Using synoptic classification to evaluate COSMOGR through Weather Dependant Verification

Gofa F.[1] and Tzeferi D.[1]

1: Hellenic National Meteorological Service (NHMS)

1 Introduction

With the multiplicity of weather prediction models and their fast-growing evolution, it is sometimes difficult for the forecaster to have an objective opinion regarding their quality. Verification analysis issued by the modelers is often not precise enough to be used as a guideline for a correct forecast. On the one hand, the forecaster would like to know when they can trust the model, on the other hand, the modeler would like to know when the model is not performing well in order to make improvements. To answer these questions, can be necessary to differentiate between different weather situations, by appropriately stratifying the verification dataset. One might suspect that the performance in winter and in summer could be different, or that, for instance, model performance in anticyclonic conditions may differ from that in a vigorous northerly flow. These differences may depend also on the geographical location, especially with respect to the presence of a land-sea border or mountains. Monthly, seasonal and annual statistical verifications are limited in that their performance is averaged over the whole spectrum of weather types the atmosphere can produce. The danger is that they can mask differences in forecast quality when the data are not homogeneous, even in terms of flow regimes.

During this study, a weather-based stratification was applied before the verification process took place. In this way, systematic model errors during the various synoptic situations could be identified.

2 Methodology of Classification

A weather type classification is a method which distinguishes between meteorological situations describing them in accordance with circulation parameters (e.g. zonality, cyclonality, position of low and high pressure systems, etc.) or local weather elements such as temperature or precipitation. Circulation parameters are often preferred since such parameters can be used very easily to relate certain features of the atmospheric circulation with local weather by statistical methods. The large number of different methods applied for classification of weather types implies open challenges to the meteorological-climatological communities [2]. The type of classification is usually adapted to a specific region and is not easily transferable to another region, or it is focused on the analysis of a specific problem so the temporal and spatial scales are adjusted accordingly.

With the aim of gaining a better understanding of model behaviour for the various types of weather that influence our area of interest, a subjective classification was adopted that is based mainly on the basic circulation patterns that the forecasters at HNMS come across in their daily experience. This tailor-made classification scheme comprises 12 different weather classes which describe the synoptic situation of the 500hPa at 12 UTC on a daily basis, with a geographical focus on the Greek region. These classes roughly separate the different weather situations into advective classes (e.g. 'northwest', 'southeast') and the accompanied convective classes 'anticyclonic' and 'cyclonic'.

Figure 1: Graphical representation of the weather classes used.

Each of these categories is related to specific weather phenomena, the intensity and amplitude of which depend greatly on the season. The categories used are presented in fig. 1 with an example of the graphical representation of the circulation. The time period covered by the study was 1 December 2009 to 30 June 2011.

Zonal C	Zonal AC	N-NW C	N-NW AC	N-NE C	N-NE AC	S-SW C	S-SW AC	S-SE C	S-SE AC	Cut-off low	Stat/ry AC
28	14	10	2	2	1	21	6	1	1	11	4

Table 1: Percentage of days in each weather regime (total number of days 577)

Table 1 shows the relative percentage of days that fall into each weather category. Particular attention must be given to gathering large enough samples to provide trustworthy verification results, i.e. interpretation of verification results for classes 'N-NE' and 'S-SE' for both convective classes is limited.

3 Results

The following section presents some of the results of the verification of the continuous and non-continuous surface parameters for the period: December 2009 - June 2011. The verification of continuous variables (e.g. T2m, Td2m, MSLP, wind speed) is typically performed using statistics that show the degree to which the forecast values differ from the observations. The Mean Error (ME) and the Root Mean Square Error (RMSE) are simple indices that provide useful information about the model's performance for a given weather parameter for a given location. Thus, for all continuous weather parameters, 3-hourly forecast values for a horizon of 72h of the 00UTC model runs were compared against the respective SYNOP data.

Looking at the overall 2m temperature verification graphs (fig. 2) for each classification class, one can identify some characteristics common to all classes. These include the distinct daily cycle of both ME and RMSE and the general trend of underestimation of temperature by the forecast model. Looking into characteristics that are related to each circulation, it can be noted that for the northern weather systems (fig. 2 c,d,e,f), there is a colder bias (underprediction) of the 2m temperature in comparison with the weather systems originating from the south. The value of ME is respectively a bit higher, and in general terms the model has a discrepancy of approximately $2 - 3^oC$.

