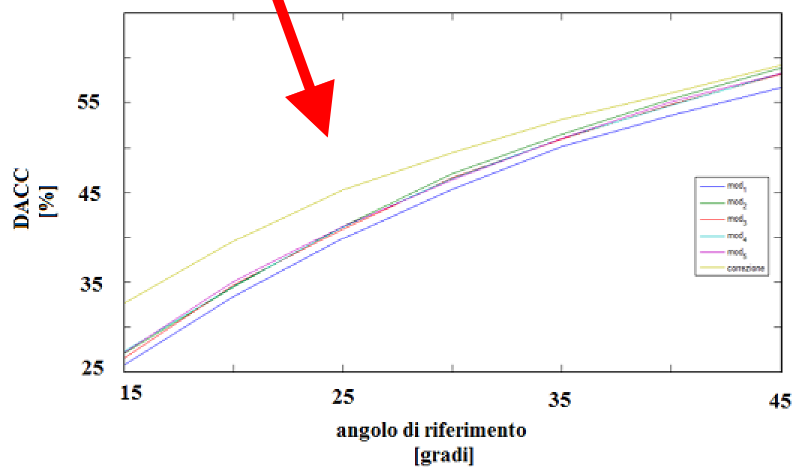


ANALYSIS 1: ENSEMBLE OF PARAMETERIZATION

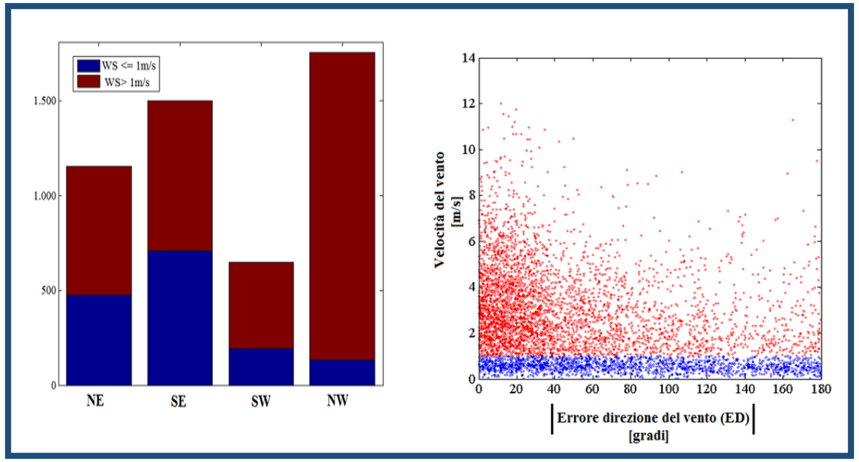
WD10

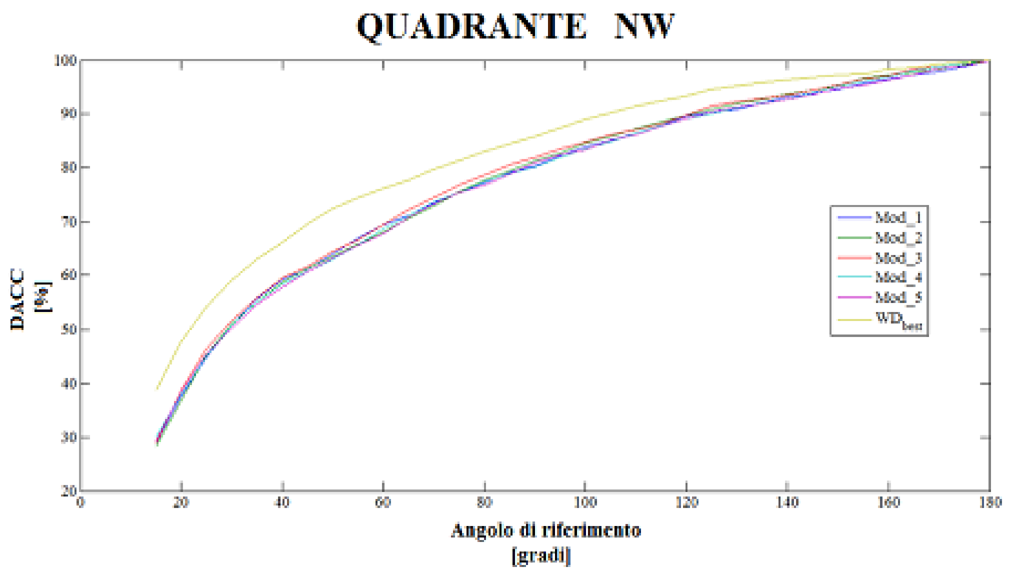
Method
Partial Average & Linear Regression





$$DACC(\theta_{Riff}) = \frac{\#\ \text{casi con}\ |ED| < \theta_{Riff}}{\#\ \text{totale casi}} \cdot 100$$

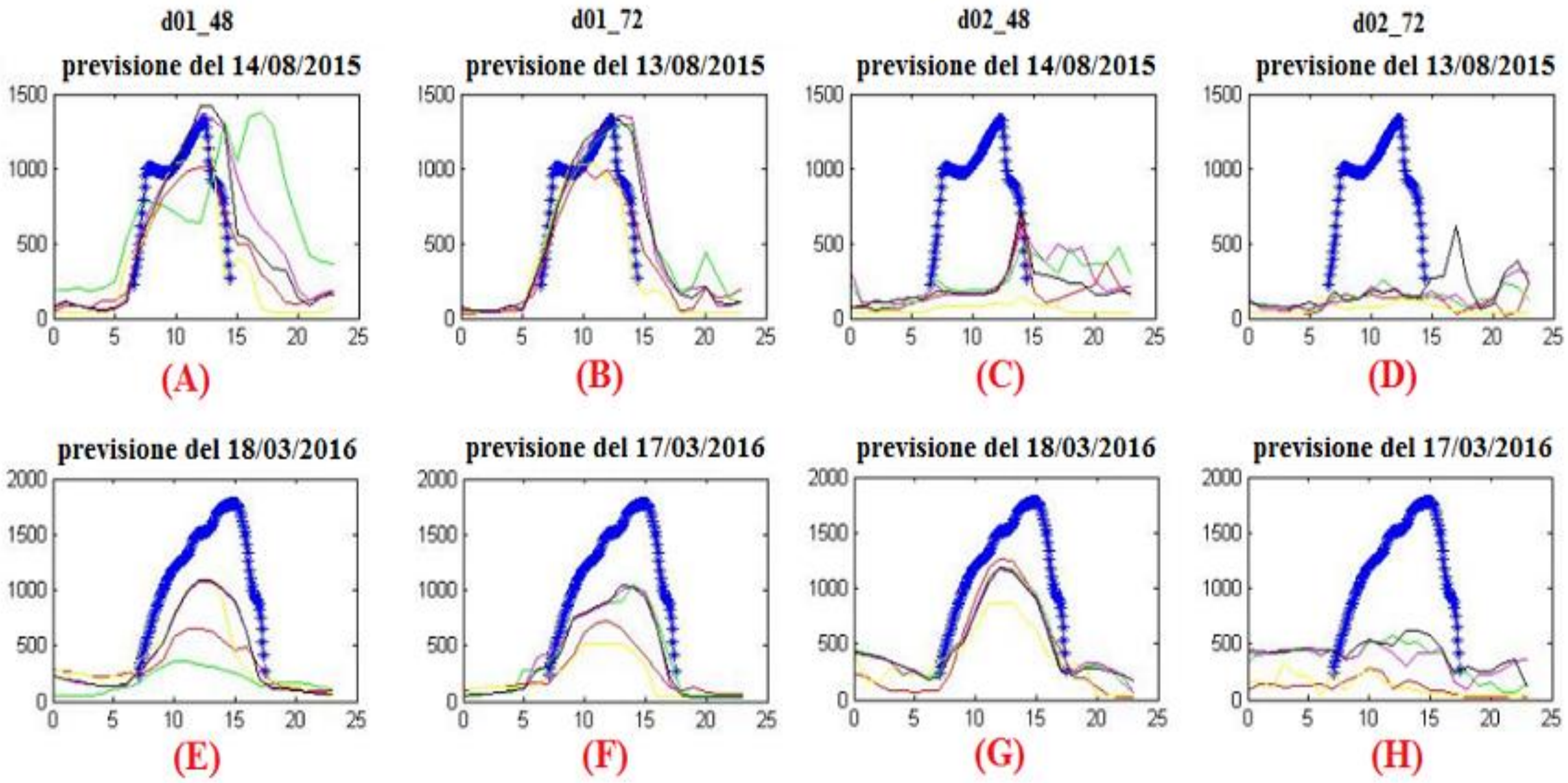




Casi con
WS > 1m/s

Analysis 1: Ensemble of Parameterization

Height PBL:



Time of analysis
12 days

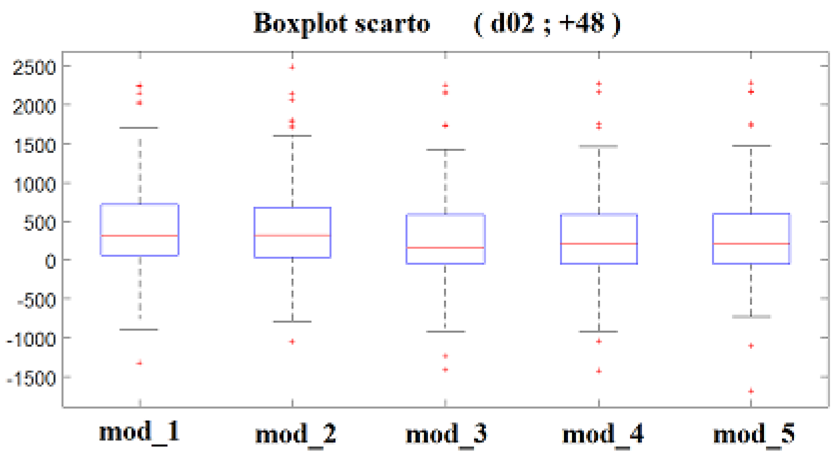
Lidar data
on
Taranto city

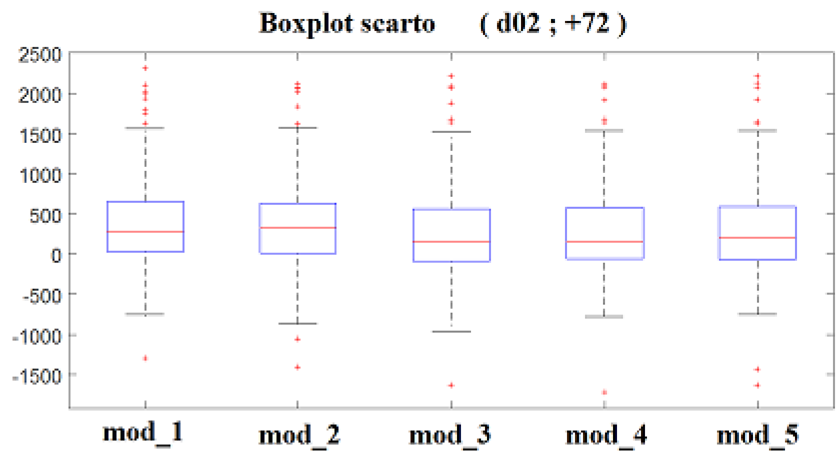
Analysis 1: Ensemble of Parameterization

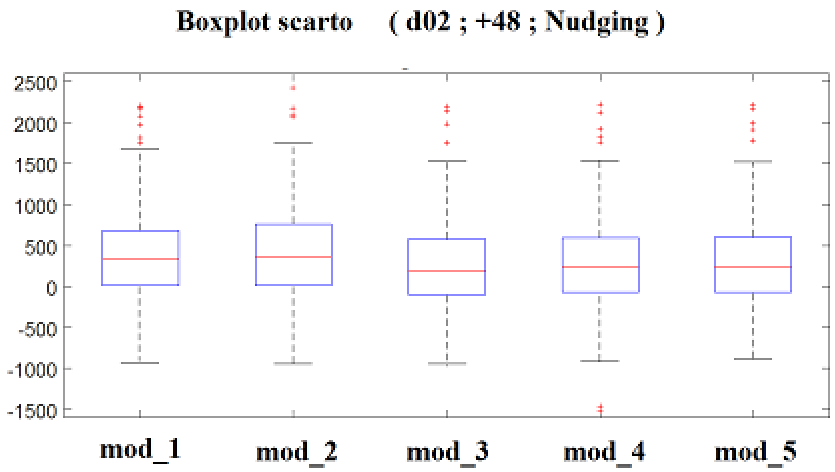
Height PBL

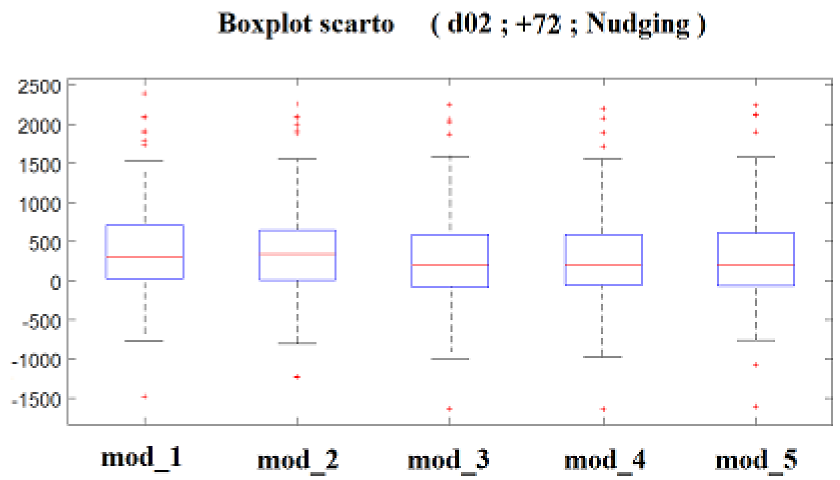
Time of analysis 08/2015-03/2016

Radiosounding on Brindisi city









Analysis 2: Ensemble of Perturbations

Wind Day Predictions

Categorical

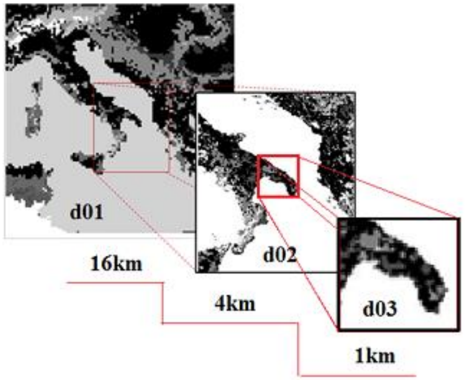
1. **GFS;**
2. **ECMWF;**
3. **c00 (control member);**
4. **ensemble mean**

