COMPARISON OF THE PERFORMANCE OF PCA-NN MODELS FOR TEC OVER THE IBERIAN PENINSULA:

PERFORMANCE OF DIFFERENT NEURAL NETWORKS CONFIGURATION

ANNA MOROZOVA,

TERESA BARATA, TATIANA BARLYAEVA

University of Coimbra, Instituto de Astrofísica e Ciências do Espaço, Portugal

OUTLINES

- PCA-NN models for TEC
- Data
  - Correlated SW predictors
- NN
  - Configuration
  - Selection of the best NN architecture & best sets of SW predictors
  - Performance & examples
- Conclusions
- Acknowledgements

PCA-NN MODELS FOR TEC

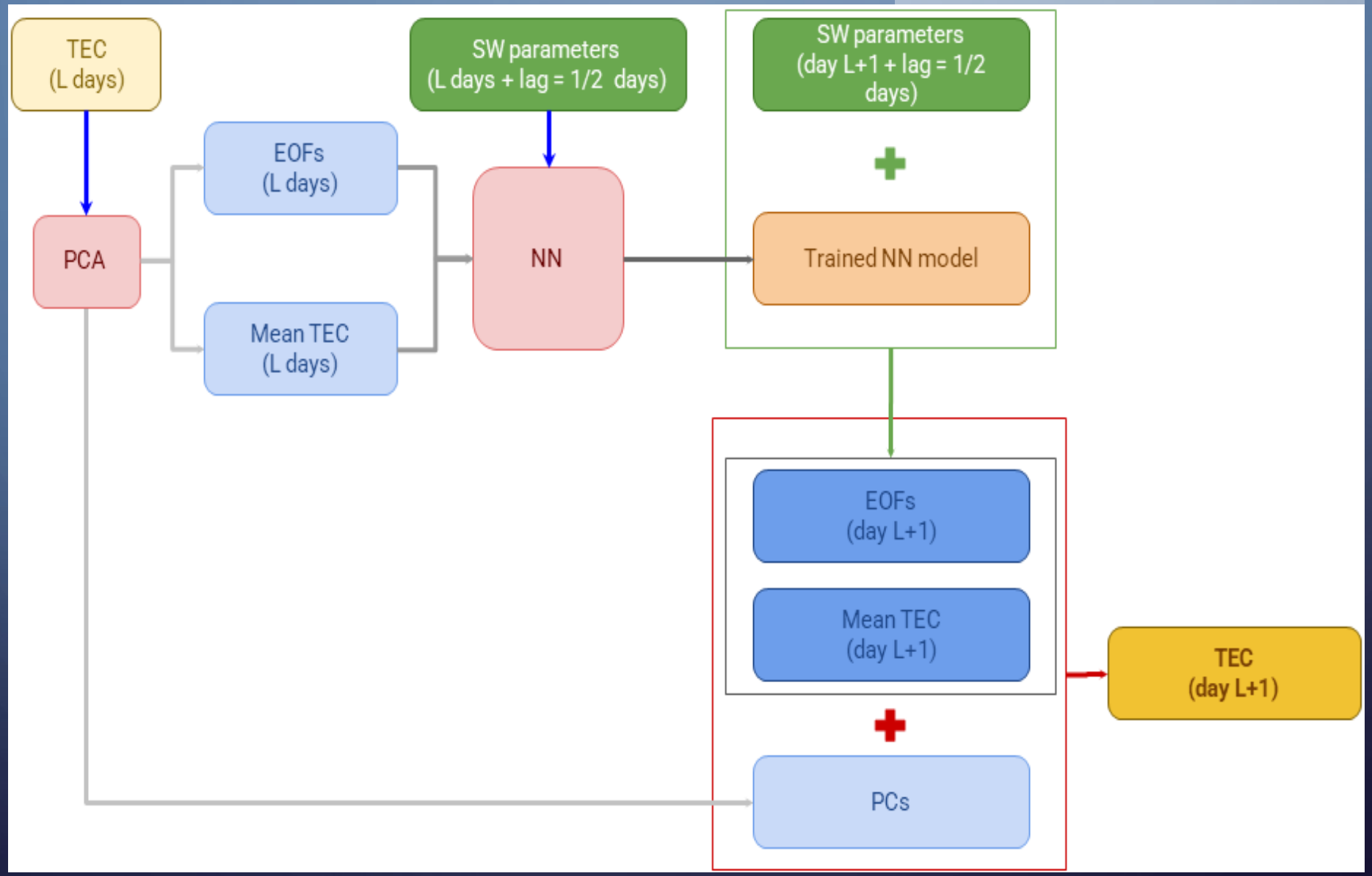
PCA-BASED MODELS

- The total electron content (TEC) over the Iberian Peninsula was modeled using a PCA-NN model based on
  1. decomposition of the observed TEC series using the principal component analysis (PCA)
  2. reconstruction of the daily mean TEC and daily PCA modes' amplitudes by. For example, regression models (PCA-MRM models*) or neural networks (PCA-NN models) using several types of space weather parameters (SW) as predictors.
- Lags of 1 and 2 days between the TEC and SW predictors are used