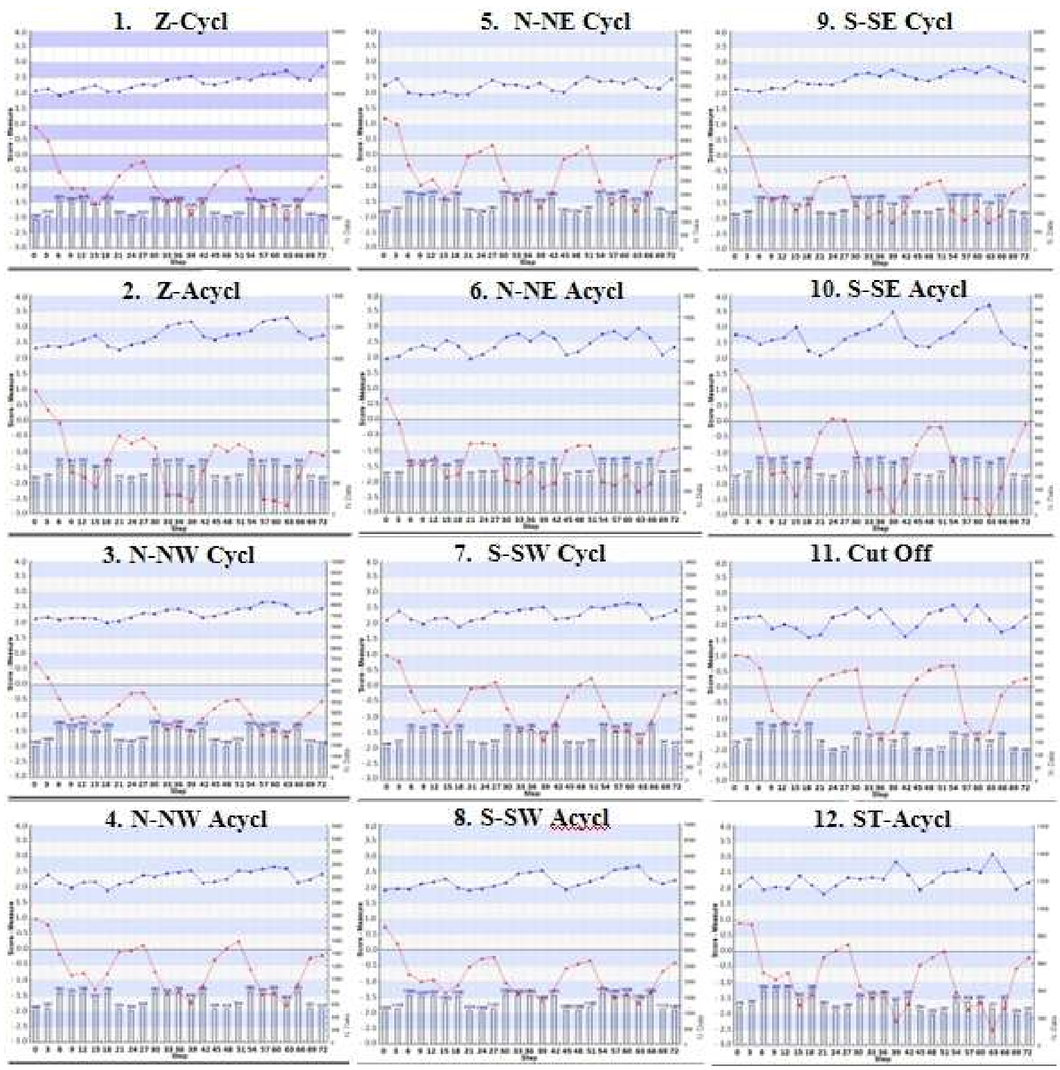

Figure 2: 2m Temp RMSE (blue) ME (red) for stratified forecasts against 80 weather stations.

A weather parameter that most if not all NWP models fail to predict correctly is the amount of clouds. COSMO-GR produces subgrid scale cloudiness using an empirical function that depends on relative humidity and height. Looking the calculated ME and RMSE for each weather type (fig. 3), we can see large differences in the ability of the model to correctly estimate the amount of clouds for each weather pattern. The error seems to be connected mainly with the cyclonality with improved performance during the passage of low pressure systems versus stable anticyclonic conditions.

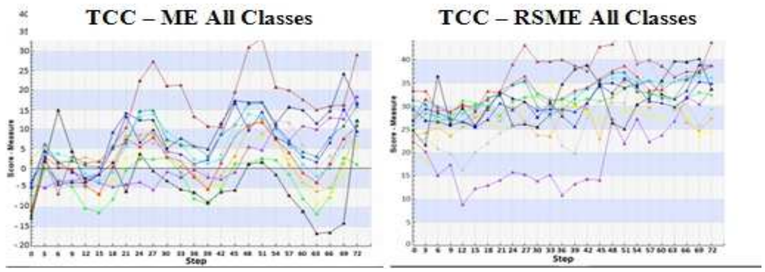

Figure 3: Cloud cover: ME (left) and RMSE (right) values for all weather classes.

