ORIGINAL RESEARCH ARTICLE

The influence of relative humidity on fires in forests of Central Poland

Krystyna Konca-Kędzierska* [ID], **Krystyna Pianko-Kluczyńska** [ID]

Institute of Meteorology and Water Management, PSHM CSPM Specialized Forecasts Section, ul. Podleśna 61, 01-673 Warsaw, Poland

* Tel. +48 22 569 4256 , e-mail: krystyna.konca@imgw.pl

Abstract. The purpose of the present work was to answer the question of how and to what extent information about low air humidity in a selected area can be used in determining the fire risk. The central region as defined in the Classification of Territorial Units for Statistics (NUTS) was selected for the study, because this classification is also used in the National Fire Information System, which provides data on the number of forest fires. Data on humidity for a given region were taken from the synoptic or climatic station of The Institute of Meteorology and Water Management – National Research Institute (IMGW-PIB) located in the region.

Two independent experiments were conducted based on various methodological assumptions. In experiment A, the main focus was on humidity deficit, as defined by nine different indicators. Using the relative risk and the Jaccard index, their connection to fire cases of third class or higher was studied. A class reconstruction of the number of fires was also carried out using the fuzzy analogues method, and the obtained reconstructions were evaluated using cross-correlation indicators in contingency tables. In experiment B, correlation relationships between the number of fires and the annual characteristics of relative humidity were studied. In order to use the association measures in the cross-tabulation tables, the number of fires and humidity characteristics were categorised using quantiles. The relationship between the number of fires and the percentage share of low-humidity (<40%) days in the year was tested using the Mann–Whitney test.

The obtained values of the examined correlation indices and the analysis of correlation relationships emphasise the important role of relative humidity dynamics in determining the fire risk.

Keywords: forest fires, humidity deficiency, correlations, association measures, fuzzy analogues

1. Introduction

One of the more dangerous and costly events occurring in Polish forests are fires. According to statistical data available in the National Forest Fire Information System (KSIPL 2017), around 85% of forested areas in Poland are susceptible to fires. Amongst the countries of the European Union, Poland belongs to those countries that are moderately threatened by forest fires (Szczygieł 2012). However, this problem should not be underestimated. As an example, in 2009, Poland was third in terms of the number of fires and in eighth place in terms of burned area (Szczygieł 2012 in Forest Fires in Europe, Report No. 10, 2009).

The factors determining the formation and spread of fires include, amongst others, meteorological conditions. In 2010, a new forecasting method was developed at the Forest Research Institute, which – based on selected meteorological parameters (temperature and relative air humidity, as well as litter, rainfall, wind speed and cloudiness) – allows the current and predicted (up to 24 hours) threat of forest fires to be determined (Prędecka 2011, DGLP 2012).

This article focuses on relative air humidity, which is one of the most important factors affecting the possible occurrence of fires. An analysis was conducted to determine the possibility of using this parameter for monthly, three-month and seasonal weather forecasts, developed as a tool to support forest fire protection as well as in studies on fire hazard scenarios resulting from climate change simulations. The long-term forecasting issue under consideration complements the work performer by Forest Reasarch Institute on risk forecasts for the next 24-hour period (Kwiatkowski 2010).

This work was a pilot study and served to develop a research method needed to conduct a planned analysis of the impact of relative humidity on the number of fires occurring in forests. The obtained results may signal the possibility of conducting further research in this direction.

The research, which was performed using two different approaches to determine the relationship between relative humidity and the number of recorded forest fires, differed significantly in the way the research problem was formulated and the methodology used. It allowed us to assess how the definition of the descriptive (or input) variables and described (or output) variables and the selection of the algorithm testing the dependencies between the studied variables influence the information contained in the amount of relative air humidity on the predictand, that is, the number of forest fires in a given region of the country.

2. Data

Information on the number of fires was obtained from the National Forest Fire Information System's annual summaries from 2007 to 2016 (KSIPL 2017). We analysed data on the number of fires in 11 subregions of the Central Region defined by the National Forest Fire Information System based on the Statistics Poland Classification of Territorial Units for Statistics – NUTS (2017 version) (Fig. 1).

Information on fires is given in two ways: as the number of recorded fires and in categories of the following number of fires: class 1, 1–31; class 2, 32–62; class 3, 63–124; class 4, 125–187; class 5, 188–250; class 6, more than 250.