Grid resolution

1. Global simulations

2. Regional simulations:

d01: 16 Km
d02: 4 Km
d03: 1 Km



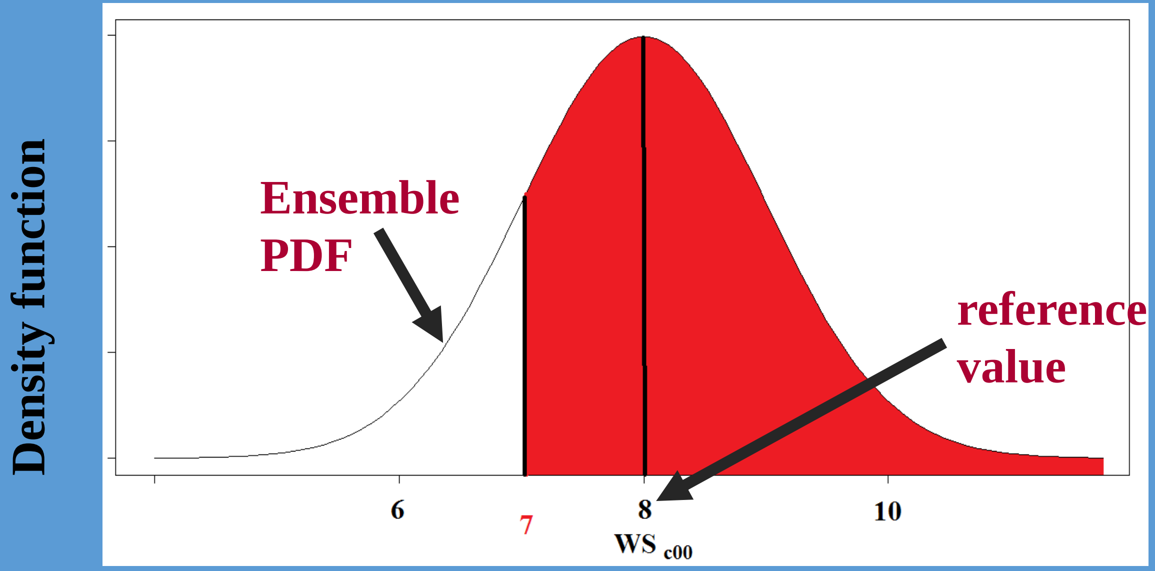
Probabilistic

GEFS – GFS *ensembles*

- **Percentage of ensemble members predicting Wind days**
- **Ensemble distribution percentiles**
- **Ensemble density function (new approach)**

ANALYSIS 2: ENSEMBLE OF PERTURBATIONS

OUR PROPOSAL



We propose an alternative method to estimate the occurrence probability

$$P(E_h) = 1 - C_h(\tilde{x})$$

$$C_h(\tilde{x}) = \int_{-\infty}^{\tilde{x}} x_h(u)du$$

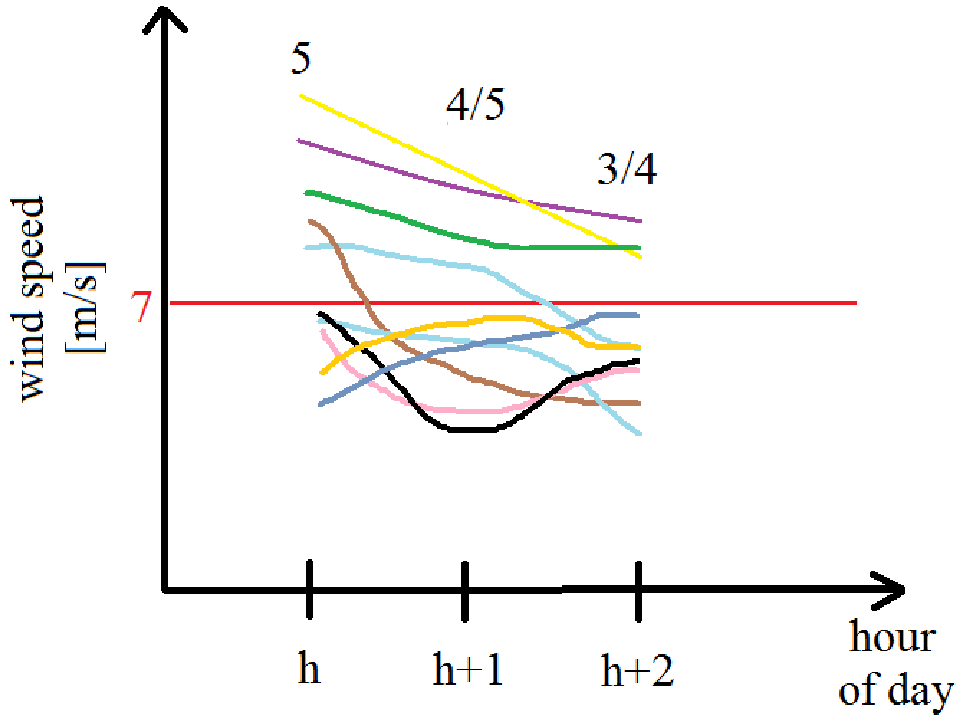
$$P(E_h \cap E_{h+1} \cap E_{h+2}) = P(E_{h+2}|E_h \cap E_{h+1}) \cdot P(E_{h+1}|E_h) \cdot P(E_h)$$

$$WDP_{pdf} = max_{h=49}^{70}[P(E_{h+2}|E_h \cap E_{h+1}) \cdot P(E_{h+1}|E_h) \cdot P(E_h)]$$

ANALYSIS 2: ENSEMBLE OF PERTURBATIONS

OUR PROPOSAL

The conditional probabilities are empirically estimated



$$P(E_{h+1}|E_h)$$

Percentage of members larger than 7m/s at the first hour, again larger than 7 m/s in the second hour [4/5]

$$P(E_{h+2}|E_h \cap E_{h+1})$$

Percentage of members larger than 7m/s for both first and second hours, again larger than 7 m/s in the third hour [3/4]

Analysis 2: Ensemble of Perturbations

WRF
(Regional model)

Categorical approach

Accuratezza [%]	*d*01	*d*02	*d*03
GFS	62	56	56
ECMWF	68	65	59
*c*00	62	56	56
media ensemble	65	62	56

Materials

Training and Test analysis:
34 possible WD in 2016

Validation analysis:
66 possible WD in 2017

Probabilistic approach

Accuratezza [%]	*d*01	*d*02	*d*03
Percentuale	71	74	71
Quantili	71	71	74
PDF	73	79	79

GFS **GFS** **(GFS, ensMean)**

Casual Error

ANALYSIS 2: ENSEMBLE OF PERTURBATIONS RESULTS FOR TRAINING

Global model

DATA	BIAS	POD	FAR	TS	ETS	accuracy
GFS	1.88	1.00	0.47	0.53	0.12	0.59
ECMWF	0.75	0.50	0.33	0.40	0.16	0.65
control member (GEFS)	1.50	0.94	0.38	0.60	0.27	0.71
ensemble mean (GEFS)	1.69	1.00	0.41	0.59	0.23	0.68
Expected value	1	1	0	1	1	1

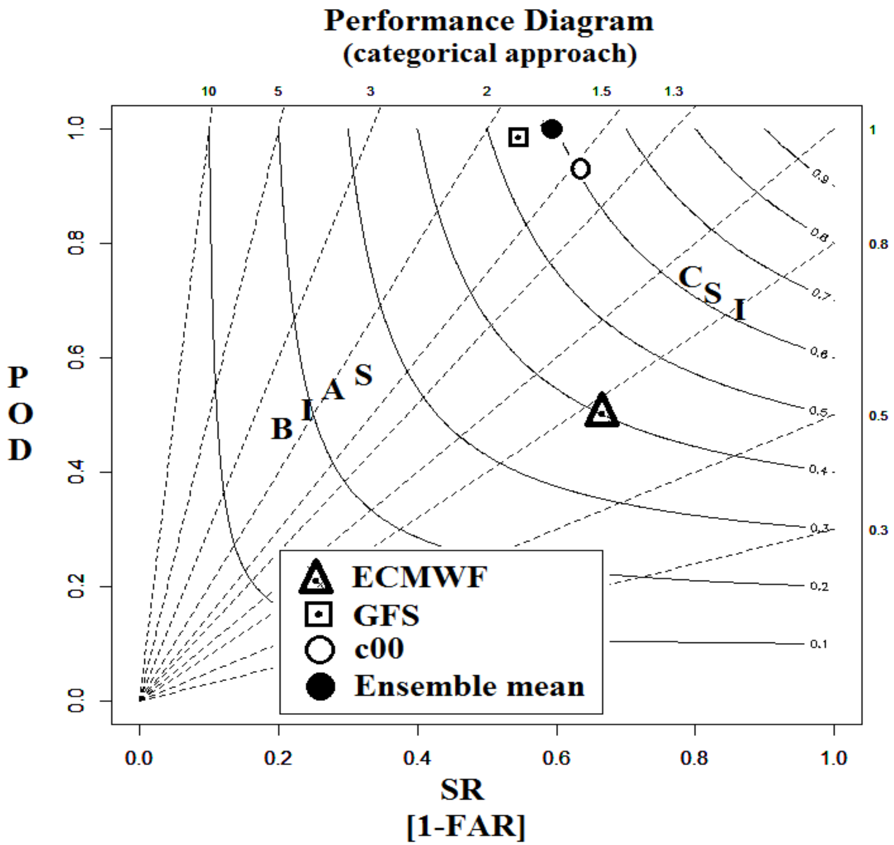
Probabilistic method	BIAS	POD	FAR	TS	ETS	accuracy	BS	AUC
percentile	1.31	0.94	0.29	0.68	0.42	0.79	0.24	0.80
Quantile	1.19	0.88	0.26	0.67	0.42	0.79	0.22	0.78
Expected value	1	1	0	1	1	1	0	1

reference of true value	distribution	BIAS	POD	FAR	TS	ETS	accuracy	BS	AUC
deterministic GFS	Normal	1.63	1.00	0.38	0.62	0.27	0.71	0.33	0.69
deterministic GFS	Gamma	1.19	0.75	0.37	0.52	0.22	0.68	0.35	0.69
control member (c00)	**Normal**	**1.31**	**0.94**	**0.29**	**0.68**	**0.42**	**0.79**	**0.25**	**0.77**
control member (c00)	**Gamma**	**1.25**	**0.94**	**0.25**	**0.71**	**0.48**	**0.82**	**0.25**	**0.80**
Average(GFS,c00)	Normal	1.56	1.00	0.36	0.64	0.32	0.74	0.29	0.73
Average(GFS,c00)	**Gamma**	**1.44**	**1.00**	**0.30**	**0.70**	**0.43**	**0.79**	**0.29**	**0.78**
Average(GFS,ensMean)	Normal	1.50	1.00	0.33	0.67	0.37	0.76	0.30	0.73
Average(GFS,ensMean)	Gamma	1.38	0.94	0.32	0.65	0.37	0.76	0.27	0.75
Expected value	// //	1	1	0	1	1	1	0	1

