Verification of Intense Precipitation over diverse climatological areas

Boucouvala D.[1], Gofa F. [1] and Kolyvas C.[1*]

1 Hellenic National Meteorological Service, Hellinikon GR-16777, Athens, Greece

Abstract Precipitation is a challenging weather forecast parameter to verify against observations as it is highly variable in space and time exhibiting sharp gradients in its value range. Many different score types and methodologies are used for precipitation verification. The ECMWF developed and applies SEEPS (Stable Equitable Error in Probability Space) as a headline verification score to monitor the accuracy of its operational forecasts. SEEPS differentiates the precipitation forecast performance into precipitation intensity categories (dry, light, heavy) based on the climatological cumulative distribution and in this way it takes into account the local characteristics of weather regimes in the areas that is applied. Similarly, the Symmetric Extremal Dependence Index (SEDI) is based on contingency tables and can be adjusted to the climatological distribution of precipitation at each location using geographically variable thresholds focused on extreme events, thus enabling the assessment of locally important aspects of the forecast while providing a reliable performance metric. In this study, the combination of these scores is suggested as a measure of the performance of a forecast system and its ability to predict relatively extreme rainfall events. SEDI and SEEPS indices are applied to a year-long dataset of 6-hour accumulated precipitation forecasts derived from high resolution NWP systems (COSMO4km-1km) over Greece. Both scores are aggregated over climatologically diverse regions and area means are obtained.

1 Introduction

Precipitation is a parameter highly variable in space and time and exhibits sharp gradients. These characteristics make the evaluation of precipitation forecasts a challenging task which is linked to the observation plurality and spatial inconsistency. On the other hand, there is a large number of possibilities with respect to the choice of score, verification method, spatio temporal aggregation which imply different approaches. Most of the verification scores are categorial and based on contingency tables by specifying appropriate thresholds.

Moreover, combining data from a larger number of stations during the evaluation process of NWP forecasts can produce false skill if climatologically diverse regions are combined. In particular, when interest is driven by the presence and implications of heavy precipitation events, one must aggregate regions of similar climatology that will be reflected in the precipitation thresholds that constitute an 'extreme' event in the specific area. Consequently, it is important for HIW events to analyse the relative strengths and weaknesses of commonly used statistical measures but also to highlight the importance of threshold choice especially during the aggregation of results of stations with different climatological characteristics.

This study is focused on the application of two forecast verification skill scores that are related to the geographical and seasonal variations and are already presented in Boucouvala et al. (2016). Short description of the methodology is also given in this paper. The first score is the Stable Equitable Error in the Probability Space (SEEPS) (Rodwell, 2010), which uses the categories ‘‘dry’’, ‘‘light precipitation,’’ and ‘‘heavy precipitation’’ based on the climatological cumulative distribution. The second one is the Symmetric External Dependency Index (SEDI) categorical score which is suitable for extreme events as it is equitable, symmetric and does not degenerate for rare events (unlike most categorical scores). It needs however to be adjusted on the climatological characteristics of a specific region by using appropriate thresholds. The combination of these two scores can contribute to the monitoring of model performance and the assistance in the decision making for rare events forecast.

In this study, SEEPS and SEDI scores already applied in the past to assess the predictability of coarser resolution models for the 24 hourly precipitation, are now adjusted (climatologically) and applied for 6h precipitation that is more related to high impact events and are used to evaluate the performance of higher resolution model (COSMO-GR4) and its finer (COSMO-GR1) for all seasons on an annual basis. The objective of this paper through these two metrics is to determine what perspectives these scores provide when climatology is taken into consideration, and focus on forecast assessment of heavy precipitation in order to underline model's ability to reliably capture challenging weather events.

2 Data and Methodology

Statistical Indices: SEEPS is designed to be as insensitive as possible to sampling uncertainty and equitability and adapts to the climate of the region in question. It is based on climatological probabilities of "light" and "heavy" precipitation calculated over a 30-year observations database (1980- 2009) for each station. The station climatology database was provided by ECMWF while the code calculating SEEPS was developed at Hellenic National Meteorological Service (HNMS) and adapted for this study for 6h accumulated precipitation.
The score involves three categories: 'dry', 'light precipitation' and 'heavy precipitation'. The boundary between the light and heavy categories depends on the relevant climatology for the station at which the score is being calculated. The overall scoring matrix for SEEPS is a function of p_1 (the observed climatological probability of dry weather) and p_2 and p_3 (the observed climatological probabilities of 'light' and 'heavy' precipitation, respectively) at the given observation station (with $p1+p2+p3=1$). Rodwell et al. (2010) assumed p3=p2/2, so the final scoring matrix is the following:

$$\mathbf{S} = \frac{1}{2}\begin{Bmatrix} 0 & \frac{1}{1-p_1} & \frac{4}{1-p_1} \\ \frac{1}{p_1} & 0 & \frac{3}{1-p_1} \\ \frac{1}{p_1}+\frac{3}{2+p_1} & \frac{3}{2+p_1} & 0 \end{Bmatrix}.$$