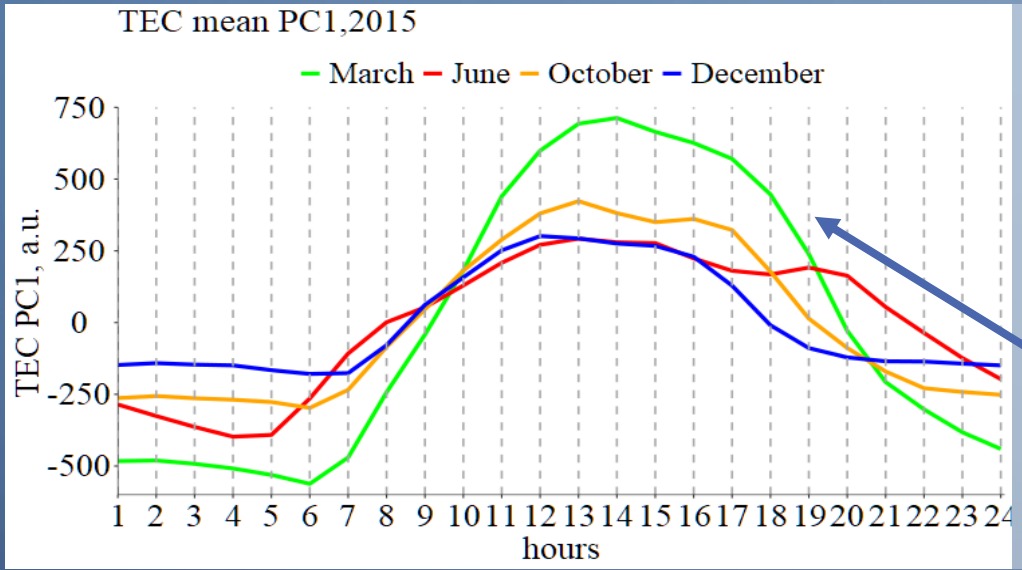
* *A. L. Morozova, T. Barata, T. Barlyaeva (2022) PCA-MRM model to forecast TEC at middle latitudes, Atmosphere, 13(2), 323; https://doi.org/10.3390/atmos13020323*

PCA-NN MODEL

L = 31 days



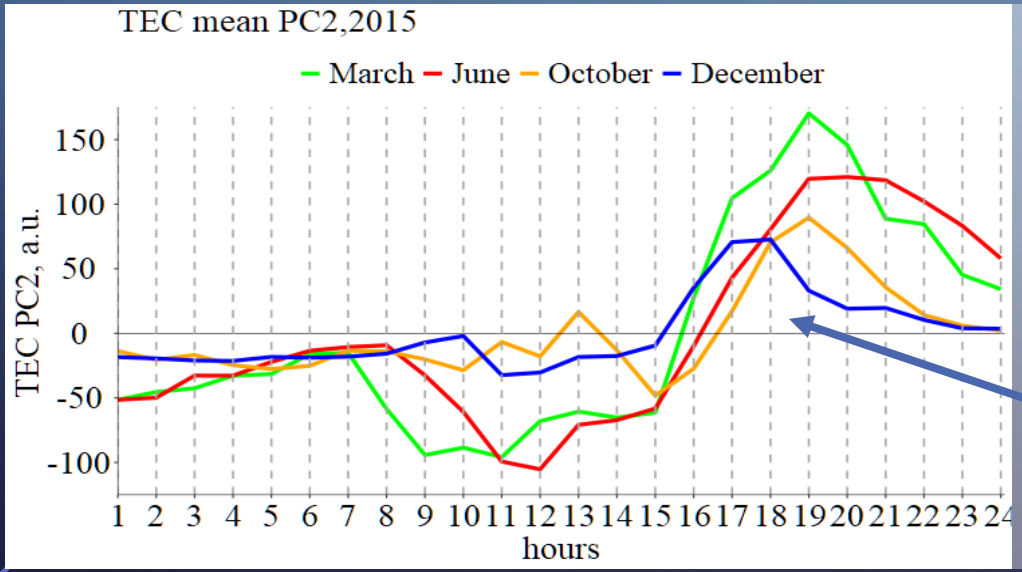
TEC PCA MODES 1 AND 2



Mode 1 (PC1 & EOF1):

Explains 77-95% of the TEC variations for different months

PC1 = regular daily variation due to the changes of the insolation



Mode 2 (PC2 & EOF2):