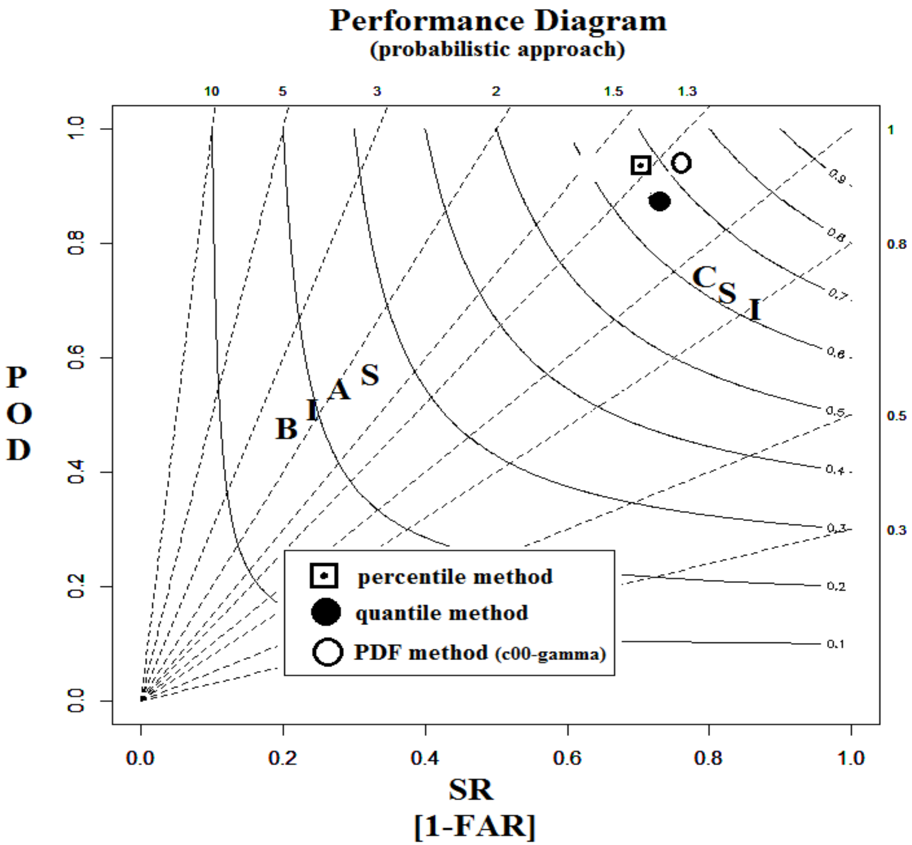
Casual Error

Analysis 2: Ensemble of Perturbations RESULTS for TRAINING

CATEGORICAL



PROBABILISTIC



Systematic Error

Analysis 2: Ensemble of Perturbations
RESULTS for Training
BY CHANGING THE WIND DAY DEFINITION

New WD definition: change wind speed threshold and number of consecutive hours.

data or method		accuracy 7m/s-3h	new threshold [m/s]	new # hours	Best accuracy	accuracy gain [%]
GFS		0.59	9	4	0.82	39
ECMWF		0.65	6	3	0.76	17
control member (GEFS)		0.71	8	5	0.88	24
Ensemble mean (GEFS)		0.68	8	3	0.85	25
percentile		0.79	8	3	0.88	11
quantile		0.79	9	5	0.88	11
PDF-deterministic GFS	Normal	0.71	9	3	0.88	24
PDF-deterministic GFS	Gamma	0.68	9	4	0.88	29
PDF-control member (c00)	Normal	0.79	8	5	0.88	11
PDF-control member (c00)	Gamma	0.82	9	2	0.88	7
PDF-Average(GFS,c00)	**Normal**	**0.74**	**8**	**5**	**0.94**	**27**
PDF-Average(GFS,c00)	**Gamma**	**0.79**	**8**	**5**	**0.94**	**19**
PDF-Average(GFS,ensMean)	**Normal**	**0.76**	**8**	**5**	**0.94**	**24**
PDF-Average (GFS,ensMean)	Gamma	0.76	9	5	0.91	20

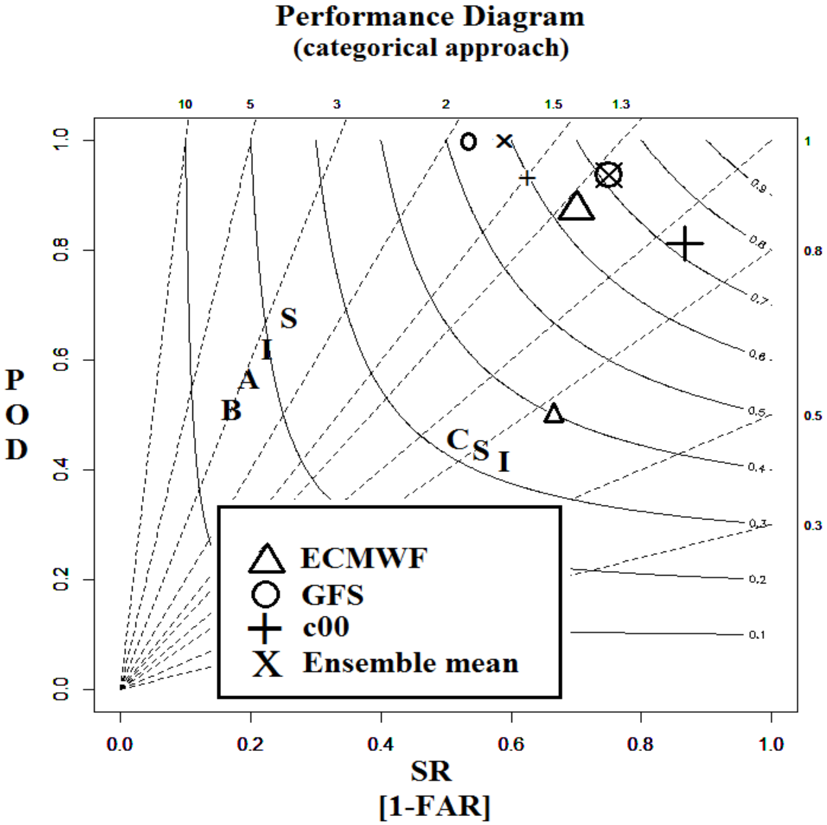
Systematic Error

ANALYSIS 2: ENSEMBLE OF PERTURBATIONS RESULTS FOR TRAINING

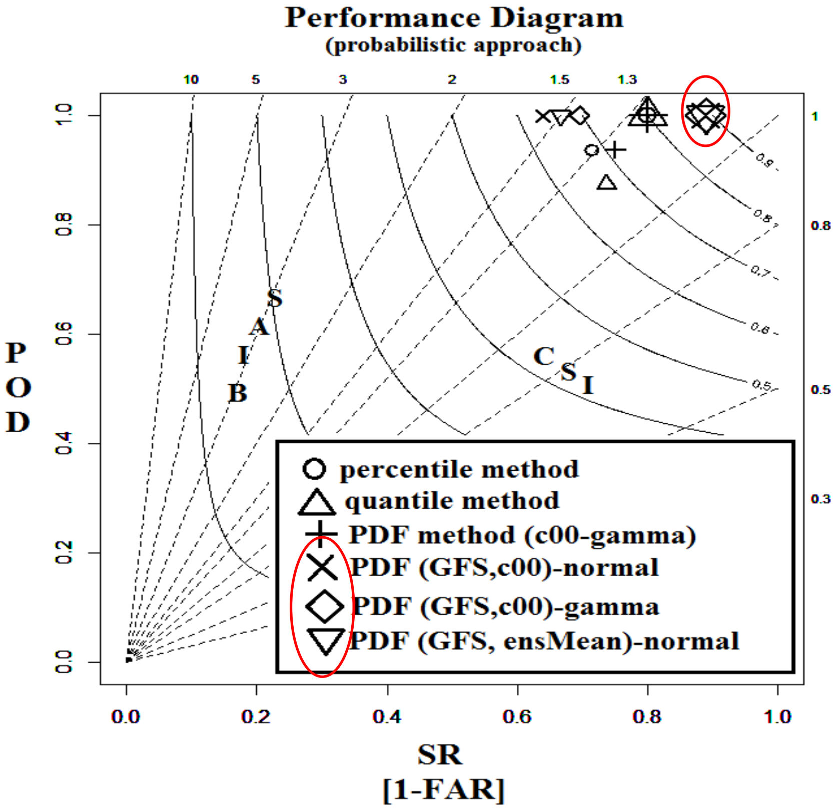
BY CHANGING THE WIND DAY DEFINITION

New WD definition: change wind speed threshold and number of consecutive hours.

CATEGORICAL



PROBABILISTIC



Result Comparison: Training, Test and Validation

data or method		Accuracy 7m/s-3h (classical WD definition)			Best accuracy (no classical WD definition)		
		Training (34 case studies - 2016)	Test (Leave-one-out 2016)	Validation (66 case studies - 2017)	Training (34 case studies - 2016)	Test (Leave-one-out 2016)	Validation (66 case studies - 2017)
GFS		0.59	0.59	0.68	0.82	0.82	0.79
control member (GEFS)		0.71	0.71	0.70	0.88	0.88	0.79
Ensemble mean (GEFS)		0.68	0.68	0.68	0.85	0.85	0.80
percentile		0.79	0.72	0.71	0.88	0.82	0.79
quantile		0.79	0.74	0.77	0.88	0.85	0.80
PDF-deterministic GFS	Normal	0.71	0.68	0.76	0.88	0.85	0.86
PDF-deterministic GFS	Gamma	0.68	0.64	0.74	0.88	0.82	0.86
PDF-control member (c00)	Normal	0.79	0.79	0.80	0.88	0.85	0.82
PDF-control member (c00)	Gamma	0.82	0.79	0.80	0.88	0.85	0.82
PDF-Average(GFS,c00)	Normal	0.74	0.75	0.79	0.94	0.85	0.85
PDF-Average(GFS,c00)	Gamma	0.79	0.77	0.79	0.94	0.85	0.85
PDF-Average(GFS,ensMean)	Normal	0.76	0.75	0.80	0.94	0.85	0.85
PDF-Average (GFS,ensMean)	Gamma	0.76	0.75	0.80	0.91	0.88	0.83

For test and validation analysis the best probability threshold, the best wind speed threshold and the best number of consecutive hours were estimated on the training set.

CONCLUSION AND FUTURE WORKS

CONCLUSION

1. The 5 considered parameterizations show a positive hourly mean bias in the WS10 prediction.
2. Among different post-processing techniques, the singular ANN shows better performance to combine the 5 considered simulations
3. The partial average and the linear regression is a good approach to reduce the error in the wind direction prediction.
4. The 5 parameterizations do not affect the PBL height prediction.
5. In general, the probabilistic approaches show better performances in the Wind Day prediction respect to the categorical approach.
6. Probabilistic predictions show better performances in the Wind Day prediction respect to the deterministic ones.
7. The new probabilistic method here proposed shows the best performances compared to the classical probabilistic approaches.

FUTURE WORK

In the last years, the data assimilation is used to improve the weather predictions.
I think interesting to test this approach on Taranto city in the Wind Day context.

CORSI ED ESAMI SOSTENUTI DURANTE IL PRIMO ANNO DI DOTTORATO

Corso	Docente
Management and knowledge of European research model and promotion of research results	Prof.ssa Alessia D'Orazio
How to prepare a technical speech in English	Prof.ssa Carmela White
LabView introductory Course	Prof. Fabio Gargano
Python course	Prof. Domenico Diacono
Complex Systems	Prof. Gianni Ferraro
Optical Sensors	Prof. Vincenzo Spagnolo
Statistical and computational models for physical data analysis	Prof.ssa Sonia Tangaro
Experimental Data Analysis and Comparisons to Theoretical Models	Prof. Alexis Pompili

Partecipazione a Scuole di Dottorato, Corsi e Seminari

- **Partecipazione alla Scuola “International Summer School EUMETSAT”**
- **Corso intensivo sulle GPU**
- **Percorso Formativo 24 CFU**
- **Corso relativo agli interventi su emissioni radon**

ATTIVITÀ DIDATTICA

1. Cultore della materia e in Commissione d’esame per il corso di “Scienze e gestione attività marittime (SGAM)” presso il Dipartimento Ionico (dal 2016 ad oggi).

2. Affiancamento per Laboratorio di Fisica Generale per Scienze Biologiche (referente del corso prof.ssa Ligonzo Teresa)

Partecipazione a Workshop e Conferenze

1. Ho partecipato al 16Th Convegno Internazionale (12-16 September 2016, Trieste) organizzato dalla EMS a Trieste presentando una ricerca dal titolo: "Ensemble of different planetary boundary layer schemes in WRF model for wind speed and direction prediction over Apulia region", attraverso un Poster ed una Presentazione orale tenuta il 15 settembre nella sessione "*Numerical aspects and physical parametrization integration in NWP (Numerical Weather Prediction) models*".

2. Ho partecipato al 17Th Convegno Internazionale (4-8 settembre 2017) organizzato dalla EMS a Dublino presentando una ricerca dal titolo: "WRF model simulations forced with GEFS: an ensemble approach for the prediction of wind speed near a complex coastline", attraverso un Abstact e un'Oral Presentation nella sessione "Applications of Meteorology and oceanography in the costal zone".

3. Ho partecipato al 11th HyMeX Workshop, 29 May - 2 June 2018, Lecce, Italy presentando una ricerca dal titolo: "A statistical method based on Ensemble probability density function for the prediction of "Wind Days", attraverso un Poster.

4. Ho partecipato all' "Intermediate Meeting nell'ambito del Progetto EphaStat" (6-7 settembre 2018) presentando l'analisi effettuata mediante un Oral presentation.

PUBBLICAZIONI

- Fedele, F., Guarnieri calò Carducci, A., Ottonelli, S., Turnone, A., Menegotto, M., TATEO, A., ... & Bellotti, R. (2016). Ottimizzazione del modello a mesoscala WRF per l'individuazione dei Wind Days nell'area di Taranto (DSpace – Digital Repository ARPA).

