Article

Categorical Forecast of Precipitation Anomaly Using the Standardized Precipitation Index SPI

Leszek Łabędzki [1]*

[1] Institute of Technology and Life Sciences, Kuyavian-Pomeranian Research Centre, Glinki 60, 85-174 Bydgoszcz, Poland; e-mail: l.labedzki@itp.edu.pl

* Correspondence: l.labedzki@itp.edu.pl; Tel.: +48-52-375-01-07

Abstract: In the paper the verification of forecasts of precipitation conditions measured by the standardized precipitation index SPI is presented. For the verification of categorical forecasts a contingency table was used. Standard verification measures were used for the SPI value forecast. The 30 day SPI moved every 10 days by 10 days was calculated in 2013-2015 from April to September on the basis of precipitation data from 35 meteorological stations in Poland. Predictions of the 30 day SPI were created in which precipitation was forecasted in the next 10 days (the SPI 10-day forecast) and 20 days (the SPI 20-day forecast). Both for the 10 and 20 days, the forecasts were skewed towards drier categories at the expense of wet categories. There was a good agreement between observed and 10-day forecast categories of precipitation. Less agreement is obtained for 20-day forecasts – these forecasts evidently "over-dry" the assessment of precipitation anomalies. The 10-day SPI value forecast accuracy is acceptable, whereas for the 20-day forecast is unsatisfactory. Both for the SPI categorical and the SPI value forecast, the 10-day SPI forecast is reliable and the 20-day forecast should be accepted with reservation and used with caution.

Keywords: precipitation deficit; precipitation surplus; standardized precipitation index SPI; forecast; verification

1. Introduction

Modern economy uses natural and at the same time highly dependent on weather conditions water resources. It needs a reliable short-, medium- and long-term forecasts of surpluses and shortages of rainfall. In agriculture knowledge of current rainfall and their forecasts over the coming days enables the prediction of soil moisture changes, which allows farmers to take appropriate mitigation measures to reduce the negative effects of adverse weather events, mainly precipitation anomalies.

Natural and climatic conditions in Poland generally conducive to agricultural production, but frequent change of weather conditions during the growing season, especially rainfall, results in crop production periods of excessive soil moisture, and more often deficient rainfall. Statistics show that the average loss in yields caused by drought ranged from 10% to 40%, and in extremely dry years (e.g. 1992 and 2000) meteorological drought covered more than 40% of Polish territory [1]. In Kujavian-Pomeranian province losses caused by natural disasters in the years 1999-2011 totalled about 3.4 billion PLN [2]. Comparative research conducted by Bojar *et al.* [3] in Kujavian-Pomeranian (western Poland) and Lublin province (eastern Poland) showed significant differences in shortage of rainfall in agricultural production and yields of some crops due to regional differences in the precipitation amount and spatio-temporal distribution.

Forecasting rainfall, especially short (1-2 days ahead) and medium-term (3-10 days ahead) is very important and significant in agriculture production. Monitoring and early warning help to reduce the impacts and to mitigate the consequences of weather and climate related natural disasters for agricultural production. Transfer of agrometeorological information to farmers can be done in different ways. Meteorological services use different options, such as periodical bulletins published on the Internet and mass media: TV, radio, newspapers. According to Stigter *et al.* [4], the

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 4 November 2016 doi:10.20944/preprints201611.0029.v1

Peer-reviewed version available at *Water* **2017**, *9*, 8; doi:10.3390/w9010008

agrometeorological services should be simple for their proper assimilation and they must be used frequently to facilitate decision-making and planning. Agrometeorological services are often exemplified by agroclimatological characterization, weather forecasting (including agrometeorological forecasting) and other advisories prepared for farmers. Agrometeorological forecasting, with special attention to rainfall, is indispensible for planning agro-technical measures e.g. ploughing, sewing, harvesting, not to mention irrigation, when rainfall amount is the main determinant of when and how much to irrigate.

Forecasting rainfall is one of the most difficult meteorological forecasts and has become one of the most important elements of forecasting weather conditions at various time scales. Powerful forecasting models have been used increasingly in recent years [5–11]. The results of forecasting are available on numerous web portals, which the majority of them presents their own interpretations of graphic copyright forecasts published by specialized research institutes, such as the European Centre for Medium-Range Weather Forecasts [12] or the National Oceanic and Atmospheric Administration [10] and by thematic portals weather, for example AgroPogoda [13], WetterOnline [14]. For planning management of water in agriculture more valuable are medium- and long-term forecasts of rainfall than the prediction of daily precipitation. However, the latter is important in operational control of irrigation.

Beside rainfall forecast giving the information if rainfall occurs or about the amount of rainfall in the forecast period, the categorical precipitation forecast is often made. Such forecast gives the information in which category (class) precipitation will occur at the given probability or as a deterministic phenomena. Moreover, for the operational purpose and to make comparative assessments of precipitation anomalies in different regions, it is indispensible to apply not precipitation alone but standardized precipitation. One of such indices is the standardized precipitation index SPI [15,16]. The SPI has been defined as a key indicator for monitoring drought by the World Meteorological Organization [17]. The SPI is a standardized deviation of precipitation in a particular period from the median long-term value of this period. It represents an event departure from the mean, expressed in standard deviation units. The SPI is a normalised index in time and space. The method ensures independence from geographical positions as the index in question is calculated with respect to average precipitation in the same place [18].