Explains 1.5-8.4% of the TEC variations for different months

PC2 = shallow minimum of TEC around the noon and a maximum in the late afternoon

PCA-BASED MODELS

- The main feature of the PCA-based models is that the TEC series is decomposed into several PCA modes which represent TEC daily variations of different types
- The amplitude of each of the mode for each day is described by the EOF coefficients
  - The EOF coefficients can be modelled using space weather parameters as predictors using, e.g., multiple regression models (MRM) or neural networks (NN)
- The advantage of the PCA-based models is that there is no need for any assumption on the phase and amplitude or seasonal/regional features of TEC daily variations: the daily variations of correct shapes are extracted automatically by PCA from the input TEC data

DATA

DATA: TEC

- Vertical **TEC** measured at Lisbon airport, Portugal (39° N, 9° W) by a GNSS receiver with SCINDA system
- Time interval: **01.01.2015- 31.12.2015**
- Time resolution: **1h data**

DATA: SPACE WEATHER PARAMETERS (SW PREDICTORS)

- Solar wind parameters:
  - Pressure (**p**), density (**n**), velocity (**v**)
- Interplanetary magnetic field:
  - Full interplanetary magnetic field (scalar **B**), GSM components (**Bx, By, Bz**)
- Geomagnetic indices:
  - **Dst, ap, AE, local K_{COI}-index** (Coimbra Geomagnetic Observatory, Portugal)
- Proxies for the solar UV & XR fluxes:
  - UV: **Mg II** composite series or **F10.7** index
  - **XR**: Solar EUV Experiment (TIMED mission)
  - Daily number of solar flares of classes ≤**C** and **M**, and the number of all flares **N**
- Time resolution: **1d data**

SPACE WEATHER PARAMETERS: CORRELATED PREDICTORS

- Some of the **SW predictors correlate with each other**
- Highest correlations:
  - n & p ($|r| = 0.75$)
  - **Dst & ap & AE & Kcoi ($|r| = 0.75 \div 0.85$)**
  - **Mg II & F10.7 & XR ($|r| \geq 0.82$)**
  - **Number of C and all flares ($|r| = 0.98$)**

SPACE WEATHER PARAMETERS: CORRELATED PREDICTORS

|r| > 0.6	B	Bx	By	Bz	n	v	p	Dst	ap	AE	Kcoi	MgII	F10.7	XR	C
B							0.6		0.66		0.62				
Bx			0.69												
By		0.69													
Bz											0.66				
n							***0.75***								
v											0.6				
p	0.6				***0.75***				0.64		0.63				
Dst									***0.74***	***0.75***	0.68				
ap	0.66						0.64	***0.74***		**0.85**	**0.86**				
AE								***0.75***	**0.85**						
Kcoi	0.62			0.66		0.6	0.63	0.68	**0.86**						
MgII													**0.79**	**0.82**	
F10.7												**0.79**		**0.84**	
XR												**0.82**	**0.84**		
N															**0.98**

NN MODELS: CONFIGURATION

PCA-NN MODEL – PREVIOUS RESULTS

- Ready-to-use package is used: **neuralnet** (R)
- NN algorithm: feedforward NN with the resilient backpropagation with weight backtracking
- The input dataset length **L = 31 days**
- **SW predictors** are submitted as **lagged series** (lag = 1 and 2 days) **together** (for X predictors there are 2*X input series)
- "**Ensemble forecast**": a number (e.g., 100) of NN models of the same architecture were trained on the same input dataset and were used to make a forecast for the day L+1; the final forecast is the arithmetic average of 100 forecasts

PCA-NN MODEL – CURRENT QUESTIONS

- ? Best NN configuration(s) that produces forecasts of reasonable quality with minimal number of SW predictors for daily mean TEC and EOFs series
- ? Best set of SW predictors for daily mean TEC and EOFs series
  - ? Can correlated SW predictors be used

SELECTION OF THE BEST NN ARCHITECTURE
&
BEST SETS OF SW PREDICTORS

PCA-NN MODEL – DAILY MEAN TEC

1SWP

- B, Bx, By, Bz
- n, v, p
- Dst, ap, AE, K
- F10.7/MgII, XR
- N, C, M flares

3 or 4 SWp

- ~~B, Bx, By, Bz~~
- ~~n, v, p~~
- **Dst**, ap, AE, ~~K~~
- ~~F10.7/~~**MgII**, XR
- N, C, ~~M~~ flares

5 SWp

- **Dst**, ap, AE
- **MgII**, XR
- **C flares**

6 SWp

- **Dst, ap, AE**
- **MgII, XR**
- **C flares**

* Selection using correlation between the observed and modelled series, RMSE and MAE

PCA-NN MODEL – EOF1 (AMPLITUDE OF THE DAILY VARIATION)