Table 1 shows the number of occurrences of particular fire classes in the subregions for the entire decade. In the case of the subregions, it is not uncommon that certain classes did not occur in the examined period. A specific situation occurred in the eastern Warsaw subregion, where only class 6 was recorded (more than 250 fires during the year). In comparison, fire class 1 (less than 32 fires) dominated in subregion 9 (city of Łódź).

Table 2 presents the number of occurrences of each fire class 1–6 for each year 2007–2016 in all the subregions together. It shows that 2015 had the most cases when the annual number of fires in the subregions exceeded 250 (class 6 occurred 7 times), followed by 2009 and 2012 (class 6 occurred five times).

Each subregion was represented by a weather station (synoptic or climatic). Because access to archival data was limited, the relative humidity recorded at the Łódź synoptic weather station simultaneously represented both the Łódzki subregion and the city of Łódź subregion.

The relative humidity of the air was analysed for available time measurements accessed from the historical database of Institute of Meteorology and Water Management – National Research Institute (IMGW-PIB) (8 measurements were performed in the case of synoptic weather stations and 3 or 4 measurements in climatic stations, whereas 24 measurements were performed per day at automatic stations).

3. Research methods

Studies devoted to the climatic determinants of forest fires indicate the significant role of the average or time (e.g. 9:00 and 13:00) of relative air humidity (Kwiatkowski, Szczygieł 2015). However, the presented study also includes the minimum and maximum relative humidity.

The data cited above have not been homogenised so as not to remove the extreme cases that are very important in forecasting processes. Only the complete data were analysed.

The study used proprietary programmes and R statistical package procedures (made available based on the Open Source principle). It consisted of two research experiments: A, using

Table 1. Number of events of particular fire classes in the subregions in the whole period of 2007–2016

Subregion	Class 1	Class 2	Class 3	Class 4	Class 5	Class 6
1. ciechanowsko-płocki	0	1	4	4	0	1
2. ostrołęcko-siedlecki	0	0	0	2	1	7
3. skierniewicki	3	2	5	0	0	0
4. warszawski zachodni	0	0	0	3	3	4
5. m. Warszawa	1	3	5	1	0	0
6. warszawski wschodni	0	0	0	0	0	10
7. sieradzki	0	2	3	1	3	1
8. łódzki	1	1	4	4	0	0
9. m. Łódź	9	1	0	0	0	0
10. piotrkowski	0	0	3	2	1	4
11. radomski	0	0	0	1	0	9

Table 2. Number of events of individual fire classes for all the subregions of the central region in 2007–2016

Year	Class 1	Class 2	Class 3	Class 4	Class 5	Class 6
2007	1	1	4	1	1	3
2008	1	1	1	2	2	4
2009	1	0	2	2	1	5
2010	2	3	1	2	1	2
2011	1	0	3	2	1	4
2012	0	2	2	1	1	5
2013	4	1	2	2	0	2
2014	2	1	4	2	0	2
2015	1	0	1	2	0	7
2016	1	1	4	2	1	2

Figure 1. Subregions of the central region (according to NUTS) and the weather stations representing them.

mainly the authors' programmes; and B, using the procedures of the R statistical package. Both experiments also differed in the way the input information about the relative air humidity was provided. Experiment A used humidity deficiency characteristics, while in experiment B, the relative humidity was mainly characterised by average and minimum values. Both concepts also differed in the treatment of information on the number of forest fires. Experiment A used the National Information System on Forest Fires KSIPL classification of the number of fires; experiment B used the direct number of occurrences of fires or the class of this number determined by means of quantiles. Both experiments also differed in the methods applied to test the strength of the relationship between the information on the number of fires and humidity. Experiment A examined the relative risk, the Jaccard index and indicators describing contingency tables and attempted to reconstruct the classes of the number of fires using fuzzy analogues. The work conducted in experiment B was based mainly on the correlation analysis and testing the significance of the relationship.

Owing to the specific way measurements are taken by IMGW-PIB, experiment A defined a day with low humidity as a day when the measurements taken at 06, 12 and 18 Greenwich Mean Time (GMT) recorded a drop in the relative air humidity below 40%. For the purpose of further work, a characteristic of humidity was determined, specifying the percentage share in the analysed period (month, season IV–VIII, year) of dry days.

Experiment A began with an analysis of the completeness of the measurement data.

For the purposes of the analysis, the data was reclassified to form two classes: new_class 0, less than 125 fires (the cases from classes 1 to 3), and new_class 1, more than 124 fires (the cases from classes 4 to 6 of the basic classification).