An important issue in the forecasting process is to assess the reliability of forecasts. The result of verification of forecasts is the answer to the question whether the discrepancy between observed and forecast precipitation or precipitation category is essential according to accepted criteria. In world literature you can find a variety of assessment methods for the verification of predictive models, including the recommended practice by the World Meteorological Organization [19]. An interesting compendium of knowledge on forecasting is a collective work "Forecast Verification. A Practitioner's Guide in Atmospheric Science" [20]. In that book, Livezey [21] discusses the assessment of conformity of the deterministic categorical forecasts with the actual situation according to the accepted multi-stage verification criteria.

There are rather few studies devoted to assessment of forecast of drought identified by SPI. Bordi *et al*. [22] used two methods for forecasting the 1-month SPI: an autoregressive model (AR) and the Gamma Highest Probability (GAHP) method. The mean-squared error (MSE) was relatively high for both methods. Mishra and Desai [23] used linear stochastic models ARIMA and multiplicative Seasonal Autoregressive Integrated Moving Average (SARIMA) models used to forecast droughts using standardized precipitation index SPI series in the Kansabati river basin in India. Cancelliere *et al*. [24] proposed methods for forecasting transition probabilities from one drought class to another and for forecasting SPI. They showed that the SPI can be forecast with a reasonable degree of accuracy, using conditional expectation based on past values of monthly precipitation. Hwang and Carbone [25] used a conditional resampling technique to generate ensemble forecasts of SPI and found reasonable forecast performance for SPI-1. Hannaford *et al*. [26] proposed a method for forecasting drought in the United Kingdom based on current occurrence of drought. Shirmohammadi *et al*. [27] carried out the research to evaluate the ability of wavelet-artificial neural network (ANN) and adaptive neuro-fuzzy inference system (ANFIS) techniques for forecasting meteorological

Peer-reviewed version available at *Water* **2017**, *9*, 8; doi:10.3390/w9010008

drought identified by SPI in southeastern part of East Azerbaijan province, Iran. The performances of the models were evaluated by comparing the corresponding values of root mean squared error, coefficient of determination and Nash–Sutcliffe model efficiency coefficient. Belayneh *et al.* [28] compared the effectiveness of five data driven models for forecasting long-term (6 and 12 months lead time) drought conditions in the Awash River Basin of Ethiopia. The Standard Precipitation Index (SPI-12 and SPI-24) was forecasted using a traditional stochastic model (ARIMA) and compared to machine learning techniques such as artificial neural networks (ANNs), and support vector regression (SVR). The performances of all models were compared using RMSE, MAE, R2 and a measure of persistence. Maca and Pech [29] compared forecast of drought indices based on two different models of artificial neural networks. The analyzed drought indices were the standardized precipitation index (SPI) and the standardized precipitation evaporation index (SPEI) and were derived for the period of 1948–2002 on two US catchments. The comparison of the models was based on six model performance measures.

Most of the methods used to forecast SPI are based purely on statistics. There are much fewer reports in the literature of an assessment of SPI forecast based on numerical prediction models of precipitation. Łabędzki and Bąk [30] conducted a verification of the 10-day forecasts of rainfall and the course of meteorological drought in 2009 and 2010 for the station Bydgoszcz ITP. The authors checked the validity of the forecasts of precipitation taken from the service WetterOnline and the forecasts of rainfall categories based on SPI using their own verification criteria. Singleton [31] analyzed the performance of the European Centre for Medium Range Weather Forecasts (ECMWF) variable resolution Ensemble Prediction System (varEPS) for predicting the probability meteorological drought. Drought intensity was measured by the SPI and forecasts of SPI-1 and SPI-3 were verified against independent observations.

Since April 2013, it is conducted in the Institute of Technology and Life Sciences nationwide monitoring and forecasting of shortage and excess of water [32]. The current assessment of precipitation anomalies and earlier 20- and 10-day forecasts are based on actual and projected values of the standardized precipitation index SPI. They are shown on the maps of the distribution of deficit and excess rainfall in Poland in real-time and forecast periods. They are available on the website of the Institute of Technology and Life Sciences (www.itp.edu.pl) – Monitoring Agrometeo (http://agrometeo.itp.edu.pl). The aim of the study is to evaluate the verifiability of these rainfall category forecasts.

2. Materials and Methods

The evaluation and forecasting of precipitation anomalies (rainfall deficit and surplus) are made using the Standardized Precipitation Index *SPI*. The *SPI* calculation for any location is based on the long-term precipitation record in a given period. SPI was calculated using the normalization method. Precipitation P is a random variable with a lower limit and often positive asymmetry and does not conform to normal distribution. Most often, periodical (monthly, half-year or annual) sums of precipitation conform to the gamma distribution. Therefore, precipitation sequence was normalized with the transformation function f(*P*):

$$f(P) = u = \sqrt[3]{P} \quad (1)$$

where *P* is the element of precipitation sequence.