1SWP

- B, Bx, By, Bz
- n, v, p
- Dst, ap, AE, K
- F10.7/MgII, XR
- N, C, M flares

2 SWp

- ~~B, Bx, By,~~ Bz
- ~~n, v, p~~
- Dst, ap, **AE**, ~~K~~
- ~~F10.7/~~**MgII**, ~~XR~~
- ~~N, C, M flares~~

3 SWp

- **Dst, AE**
- **MgII**

4 or 5 SWp

- Bz
- Dst, ap, **AE**
- **MgII**

* Selection using correlation between the observed and modelled series, RMSE and MAE

PCA-NN MODEL – EOF2 (AMPLITUDE OF THE 2ND DAILY MAX)

1SWP

- B, Bx, By, Bz
- n, v, p
- Dst, ap, AE, K
- F10.7/MgII, XR
- N, C, M flares

2 or 3 SWp

- ~~B, Bx, By,~~ Bz
- ~~n~~, v, ~~p~~
- Dst, ap, **AE**, K
- F10.7/MgII, ~~XR~~
- ~~N, C, M flares~~

4 SWp

- **v**
- **Dst, AE**
- **F10.7**

5 SWp

- Bz
- v
- **Dst**, ap, **AE**
- **F10.7**

* Selection using correlation between the observed and modelled series, RMSE and MAE

PERFORMANCE OF THE NN MODELS

BEST NN ARCHITECTURE AND BEST SETS OF SW PREDICTORS

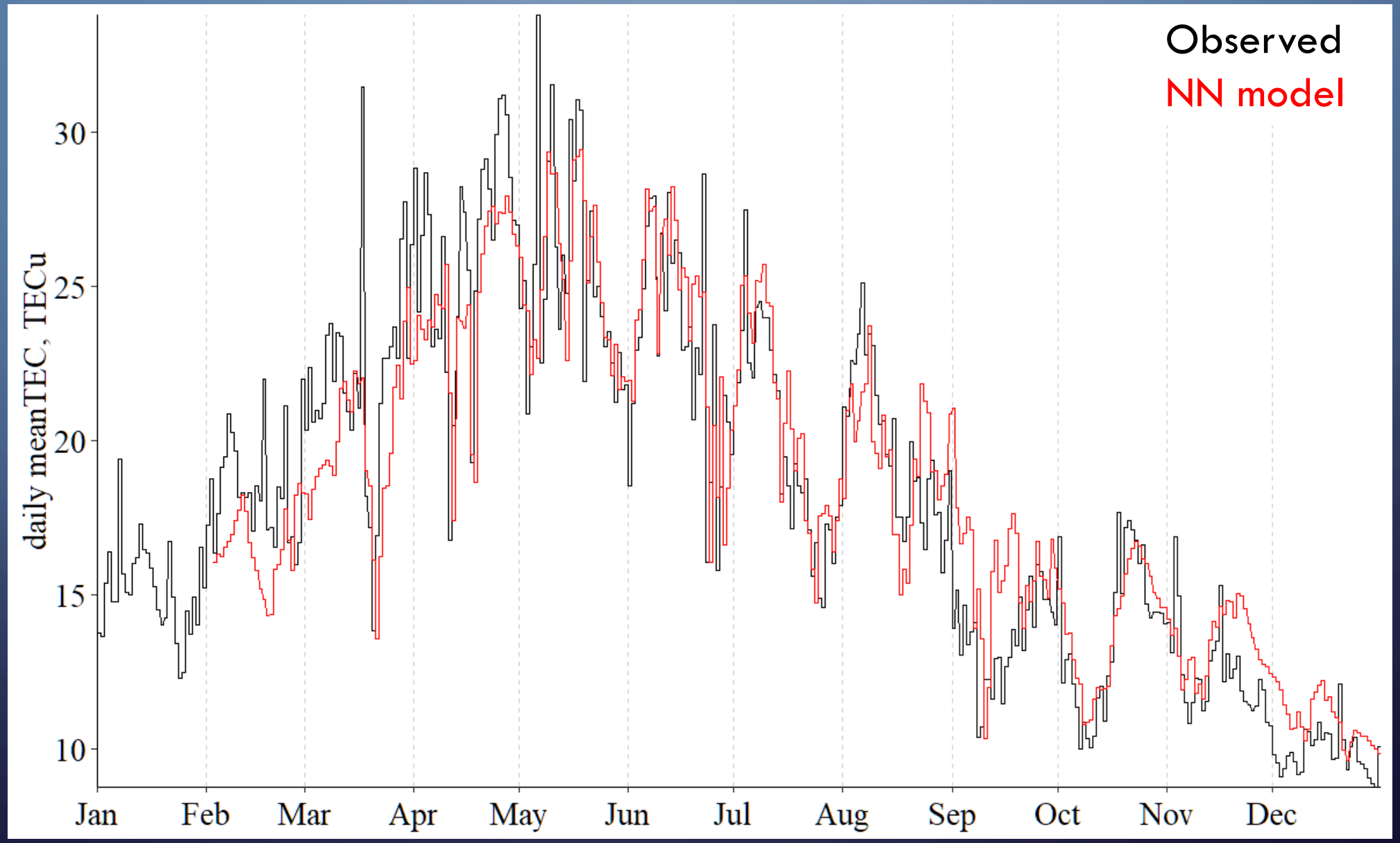
	Daily mean TEC	EOF1	EOF2
NN (layers & nodes)	(12,6,4)	(6,4)	(8,4,2)
Predictors	6 predictors	3 predictors	4 predictors
Best predictors	MgII Dst N. C flares ap AE XR	AE MgII Dst	AE Dst F10.7 (!) V

