

Comparison of selected methods of analysis for reconstructed fields of precipitation in climate scenarios over Poland

Krystyna Konca-Kedzierska¹ 

Received: 22 July 2014 / Accepted: 20 August 2015 / Published online: 11 September 2015
© The Author(s) 2015. This article is published with open access at Springerlink.com

Abstract Obtaining information about future climate change is essential to the development of adaptation strategies, but no less a significant problem than gathering the data is assessing its credibility. The task of assessing the credibility of scenarios is solved by evaluating climate models in the referenced period. Methods of comparing meteorological fields in the grid points often lead to an overestimation of the error for precipitation. The way precipitation affects the environment—a much wider concern than just where the precipitation occurs—allows for the assumption of some tolerance in the verification of its location. The main objective of this study is to find a way to evaluate climate scenarios without taking into account the details of location and intensity of precipitation. One of the methods for a less restrictive examination of compatibility of precipitation fields is to compare their properties in the neighborhood of the point rather than the value at the point. Maps of climate indices and probabilities of exceeding some threshold are evaluated. The second method is a cluster approach, and each field of precipitation is replaced by a set of clusters found in the phase space $\{(lon,lat,cl)\}$, where cl is the class of precipitation. Instead of comparing the fields of precipitation, it deals with the set of exemplars of the clusters. Further analysis concerns the spatial distribution of exemplars within a selected class of precipitation. This generalized description allows for comparison between precipitation fields with tolerance in respect to exact location and intensity. Comparison with such descriptions allows one to assess whether

two given fields have a similar property in a very general sense.

1 Introduction

The credibility of climate scenarios is based largely on the ability of climate models to reproduce the observed climate. Evaluation of the quality of the reconstruction of the climate models is to compare the fields of meteorological parameters. Fields of continuous parameters, such as temperature, can be evaluated by comparing the values at the grid points. Statistical evaluation of these point differences gives a realistic view of the quality of the reconstruction of the climate. The situation is different in the case of phenomena such as precipitation. This phenomenon occurs in a limited area and is usually short-lived, so even in high-resolution models, it is difficult to precisely reconstruct. However, minor deviations in the duration and location of the phenomenon should not be crucial in assessing the adequacy of fields. This problem also occurs in the verification of numerical weather forecasts and has been studied in the project “Spatial Forecast Verification Methods Inter-Comparison Project” (Ahijevych et al. 2009a, b). In the view of many publications related to this project, there are more appropriate methods to analyze the compatibility of precipitation fields (Gilleland et al. 2010; Ahijevych et al. 2009a, b; Gilleland et al. 2009). The methods and ideas adopted from this project can be applied to the evaluation of climate models. An example of the application of techniques from this project to climate data can be found in the work of Moise and Delange (2011). In this work, an object-oriented technique for contiguous rain areas (CRAs) was developed.

The aim of our research is to present methods for comparing very general properties of precipitation fields, which does not take into account the details of location and intensity of

✉ Krystyna Konca-Kedzierska
krystyna.konca@imgw.pl

¹ Institute of Meteorology and Water Management, National Research Institute, 61, Podlesna str., 01-673 Warsaw, Poland

precipitation. This assumption allows the use of climate indices to evaluate climate scenarios. The analysis using the climate indices is carried out in two ways. First, absence or occurrence of precipitation is analyzed by using the principle of the nearest neighborhood. The properties of precipitation at the grid point are replaced by properties of nearby fields at this point. Then, the percentage of days with precipitation exceeding the threshold of 1 mm from EOBS data—ENSEMBLES Observations gridded dataset (Haylock et al. 2008)—is compared with that for ERA40—reanalysis of meteorological observations produced by the European Centre for Medium-Range Weather Forecasts (ECMWF) (Uppala et al. 2005) and model scenarios. Evaluation is done comparing the field of percentage of wet days. The maximum daily total precipitation is also analyzed.

The second method is based on a cluster analysis where the value of precipitation is replaced by a class of precipitation defined by the thresholds. This method identifies certain objects consisting of points of precipitation fields with a similar precipitation class that are located close to each other. Each such object is represented by the exemplar, its location, and its class of precipitation. Then, we compare the characteristics of the occurrence of these objects in the subregions of the domain.

2 Methods

The method based on the principle of the nearest neighborhood analyses precipitation at the grid point, considering values in its immediate vicinity, consisting of nine points. Due to the resolution (0.25°) of the data, this neighborhood is a $50\text{-km} \times 50\text{-km}$ rectangle and is the smallest symmetric neighborhood of the point in the net. This method compares the precipitation fields using the climate indices eca_rr1 —wet day index per time period (the number of days where RR is at least 1 mm)—and eca_rx1day —highest 1-day precipitation amount per time period (the maximum of one day sum of precipitation in the period)—both calculated by the program Climate Data Operators (CDO 2015). The wet day is defined as RR—daily sum of precipitation exceeding the threshold of 1 mm ($RR \geq 1$). The index eca_rr1 is used to evaluate the percentage of wet days. Wet days are defined on the basis of the average values in the neighborhood. Further analysis concerns the comparison of the possibility of occurrence of areas in which the number of wet days N_{wet} exceeds the threshold of 50 %, or 33 % of the total number of days N . In the following step, the maps of probabilities of exceeding p_{50} or p_{33} are visually evaluated. The probabilities at each grid point are calculated on the basis of the value N_{wet} at K points in the neighborhood. It is assumed that $p_{50} = k_{50}/K$ where k_{50} is the number of grid in the neighborhood for which $N_{wet} > 50 \% N$ (p_{33} and k_{33} are defined analogously). The next step of analysis is to compare maps of probabilities p_{33} and p_{50} . The index eca_rx1day is used to assess the

probability of maximum precipitation greater than 50 mm. In this case, the analysis refers to the value of $exca_rx1day$ at K points in the neighborhood. The probability of exceeding by RR_{max} the threshold 50 mm ${}^{50}p_{max} = {}^{50}k_{max}/K$, where ${}^{50}k_{max}$ is the number of points for which $exca_rx1d > 50$. The maps of spatial distribution of p_{33} , p_{50} , and ${}^{50}p_{max}$ are evaluated visually and by using a fractional skill score FSS (Ebert 2008) (Fig. 1). Comparison of precipitation fields using these tools are presented in Section 4.