Threshold between 'dry' and 'light' category is assumed constant at 0.2mm/6h for all time periods and all stations taking into account World Meteorological Organization (WMO) guidelines (Rodwell et al., 2010). Thresholds between 'light' and 'heavy' category are extracted from the database for every station and every month. Therefore, for every month of our dataset, a 3x3 contingency table with the sum of the daily combination of modeled/observed occurrences of each of the 3 categories ('dry', 'light', 'heavy') was computed for each station. The resulting SEEPS index matrix was calculated as the scalar product of the SEEPS weights matrix and the contingency table of total available model/observation pairs for each station averaged over the number of the days of the month. The SEEPS index matrix elements represent the HD (modeled Heavy-observed Dry), LD (modeled Light, observed Dry), LH (modeled Light, observed Heavy), DH (modeled Dry, observed Heavy).

In this study, a weighting distance factor (Rodwell et al., 2010) was also applied in order to avoid over-emphasis of regions with high density. The sum of these components is the total SEEPS value for each month. For our study, the monthly values were also averaged for each season of the whole analyzed period. A perfect forecast has a SEEPS score of 0.

SEDI Symmetric Extremal Dependence Index (Ferro, 2011) is a verification index suitable for low-base (rare) events. It is a function of hit rate (H), and false alarm (F), is complement symmetric, and has a fixed range [-1,1]. It is maximized when H ->1 and F ->0 and minimized when H=0 and F=1. All contingency tables must be non-zero. It is asymptotically equitable, and values >0 imply a forecast that is better than random.

$$SEDI = \frac{\ln F - \ln H + \ln(1-H) - \ln(1-F)}{\ln F + \ln H + \ln(1-H) + \ln(1-F)}$$

Observational and Forecast data: The monthly climatological values of the stations used in this analysis are presented in Fig.3 and were extracted from the climatological map of Greece (www.climatlas.gr). The complex topography of Greece, which is dominated by both sea and orography, creates variability in both precipitation amounts and frequency, as factors such as elevation, synoptic conditions as well as the region’s exposure to wind lead to small scale climatological patterns (Gofa et al., 2019). A dataset of 6h accumulated precipitation values for 12 months (June 18 to May 2019) were used for 19 stations from various locations (continental, coastal, mountainous) (Fig.1).

With respect to forecast precipitation data, NWP data from operational at HNMS COSMO models were evaluated. Two one-way nested domains were utilized, the coarse domain (4km resolution) covered a wider Mediterranean area, while the inner domain (1km resolution) was set up over the wider geographical domain of Greece. ECMWF operational analysis is used as initial and lateral boundary conditions of the coarse domain.

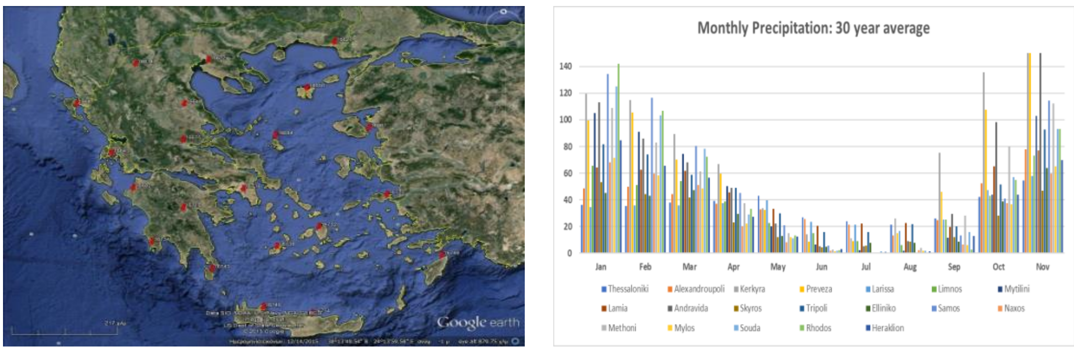

Fig. 1. Map of stations that were used for the analysis (left), monthly accumulated precipitation for all used stations (right).