- Bellotti, R., Lombardi, A., Amoroso, N., TATEO, A., & Tangaro, S. (2016). Semi-unsupervised prediction for mild TBI based on Both Graph and K-nn methods. In Proceedings of the International Workshop on Brainlesion: mTOP. (attività collaterale)

- Allen, G. I., Amoroso, N., Anghel, C., Balagurusamy, V., Bare, C. J., Beaton, D., ... & Caberlotto, L. (2016). Crowdsourced estimation of cognitive decline and resilience in Alzheimer's disease. Alzheimer's & Dementia, 12(6), 645-653. (attività collaterale)

- Maglietta, R., Amoroso, N., Boccardi, M., Bruno, S., Chincarini, A., Frisoni, G. B., ... & Bellotti, R. (2016). Automated hippocampal segmentation in 3D MRI using random undersampling with boosting algorithm. Pattern Analysis and Applications, 19(2), 579-591. (attività collaterale)

- Tateo, A., Iurino, A., Settanni, G., Andrisani, A., Stifanelli, P. F., Larizza, P., ... & Bellotti, R. (2016). Hybrid x-space: a new approach for MPI reconstruction. Physics in Medicine & Biology, 61(11), 4061. (attività collaterale)

Pubblicazioni

- Tanzarella A., Morabito A., Schipa A., Intini F., Menegotto M., TATEO A., Pastor T., Tinarelli G., D'Allura A., Costa M.P., Giua R., Messa a punto di un Sistema modellistico previsionale della qualità dell'aria, progettato ad hoc sull'Area di Taranto per la previsione dei wind days. Rapporto Ambiente – SNPA, Ambiente in primo piano, 2017 (attività collaterale)

- Tateo, A., Miglietta, M. M., Fedele, F., Menegotto, M., Monaco, A., and Bellotti, R.: Ensemble using different Planetary Boundary Layer schemes in WRF model for wind speed and direction prediction over Apulia region, Adv. Sci. Res., 14, 95-102, https://doi.org/10.5194/asr-14-95-2017, 2017.

- La Rocca, M., Amoroso, N., Bellotti, R., Diacono, D., Monaco, A., Monda, A., ... & Tangaro, S. (2017). A multiplex network model to characterize brain atrophy in structural MRI. In Emergent Complexity from Nonlinearity, in Physics, Engineering and the Life Sciences (pp. 189-198). Springer, Cham. (attività collaterale)
- Tateo, A., Miglietta, M. M., Fedele, F., Menegotto, M., Pollice, A., & Bellotti, R. (2018). A statistical method based on the ensemble probability density function for the prediction of "Wind Days". Atmospheric Research.

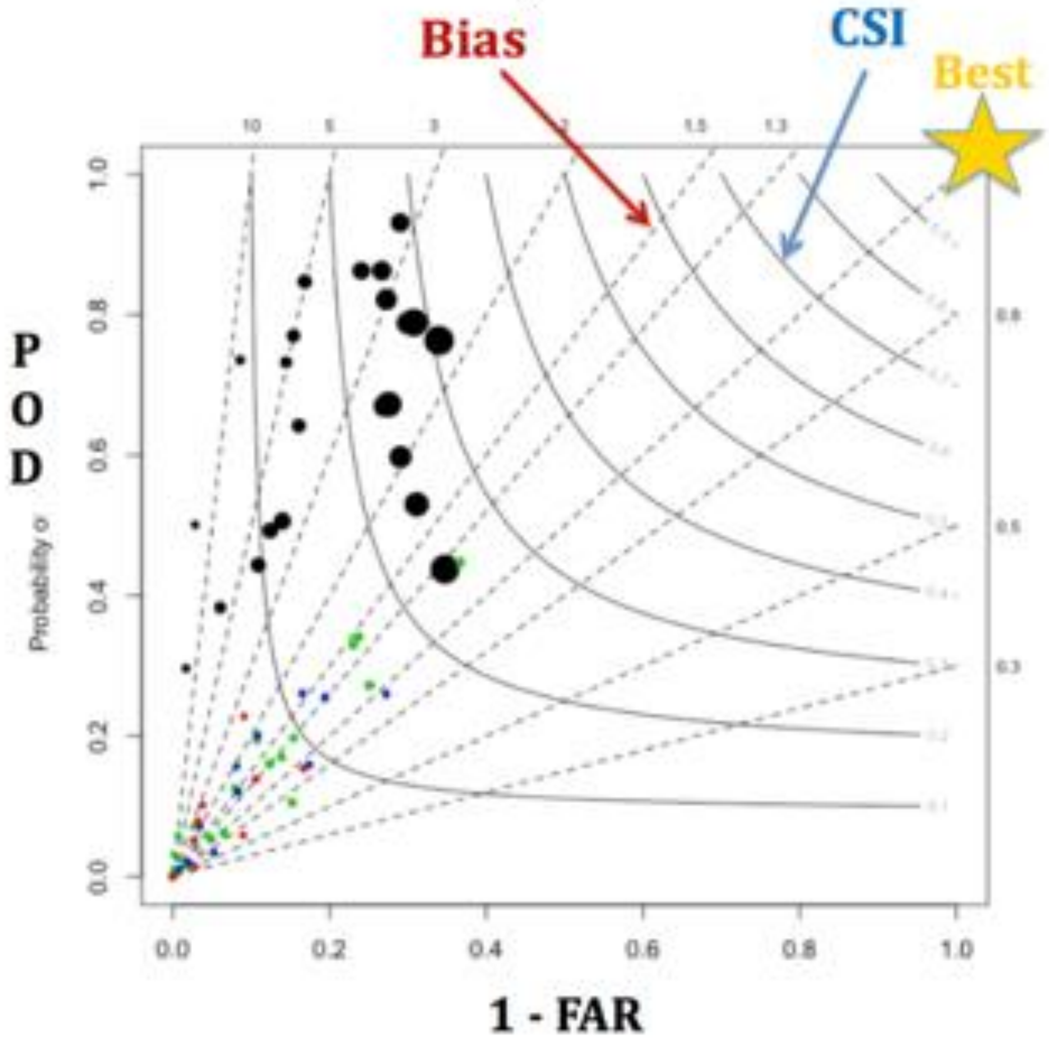
Thank you
for
attention

Evaluation indices

BIAS	POD	FAR	TS	ETS	accuracy
$\frac{Tp+Fp}{Tp+Fn}$	$\frac{Tp}{Tp+Fn}$	$\frac{Fp}{Tp+Fp}$	$\frac{Tp}{Tp+Fp+Fn}$	$\frac{Tp+H_Z}{Tp+Fp+Fn-H_Z}$ $H_Z=\frac{(Tp+Fn)(Tp+Fp)}{Tp+Tn+Fp+Fn}$	$\frac{Tp+Tn}{Tp+Tn+Fp+Fn}$

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

Performance diagram



STATISTICAL APPROACH: the ensembles

Since there are different typologies of error, there are also different criteria for estimating the ensembles:

a) Use different processors;

b) Use the same NWP system with different parameterizations;

c) Use different NWP systems;

d) Increase the spatial resolution (addressed in a previous work with the present case study).

Chaotic system simulation

Toy model by Edward Norton Lorenz

(Lorenz, Edward N., and K. Haman. "The essence of chaos." *Pure and Applied Geophysics* 147.3 (1996): 598-599)

System of differential equations

$$\frac{dx}{dt} = \sigma(y - x)$$

$$\frac{dy}{dt} = rx - y - xz$$

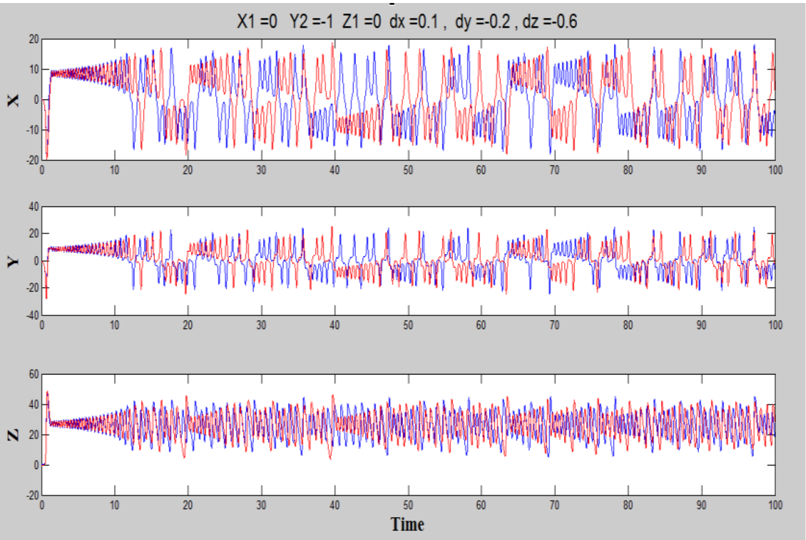
$$\frac{dz}{dt} = xy - bz$$

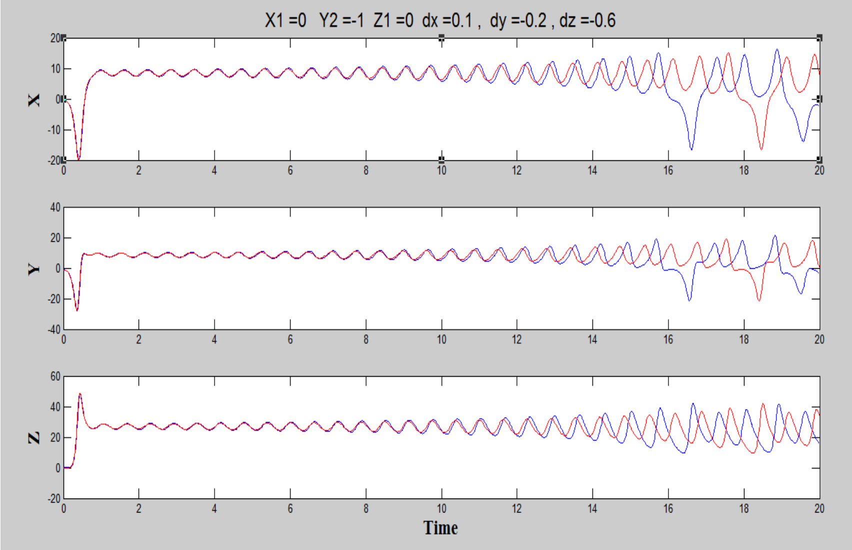
X → rotational speed

Y → horizontal temperature gradient of the cell extremes

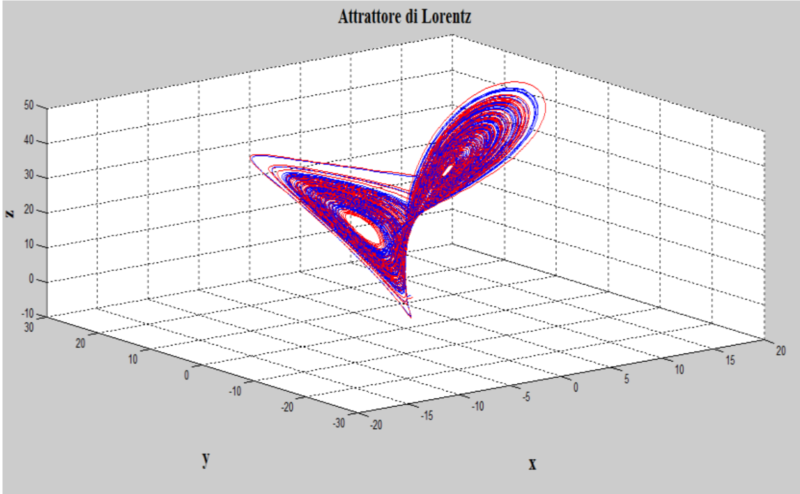
Z → vertical temperature gradient of the cell extremes

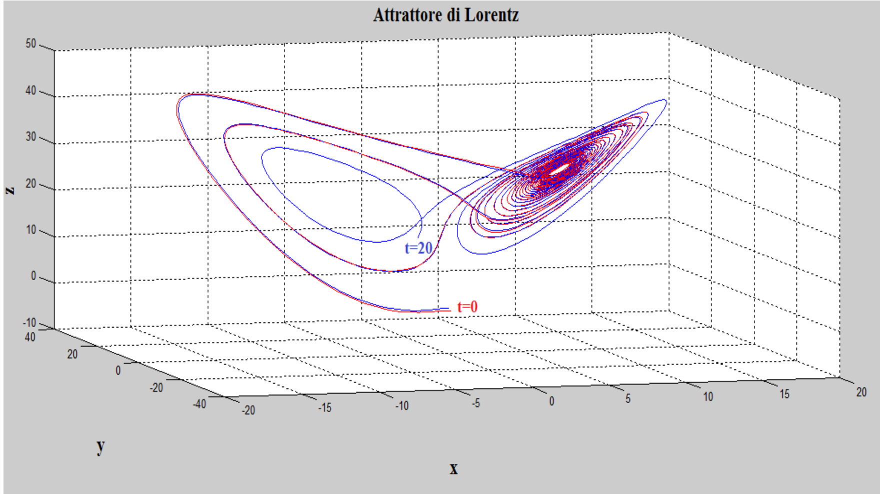
σ=10 b = 8/3 r = 28.