Motivation for the use of cluster analysis to verify the precipitation fields, reconstructed by RCM models, is to find a way to compare the very general properties of these fields. First, generalization refers to the daily total precipitation. Studies are subject to the data categorized by the use of threshold values. The values of daily total precipitation RR are divided into five categories:

- No rain: $RR < 0.1$ mm (for this case variable $cl=0$)
- Very light rain: $0.1 \text{ mm} \leq RR < 1$ mm ($cl=10$)
- Light rain: $1 \text{ mm} \leq RR < 5$ mm ($cl=20$)
- Moderate rain: $5 \text{ mm} \leq RR < 10$ m ($cl=30$)
- Strong rain: $RR \geq 10$ mm ($cl=40$)

In this way, the field (lon, lat, RR) is transformed into (lon, lat, cl). Then, the cluster analysis method applied to these data allows us to create clusters—groups of points lying close to each other with the same precipitation class. It is difficult to identify in advance the appropriate number of clusters for every precipitation field. Marzban and Sandgathe (2008) used a hierarchical cluster analysis method for forecast verification. Application of hierarchical cluster analysis resulted in the number of clusters varied from the total number of grid points in the data down to 1. These numbers were interpreted as a resolution of analysis. In Konca-Kedzierska (2012) were some comments on the selection of the clustering procedure, which showed the need for individual choice of a number of clusters for each field. The method chosen for calculation is a method called by Frey and Dueck (2007) as “affinity propagation,” which takes the measures of similarity between pairs of data points as input and simultaneously considers all data points as potential exemplars. The affinity propagation method was implemented in System R as a package APCluster (Bodenhofer et al. 2011). This method provides individual selection of the number of clusters, suitable for the particular field of precipitation classes.

Based on the results of clustering, we have assessed a very overall view of the spatial distribution of precipitation classes. The result of clustering procedure is a set of clusters represented by their exemplar (lon_{exp} , lat_{exp} , cl_{exp}), i.e., their localisation and precipitation class. As shown in Fig. 2, the domain has been split into four areas: northeast (NE), northwest (NW), southeast (SE), and southwest (SW). Of course, there are other possible divisions one could take into account

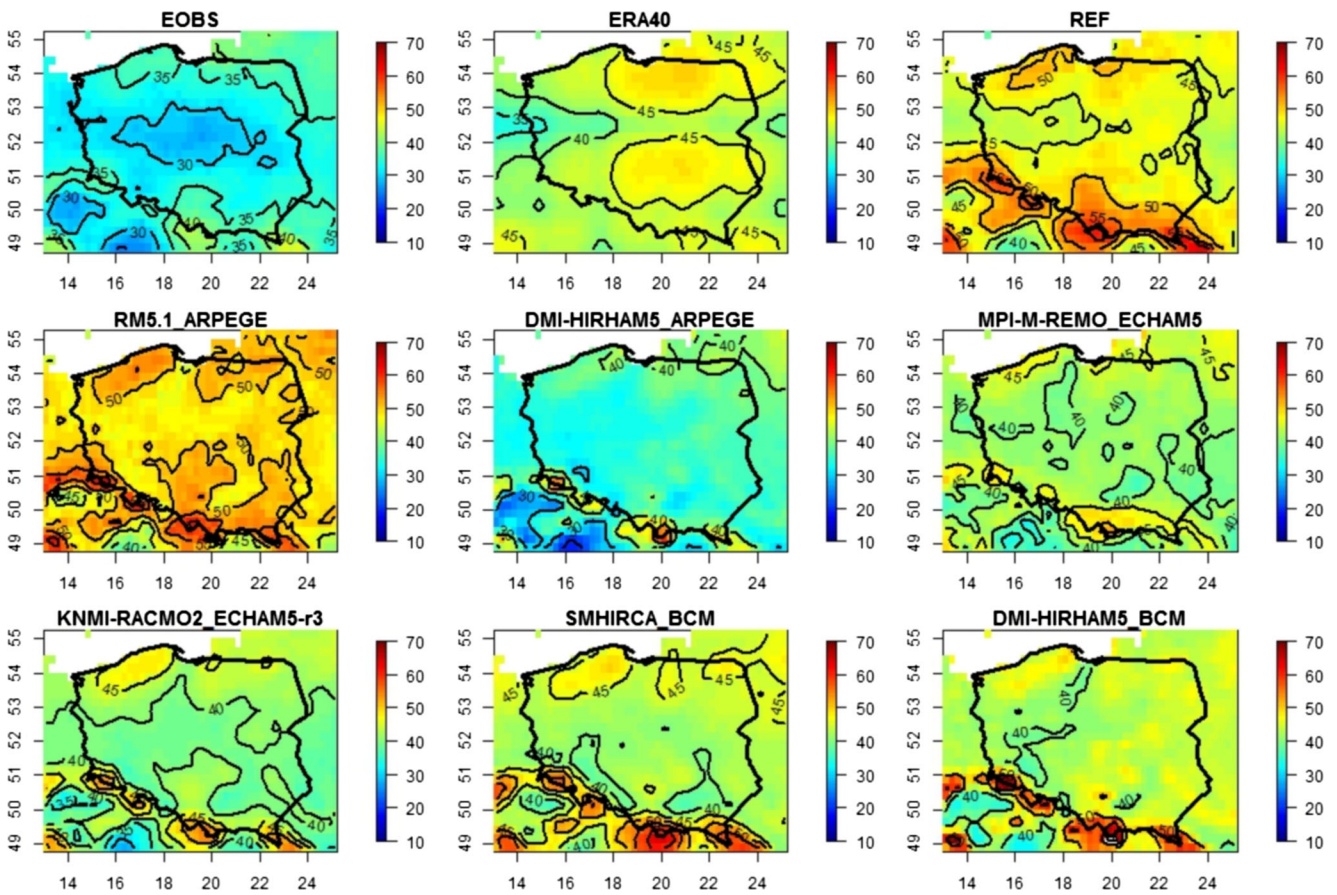


Fig. 1 Spatial distribution of the percentage of wet days. N_{wet} : $RR \geq 1$ mm