Values of the *SPI* for a given *P* are calculated with the equation:

$$SPI = \frac{f(P) - \bar{u}}{d_u} \quad (2)$$

where *SPI* is the standardised precipitation index, f(*P*) is the transformed sum of precipitation, $\bar{u}$ is the mean value of the normalised precipitation sequence, d_u is the standard deviation of the normalised precipitation sequence.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 4 November 2016 doi:10.20944/preprints201611.0029.v1

Peer-reviewed version available at Water 2017, 9, 8; doi:10.3390/w9010008

The values of *SPI* are compared with the boundaries of different classes. Because the *SPI* is normalized, wet and dry periods can be classified symmetrically. There are many classifications used by different authors. Originally, McKee *et al*. [15] distinguished four classes of drought and four classes of wet periods: mild, moderate, severe and extreme. The threshold value of *SPI* for the mild drought and mild wet category equals to *SPI* = 0. Agnew [33] writes that in this classification all negative values of *SPI* are taken to indicate the occurrence of drought – this means that for 50% of the time drought is occurring. He concluded that it was not rational and suggested alternative, more rational thresholds. He recommended the *SPI* drought thresholds corresponding to 5% (moderate drought), 10% (severe drought), and 20% (extreme drought) probabilities (*SPI* = -0.84, -1.28 and -1.65, respectively). Vermes [34] proposed seven categories, with the first class of a dry period starting at *SPI* = -1 and with the wet period at *SPI* = 1. In this study, this classification was applied (Table 1).

Table 1. Precipitation categories according to *SPI*.

Category	*SPI*
Extremely dry	≤ –2.0
Very dry	–2.0 < *SPI* ≤ –1.5
Moderately dry	–1.5 < *SPI* ≤ –1.0
Normal	–1.0 < *SPI* ≤ 1.0
Moderately wet	1.0 < *SPI* ≤ 1.5
Very wet	1.5 < *SPI* ≤ 2.0
Extremely wet	> 2.0

The *SPI* values are calculated on the basis of precipitation data from 35 meteorological stations of the Institute of Meteorology and Water Management (IMGW) - National Research Institute in Poland (Figure 1). Series of precipitation records from the period 1961-2012 at each station, were used as historical data.

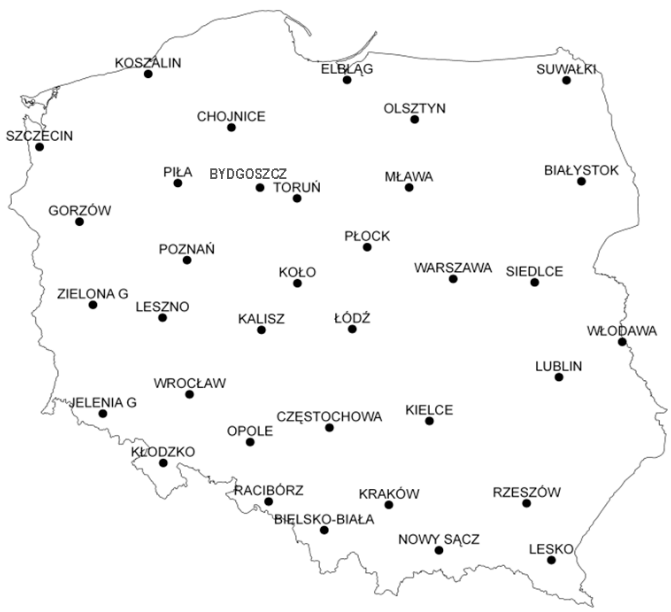

Figure 1. Location of precipitation stations.

Peer-reviewed version available at Water 2017, 9, 8; doi:10.3390/w9010008

The *SPI* was calculated in 2013-2015 from April to September and for the 30(31)-day periods moved every 10(11) days by 10(11) days. Using the forecasted precipitation, predictions of the 30(31)-day *SPI* are created in which precipitation is forecasted in the next 10(11) (the *SPI* 10-day forecast) and 20(21) days (the *SPI* 20-day forecast). Altogether there were 1330 observed-forecast pairs. The period of 10, 20 and 30 days refers to the calendar decade with 10, 20 and 30 days and the period of 11, 21 and 31 to the calendar decade with 11, 21 and 31 days.

Rainfall forecasts necessary to develop predictions of precipitation anomalies in the next 10 and 20 days, come from the meteorological service of MeteoGroup [9]. MeteoGroup has developed its own system of forecasting called Multi-Model MOS (Model Output Statistics) which is based on numerical model calculations of the most respected European meteorological centres: ECMWF model (European Centre for Medium-Range Weather Forecasts), EPS model (Ensemble Prediction System), GFS model (National Centers for Environmental Prediction), UKMO model (British Meteorological Institute) as well as on the measurement and observation data from all available sources (national synoptic meteorological stations, aerodrome meteorological stations, satellite images, radar images). The calculation results of each model are included with different weights. For each location, where historical measurements are available (with at least 1 year), for each meteorological element are assigned appropriate weights based on the degree of testability of each of the models in the past. Weighting is held every year with the new data. Major updates of MOS forecasts are held 4 times a day (7, 9, 19 and 21 UTC) based on the new model results (2 to 4 times a day depending on the model). In addition, MOS forecast is updated continuously as the inflow of the measurement data (1-3 hours). Also a special tool Meteobase is developed that, if necessary, allows meteorologists to enter manual adjustments to the forecasts at any time. MeteoGroup can provide forecast for any location specified by the user. For this purpose, the method of so-called "smart interpolation" is used, taking into account the results of the forecasts for the neighbouring measuring stations with weights dependent on their distance from the location, degree of similarity in terms of location (height above sea level, distance from the sea, location in a mountain valley, etc.). There is also the possibility of including measurement data supplied by the user, which further improves the quality of predictions for the location.