3 Results

The daily distribution of 6-hourly analysis of precipitation differs for each season as shown for the months of February and June (representative months for DJF and JJA season) (Fig. 2). In JJA the precipitation in the afternoon is more intense as it is has mainly convective nature, while it is relatively equally distributed in the day in winter period. In addition (not shown), the months with the highest precipitation events were June and January, while the season with no precipitation extremes was MAM for the examined year.

Because of its linearity, the SEEPS score can be broken down into the individual contributions from the six off-diagonal elements of the 3×3 contingency table. This provides some insight into the source of error and also facilitates a comparison of the strengths and weaknesses in model intercomparison. In this study, the emphasis is given on ‘Heavy’ observed which is related to extreme precipitation events. On a seasonal basis (Fig.3), it is shown that for JJA, the largest SEEPS error contribution comes from predicting the ‘dry’ category, when ‘heavy’ was observed (HD component - orange in Figure 3). Therefore, summertime heavy precipitation events are significantly underestimated from the model. The study of the 6-hourly precipitation allows to identify that the maximum error is in the 12-18h interval, when convection mainly occurs in this season. During DJF however, the contribution of HL ('Heavy' observed 'Light' predicted) is the dominant component (purple), so the intense precipitation events are also underestimated but less than in JJA period. In addition, during winter, the daily 6h error distribution exhibits only slightly higher values at night and early morning, a sign of possible underestimation of events at this period of the day. SEEPS values for MAM are the lowest, possibly due to the lack of heavy precipitation events. The differences between COSMO-GR4 and COSMO-GR1 are not so significant on a seasonal basis; therefore COSMO-GR1 results are not shown in this paper.

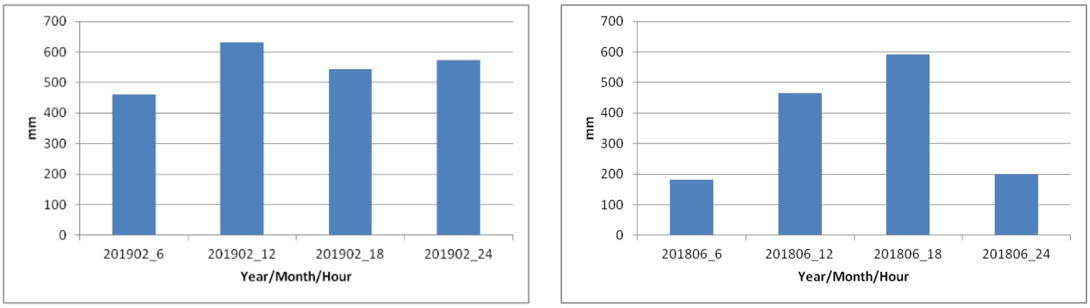

Fig. 2. Daily mean 6-h precipitation values for all stations for February (left) and June (right) (hours in UTC).