In experiment A, an attempt was made to link information about the classes of the number of fires recorded in the subregions in particular years with the humidity characteristics of the year, season IV-VIII and individual months.

The methodology described in subsections AA, AB and AC concerned nine variants for which analyses of the relationship between the lack of humidity at the subregion stations and the number of fires in individual central subregions for the following defined humidity conditions were conducted:

1. There was at least 1 month during the year with more than 10% of dry days (RH_MC_10).

2. There was at least 1 month during the year with more than 20% of dry days (RH_MC_20).

3. There was at least 1 month during the year with more than 40% of dry days (RH_MC_40).

4. The share of dry days exceeded 10% in the months of IV–VIII (RH_IV_VIII_10).

5. The share of dry days exceeded 15% in the months of IV–VIII (RH_IV_VIII_15).

6. The share of dry days exceeded 20% in the months of IV–VIII (RH_IV_VIII_20).

7. The share of dry days exceeded 10% in the year (RH_ROK_10).

8. The share of dry days exceeded 15% in the year (RH_ROK_15).

9. The share of dry days exceeded 20% in the year (RH_ROK_20).

The following measures were used to estimate the strength of the relationship:

AA. Relative risk, meaning that if the meteorological conditions of interest to us occurred, the likelihood of exceeding fire class 3 increased in relation to the situation when the assumed meteorological situation did not occur.

$$RR = \frac{Ra}{Rb}, \quad Ra = \frac{m11}{m}, \quad Rb = \frac{b}{n} \tag{1}$$

where
$m11$ is the number of cases when drought defined by the relative humidity occurred and the number of fires was within the range of new_class 1,
m is the number of cases when drought occurred as defined by the relative humidity,
b is the number of cases when the relative humidity did not indicate drought and the number of fires were in new_class 1,
n is the number of cases when the relative humidity did not indicate drought.

AB. The Jaccard index (Real, Vargas 1996) is a measure of the similarity of the compared sets (A describes the meteorological situation and B reflects the fire situation). This index is the quotient of the cardinality of the common parts of sets A and B divided by the power of the sum of these sets.

$$J(A, B) = \frac{|A \cap B|}{|A \cup B|} \qquad (2)$$

$$J(A, B) = \frac{m11}{m11 + m10 + m01} \qquad (3)$$

where
$m11$ is the number of cases when drought defined by the relative humidity occurred and the number of fires was classified to new_class 1,
$m10$ is the number of cases when drought defined by the relative humidity occurred, and the number of fires was classified to new_class 0,
$m01$ is the number of cases when the relative humidity did not indicate drought, and the number of fires was classified to new_class 1.

For the purpose of this study, the strength of the relationship is expressed as a percentage:

$$\text{Jacc}(A, B) = 100\% * J(A, B) \qquad (4)$$

AC. To assess the strength of the relationship between the occurrence of one of the nine previously defined humidity deficiency variants and the occurrence of a large number of fires, TAB quality indexes in contingency tables were used. It was assumed that the occurrence of a deficiency in relative air humidity is synonymous with forecasting more than 124 fires. The following designations were introduced:
TAB [1,1] – the number of times more than 124 fires were forecasted and the observed number of more than 124 fires,
TAB [1,2] – the number of times more than 124 fires were forecasted and the observed number of less than 125 fires,
TAB [2,1] – the number of times less than 125 fires were forecasted and the observed number of more than 124 fires,
TAB [2,2] – the number of times less than 125 fires were forecasted and the observed number of less than 125 fires.

The hit rate (HR), false alarm rate (FAR) and Hanssen and Kuipers score (KS I KS 2) indices were used, recommended by the World Meteorological Organization (2002).

HR is a measure of success, that is, an accurate prediction that the fire class will be greater than 3. HR values range from 0 to 1. As the value of this index increases, so does the accuracy of the model.

FAR is a measure of how often a warning has been issued about a large number of fires, and the observed class was not greater than 3. FAR values range from 0 to 1, and a lower value is obtained in the absence of false alarms. KS = HR-FAR and has values between −1 and 1. The value of 1 indicates a correct forecast.

The KS2 index is scaled according to the formula:

$$KS2 = \frac{KS + 1}{2} \qquad (5)$$

and ranges from 0 to 1; it is difficult to determine the accuracy of a forecast when the value is 0.5 because FAR = HR, whereas the forecast is perfect when the value is 1.