The forecasts, presented and analysed in the paper, are deterministic forecasts of a nominal variable. The variable is the standardized precipitation index *SPI* which value in a given period is qualified to the one of the *SPI* categories. The short-range forecast of *SPI* issued 10 days ahead and medium-range forecast covered the next 20 days were made.

Verification of two types of the *SPI* forecast is made: the *SPI* category forecast and the *SPI* value forecast.

For the verification of categorical forecasts and the analysis of the joint distribution for forecasts and observations a contingency table was used which is considered a good tool for this purpose [21]. A contingency table is a type of table in a matrix format that displays the multivariate frequency distribution of the variables. It provides a basic picture of the interrelation between two variables and can help find interactions between them.

A contingency table shows the distribution of one variable in rows and another in columns, used to study the association between the two variables. The tow-way contingency table is a two-dimensional table that gives the discrete joint sample distribution of deterministic forecasts and

Peer-reviewed version available at Water 2017, 9, 8; doi:10.3390/w9010008

categorical observations in cell counts [21]. The contingency table is a combination of two or more frequency tables arranged in such a way that each cell in the table represents clearly a combination of specific values of the analysed variables. Such a multi-way table enables the analysis of the frequencies corresponding to the categories designated by more than one variable. By analyzing these frequencies you can identify the relationships that exist between the variables.

Each sell of the contingency table contains the relative frequency p_{ij} of forecast category i and observed category j. It is calculated as the cell count n_{ij} divided by the total forecast-observation pair sample size n. The sums of p_{ij} for a given forecast category i and observed category j is called marginal frequencies.

To test if frequencies in each category of observed and forecasted SPI values are strongly dependent (it means the significant relationship between them) the Pearson Chi-squared test (χ^2) was used. The null hypothesis is that they are not dependent (there is no relationship between them) and the contingency table is the result of independent forecast-observation pairs for categorical events. High statistically significance of the dependence of observed and forecasted SPI category indicates high forecast accuracy. Test χ^2 consists of comparing observed frequencies with expected frequencies with the assumption of the null hypothesis (no association between observed and predicted values). Expected frequency E_{ij} is calculated using the empirical marginal distributions as:

$$E_{ij} = \sum_{j=1}^{k} p_{ij} \sum_{i=1}^{k} p_{ij} \Bigg/ \sum_{i=1}^{k} \sum_{j=1}^{k} p_{ij} \qquad i,j = 1, \dots, k \tag{3}$$

where:

p_{ij} – relative frequency of forecast category i and observed category j

k – number of observed and forecast categories

The test statistic, called the Pearson Chi-squared statistic, takes the form:

$$\chi^2 = \sum_{i=1}^{k} \sum_{j=1}^{k} \frac{\left(p_{ij} - E_{ij}\right)^2}{E_{ij}} \tag{4}$$

Assuming the veracity of the null hypothesis this statistic has the asymptotic χ^2 distribution with the degrees of freedom df equal to:

$$df = (k-1)^2 \tag{5}$$

The results of observed-forecast frequencies depend on the relation of the number of categories and the sample size. For more than two categories forecast, a sample size required for proper estimates should be of the order of $10k^2$ [21]. In the presented study $k = 7$ and the sample size of 1330 forecast-observation pairs is completely sufficient.

If the values of the computed statistic according to Eq. (4) exceed the critical χ^2_{cr} for their chance probabilities to be less than e.g. 0.05, 0.01, 0.001 ($\chi^2 > \chi^2_{cr}$) the null hypothesis can be rejected at a given probability level. The asymptotic distribution of χ^2 for different degrees of freedom is tabulated in different sources from which χ^2_{cr} can be determined for a given probability and the sample size n.

For categorical forecasts presented in the form of a contingency table the following measures of accuracy were used based on the frequencies and the marginal distributions:

1) Proportion correct PC

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 4 November 2016 doi:10.20944/preprints201611.0029.v1

Peer-reviewed version available at *Water* **2017**, *9*, 8; doi:10.3390/w9010008

$$PC = \sum_{i=1}^{k} p_{ii} \tag{6}$$

2) Bias *B*

$$B_i = \sum_{j=1}^{k} p_{ij} \Big/ \sum_{j=1}^{k} p_{ji} \qquad i = 1, \dots, k \tag{7}$$

3) Probability of detection *POD*

$$POD_i = \sum_{j=1}^{k} p_{ij} \Big/ \sum_{j=1}^{k} p_{ji} \qquad i = 1, \dots, k \tag{8}$$

Besides the verification of the *SPI* category forecasts on the basis of the contingency table, the verifiability of the SPI value forecasts was assessed. The following measures of goodness of fit were used to evaluate the forecast performance:

1) Ratio of the number of the periods in which the criterion

$$\left|SPI_{forecast} - SPI_{observed}\right| \leq 0.5 \tag{9}$$

was met to the number of all periods.

2) Mean systematic error (bias) *b*

$$b = \frac{1}{n}\sum_{i=1}^{n}\left(SPI_{forecast} - SPI_{observed}\right) \tag{10}$$

where n is the number of forecast-observation pairs.