other than the geographical aspects of the domain, for example, an administrative division. It is assumed that precipitation $RR > 10$ mm occurs in subregions of the division, if there is an exemplar with $cl=40$. Considering the position of exemplar, each day of the period 1971–1990 can be classified into one of 16 situations:

- 00—exemplars with $cl=40$ does not appear in any of the regions
- NE—exemplars with $cl=40$ are found only in the area NE
- NW—exemplars with $cl=40$ found only in the area NW
- NESE—exemplars with $cl=40$ found in the areas NE and SE
- NENWSESW—exemplars with $cl=40$ are found in all areas

The next step is to compare the frequency of these situations in the observational and model data. This enables the possibility to use a methodology for examining the compatibility of contingency tables. Furthermore, as a measure of the differences between histograms, we calculate the sum of the percentage differences in the frequencies in the histogram

classes TSD^m defined by the following formula:

$$TSD^m = \frac{1}{N} \sum_{i=1}^{16} n_i^m \times 100$$

where

$i=1, \dots, 16$ possible locations for $cl=40$ (“00,” “NE,” “NW,” ..., “NENWSESW”)

$m=ERA40, REF, DMI$ —indicator of the analyzed model
 $n_i^m = \text{abs}(N_i^{EOBS} - N_i^m)$ —absolute value of the difference in the number of cases

N_i^{EOBS} —the number of cases for the i th location in the EOBS data

N_i^m —the number of cases for the i th location in the model m data

This method does not take into account co-occurrence of events; it only compares the fit of distributions sufficient to evaluate the climate model. Detailed analysis of the occurrence of clusters with precipitation class $cl=40$ is discussed in Section 5.

3 Data

The study concerns precipitation fields that were reconstructed by regional climate models RCM for the period of 1971–

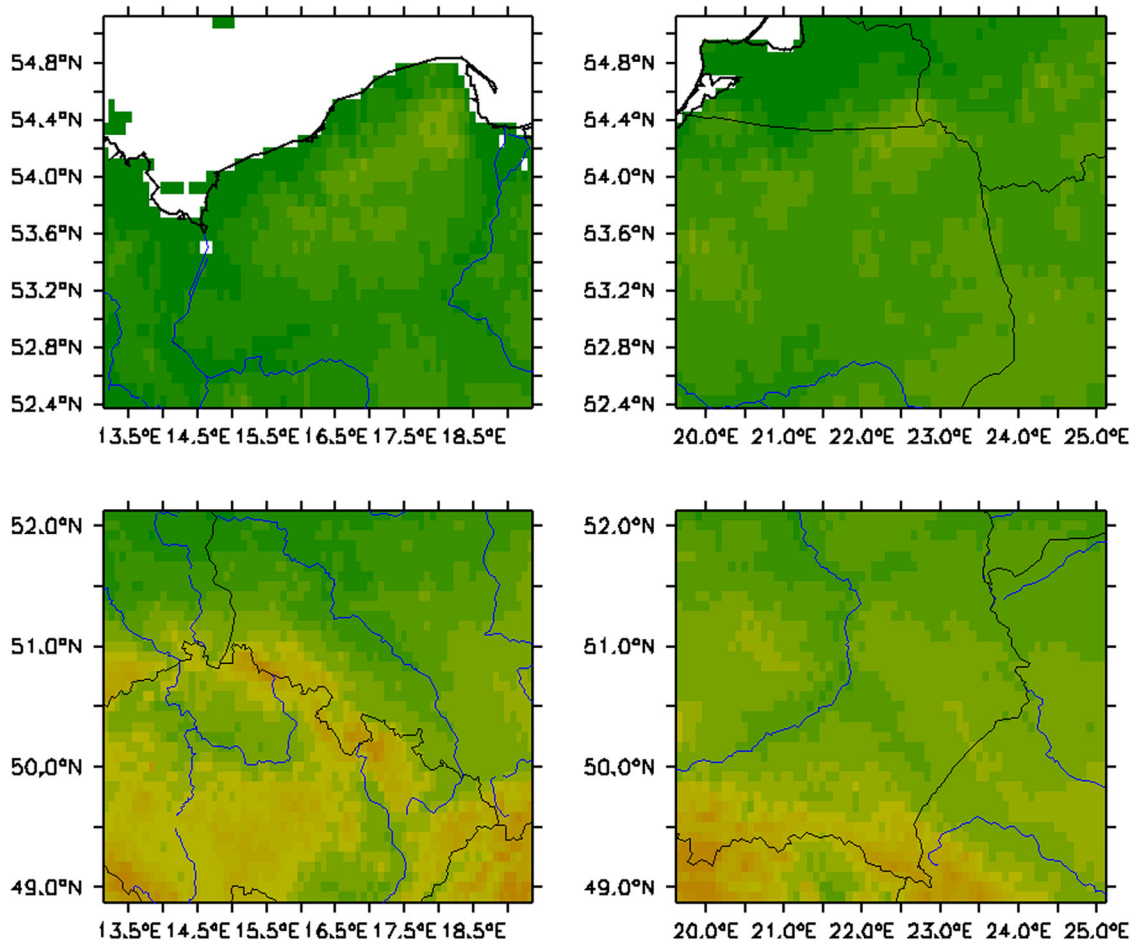


Fig. 2 The domain division into four regions NE, NW, SE, and SW which correspond to the geographical parts of it