AD. An attempt was made to reconstruct the fire class based on the relative humidity characteristics using fuzzy analogues (Matteucci 2017). The Euclidean distance between the X_i and C_k vectors was used, written as dist (X_i, C_k). These vectors correspond to the two compared years. Three variants of the input information were analysed:
- The share of dry days in individual months,
- The share of dry days in the year,
- The share of dry days in the IV–VIII season.

In the first case, the X_i and C_k vectors have 12 components; in the other cases, one component each.

On the basis of these distances, for the i-th year, a measure of its similarity (Uij) to humidity was determined to the remaining 9 analogues, $i = 1,2,3, ..., 10, j = 1,2, ..., 10, i \neq j$.

$$Uij = \frac{1}{\sum_{k=1}^{10} \frac{dist(Xi, Cj) * dist(Xi, Cj)}{dist(Xi, Ck) * dist(Xi, Ck)}} \qquad (6)$$

Each analogue also carries information about the fire class recorded in a given year. On the basis of the Uij similarity measures of humidity conditions, it is possible to determine the probability of the occurrence of each fire class. The following is an example illustrating how the Uij measures can take us to the probability of the occurrence of the fire class we are interested in.

For the subregion of the city of Warsaw, the following fire classes were noted in successive years:
year 1 (2007), class 3; year 2 (2008), class 2; year 3 (2009), class 3; year 4 (2010), class 2; year 5 (2011), class 3; year 6 (2012), class 2; year 7 (2013), class 1; year 8 (2014), class 3; year 9 (2015), class 4; year 10 (2017),– class to be determined based on the data from the previous 9 years.

Class 1 was recorded only in year 7 (2013); therefore, the probability of the occurrence of this class P(kl1) in 2016 is

$$P(k11) = \frac{U10,7}{U10,1 + U10,2 + U10,3 + U10,4 + U10,5 + U10,6 + U10,7 + U10,8 + U10,9} \quad (7)$$

Then, the conformity of the most probable fire class predicted with the observed one was analysed. This consistency was determined by the percentage share of underestimated, accurate and overestimated forecasts. The research tool described above was applied both to the basic 6 fire classes and to their division into two new classes.

In experiment B, the number of occurrences of forest fires in a year and the number of fires defined by the quantile classes were analysed. Research was conducted to confirm the existence of a relationship between the number of fires occurring in a given subregion of the central region and the relative humidity characteristics at the station representing that subregion. Owing to the lack of humidity observations at the Puczniew station, observations for the city of Łódź were used in the analysis for the Łódzki subregion.

Three characteristics of humidity were taken into account: average monthly relative humidity (RH_{SR_MC}), the minimum monthly value (RH_{MIN_MC}) and the percentage of days in the year with RH below 40% ($^{\%}RH_{40}$), which was based on relative humidity values from three periods (see experiment A). As only information on fires in particular years was available, in order to conduct a correlation analysis for the monthly mean and minimum RH values, we used annual or calculated characteristics for selected sub-periods of the year. Three sub-periods were selected for RH_{MIN_MC}: summer, from June to August; summer extended by the neighbouring spring months of April and May; and the warm half of the year (April–September). The following RH characteristics obtained in an annual time resolution were analysed:

$^{SR}RH_{SR\_MC}$	Annual average of monthly mean relative humidity;
^{MIN}RH	Annual minimum relative humidity;
$^{SR4-9}RH_{MIN\_MC}$	Warm half-year (from April to September) average monthly minimum relative humidity;
$^{SR4-8}RH_{MIN\_MC}$	Average monthly minimum relative humidity for the period of April–August;
$^{SR6-8}RH_{MIN\_MC}$	Average monthly minimum relative humidity in the summer season (from June to August);
$^{\%}RH_{40}$	Annual percentage of days with low relative humidity (below 40%).

The analysis used the correlation coefficient values of the above relative humidity characteristics and the number of fires. In addition, the correlation analysis in the contingency tables and the Mann–Whitney test were used (Bauer 1972, Hollander, Wolfe 1973) to determine the significance of the relationship between the number of fires and the percentage of days of the year with low relative humidity. For the Mann–Whitney tests, a sample of the number of fires for all subregions was used, assuming its stratification with the categorical variable for the classes of the share of days with humidity below 40%.

The calculations were performed in 'R' for statistical calculations and the graphical presentation of data and results, using the procedures of the stats and vcd package (R Core Team 2017).