3) Mean absolute error *MAE*

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|SPI_{forecast} - SPI_{observed}\right| \tag{11}$$

4) Mean squared error *MSE*

$$MSE = \frac{1}{n}\sum_{i=1}^{n}\left(SPI_{forecast} - SPI_{observed}\right)^2 \tag{12}$$

5) Root mean squared error *RMSE*

$$RMSE = \sqrt{MSE} \tag{13}$$

6) Pearson's linear correlation coefficient *r*

$$r = \frac{\sum_{i=1}^{n}\left(SPI_{forecast} - \overline{SPI}_{forecast}\right)\left(SPI_{observed} - \overline{SPI}_{observed}\right)}{\sqrt{\sum_{i=1}^{n}\left(SPI_{forecast} - \overline{SPI}_{forecast}\right)^2}\sqrt{\sum_{i=1}^{n}\left(SPI_{observed} - \overline{SPI}_{observed}\right)^2}} \tag{14}$$

In the above equations $SPI_{forecast}$ denotes the forecast *SPI* value in the 30(31)-day period in which the 20(21)-day rainfall sum was measured and the 10(11)-day rainfall sum was forecast in the case of the 10-day forecast and the 10(11)-day rainfall sum was measured and the 20(21)-day rainfall was forecast in the case of the 20-day forecast. $SPI_{observed}$ denotes the observed SPI value in the same 30(31)-day period on the base of the measured rainfall sum in this period.

3. Results

3.1. SPI category forecast

The joint distribution of forecast and observed SPI is presented in the contingency tables for the 10-day forecasts (Table 2) and for the 20-day forecasts (Table 3). The contingency tables show the relative frequencies and the empirical margins distributions in seven categories of precipitation. The forecasts were made for 35 stations and for the years 2013-2015 for April through September. Each table is constructed from a sample of 1330 forecasts-observations.

Base on the distribution of the observed *SPI* it can be concluded that in 2013-2015 the periods drier than normal dominated (23%) in comparison with the wetter periods (11%). Normal periods occurred most often (66%). The similar frequency distribution was found for the forecasts, both for 10 and 20 days ahead. These forecasts are skewed towards forecasts of drier categories at the expense of wet categories – 27% periods were predicted as drier than normal in the case of 10-day forecasts and 30% in the case of 20-day forecasts. Comparing the distribution of observations and forecasts it seem reasonable to conclude that there is a good agreement between observed and 10-day forecast categories of precipitation. Less agreement is obtained for 20-day forecasts – these forecasts evidently "over-dry" the assessment of precipitation anomalies. The observed normal category of precipitation is almost as often as the 10-day forecast of this category (66% and 63%, respectively). The 20-day forecast of normal category is less frequent (55%) than the observed normal category. The frequency of 20-day forecast of dry periods distinctly increased while of normal and wet periods decreased.

Peer-reviewed version available at Water 2017, 9, 8; doi:10.3390/w9010008

Table 2. Relative frequency (in percent) for *SPI* 10-day forecasts (n = 1330).

Forecast	Observed							
	Extremely dry	Very dry	Moderately dry	Normal	Moderately wet	Very wet	Extremely wet	Forecast distribution
Extremely dry	3	1	1	1	0	0	0	6
Very dry	1	3	3	1	0	0	0	8
Moderately dry	1	2	5	5	0	0	0	13
Normal	0	0	3	56	3	1	0	63
Moderately wet	0	0	0	2	3	1	0	6
Very wet	0	0	0	1	1	1	0	3
Extremely wet	0	0	0	0	0	0	1	1
Observed distribution	5	6	12	66	7	3	1	100

n – the number of observation-forecast pairs

Table 3. Relative frequency (in percent) for *SPI* 20-day forecasts (n = 1330).

Forecast	Observed							
	Extremely dry	Very dry	Moderately dry	Normal	Moderately wet	Very wet	Extremely wet	Forecast distribution
Extremely dry	3	2	3	4	0	0	0	12
Very dry	1	2	2	7	0	0	0	12
Moderately dry	1	1	3	11	0	0	0	16
Normal	0	1	4	42	6	1	1	55
Moderately wet	0	0	0	2	1	1	0	4
Very wet	0	0	0	0	0	1	0	1
Extremely wet	0	0	0	0	0	0	0	0
Observed distribution	5	6	12	66	7	3	1	100

To answer the question whether the constructed contingency tables are the result of dependent forecast-observations pairs for categorical events, a Chi-squared test (χ^2) was performed with the assumption of the null hypothesis that no association between observed and predicted values occurred. Both for 10-day and 20-day forecast the test statistics χ^2 are greater than the critical values of χ^2_{cr}. It means that the null hypothesis should be rejected. The relation between the frequency distribution in *SPI* categories is statistically significant at least at the 0.001 level for 10-day forecast and at the 0.05 level for 20-day forecast (Table 4). These results show that categorical forecasts of *SPI* are highly accurate and reliable for 10 days ahead and reasonably accurate for 20 days ahead.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 4 November 2016 doi:10.20944/preprints201611.0029.v1

Peer-reviewed version available at *Water* **2017**, *9*, 8; doi:10.3390/w9010008

Table 4. χ^2 values for *SPI* forecasts in seven categories (n = 1330; df = 36).