1990. The research refers to the domain 13 E–24 E and 48 N–55 N, which surrounds Poland. The resolution of the data is 0.25°, meaning that the distance between the grid points is about 25 km. Climate during the reference period is described by the E OBS (Haylock et al. 2008) data downloaded from the European Climate Assessment & Dataset project ([http://eca.](http://eca.knmi.nl/dailydata/index.php)

<http://www.ensembles-eu.org/>) and the project KLIMAT (<http://www.ensembles-eu.org/>). Also used as reference data is the reanalysis from project ERA-40 (<http://data-portal.ecmwf.int/>). The climate scenarios are taken from the Polish project “Climatic service” (http://klimat.icm.edu.pl/serv_climate.php) (based on data published in project ENSEMBLES <http://www.ensembles-eu.org/>) and the project KLIMAT

Table 1 The list of models being compared in the method of the nearest neighborhood

	Regional model	Reference	Global model
1	REF	Wibig et al. 2012	ECHAM5
2	RM5.1_ARPEGE	Radu et al. 2008	ARPEGE
3	DMI-HIRHAM5_ARPEGE	Christensen et al. 2007	ARPEGE
4	MPI-M-REMO_ECHAM5	Jacob 2001; Jacob et al. 2001	ECHAM5
5	KNMI-RACMO2_ECHAM5-r3	Meijgaard et al. 2008	ECHAM5
6	SMHIRCA_BCM	Kjellström et al. 2005	BCM
7	DMI-HIRHAM5_BCM	Christensen et al. 2007	BCM

carried out in IMWM. Table 1 contains a list of the selected RCM scenarios, together with information on references and global circulation models, GCM, from which the boundary conditions were taken.

4 The principle of the nearest neighborhood

Percentage of wet days is calculated on the basis of climate indices *eca_rr1* from the program Climate Data Operators. The results of calculation presented as maps of percentage of wet days, are shown in Fig. 1. Most similar to the spatial distribution of the observational EOBS data is the distribution of the model DMI-HIRHAM5_ARPEGE. From Fig. 1, it is easy to see that the models overestimate the percentage of wet days. For observational EOBS data, it is verified that the number of wet days does not exceed 50 % at any point in the domain. A similar situation is in the case of ERA40 data, while the models allow for the occurrence of areas where the number of wet days exceeds 50 %. In the cases of models DMI-HIRHAM5_ARPEGE and MPI-M-REMO, these unreal areas are very small and occur in mountainous terrain in the southern Polish border. The largest areas with the number of wet days exceeding 50 % are found for models REF and RM5.1. The fractional skill score FSS is used to objectively compare the maps of p_{50} . The values of FSS are shown in Table 2. The highest value of FSS, which means the best compatibility with EOBS data, is achieved by model DMI-HIRHAM5_ARPEGE.

The probability that the number of wet days exceeds the threshold of 33 % p_{33} is also examined. Nearly all used models and ERA40 data show $p_{33}=1$ throughout the territory of Poland. The exception is the DMI-HIRHAM5_ARPEGE model, which represents a large area in the western part of Poland with $p_{33}=0$ which is also found in the EOBS data.

Spatial distributions of RR_{max} are visually compared on maps of *eca_rx1day*. On this basis, we conclude that the maps for the models DMI_HIRHAM5_ARPEGE, DMI_HIRHAM5_BCM, and KNMI-RACMO2 are most similar to that for observational EOBS data. However, for both models DMI_HIRHAM5, there is no area with $RR_{max}>100$ mm in the western Polish border which occurs in the observational EOBS data. This area is slightly reconstructed by a KNMI-RACMO2 model. The next step is to determine ${}^{50}p_{max}$ the probability of exceeding the threshold 50 mm by maximum daily precipitation RR_{max} . It is done on the basis of *eca_rx1day* using the principle of the nearest neighborhood. Then, the maps of the probability of occurrence of maximum precipitation over the threshold of 50 mm ${}^{50}p_{max}$ are evaluated. The size of the area in which ${}^{50}p_{max}$ is close to 1 is too high for all models. From the visual assessment, it follows that the spatial distribution of the DMI-HIRHAM5_BCM model is

Table 2 The values of fractional skill score FSS for maps of probability of exceeding the threshold by climate indices for analyzed models compared to the EOBS data

Model	p_{50} (eca_rr1)	${}^{50}p_{max}$ (eca_rx1day)
1 ERA40	0.678	0.236
2 REF	0.677	0.568
3 RM5.1_ARPEGE	0.683	0.588
4 DMI-HIRHAM5_ARPEGE	0.783	0.624
5 MPI-M-REMO_ECHAM5	0.683	0.541
6 KNMI-RACMO2_ECHAM5-r3	0.689	0.635
7 SMHIRCA_BCM	0.683	0.557
8 DMI-HIRHAM5_BCM	0.687	0.622

most similar to the distribution of the observational EOBS data. For the ERA40 data almost in the entire domain, the likelihood ${}^{50}p_{max}$ is equal to zero, while for EOBS data in the north, west, and south parts of the domain, there are significant areas with ${}^{50}p_{max}=1$. The objective evaluation of the field ${}^{50}p_{max}$ is set using a fractional skill score FSS, and the result are shown in Table 2. FSS highest values (FSS>0.6) are achieved for the aforementioned three models, while for the data ERA40, it is the lowest and FSS=0.24. Application of climate indices, a principle of the nearest neighborhood, and a fractional skill score FSS allows you to specify the method of comparing climate scenarios. Results obtained using a fractional skill score FSS are consistent with the subjective assessment of climate indices maps.