COMPARISON OF THE NN FORECASTS AND OBSERVATIONS (SCORES)

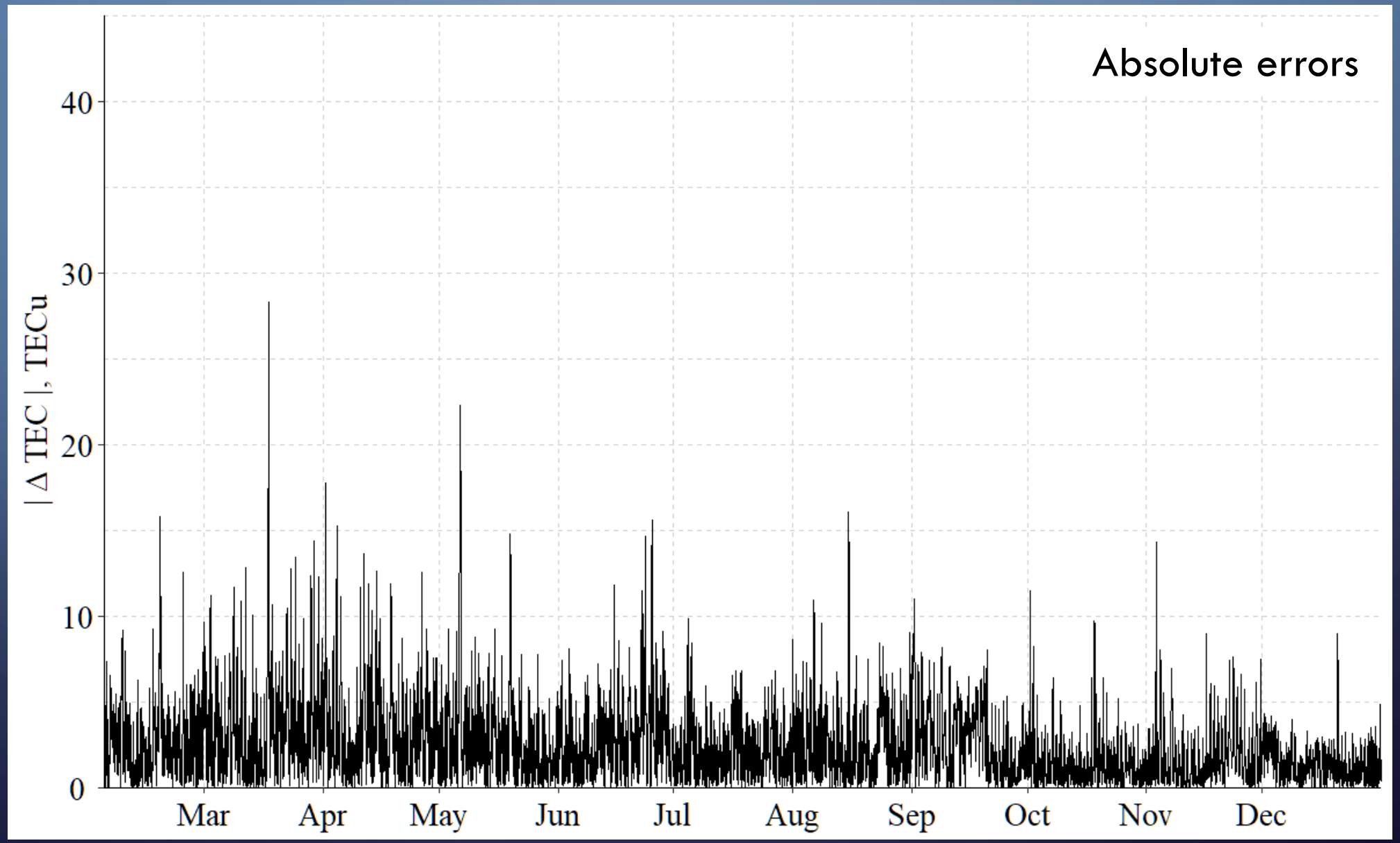
	1d mean TEC	1h TEC	1h (TEC – $TEC_{1d\ mean}$)
r	0.91	0.93	0.94
MAE, TECu	1.8	2.5	1.7
RMSE, TECu	2.4	3.4	2.4
MaxE, TECu	11.1	28.3	18.8