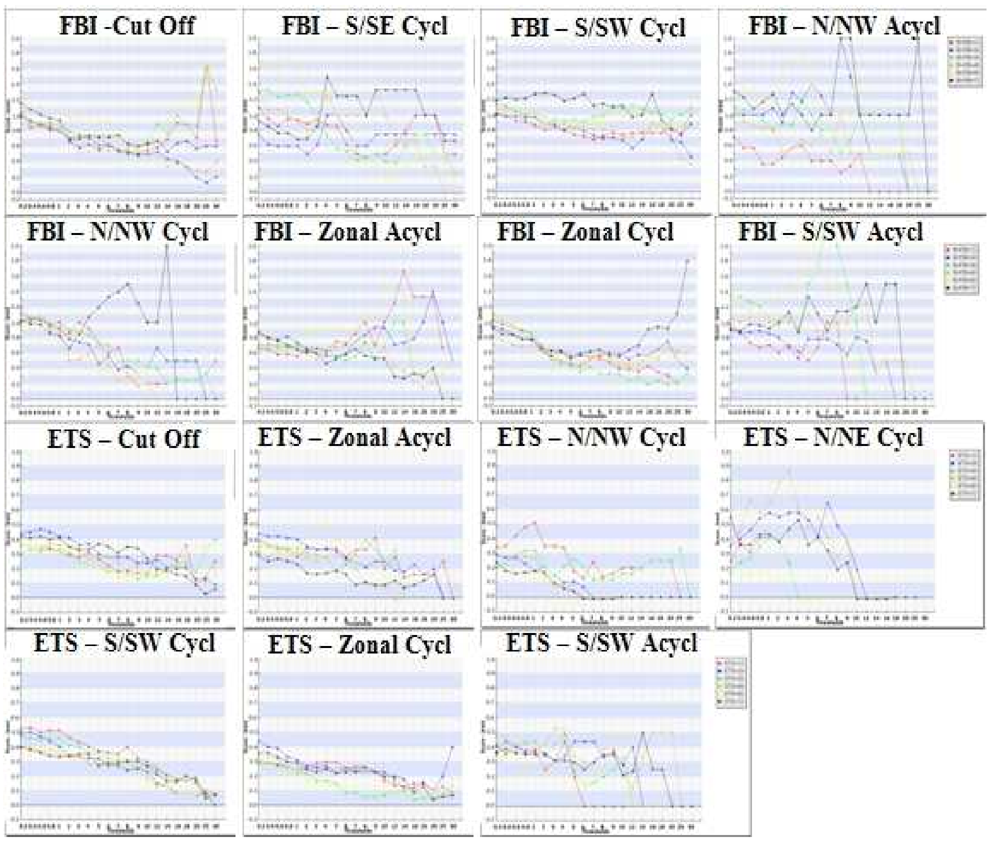

Figure 4: FBI (top rows) and ETS (bottom rows) for 12h precipitation forecasts (selected cases).

Precipitation is commonly accepted as the most difficult weather parameter to correctly predict in terms of its spatial and temporal structure due to its stochastic behaviour and any connection with specific weather systems is greatly appreciated by forecasters. The 12h-hour precipitation amounts were verified for this study and the thresholds for the precipitation amounts ranged from 0.2mm up to 30mm accumulated over each time interval. For each threshold a number of scores were calculated that provide insight into model behaviour, the most representative of which are shown in fig. 4.

The Frequency Bias (FBI) is a measure of comparison between the frequency of forecasts to the frequency of occurrences (range: $0-\infty$, perfect score= 1, $FBI > 1$ indicates over-forecast) while the Equitable Threat Score (ETS) is a measure of the fraction of correctly predicted events, adjusting for random hits (range: $-1/3-1$, perfect score= 1). In the case of precipitation, statistical indexes worsen when model resolution is increased as it produces better defined mesoscale structures, higher amplitude features and larger gradients, and inevitably leads to increased spatial and temporal errors. The results indicate that the COSMO-GR model performs well for the thresholds corresponding to small amounts of precipitation, but it fails to accurately predict large rainfall events. In cases that there was precipitation during a substantial number of days, FBI index results indicate that there is an overprediction for the lower thresholds during all cyclonic circulations, independent of the origin of the system, meaning that the model was giving us more often precipitation than truly occurred. On the other hand, the model underforecasts precipitation during heavy rainfall events ($> 8mm$), especially during anticyclone circulations. The ETS index, which provides a measure of the general performance of the model, reduces dramatically as the precipitation threshold increases. After measuring this index for all the statistically significant weather classes, it was discovered that precipitation forecasts were more successful for weather systems originating from the south-west, but this behaviour can only be better understood if a seasonal analysis is performed.

4 Conclusions

A systematic weather dependent comparison of forecast weather parameters with synoptic station measurements has been presented for the period of 2010-2011. In summary, the analysis identified: a colder bias of the 2m temperature during the passage of weather systems originating from the north, a reduced 10m wind speed error for all anticyclonic convective classes, an improved performance of cloud cover prediction when low pressure systems are present and, finally, an overprediction of the precipitation for all cyclonic circulations for the lower thresholds of precipitation, independent of the origin of the system. The limitations of this study are related to the lack of large samples for every weather class. Moreover, the weather classification scheme that was followed is not specifically geared to a specific weather parameter and may, not be the optimal choice every time.

References

[1] Gofa F., A. Raspanti, T. Charantonis, 2010: Assessment of the performance of the operational forecast model COSMOGR using a conditional verification tool. In Proc 10^{th} COMECAP Conference of Meteorology, Patras. Greece, Hellenic Meteorological Society.

[2] Maheras P., 1988: The synoptic weather types and objective delimitation in the winter period in Greece. Weather 43: 40-45.