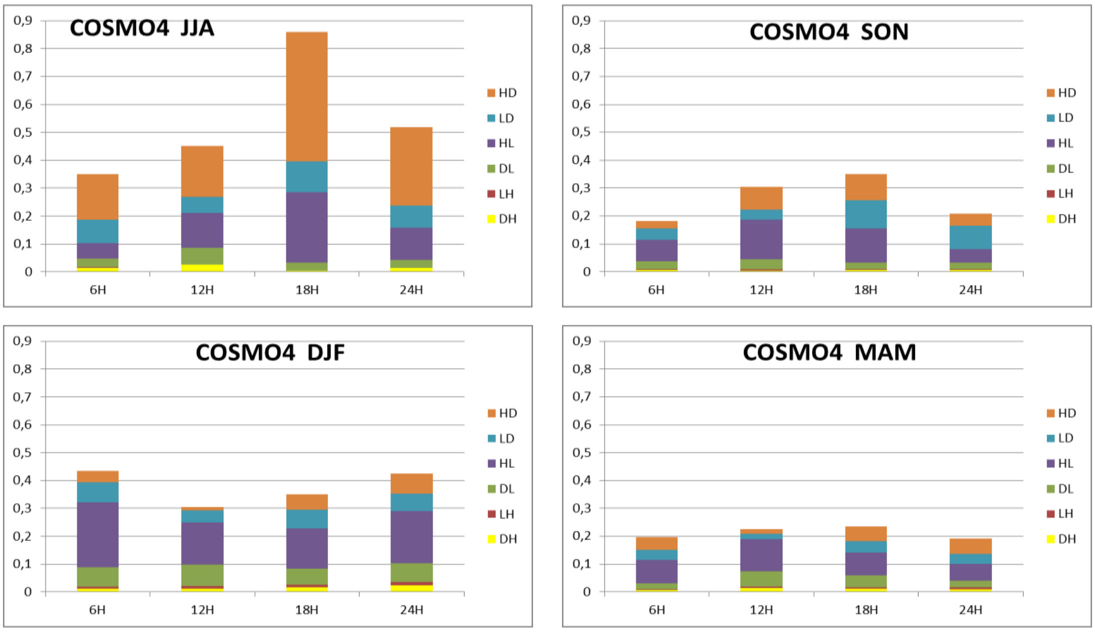

Fig. 3. Seasonal SEEPS decomposition on a 6-hourly basis (COSMO-GR4). Colors denote the different components of the index.

Monthly graphs for 12-18UTC and 18-24UTC 6h precipitation are also calculated for the SEEPS attributes HD and HL for the whole period (Fig.4). HD (Heavy observed, Dry modeled) is higher in JJA months and drops afterwards. Small secondary maxima are also exhibited in January and April. COSMO-GR1 HD error is slightly higher than that of COSMO-GR4 in JJA. One possible reason is that higher resolution models locate convective precipitation in smaller scale and point verification approach that is used in this methodology, favors the double penalty effect for small spatial misses. The component HL (Heavy observed, Light modeled) is also higher in JJA but only for the 12-18h interval. For this component, COSMO-GR4 values are slightly higher than those of COSMO-GR1, possibly due to the lower predicted values than observed as a result of smoothing related to the lower grid resolution. A secondary significant maximum is shown in January (a month with intense precipitation) implying that in winter, especially during night periods (18-24h), the heavy rain events are underestimated.

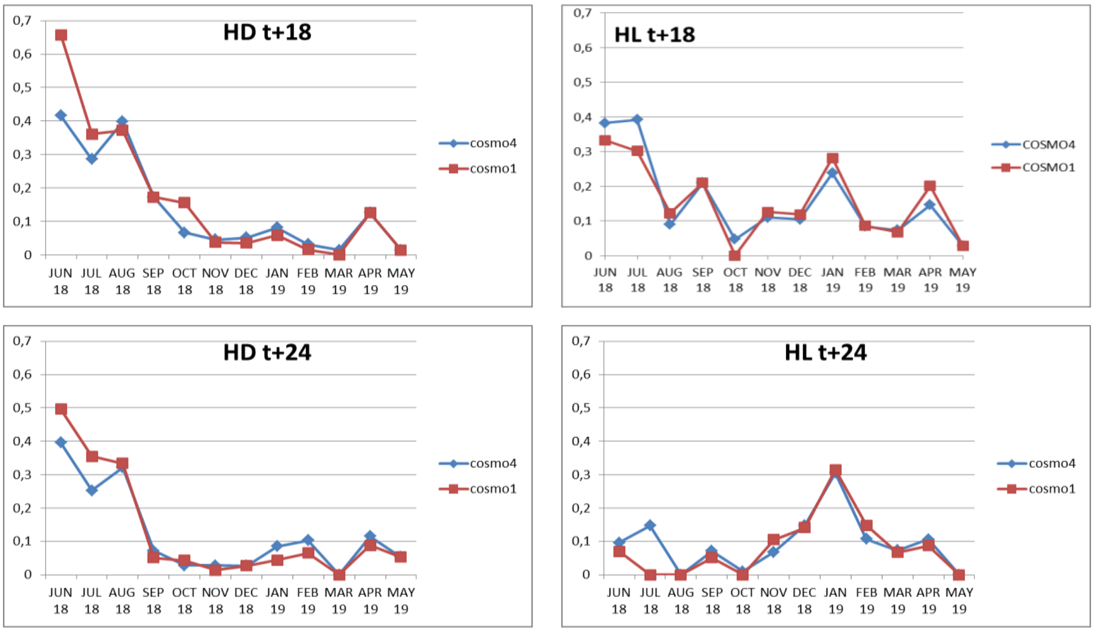

Fig. 4. HD and HL components of SEEPS on a monthly basis (COSMO-GR4, COSMO-GR) for 12-18h (upper) and 18-24h (lower)