Test statistic	10-day forecast	20-day forecast
χ^2 calculated	155.7	51.5
χ^2_{cr} for $\alpha = 0.05$	51.0	
χ^2_{cr} for $\alpha = 0.01$	58.6	
χ^2_{cr} for $\alpha = 0.001$	68.0	

df – degree of freedom

For categorical forecasts the measures of accuracy based on the frequencies and the marginal distributions are shown in Table 5. The proportion correct *PC* shows the proportion of correct categorical forecasts, the bias *B* reveals whether some forecast categories are over- or under-forecast while the probability of detection *POD* quantifies the success rate for detecting different categorical events. *PC* is rather high for 10-day forecasts (72%) and less for 20-day forecasts (52%). In case of the 10-day forecasts, the forecast-observation set has little bias for the normal as well as for the moderately and very dry and wet categories (value close to 1). The forecasts and observations are rather dissimilar for the extreme category. The values of bias *B* are worse for the 20-day forecasts. Both for the 10-day and 20-day forecasts, the dry categories are above-forecast ($B > 1$) and the wet categories are under-forecast ($B < 1$). The probability of detection is only satisfactory for the 10-day normal category forecast (*POB* = 0.83); other forecasts are modestly under-detected.

Table 5. Measures of accuracy for *SPI* forecasts in seven categories.

Measure	Extremely dry	Very dry	Moderately dry	Normal	Moderately wet	Very wet	Extremely wet
				10-day forecast			
PC				0.72			
B	1.63	1.15	1.08	0.94	0.85	0.92	1.38
POD	0.64	0.52	0.46	0.83	0.42	0.42	0.69
				20-day forecast			
PC				0.51			
B	3.13	1.89	1.49	0.80	0.50	0.45	0.56
POD	0.67	0.28	0.27	0.62	0.13	0.18	0.19

3.2. SPI value forecast

In this chapter the verification of SPI value forecast is done (Table 6).

Table 6. Measures of accuracy for *SPI* value forecasts.

Measure	10-day forecast	20-day forecast
Ratio	72%	40%
Bias *b*	-0.10	-0.53
MAE	0.39	0.80
MSE	0.30	1.08
RMSE	0.543	1.037
Correlation coefficient *r*	0.870	0.648

The first measure of the accuracy – the ratio of the number of the periods in which the absolute value of the difference between the forecast and observed *SPI* was not greater than 0.5 to the number of all periods, averaged for all stations, was 72% for the 10-day forecast and 40% for the 20-day forecast. At different stations the ratio changes from 54 to 85% for the 10-day forecast and from 18 to 58% for the 20-day forecast.

Peer-reviewed version available at *Water* **2017**, *9*, 8; doi:10.3390/w9010008

The mean systematic error (bias) is negative (-0.10 for 10-day forecast and -0.53 for 20-day forecast). It means that the forecasts are too dry on average. This verification measure in not fully adequate because negative errors can be compensated by positive errors. The mean absolute error *MAE* avoids this disadvantage since it takes into the account absolute values of the individual forecast error. The *MAE* is used to measure how close forecasted values are to the observed values. It is the average of the absolute errors. In our study it shows that the positive and negative errors of the SPI forecast are twice greater for 20-day forecast than for 10-day forecast. However, the *MAE* of 10-day forecast (0.4) is relative small (10%) compared to the range of the most often observed *SPI* values (from -2 to 2).

The mean squared error (*MSE*) of a forecast measures the average of the squares of the errors, that is, the difference between the forecast and observed *SPI*. The *MSE* is the second moment of the error, and thus incorporates both the variance of the forecast and its bias. Taking the square root of *MSE* yields the root-mean-squared error (*RMSE*), being the square root of the variance. The values *MSE* = 0.30 and *RMSE* = 0.54 for the 10-day forecast seem to be acceptable taking into account the possible range of *SPI*; for 20-day forecast they are unsatisfactory (*RMSE* > 1).

The last measure most often used for evaluation of the forecasts is simply the correlation coefficient between forecast and observed values. This coefficient measures the degree of association among the forecast and observed values. It is acceptable for 10-day forecast (0.87) and rather unsatisfactory for 20-day forecast (0.65).

4. Conclusions

This study investigated the accuracy of forecasts of precipitation conditions measured by the standardized precipitation index *SPI*. Verification of two types of the *SPI* forecast was performed: the *SPI* category forecast and the *SPI* value forecast. For the verification of categorical forecasts a contingency table was used. Standard verification measures were used for the *SPI* value forecast. The *SPI* was calculated for the 30(31)-day periods moved every 10(11) days by 10(11) days. Using the forecasted precipitation, predictions of the 30(31)-day *SPI* were created in which precipitation was forecasted in the next 10(11) and 20(21) days.

In 2013-2015 both for the 10 and 20 days, the forecasts were skewed towards forecasts of drier categories at the expense of wet categories. Comparing the distribution of observations and forecasts there was a good agreement between observed and 10-day forecast categories of precipitation. Less agreement is obtained for 20-day forecasts – these forecasts evidently "over-dry" the assessment of precipitation anomalies. The observed normal category of precipitation was almost as often as the 10-day forecast of this category. The 20-day forecast of normal category was less frequent than the observed normal category. The frequency of 20-day forecast of dry periods distinctly increased while of normal and wet periods decreased

Considering the *SPI* values, the ratio of the number of the periods in which the absolute value of the difference between the forecast and observed *SPI* was not greater than 0.5 to the number of all periods, averaged for all stations, was 72% for the 10-day forecast and 40% for the 20-day forecast. Considering the measures of the *SPI* value forecast accuracy, the reliability of the 20-day forecast was shown to be weaker than of the 10-day forecast. The mean absolute error *MAE* of the *SPI* forecast was twice greater for 20-day forecast than for 10-day forecast. The *MAE* of 10-day forecast was relative small compared to the range of the most often observed *SPI* values. Other measures (the mean squared error *MSE*, the square root of *MSE*, the correlation coefficient) shows that the 10-day forecast accuracy is acceptable taking into account the possible range of *SPI*, whereas for the 20-day forecast is unsatisfactory.