Chaotic system simulation





SUMMARIZING

Causes of errors on the initial condition

Direct and indirect, inhomogeneous, asynchronous measures that provide only an estimate of the different real state [in the phase space, the distance from real and measured state is $d\varepsilon$]

Other sources of error

Parametrizations
Spatial and temporal resolution too low

Parameterization schemes

Fixed parameters

Opzioni fisiche del WRF	Nome dello schema
Radiation Processes (*ra_lw_physics and ra_sw_physics*)	rrtm scheme [38] and Dudhia scheme [39]
Surface Processes (*sf_surface_physics*)	Noah Land Surface Model [40]
Cumulus Processes (*cu_physics*)	Kain-Fritsch scheme [41]
Microphysics Processes (*mp_physics*)	Thompson scheme [42, 43]

Variable parameters

Modello	Schema PBL	Schema SL
1	YSU	Revised MM5 Monin-Obukhov (nella v.6 Jimenez)
2	MYJ	Monin-Obukhov (Janjic Eta)
3	MYNN	Revised MM5 Monin-Obukhov (nella v.6 Jimenez)
4	MYNN	Monin-Obukhov (Janjic Eta)
5	MYNN	Mellor-Yamada-Nakanishi-Niino

Taylor Diagram

Correlation Coefficient

$$R = \frac{\frac{1}{N}\sum_{i=1}^{N}[(f_n - \bar{f}) \cdot (r_n - \bar{r})]}{\sigma_f \sigma_r}$$

Centered RMSE

$$E'^2 = \frac{1}{N}\sum_{i=1}^{N}[(f_n - \bar{f}) - (r_n - \bar{r})]^2$$

Standard Deviation of test model

$$\sigma_f^2 = \frac{1}{N}\sum_{i=1}^{N}(f_n - \bar{f})^2$$

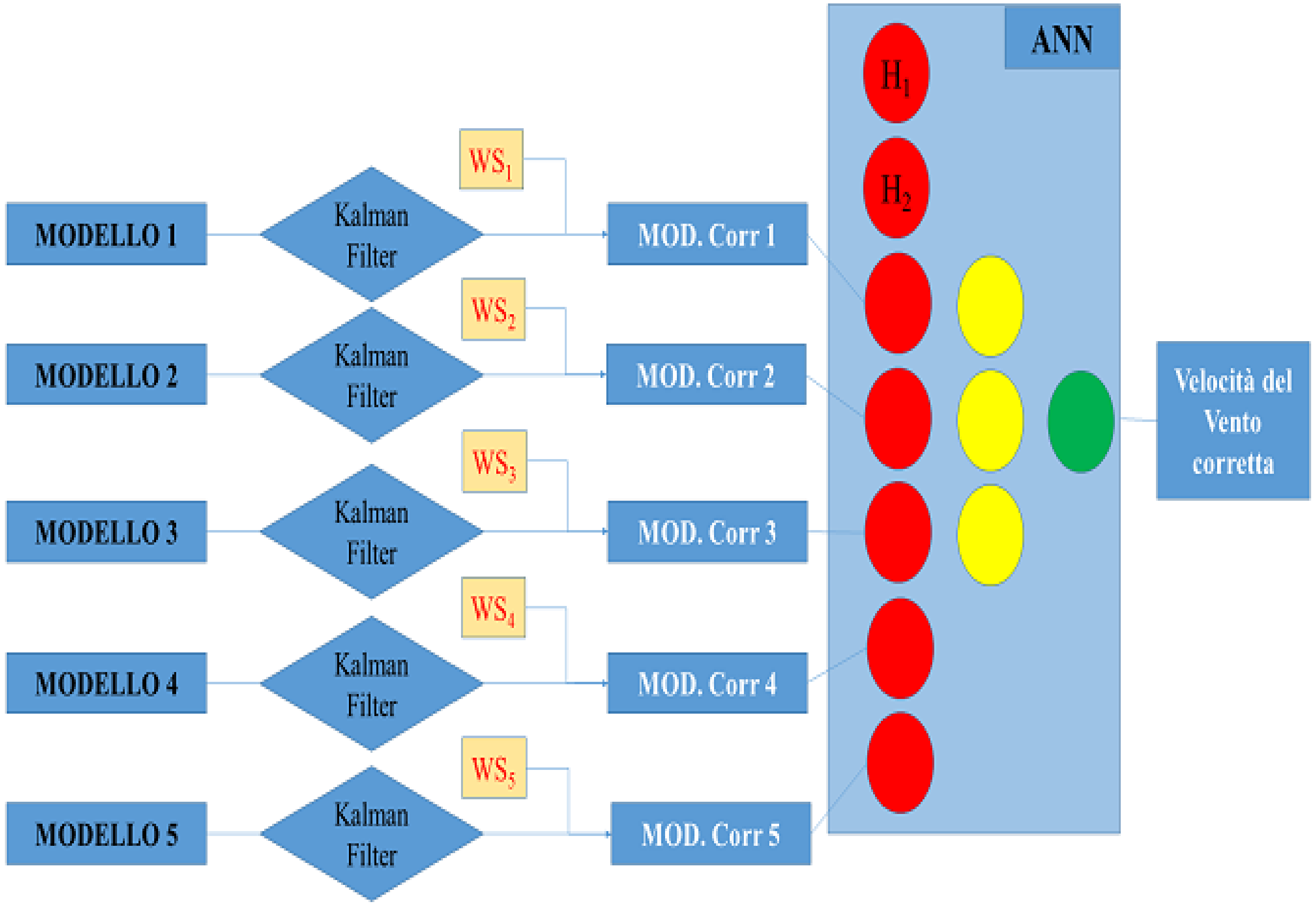
Standard Deviation of reference model

$$\sigma_r^2 = \frac{1}{N}\sum_{i=1}^{N}(r_n - \bar{r})^2$$

WS10 corrections: Kalman Filters & ANN

Hourly Cyclic components

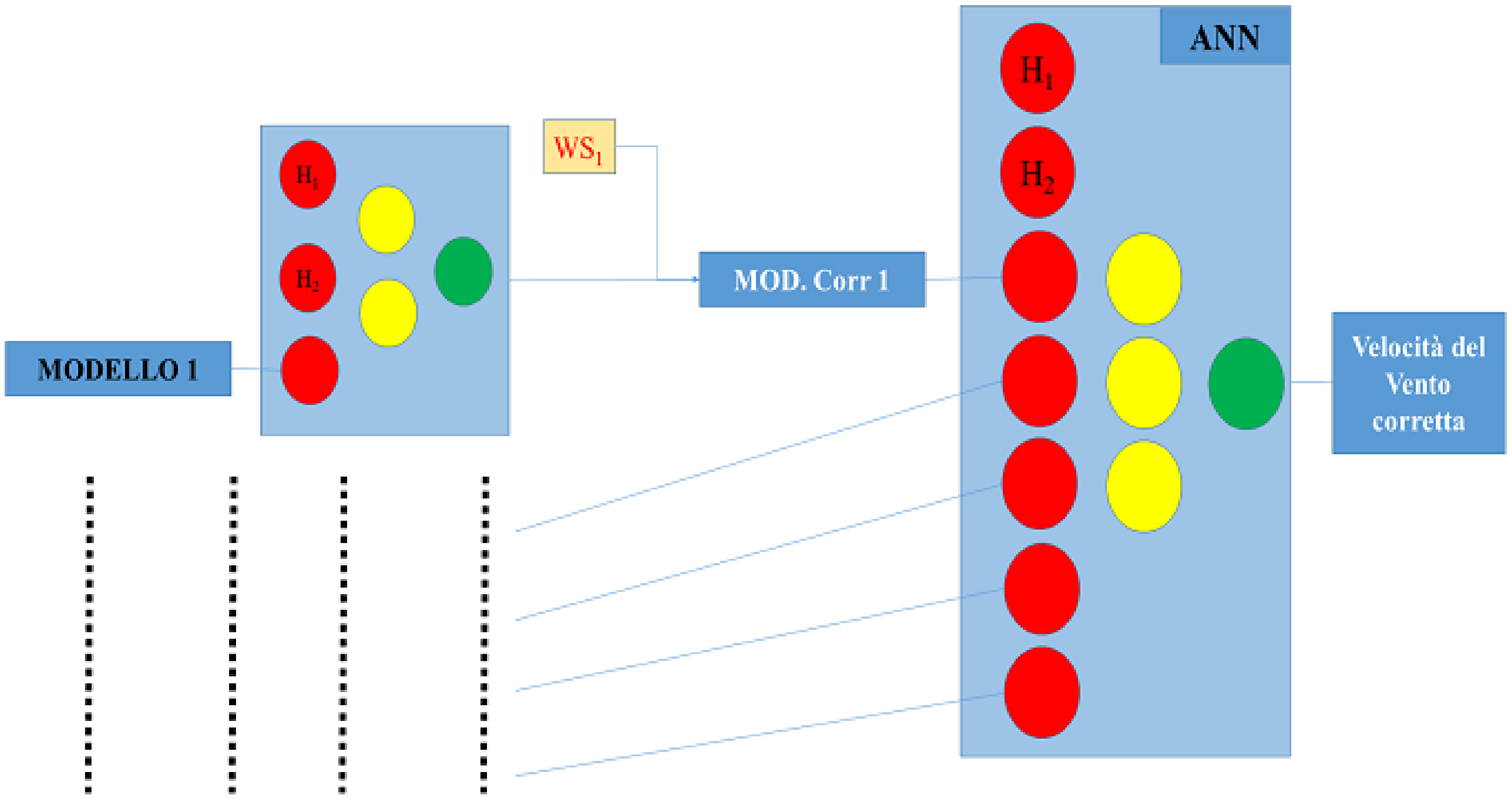
$$H_1 = sin(\frac{h \cdot 2\pi}{24})$$
$$H_2 = cos(\frac{h \cdot 2\pi}{24})$$



WS10 corrections: ANNs & ANN

Hourly Cyclic components

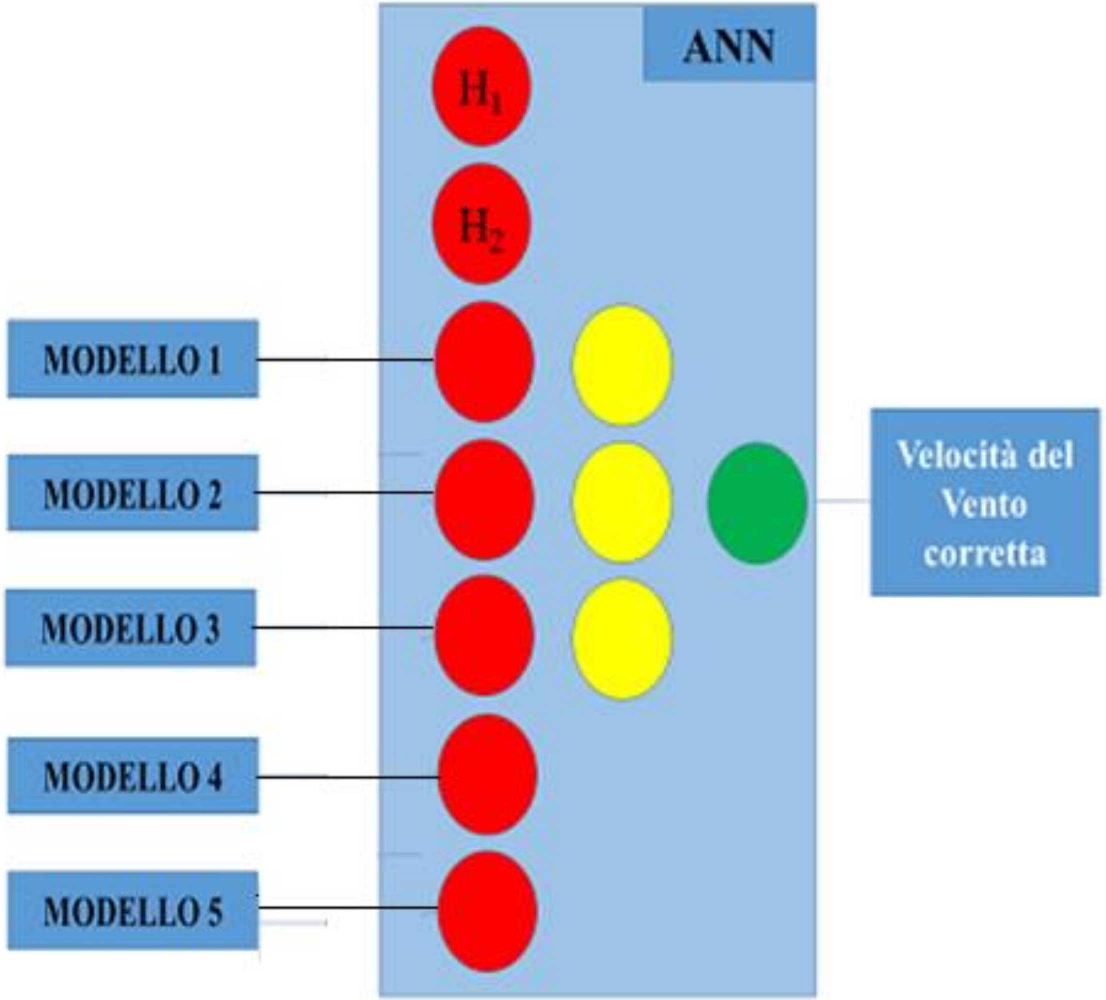
$$H_1 = sin(\frac{h \cdot 2\pi}{24})$$
$$H_2 = cos(\frac{h \cdot 2\pi}{24})$$



WS10 corrections: Single ANN

Hourly Cyclic components

$$H_1 = sin(\frac{h \cdot 2\pi}{24})$$
$$H_2 = cos(\frac{h \cdot 2\pi}{24})$$



WS10 corrections: Single ANN

Average

$$WS_{corr} = \frac{\sum_{i=1}^{5} WS_i}{5}$$

Weighted Average

$$WS_{corr} = \sum_{i=1}^{5} WS_i \cdot P_i$$

$$P_{hi} = \frac{M'_{hi}}{\sum_{i=1}^{5} M'_{hi}}$$

$$M'_{hi} = (\sum_{i=1}^{5} M_{hi}) - M_{hi}$$

$$P_{hi} = \frac{M'_{hi}}{\sum_{i=1}^{5} M'_{hi}}$$

$$M'_{hi} = \frac{1}{M_{hi}}$$

Linear Regression

$$WS_{corr} = K_1 \cdot WS_1 + K_2 \cdot WS_2 + K_3 \cdot WS_3 + K_4 \cdot WS_4 + K_5 \cdot WS_5 + K_6$$