4.1 The cluster analysis method

ERA40 reanalysis and REF data, obtained in IMWM from RegCM model, are selected for comparative analysis.

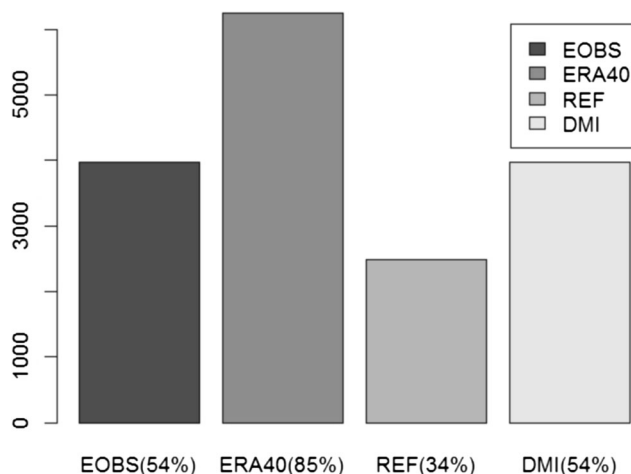


Fig. 3 The number of cases for the absence of significant precipitation in the domain

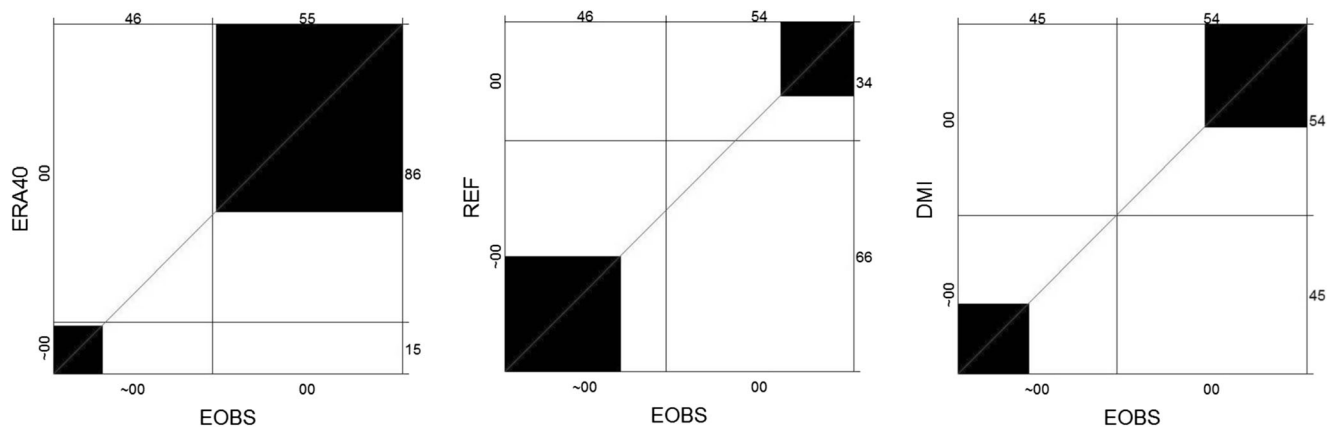


Fig. 4 Bangdiwala's agreement plots for occurrence (~00) and absence (00) of class cl=40. The numbers correspond to the percentage of cases

DMI_HIRHAM5_ARPEGE model is also selected from the other models because for him the value of FSS obtained in Section 4 is the highest. In this section, the model DMI_HIRHAM5_ARPEGE is denoted by the acronym DMI.

The cluster analysis method is used to analyze the spatial occurrence of the significant precipitation. It is defined that in the subregion of the domain, the selected class of precipitation occurs if the exemplar with this class is located in this region. It is analyzed that the occurrence of class cl=40 which means the precipitation is above 10 mm. Such situations, where the whole area has no significant precipitation (no exemplar with class cl=40) are numerous, and there are significant differences in the number of such cases for each data set. As shown in Fig. 3, the highest number of such cases is for the ERA40 data, nearly two times more than for EOBS. The number of such situations for the DMI model (3965) is almost equal to that for the EOBS data (3972); the difference is only 7 days.

Analysis of the occurrence of class cl=40 is carried out using the method of contingency table. Pearson's chi-squared χ^2 test is performed, and indicators such as contingency coefficient, Phi-coefficient, and Cramér's V are calculated (Meyer et al. 2012).

Three selected situations of coexistence of class cl=40 in observed EOBS data and the model data are explored. The first concerns the absence of class cl=40 in the whole domain (situation denoted by "00"), and the opposite event is the appearance of this class anywhere in the domain (situation denoted by "~00"). Figure 4 shows Bangdiwala's agreement plots, which provide a simple graphic representation of the strength of agreement in a contingency table. It can be noted that the biggest adjustment is for the ERA40 data because, in

this case, the black areas in the figure, which are the measure of co-occurrence, are the largest. Numbers describing Bangdiwala's agreement plots in Fig. 4 are the rounded marginal percentage of each of the situations "00" and "~00." Table 3 contains the exact percentage distribution for contingency tables that correspond to agreement plots in Fig. 4. The marginal homogeneity is shown directly by the relation of the dark squares to the diagonal line. In Fig. 4, we see that the ERA40 data overestimate situation "00" (85.4 %), whereas the REF data underestimate it (34.1 %). The DMI model data shows the marginal homogeneity, because in this case all squares are located on the diagonal line. This means that the percentage of data EOBS and DMI are identical. The highest values of several measures of strength of agreement are obtained for ERA40 data, because in this case the co-occurrence of ERA40 and EOBS reaches 67.8 %. The lowest values were obtained for DMI data because, in this case, the co-occurrence of classes "00" and "~00" is very rare (49.4 %). The χ^2 test (at the level of significance $p=0.01$) only for DMI does not give reason to reject the hypothesis of independence with EOBS data. Analogous analysis is performed for the division of the data into three categories: "00" for absence of cl=40 in the whole domain, "S" for cl=40 located in south part of the domain (this part includes mountains), and "~S" for cl=40 located in the rest of the domain. Bangdiwala's agreement plots for this case are shown in Fig. 5. The number describing Bangdiwala's agreement plots in Fig. 5 are the numbers of cases. On the chart for the ERA40, one can see the strong underestimation of the number of occurrences of cl=40 in the southern part of the domain. This is correct, but only 24 % of the occurrence of cl=40 in this part of the domain