EXAMPLES OF THE PCA-NN FORECASTS

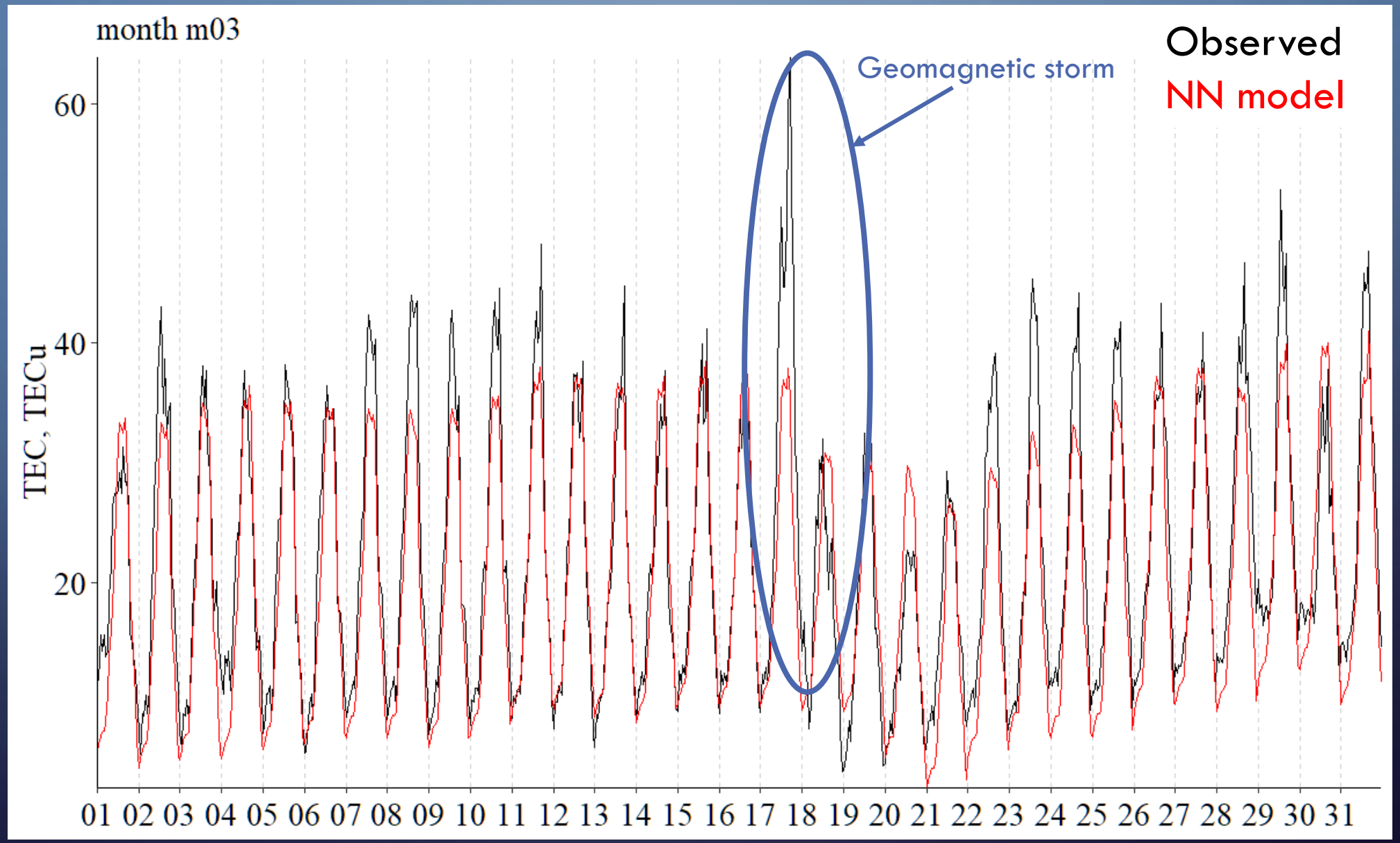
COMPARISON OF PC-NN FORECASTS AND OBSERVATIONS: daily mean TEC