WD10 corrections

Partial Average

$\overline{WD}_{par}$

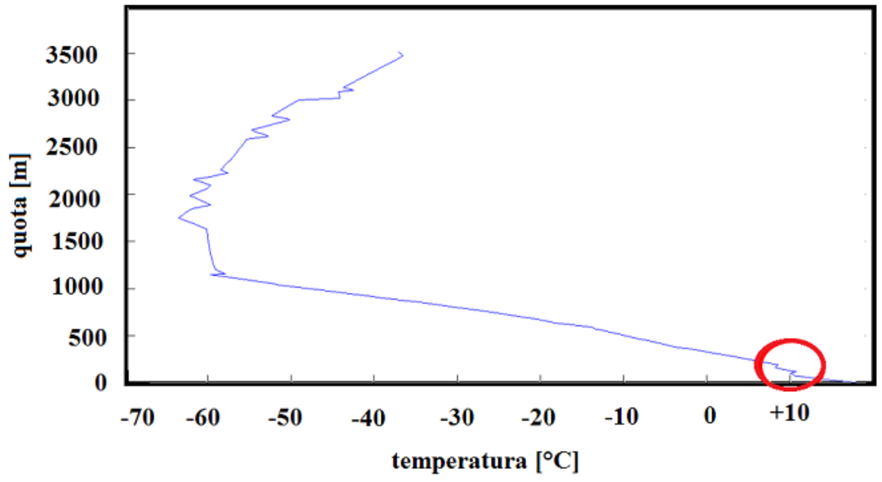
&

Linear Regression

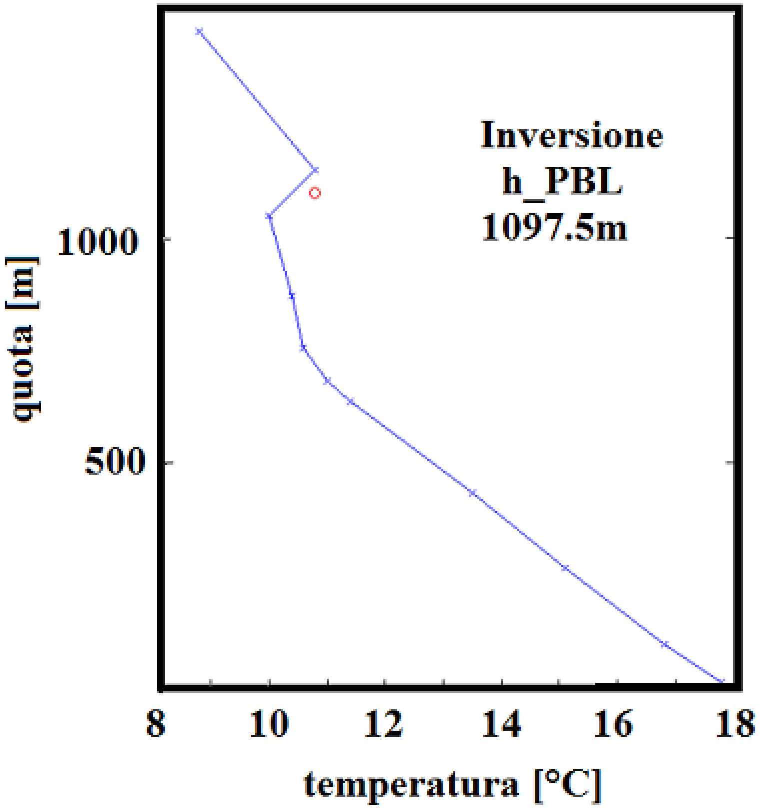
Average among only members with a distance less then 1 standard deviation from total mean

$$ED_{best} = A_1 \cdot \overline{WD}_{par} + A_2$$

Height PBL: temperature profile

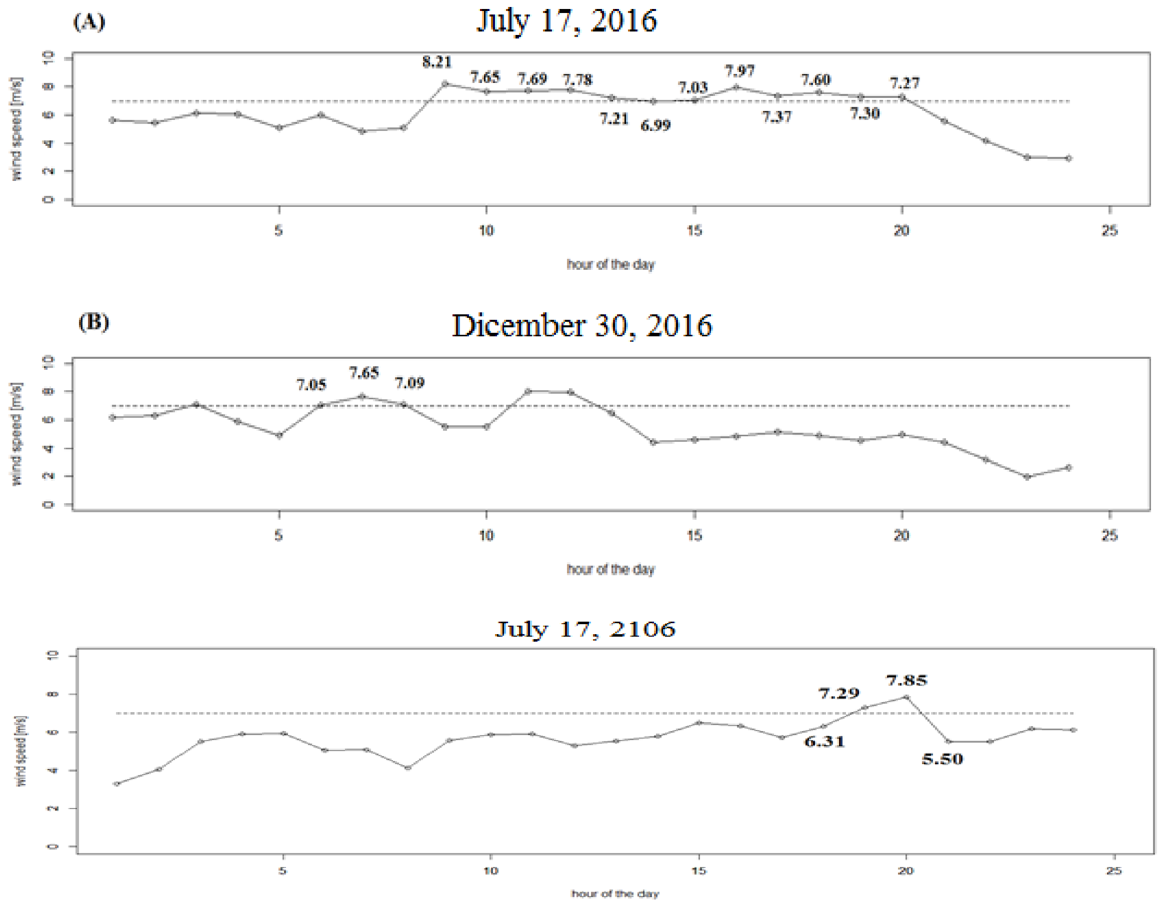


first thermal inversion

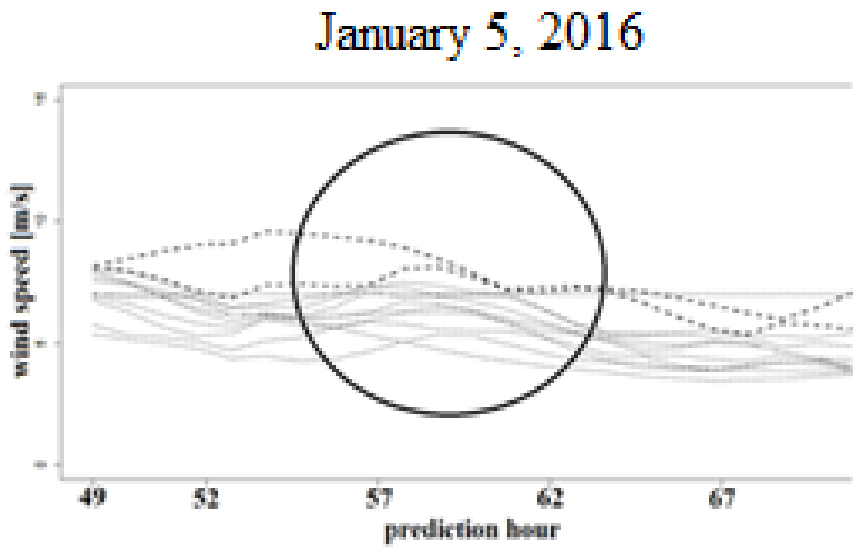


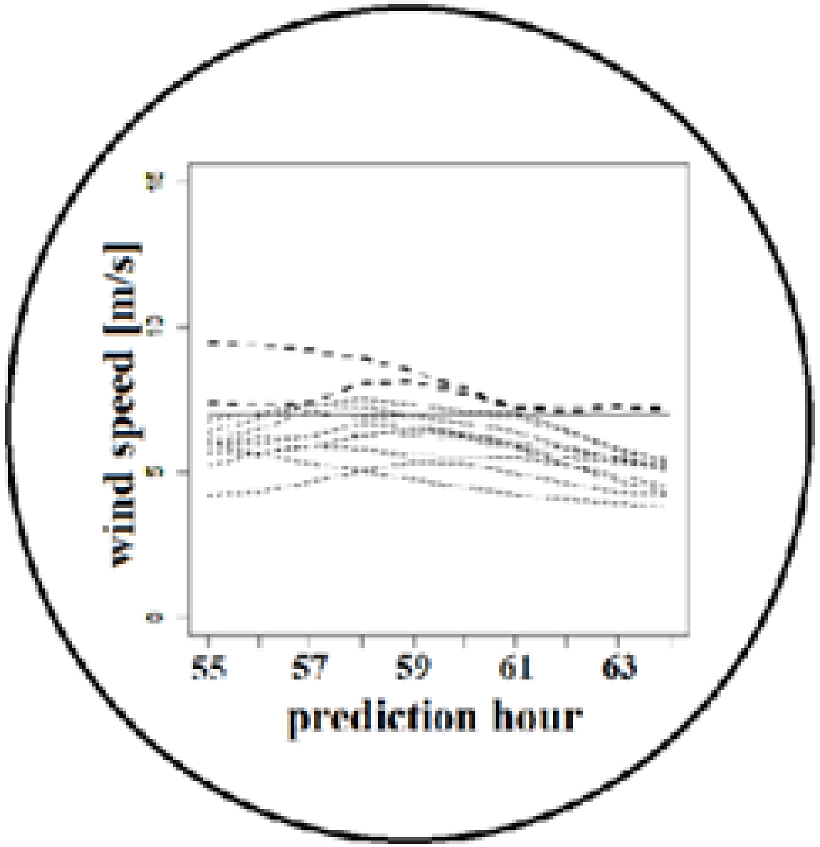
SELECTION OF CASES

For the analysis only borderline WD are selected where borderline means difficult to predict