Table 3 Contingency tables with the percentage for analysis of the occurrence or absence of class cl=40

	EOBS 00	EOBS ~00		EOBS 00	EOBS ~00		EOBS 00	EOBS ~00	
ERA40 00	53.8	31.6	85.4	21.2	12.9	34.1	29.0	25.3	54.3
ERA40 ~00	0.6	14.0	14.6	33.2	32.8	65.9	25.4	20.4	45.7
	54.4	45.6		54.4	45.6		54.4	45.6	

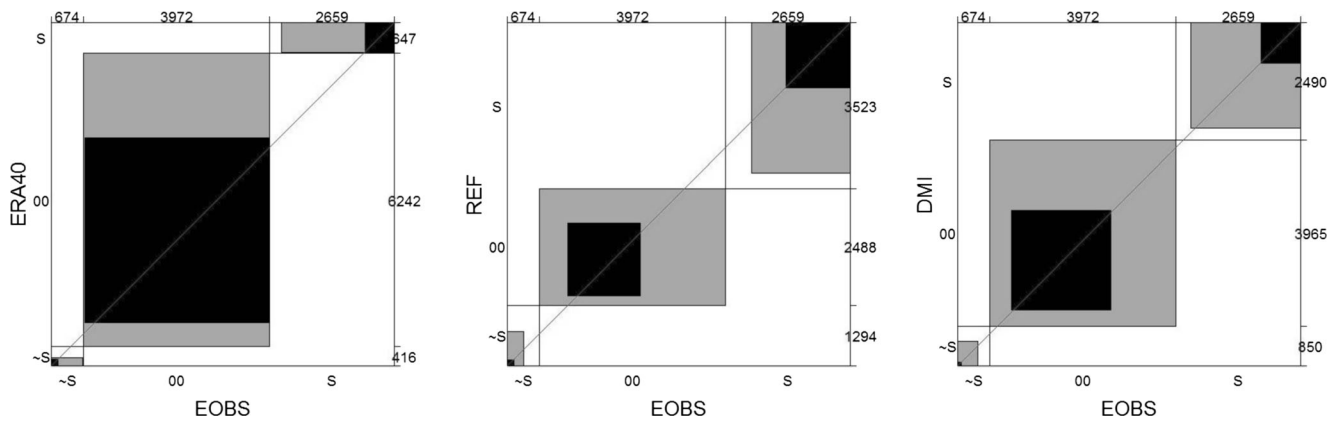
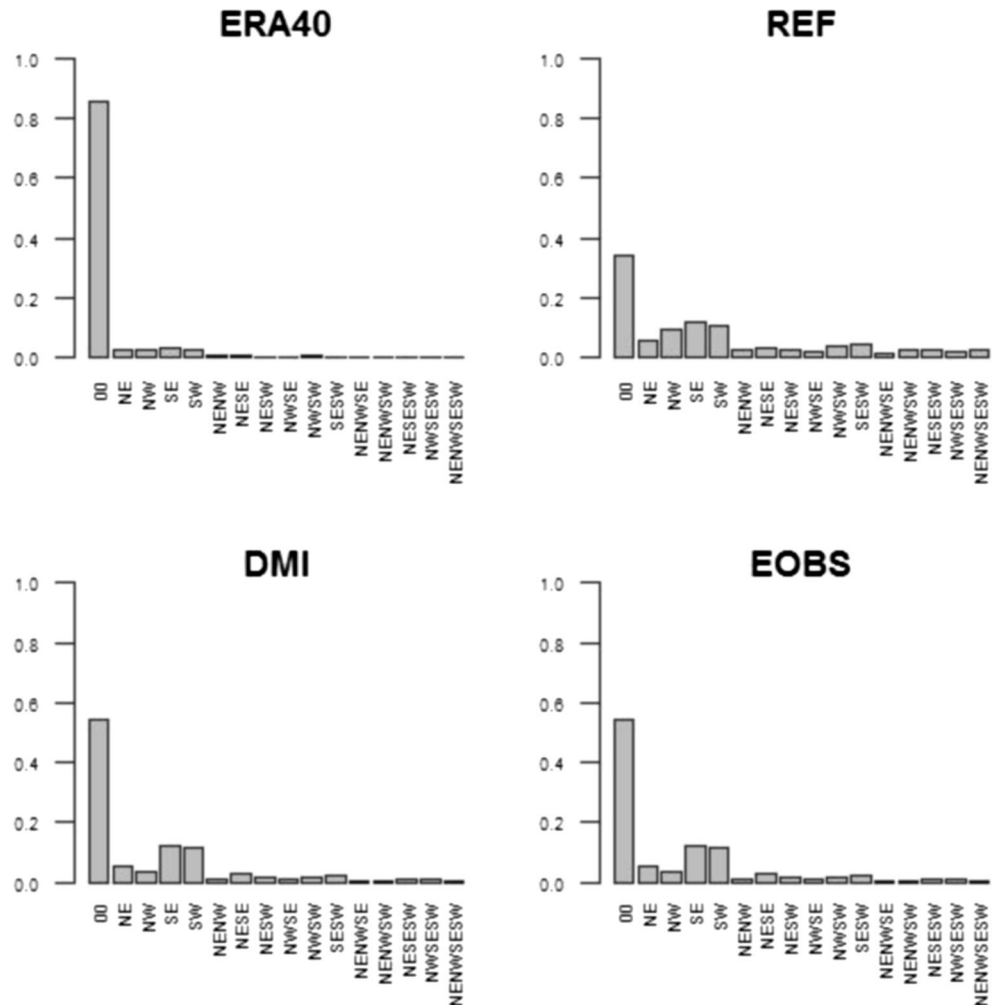