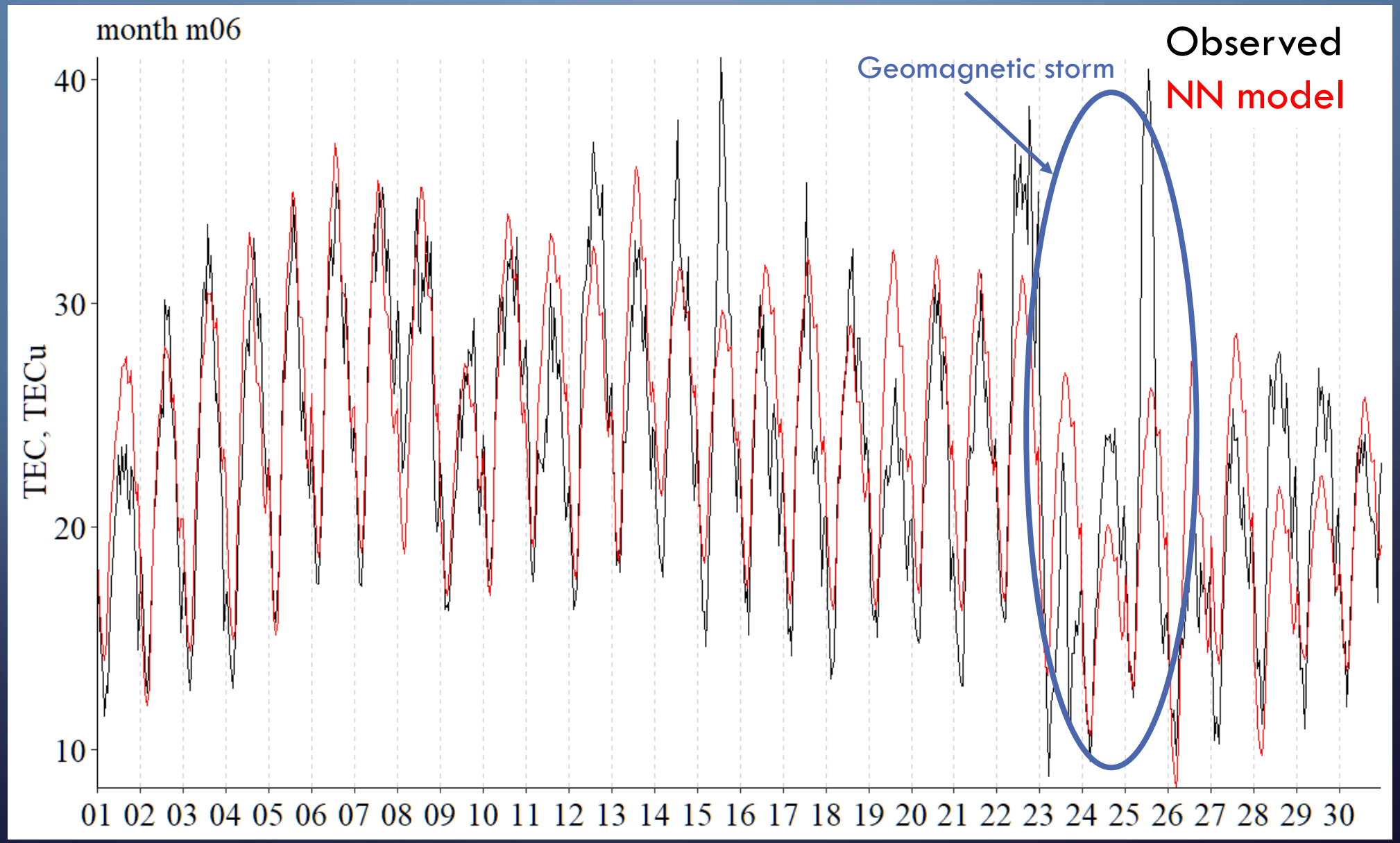
COMPARISON OF PC-NN FORECASTS AND OBSERVATIONS: 1h model's errors