The performed analysis shows that, both for the *SPI* categorical and the *SPI* value forecast, the 10-day *SPI* forecast is reliable and the 20-day forecast should be accepted with reservation and used with caution.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 4 November 2016 doi:10.20944/preprints201611.0029.v1

Peer-reviewed version available at *Water* **2017**, *9*, 8; doi:10.3390/w9010008

Acknowledgments: The results presented in the paper were obtained within the Programme "Standardization and monitoring of environmental projects, agricultural technology and infrastructure solutions for security and sustainable development of agriculture and rural areas", the activity 1.2 "Monitoring, predicting of progress and risk of water deficit and surplus in the rural areas", conducted by the Institute of Technology and Life Sciences in 2011-2015 and financed by the Polish Ministry of Agriculture and Rural Development. The author is greatly appreciated for financing this project. The author would like to thank Mrs. Ewa Kanecka-Geszke, Mr. Bogdan Bak and Mr. Tymoteusz Bolewski from the Institute of Technology and Life Sciences, Poland, for their contribution in creating the database and performing the calculations of SPI.

Conflicts of Interest: The author declares no conflict of interest.

References

1. Łabędzki, L. Estimation of local drought frequency in central Poland using the standardized precipitation index SPI. *Irrig. Drain.* **2007,** *56(1)*, 67-77. DOI: 10.1002/ird.285.
2. Bąk, B.; Łabędzki, L. Prediction of precipitation deficit and excess in Bydgoszcz Region in view of predicted climate change. *J. Water Land Dev.* **2014,** *23*, 11-19. DOI: 10.1515/jwld-2014-0025.
3. Bojar, W.; Knopik, L.; Żarski, J.; Sławiński, C.; Baranowski, P.; Żarski, W. Impact of extreme climate changes on the predicted crops. *Acta Agrophys.*, **2014**, *21(4)*, 415-431.
4. Stigter, K.; Walker, S.; Das, HP.; Huda, S; Dawei, Z; Jing, L.; Chunqiang, L.; Hurtado, IHD.; Mohammed, A.E.; Abdalla, A.T.; Bakheit, NI.; Al-Amin, N.K.N.; Yurong, W.; Kinama, J.M.; Nanja, D.; Haasbroek, P.D.; Sudan, K. 2010. Meeting farmers' needs for agrometeorological services: an overview and case studies. (Second draft of June 2010). Available online: http://www.researchgate.net/publication/228402080_Meeting_farmers'_needs_for_agrometeorological_services_An_overview_and_case_studies (accessed on 12 March 2014).
5. Acharya, N.; Kulkarni, M.A.; Mohanty, U.C.; Singh, A. Comparative evaluation of performances of two versions of NCEP climate forecast system in predicting Indian summer monsoon rainfall. *Acta Geophys*. **2014,** *62(1)*, 199-219. DOI: 10.2478/s11600-013-0145-x.
6. Chattopadhyay, S. Feed forward artificial neural network model to predict the average summer-monsoon rainfall in India. *Acta Geophys*. **2007**, *55(3)*, 369-382. DOI 10.2478/s11600-007-0020-8.
7. Feng, G.; Cobb, S.; Abdo, Z.; Fisher, D.K.; Ouyang, Y.; Adeli, A.; Jenkinsa, J.N. Trend analysis and forecast of precipitation, reference evapotranspiration, and rainfall deficit in the Blackland Prairie of Eastern Mississippi. *J. Appl. Meteor. Climatol*. **2016**, *55*, 1425–1439. DOI: 10.1175/JAMC-D-15-0265.1.
8. Lavers, D.; Luo, L.; Wood, E.F. A multiple model assessment of seasonal climate forecast skill for applications. *Geophys. Res. Lett.* **2009**, *36*, L23711. DOI:10.1029/2009GL041365.
9. MeteoGroup. Multi-model approach. Available online : http://www.meteogroup.com/pl/gb/research/multi-model-approach.html (accessed 20 April 2013).
10. NOOA's National Weather Service. Current MOS forecast products. Available online: http://www.nws.noaa.gov/mdl/synop/products.php (accessed 15 March 2013).
11. Saha, S.; Moorthi, S.; Pan, H.L.; Wu, X.; Wang, J.; Nadiga, S.; Tripp, P.; Kistler, R.; Woollen, J.; Behringer, D.; Liu, H.; Stokes, D.; Grumbine, R.; George, G.; Hou, Y.T.; Chuang, H.; Juang, H.; Sela, J.; Iredell, M.; Treadon, R.; Kleist, D.; van Delst, P.; Keyser, D.; Derber, J.; Ek, M.; Meng, J.; Wei, H.; Rongqian, Y.R.; Lord, S.; van den Dool, H.; Kumar, A.; Wang, W.; Long, C.; Chelliah, M.; Xue, Y.; Huang, B.; Schemm, J.K.; Ebisuzaki, W.; Lin, R.; Xie, P.; Chen, M.; Zhou, S.; Higgins, W.; Zou, C.Z.; Liu, Q.; Chen, Y.; Han, Y.; Cucurull, L.; Reynolds R.W.; Rutledge, G.; Goldberg, M. The NCEP climate forecast system reanalysis. *Bull. Amer. Meteor. Soc.* **2010**, *91*, 1015-1057. DOI: 10.1175/2010BAMS3001.1.
12. European Centre for Medium-Range Weather Forecasts. Available online: http://www.ecmwf.int/ (accessed 01 April 2013).
13. AgroPogoda. MeteoGroup service for agriculture. Available online: http://www.agropogoda.meteogroup.pl (accessed 01 March 2013).
14. WetterOnline. Available online: http://www.wetteronline.de/ (accessed 01 March 2013).
15. McKee, T.B.; Doesken, N.J.; Kleist. J. The relationship of drought frequency and duration to time scales. Proc. 8th Conf. Applied Climatology, 17-22 January 1993. Anaheim, California, 179-184.
16. McKee, T.B.; Doesken, N.J.; Kleist, J. Drought monitoring with multiple time scales. Preprints 9th Conf. Applied Climatology, 15-20 January 1995. Dallas, Texas, 233-236.
17. WMO (World Meteorological Organization). Standardized Precipitation Index. User Guide. **2012;** p. 24