Fig. 5 Bangdiwala’s agreement plots for three locations for $cl=40$: “00”—absence, “S”—in the south part, and “~S”—in the rest part of domain. The numbers in the graph correspond to the number of cases

in EObs data is reconstructed. For EObs data, the largest group is the case of “00” (55 %), while for REF data, the largest group is the case of “S” (48 %). Overestimation of the number of cases of significant precipitation in the southern part of the domain by the REF model is about 33 %. The number of situations “S” for the DMI model is underestimated

by 6 %. Measures of association of these contingency tables have values similar to the values for the two-class tables “00” and “~00.” Similar calculations are also performed for all 16 classes of possible locations of significant precipitation $cl=40$. The contingency coefficient in this case is 0.7 for ERA40, 0.23 for REF, and 0.18 for the DMI model. ERA40 appears

Fig. 6 Histograms for the distribution of frequency of precipitation class $cl=40$ in the regions to which the domain was divided. 00, NE, NW, ..., NENWSESW describe the possible coexistence of class $cl=40$ in the subregions



to be the best model, but in this case there is a very big overestimation of the absence of significant precipitation “00” by approximately 57 %. Taking into account the co-occurrence of events, the percentage of conformity with EOBS data (in Bangdiwala’s agreement plots, this corresponds to a total of the surface of the black rectangles) are as follows: ERA40 59 %, model DMI 32 %, and model REF 25 %. On the other hand, from the analysis, it follows that the highest marginal homogeneity occurs for model DMI, and such evaluation is more important from the point of view of climatology.

In Fig. 6, histograms of location of precipitation events for the class $cl=40$ in all 16 locations are shown. The histograms for both models REF and DMI have a better fit than histogram for ERA40 reanalysis. The visual assessment of the similarity of these histograms shows that the histogram most similar to the EOBS histogram is that for DMI model. In cases of histograms for EOBS and DMI data, there are the two clearly predominant locations (except in the case of “00”) that means the southern parts of the domain SE and SW. Values of TSD index for the analyzed data are as follows: $TSD^{ERA40}=64\%$, $TSD^{REF}=42\%$, and $TSD^{DMI}=14\%$ and confirm that the best fit is for the model DMI.

5 Conclusion

This paper considers two ways to assess whether the precipitation fields are compatible without comparing their values at the grid points. The first method (Section 4) compares the properties of the fields in the neighborhood of the grid point instead of its value at that point. The probability of exceeding the threshold of number of wet days (daily total $RR>1$ mm) and maximum precipitation threshold are specified based on climate indices in the neighborhood. The correctness of the reconstruction of past climate is assessed by comparing the maps for EOBS data and regional models. The results of visual assessment are confirmed by the values of fractional skill score FSS. The best rating in this analysis is obtained by DMI_HIRHAM5_ARPEGE model. Furthermore, in the analysis of the number of wet days, it is found that models overestimate this number, especially in mountainous areas in the southern Polish border. In contrast to the observed EOBS data, models allow for regions in which the number of wet days exceeds 50 %. The analysis with the climate index eca_ex1day shows that regional models RCM better estimate the RR_{max} than ERA40 reanalysis.

The second method (Section 5) examines the location of the cluster representatives in the assumed division of the domain. Analysis of the occurrence of significant precipitation (daily sum of precipitation $RR>10$ mm) shows that for ERA40 data, the absence of such precipitation in the entire domain is overstated and is 86 %. For EOBS data, cases where there is no significant precipitation are 54 %. All these cases

are accurately reconstructed in the ERA40 data. For this reason, measurement of compliance in the contingency tables and Bangdiwala’s agreement plots, which are sensitive to the simultaneity of events, are best for the ERA40 data. Co-occurrence of events is not relevant for climate assessment, so it is sufficient to evaluate the histograms of the exemplar location. In this evaluation, closest to the histogram for EOBS is the histogram for model DMI_HIRHAM5_ARPEGE. The difference between these histograms can be estimated at 14 %. Categorized data allowed for comparing the scenarios and the reference fields for different levels of precipitation. Further research is planned on the use of cluster analysis for the other classes of precipitation that will allow for a full analysis of the model fit.

Presented examples allow us to obtain a very general description of the time series of precipitation fields reconstructed by RCM model. However, it is possible on this basis to assess which models better reconstruct precipitation fields for the reference period.

Acknowledgments This article is part of the sustainability of Project CLIMATE, Impact of climate change on environment, economy, and society (changes, effects, and methods of reducing them, conclusions for science, engineering practice, and economic planning) No Project POIG. 01.03.01-14-011/08-00-Innovative Economy Operational Programme.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Ahijevych DA, Gilleland E, Brown BG, Ebert EE (2009) The spatial verification method inter conversion project. 23rd Conference on Weather Analysis and Forecasting/19th Conference on Numerical Weather Prediction (1-5 June 2009) (Omaha, NE) , <http://ams.confex.com/ams/pdfpapers/153061.pdf> Accessed 1 March Aug 2012
- Ahijevych DA, Gilleland E, Brown BG, Ebert EE (2009b) Application of spatial verification methods to idealized and NWP-gridded precipitation forecasts. *Weather Forecast* 24:1485–1497. doi:10.1175/2009WAF2222298.1
- Bodenhofer U, Kothmeier A, Hochreiter S (2011) APcluster: an R package for affinity propagation clustering. *Bioinformatics* 27(17):2463–2464. doi:10.1093/bioinformatics/btr406
- CDO (2015) Climate data operators <http://www.mpimet.mpg.de/cdo> Accessed 12 Aug 2015
- Christensen OB, Drews M, Christensen JH, Dethloff K, Ketelsen K, Hebestadt I, Rinke A (2007) The HIRHAM regional climate model version 5 (beta). Danish Meteorological Institute: <http://www.dmi.dk/fileadmin/rapporter/tr/tr06-17.pdf> Accessed 10 April 2012
- Ebert EE (2008) Fuzzy verification of high-resolution gridded forecast: a review and proposed framework. *Meteorol Appl* 15:51–64. doi:10.1002/met.25