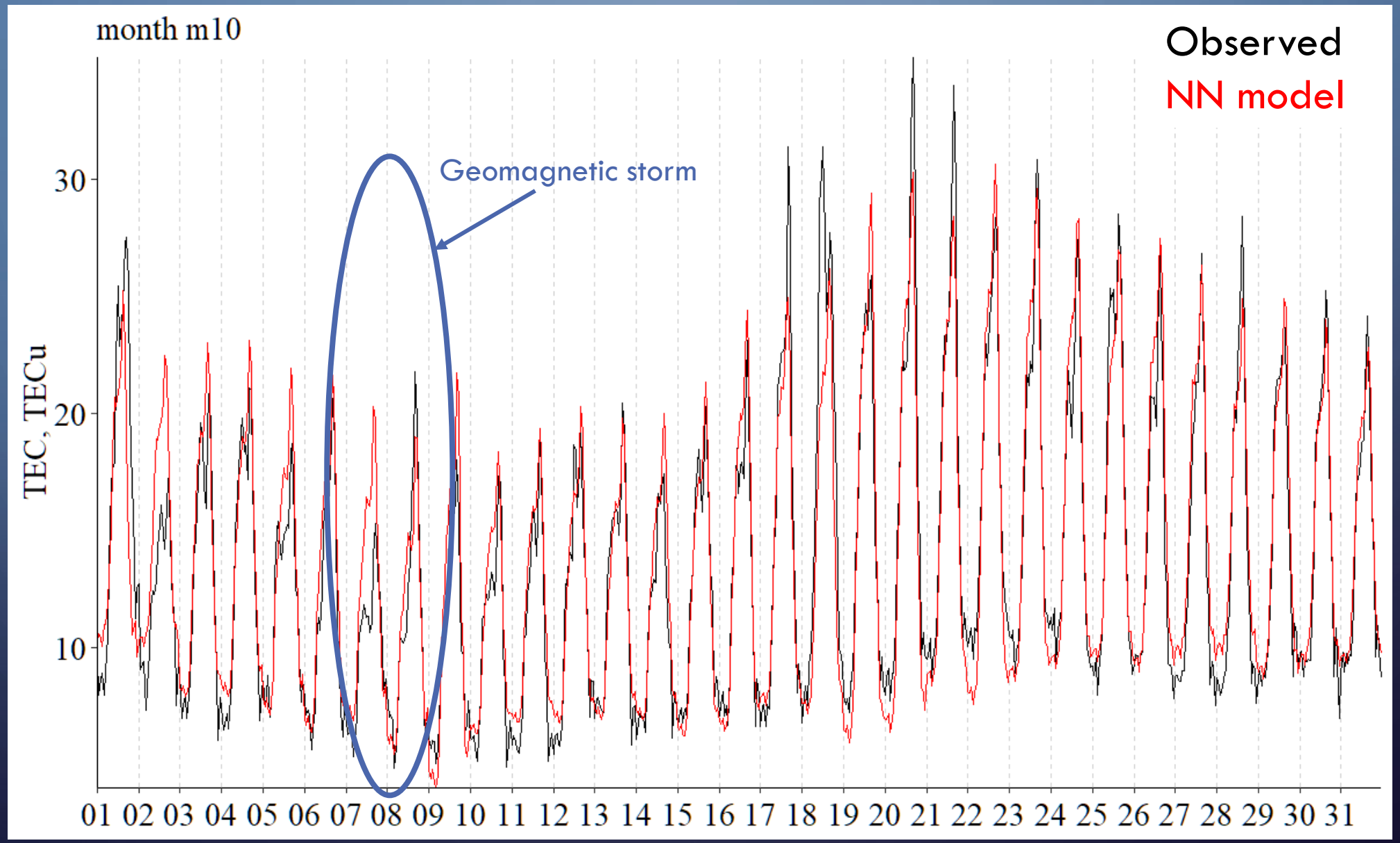
COMPARISON OF PC-NN FORECASTS AND OBSERVATIONS: 1h TEC, March 2015



COMPARISON OF PC-NN FORECASTS AND OBSERVATIONS: 1h TEC, June 2015



COMPARISON OF PC-NN FORECASTS AND OBSERVATIONS: 1h TEC, October 2015



CONCLUSIONS

CONCLUSIONS (1)

- **Best sets of SW predictors for PCA-NN models:**
  - Daily mean TEC – 6 predictors
  - EOF1 – 3 predictors
  - EOF2 – 4 predictors
- NN with small number of layers perform better (only 2 or 3 layers for 6 to 12 input SW series)

CONCLUSIONS (2)

- Most important predictors:
  - **Dst, AE, solar UV proxies** – used for all TEC parameters
- NN model for the **daily mean TEC** and **EOF1** perform better with **Mg II** as a proxy for the solar UV, but models for **EOF2** perform better with **F10.7**
- The use of correlated predictors may improve the prediction quality:
  - ***Dst & AE*** for all TEC series ***(|r| = 0.75)***
  - **MgII & XR** for the daily mean TEC (**|r| = 0.82**)
  - **ap & AE** for the daily mean TEC (**|r| = 0.85**)

ACKNOWLEDGEMENT

- This research is supported through the projects
  - "**SWAIR** - Space weather impact on GNSS service for Air Navigation", ESA Small ARTES Apps
  - "**PRIME**: Portuguese Regional Ionosphere Model " (EXPL/CTAMET/0677/2021), FCT
- IA is supported through the projects
  - UIDB/04434/2020 and UIDP/04434/2020 , FCT

SWAIR esa

ARTES IAP Demonstration Projects

prime

ia instituto de astrofísica e ciências do espaço

1 2 9 0 UNIVERSIDADE D COIMBRA

Cofinanciado por:

FCT Fundação para a Ciência e a Tecnologia
MINISTÉRIO DA CIÊNCIA, TECNOLOGIA E ENSINO SUPERIOR

COMPETE 2020

PORTUGAL 2020

UNIÃO EUROPEIA
Fundo Europeu de Desenvolvimento Regional