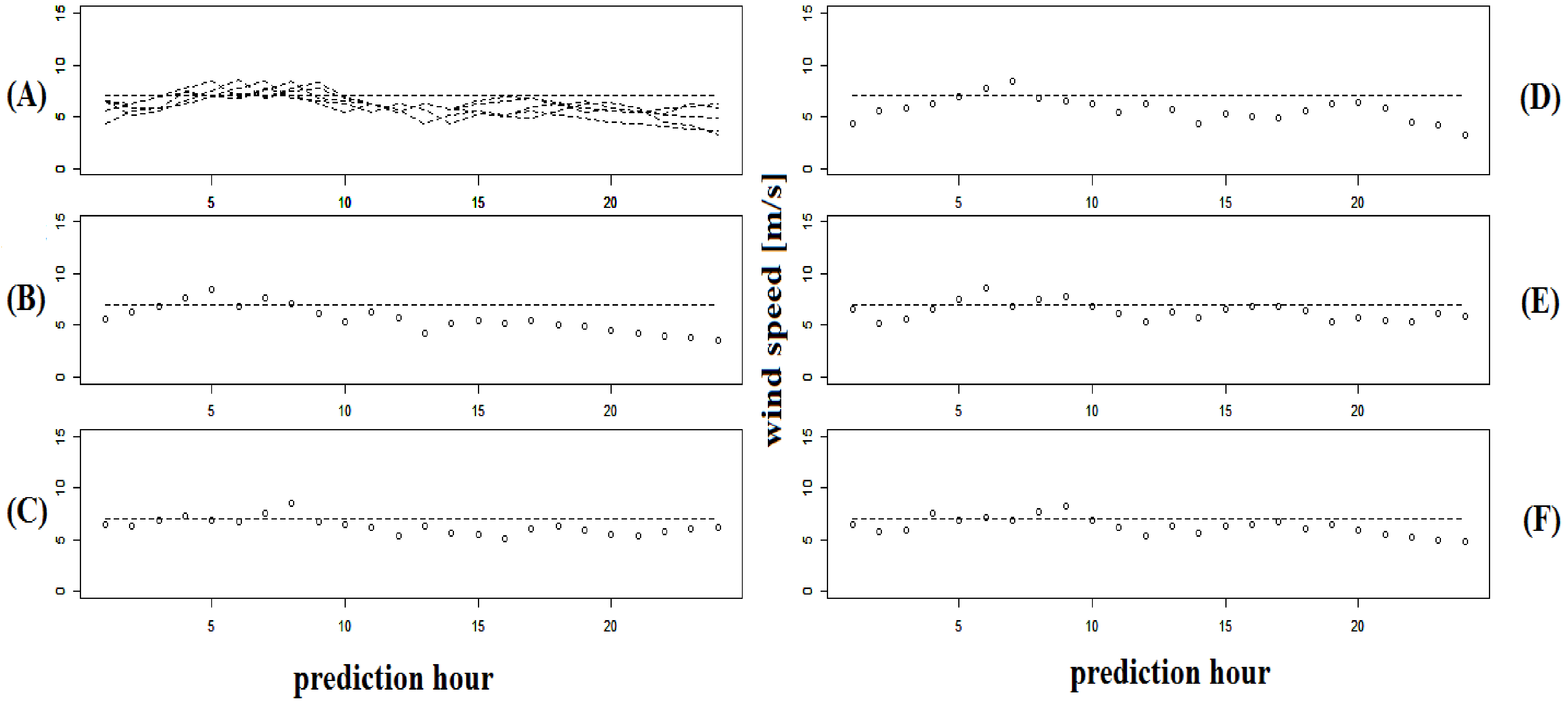
ENSEMBLE DISTRIBUTION PERCENTILES



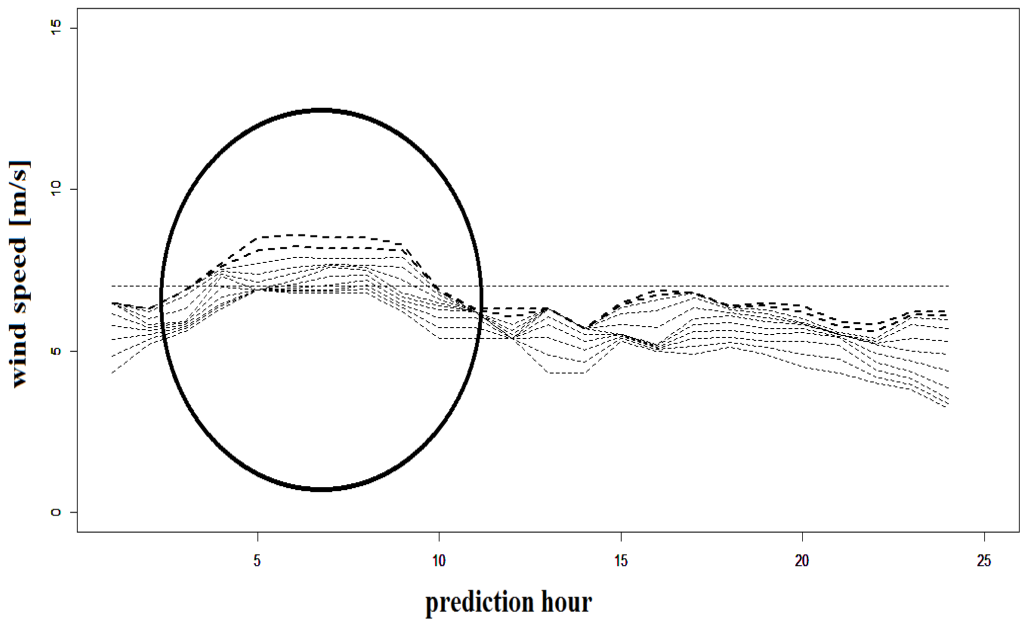


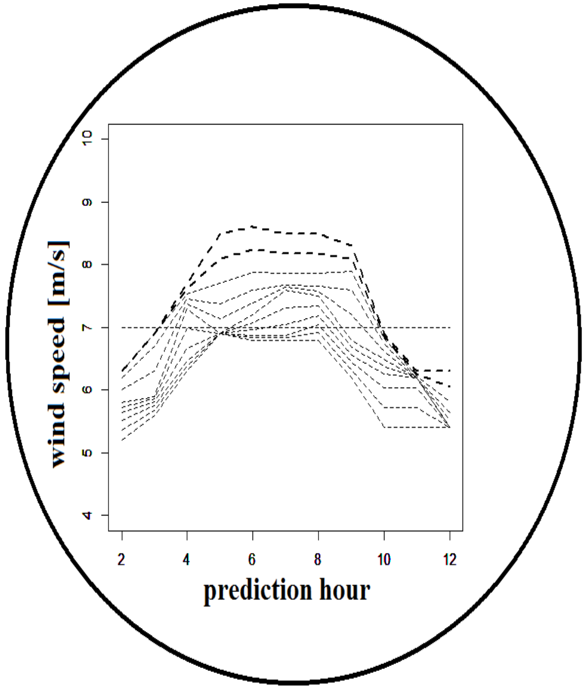
$$WDP_q = 100 - [\text{lowest hourly percentile curve predicting the wind day}]\ \%$$

Ensemble distribution Percentile VS percentage

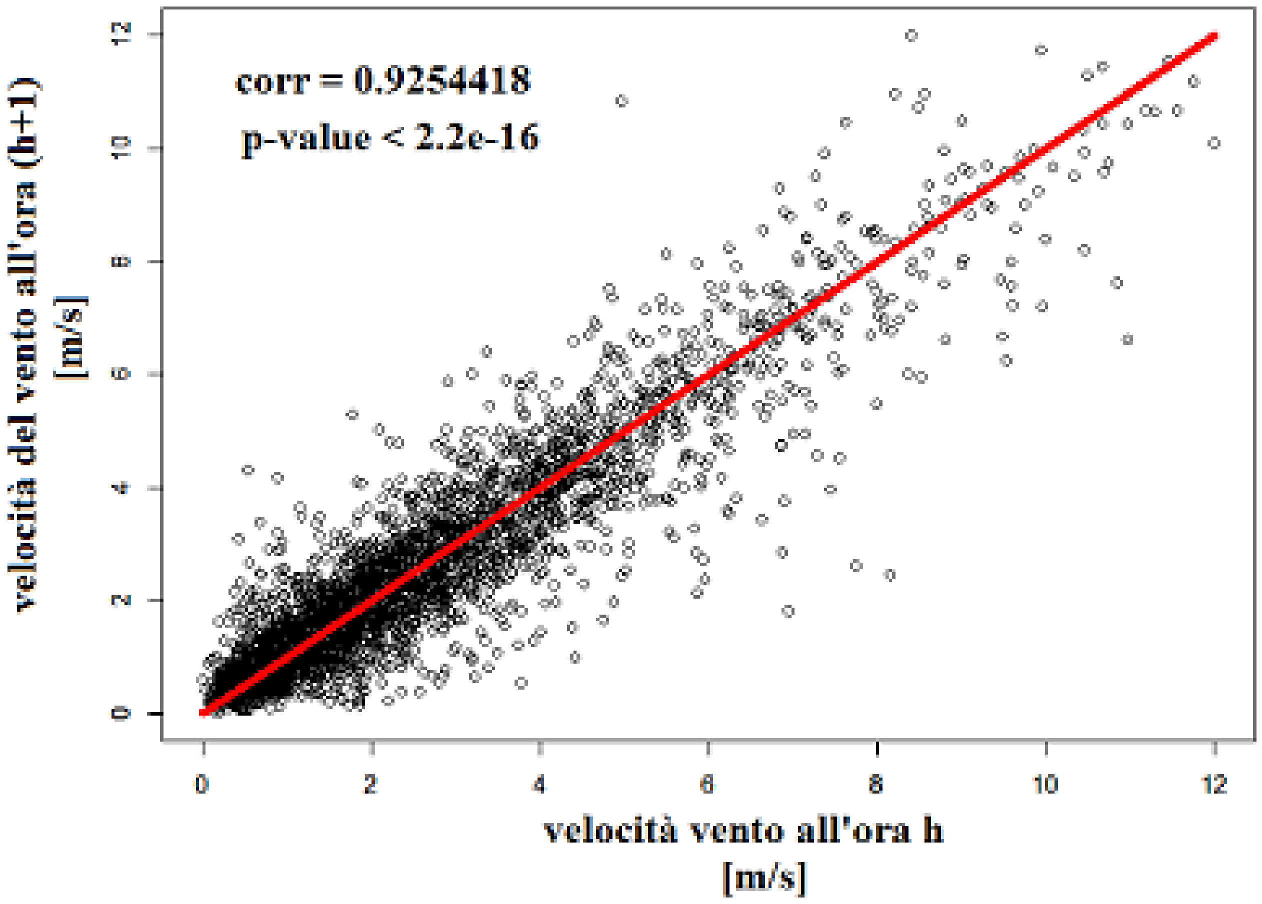


Ensemble distribution Percentile VS percentage

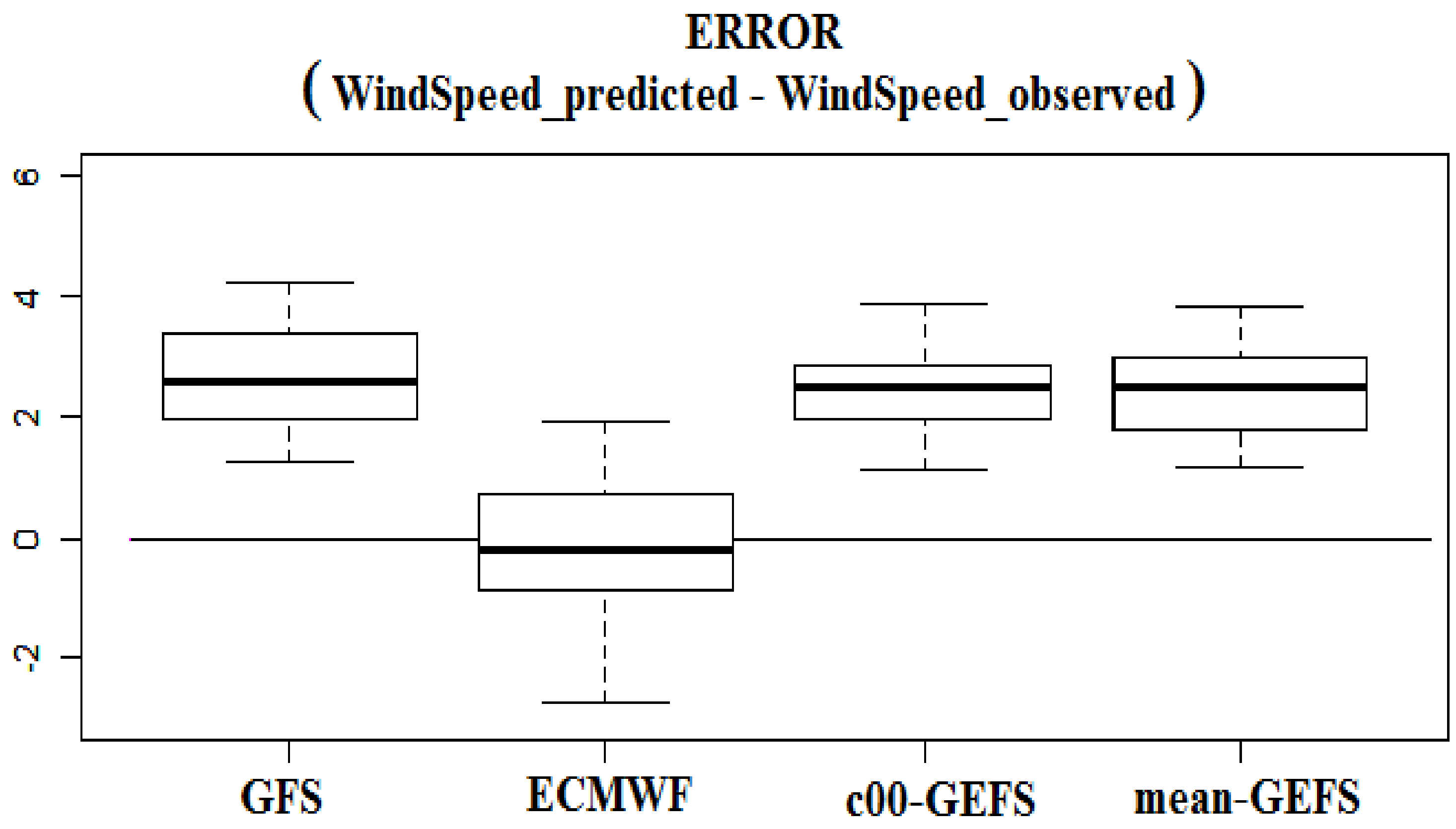




LINEAR DEPENDENCE OF TWO CONSECUTIVE HOURS



SYSTEMATIC ERROR AFFECTING GLOBAL MODEL



DEFINITION OF WIND DAYS

PM10

benzo(a)pyrene

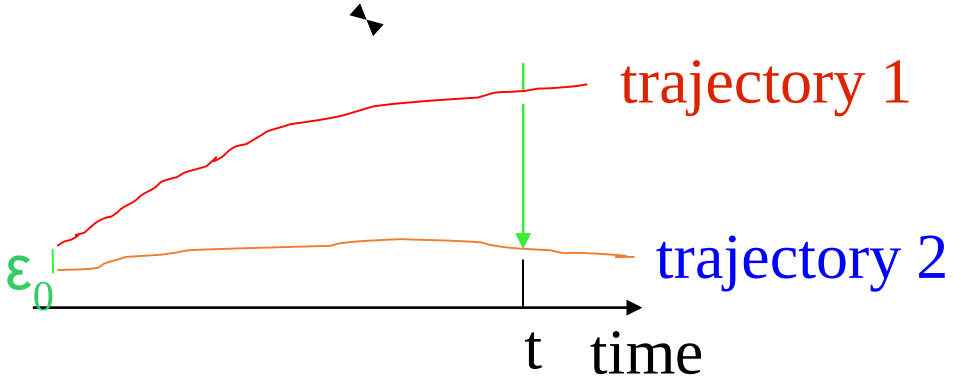
ATMOSPHERIC SYSTEM

The Atmosphere is

- **a deterministic dynamical physic system** → spatial and temporal evolution
- **nonlinear** → non linear differential equation system
- **many variables** → multidimensional phase space

In the "Chaos theory" of E. Lorenz[1], the Atmosphere is defined as "chaotic system", i.e.

very small changes in the initial state could greatly change the evolution in the following days.