- Frey JB, Dueck D (2007) Clustering by passing messages between data points. *Science* 315(no. 5814):972–976. doi:10.1126/science.1136800
- Gilleland E, Ahijevych DA, Brown BG, Ebert EE (2010) Verifying forecasts spatially. *Bull Am Meteorol Soc* 91:1365–1373. doi:10.1175/2010BAMS2819.1
- Gilleland E, Ahijevych DA, Brown BG, Casati B, Ebert EE (2009) Intercomparison of spatial forecast verification methods. *Weather Forecast* 24:1416–1430. doi:10.1175/2009WAF2222268.1
- Haylock MR, Hofstra N, Klein Tank AMG, Klok EJ, Jones PD, New M (2008) A European daily high-resolution gridded data set of surface temperature and precipitation for 1950–2006. *J Geophys Res* 113: D20119. doi:10.1029/2008JD010201
- Jacob D (2001) A note to the simulation of the annual and inter-annual variability of the water budget over the Baltic Sea drainage basin. *Meteorol Atmos Phys* 77(1–4):61–73
- Jacob D, Van den Hurk BJJM, Andrae U, Elgered G, Fortelius C, Graham LP, Jackson SD, Karstens U, Chr K, Lindau R, Podzun R, Rockel B, Rubel F, Sass BH, Smith RNB, Yang X (2001) A comprehensive model intercomparison study investigating the water budget during the BALTEX-PIDCAP period. *Meteorol Atmos Phys* 77(1–4):19–43
- Kjellström E, Barring L, Gollvik S, Hansson U, Jones C, Samuelsson P, Rummukainen M, Ullerstig A, Willén U, Wyser K (2005) A 140-year simulation of European climate with the new version of the Rossby Centre regional atmospheric climate model (RCA3). SMHI Rep Meteorol Climatol No. 108, SMHI, SE-60176 Norrköping, Sweden, 54 pp
- Konca-Kedzierska K (2012) The method of defining the objects that describe the intensity and location in the precipitation field. In: Bielec-Bakowska Z, Lupikasz E, Widawski A (Eds.) Role of atmospheric circulation in climate change. The work of Department of Earth Sciences, University of Silesia 74:443–451. ISSN1895-6777 (in Polish, abstract in English)
- Marzban C, Sandgathe S (2008) cluster analysis for object-oriented verification of fields: a variation. *Mon Weather Rev* 136:1013–1025. doi:10.1175/2007MWR1994.1
- Meijgaard E, Ulft LH, Berg WJ, Bosveld FC, Hurk BJJM, Lenderink G, Siebesma AP (2008) The KNMI regional atmospheric climate model RACMO, version 2.1. KNMI Technical Report 302. KNMI Publications http://www.knmi.nl/publications/fulltexts/tr302_racmo2v1.pdf Accessed 14 Jan 2013
- Meyer D, Zeileis A, Hornik K, Friendly M (2012) Visualizing categorical data.—: <http://cran.r-project.org/web/packages/vcd/vcd.pdf> Accessed 14 Jan 2013
- Moise AF, Delange FP (2011) New climate model metrics based on object-oriented pattern matching of rainfall. *J Geophys Res* 116(D12108):1–7. doi:10.1029/2010JD015318
- Radu R, Deque M, Somot S (2008) Spectral nudging in a spectral regional climate model. *Tellus A* 60:898–910. doi:10.1111/j.1600-0870.2008.00341.x
- Uppala SM, KÅllberg PW, Simmons AJ, Andrae U, Bechtold VDC, Fiorino M, Gibson JK, Haseler J, Hernandez A, Kelly GA, Li X, Onogi K, Saarinen S, Sokka N, Allan RP, Andersson E, Arpe K, Balmaseda MA, Beljaars ACM, Berg LVD, Bidlot J, Bormann N, Caires S, Chevallier F, Dethof A, Dragosavac M, Fisher M, Fuentes M, Hagemann S, Hólm E, Hoskins BJ, Isaksen I, Janssen PAEM, Jenne R, McNally AP, Mahfouf JF, Morcrette JJ, Rayner NA, Saunders RW, Simon P, Sterl A, Trenberth KE, Untch A, Vasiljevic D, Viterbo P, Woollen J (2005) The ERA-40 re-analysis. *QJR Meteorol Soc* 131:2961–3012. doi:10.1256/qj.04.176
- Wibig J, Brzóška B, Curylo A, Jaczewski A, Konca-Kedzierska K, Liszewska M, Pianko_Kluczyńska K (2012) Dynamic scenarios of climate change for Poland for the years 2011–2030. In: Wibig J, Jakusik E (eds) Climate change influence on environment, economy and society. Changes, impact results, and ways of their limiting. Conclusions for science, practice, engineering and economical planning. Vol. 1: climate and oceanographic conditions in Poland and over the Southern Baltic Sea. Institute of Meteorology and Water Management – National Research Institute, Warsaw, pp. 93–123, ISBN 978-83-61102-65-6 (in Polish)