SEDI score was also calculated for thresholds based on percentiles-values with low probability to occur (extreme). For example, the 90% percentile value means that according to climatology there is 5% chance that precipitation higher than this occurs. This threshold-based approach is more suitable when stations of different climatology are taken into account for the extraction of average scores. The monthly percentile values for each station were extracted from the 30-year database that was mentioned eralier. SEDI values for 6-12UTC and 18-24UTC intervals are plotted for each season for COSMO-GR4 (Fig. 5).

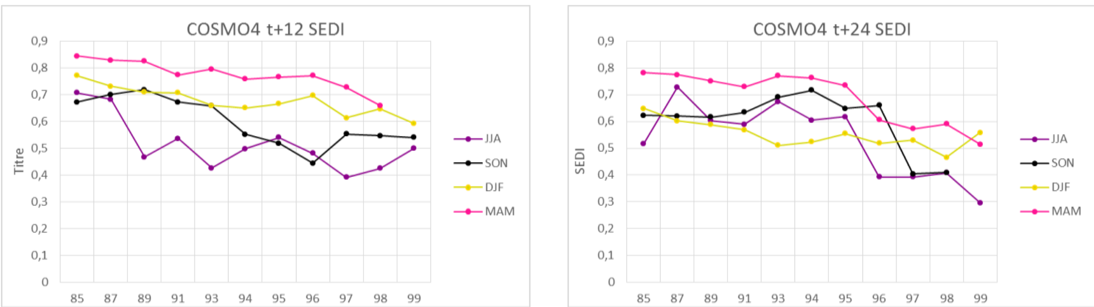

Fig. 5. Seasonal SEDI index for each season for COSMO-GR4 for 06-12UTC (left) and 18-24UTC (right) intervals.

SEDI score values (best is 1), generally reduce with increasing percentile values especially for 18-24UTC precipitation. Worse SEDI values during daytime are worse in JJA season, while score is improving in MAM. This result is consistent with the analysis when SEEPS index was considered for the same season. Moreover, SEDI score values for DJF nighttime period, are worse than daytime and this also confirms what was previously found for SEEPS score.

4 Conclusions

In this study, effort was given to include in the evaluation process of NWP precipitation forecasts, the aspect of climatology by making regions comparable using variable thresholds depending on precipitation climatology. SEEPS and SEDI scores were adjusted and applied on 6h precipitation intervals, as the focus was on the model's ability to capture in a timely manner intense precipitation events. SEEPS is based on a 3×3 contingency table and measures the ability of a forecast to discriminate between 'dry', 'light precipitation', and 'heavy precipitation', while SEDI is a verification index suitable for low-base (rare) events.

The analysis of one-year period allowed to identify the source of forecast errors for two high resolution models (COSMO-GR4 and COSMO-GR1) on a seasonal and monthly basis. The methodology that was developed, reveals the relative contribution and source of error of each model. Furthermore, it permits a more fair evaluation of forecast performance during intense precipitation events, when a model domain of variable climatology is considered. Climatologically-derived and site-specific percentile thresholds, combined with large time-windows, give large enough sample to make SEDI and SEEPS robust and informative, both suggesting that the higher resolution model is more capable (in most cases) to represent high intensity precipitation events.

References

Boucouvala D, Gofa F, Fragkouli P (2017) Complimentary Assessment of Forecast Performance with Climatologically Based Approaches. In: Karacostas T., Bais A., Nastos P. (eds) Perspectives on Atmospheric Sciences. Springer Atmospheric Sciences. Springer, Cham. https://doi.org/10.1007/978-3-319-35095-0_107

Ferro CAT, Stephenson DB (2011) Extremal dependence indices improved verification measures for deterministic forecasts of rare events. Weather Forecast 26: 699-713.

Gofa F, Mamara A.; Anadranistakis M, Flocas H (2019) Developing Gridded Climate Data Sets of Precipitation for Greece Based on Homogenized Time Series. Climate 7: 68. https://doi.org/10.3390/cli7050068

Rodwell MJ, Richardson DS, Hewson TD, Haiden T (2010): A new equitable score suitable for verifying precipitation in NWP. Q. J. R. Meteorol. Soc.: 136, 1344-1363.