[1]Lorenz, E. N. (1995). *The essence of chaos*. University of Washington Press.

STATISTICAL APPROACH: the ensembles

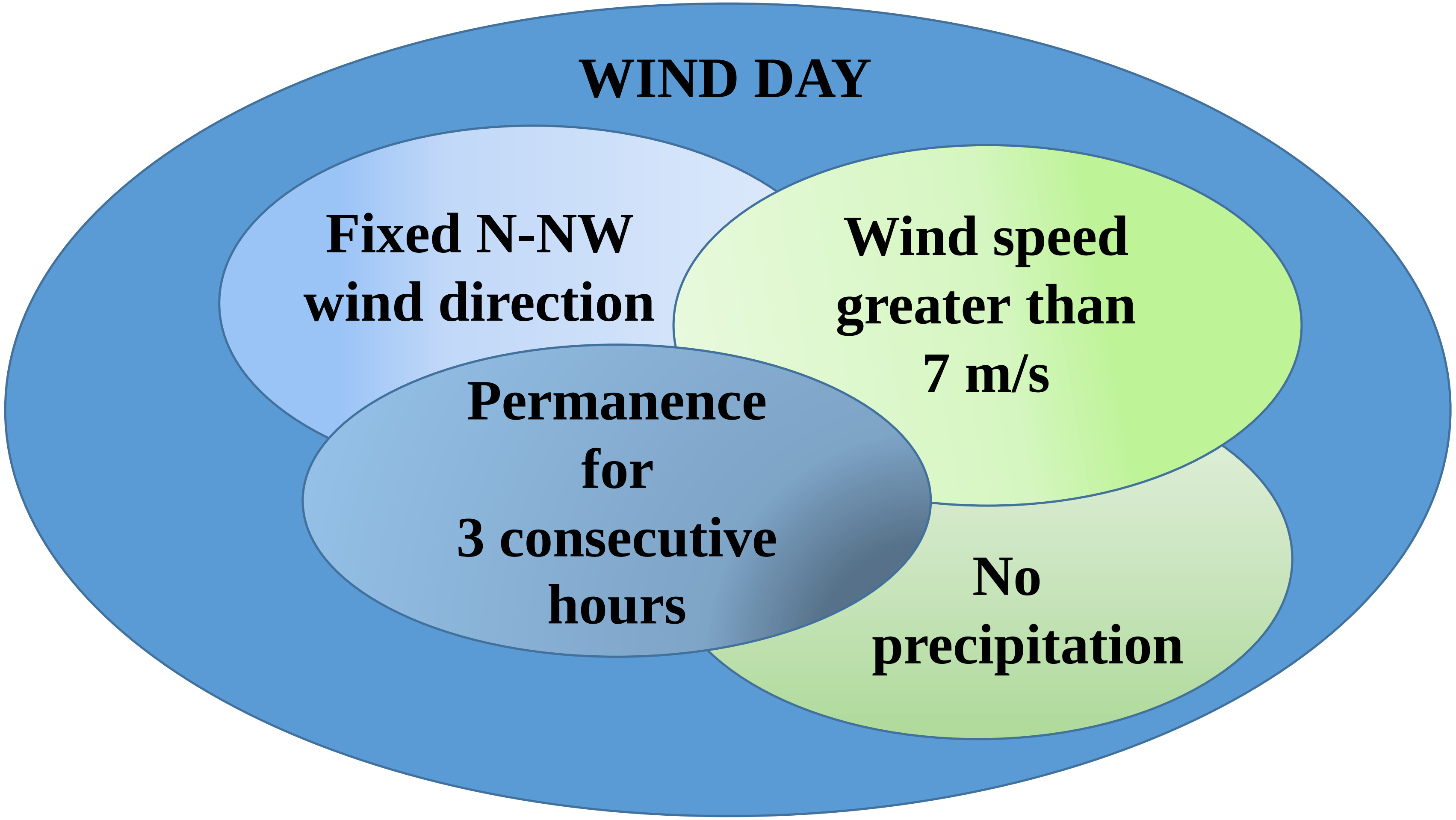
EPS's consider many possible initial conditions (each global model uses a different approach) by perturbing the measured state.

For each possible initial condition the deterministic runs are performed.

During the run, additional perturbations are considered.

The statistical analysis provides a better prediction with an associated uncertainty.

DEFINITION OF WIND DAYS



Apulia Government adopted a regional air quality plan (2012)

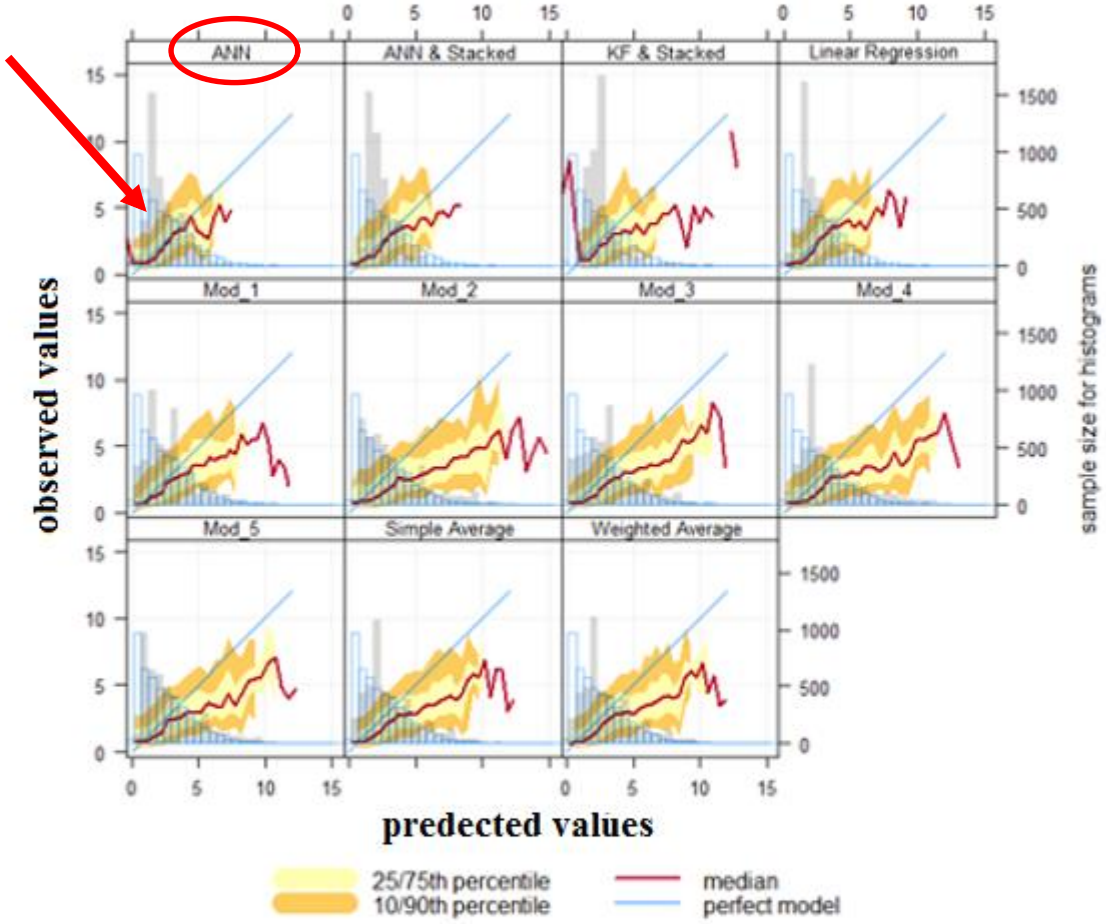
ANALYSIS 1: ENSEMBLE OF PARAMETERIZATION

WS10

TIME OF ANALYSIS
08/2015-03/2016

DATA
WRF A 4 KM CON INPUT GFS
+72H (49TH-72TH)

Conditional Quintile Plot



Analysis 2: Ensemble of Perturbations

Currently, the deterministic and probabilistic approaches are considered separately.

The probabilistic forecast is given by the ensemble average or the median.

The occurrence probability is given by the percentage of ensemble members predicting the event.

Analysis 2: Ensemble of Perturbations

Materials

Training and Test analysis:

34 possible WD in 2016

Validation analysis:

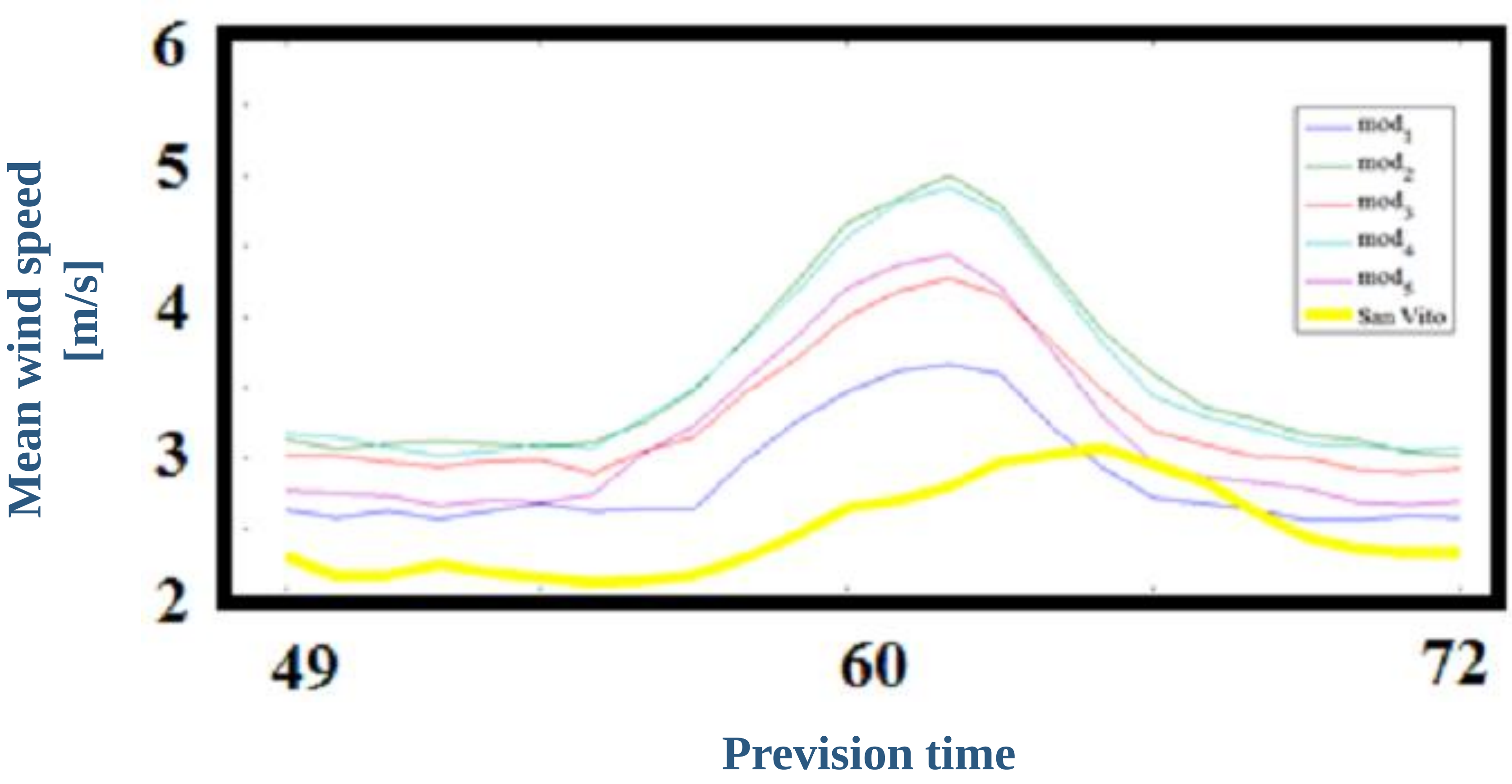
66 possible WD in 2017

Analysis 1: Ensemble of Parameterization

WS10

Time of analysis
08/2015-03/2016

Data
WRF a 4 Km con input GFS
+72h (49th-72th)



CONCLUSION

ACCURACY

	Classical definition WD		No Classical definition WD	
	Deterministic	Proposed method	Deterministic	Proposed method
Training	~60%	~82%	~82%	~94%
Test	~60%	~80%	~82%	~85%
Validation	~68%	~80%	~79%	~85%