A statistical method based on the ensemble probability density function for the prediction of "Wind Days"

A. Tateo, M.M. Miglietta, F. Fedele, M. Menegotto, A. Pollice, R. Bellotti

PII: S0169-8095(18)30964-5
DOI: doi:10.1016/j.atmosres.2018.10.001
Reference: ATMOS 4382

To appear in: *Atmospheric Research*

Received date: 31 July 2018
Revised date: 24 September 2018
Accepted date: 4 October 2018

Please cite this article as: A. Tateo, M.M. Miglietta, F. Fedele, M. Menegotto, A. Pollice, R. Bellotti , A statistical method based on the ensemble probability density function for the prediction of "Wind Days". Atmos (2018), doi:10.1016/j.atmosres.2018.10.001

A statistical method based on the Ensemble probability density function for the prediction of "Wind Days"

A. Tateo[a], M. M. Miglietta[d], F. Fedele[b], M. Menegotto[b], A. Pollice[e], R. Bellotti[a,c]

[a]Dipartimento Interateneo di Fisica, Università deli Studi di Bari "A. Moro", Via G. Amendola 173, 70126, Bari, Italy,
[b]Apulia Region Environmental Protection Agency (ARPA Puglia), C.so Trieste 27, 70126, Bari, Italy,
[c]Istituto Nazionale di Fisica Nucleare (INFN), Sezione di Bari, Via Orabona 4, 70125 Bari, Italy,
[d]National Research Council of Italy (CNR), Institute of Atmospheric Sciences and Climate (ISAC), Lecce s.p. Lecce-Monteroni km 1.2 73100, Lecce, Italy,
[e]Dipartimento di Economia e Finanza, Università degli Studi di Bari, "A. Moro", Largo Abbazia S. Scolastica (già via C. Rosalba, 53) 70124 - Bari, Italy

Abstract

Numerical Weather Prediction (NWP) models are often used to predict meteorological events in a deterministic way. In recent years, operational Ensemble Prediction Systems are able to take into account some of the errors affecting the NWP models, and allow to estimate the probability of occurrence. In the traditional approach, this probability is given by the percentage of ensemble members predicting the event. In this study, we propose an alternative method to estimate the probability of occurrence, based on the ensemble probability density function (PDF), which takes into account only random errors unavoidably affecting the model. To estimate its reliability, we compare this method with classical categorical and probabilistic approaches by using different global models: ECMWF, GFS, and GEFS.

In particular, we focus on wind speed forecasts in the area around the city of Taranto, located in Apulia region (southeastern Italy), to simulate the events called "Wind Days", i.e. northwesterly wind above 7 m/s for 3 consecutive hours. Our analysis concerns 34 case studies covering 2016, opportunely chosen to have a balanced dataset of WD and no WD, the latter category mainly including cases that are very difficult to predict, at the border of the two categories. The results show that the probabilistic approaches have a better skill than the categorical ones. Among the probabilistic approaches, the best result (accuracy of 82%) is obtained using the method proposed here, with the control run of GEFS used to estimate the true value and the gamma distribution to model the error distribution.

To reduce the systematic error, we test different thresholds and numbers of consecutive hours when the definition of WD is applied to model outputs. All the models show remarkably better performances after these parameters are changed. In particular, our method shows the best performance, with an accuracy of 94%. The analysis on test (leave-one-out strategy in 2016) and validation datasets (66 cases in 2017) confirms the previous outcomes. We test our procedures considering the forecast time intervals of 49-72 and 25-48 hours, where similar performances are found. In conclusion, our analysis show that the proposed method presents better performances compared to the traditional approaches for different statistical performance indicators.

Keywords: probabilistic prediction approaches, GEFS, wind day, heavy events prediction

1.1 Introduction

Despite the exponential growth of computational power in the last years, the numerical models for weather prediction are still affected by errors (Boisserie et al., 2014; Wang, 2015). These errors are due to different causes, such as the inability of the models to represent correctly both the atmospheric dynamics and the relevant physical processes, and the sensitivity of simulations to the initial conditions, which is unavoidable.

Although this kind of errors cannot be prevented, due to the intrinsic limitation in the observing systems and in numerical models, some efforts can be done to reduce the limitation of deterministic numerical systems. A typical way to surmount this problem is to develop a probabilistic approach,