18. Cacciamani, C.; Morgillo, A.; Marchesi, S.; Pavan, V. Monitoring and forecasting drought on a regional scale: Emilia-Romagna Region. In: Methods and tools for drought analysis and management. *Water Science and Technology Library*, 2007, *62*, 29-48.
19. WMO (World Meteorological Organization) Recommendations for the verification and intercomparison of QPFs and PQPFs from operational NWP models. Atmospheric Research and Environment Branch. 2008, 7-23.
20. Forecast verification. A practitioner's guide in atmospheric science. Jolliffe, I.T.; Stephenson, D.B., Eds.; Willey-Blackwell: West Sussex, UK, 2012.
21. Livezey, R.E. Deterministic forecasts of multi-category events. In: *Forecast verification. A practitioner's guide in atmospheric science*. Jolliffe, I.T.; Stephenson, D.B., Eds.; Willey-Blackwell, 2012, pp. 61-75.
22. Bordi, I.; Fraedrich, K.; Petitta, M.; Sutera, A. Methods for predicting drought occurrences. Proc. 6th Inter. Conf. of the European Water Resources Association. Menton, France, 7-10 September 2005, 7–10.
23. Mishra, A.K.; Desai, V.R.) Drought forecasting using stochastic models. *Stoch. Environ. Res. Risk. Assess.* **2005**, *19*, 326–339. DOI 10.1007/s00477-005-0238-4.
24. Cancelliere, A.; Di Mauro, G.; Bonaccorso, B.; Rossi, G. Drought forecasting using the Standardized Precipitation Index. *Water. Resour. Manage*. **2007,** *21*, 801–819.
25. Hwang, Y.; Carbone, G.J. Ensemble forecasts of drought indices using a conditional residual resampling technique. *J. Appl. Meteor. Climatol.* **2009**, *48*, 1289–1301.
26. Hannaford, J.; Lloyd-Hughes, B.; Keef, C.; Parry, S.; Prudhomme, C. Examining the large-scale spatial coherence of European drought using regional indicators of rainfall and streamflow deficit. *Hydrol. Process.* **2011,** *25*, 1146-1162. DOI: 10.1002/hyp.7725.
27. Shirmohammadi, B.; Moradi, H.; Moosavi, V.; Semiromi, MT.; Zeinali, A. Forecasting of meteorological drought using Wavelet-ANFIS hybrid model for different time steps (case study: southeastern part of east Azerbaijan province, Iran). *Nat. Haz.* **2013**, *69*, 389–402. DOI: 10.1007/s11069-013-0716-9.
28. Belayneh, A.; Adamowski, J.; Khalil, B.; Ozga-Zielinski, B. Long-term SPI drought forecasting in the Awash River Basin in Ethiopia using wavelet neural network and wavelet support vector regression models. *J. Hydrol.* **2014**, *508*, 418–429.
29. Maca, P.; Pech, P. Forecasting SPEI and SPI drought indices using the integrated artificial neural networks. *Comput. Intellig. Neurosci.* **2016,** Vol. 2016, Article ID 3868519, p. 17. DOI: 10.1155/2016/3868519.
30. Łabędzki, L.; Bąk, B. Predicting meteorological and agricultural drought in the system of drought monitoring in Kujawy and the Upper Noteć Valley. *Infrastructure and Ecology of Rural Areas*. **2011,** *5*, 19-28.
31. Singleton, A. Forecasting drought in Europe with the Standardized Precipitation Index. Luxembourg: Publications Office of the European Union. **2012**, p. 68. DOI:10.2788/16522.
32. Łabędzki, L.; Bąk, B. Indicator-based monitoring and forecasting water deficit and surplus in agriculture in Poland. *Ann. Warsaw Univ. of Life Sci. – SGGW, Land Reclam.* **2015**, *47(4)*, 355-369.
33. Agnew, C.T. Using the SPI to Identify Drought. Drought Network News. **2000,** *12(1)*, 6-11.
34. Vermes, L. How to work out a drought mitigation strategy. An ICID Guide. DVWK Guidelines for water management. **1998,** 309, p. 29.