In order to analyse the interdependencies in the contingency tables, numeric variables (of both relative humidity characteristics and data on the number of fires) were transformed into categorical variables using the quantiles from the sample:

Q25, 25% quantile; Q50, median; Q75, 75% quantile. The obtained categorical variables had the following values:

1 – for RH ≤ Q25,
2 – for Q25 < RH and RH ≤ Q50,
3 – for Q50 < RH and RH ≤ Q75,
4 – for RH > values above Q75.

For the five humidity characteristics, the values of standardised variables and normal distribution quantiles were used: Q25 = −0.76, Q50 = 0.0, Q75 = 0.67. The number of fires was divided into four classes using quantiles from the sample Q25 = 72, Q50 = 146 and Q75 = 299. The division into four classes of the percentage of days with low humidity was performed using quantiles $Q^U25 = 7.7\%$, $Q^U50 = 10.8\%$ and $Q^U75 = 13.05\%$.

4. Results

Completeness of data

The information on fires and humidity for 2007–2016 is complete. Unfortunately, there was no representative measurement point for the Łódź subregion, and as a result, the Łódź station represents two subregions – Łódzki and the city of Łódź.

Experiment A

For the Ostrołęcko-siedlecki subregion, humidity measurements from the climatic station in Pułtusk were used once, and for the second time, those from the synoptic station in Siedlce were used. The final results did not show differences because of the selection of the measuring point. This means that in the case of the research apparatus used in the study, both measuring points are equally representative for this subregion.

An interesting phenomenon is the adoption of the same values by indices for a year with the threshold of 10% and for the IV–VIII period with the threshold of 20%. Additional performed tests confirmed that doubling the threshold value for the IV–VIII period in relation to the threshold for the whole year increases the agreement between the designated indices.

AA. Relative risk

The value '−1' in Figure 2 means that the index value cannot be determined because the data were analysed from too short a period. In the remaining cases, the relative risk is greater than 1, which means that the lack of humidity is conducive to having a large number of forest fires.

The highest relative risk of forest fires at low relative humidity was observed for the IV–VIII period with the 10% threshold.

AB. Jaccard index

The Jaccard index exceeded the threshold of 50% on several occasions (Fig. 3), which means that the defined deficiency of humidity is an important factor not only accompanying but also shaping the number of forest fires. Also in this case, the index showed the strongest relationship
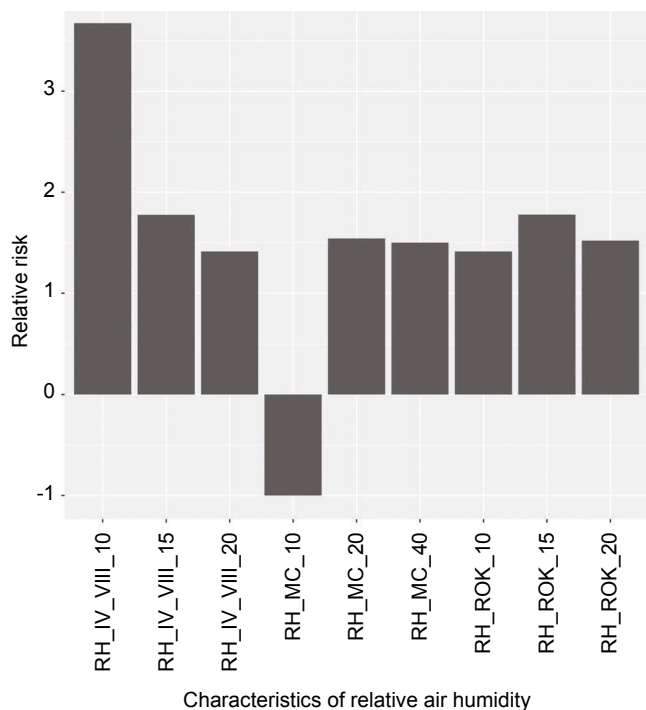
between humidity and the number of fires for the IV–VIII period with the threshold of 10%.

AC. Cross-tabulation tables

The obtained results indicated that the best approach is to use the relative humidity data for the IV–VIII period. When using monthly data and adopting the 10% threshold, specific values were related to the length of the data series (10 years), in which certain combinations of variables in the subregions did not occur (e.g. high humidity and a large number of fires).

In the case of the 10% threshold for days with low humidity in the month or in the warm period of IV–VIII, the HR indicator is close to 1, which means that the lack of humidity
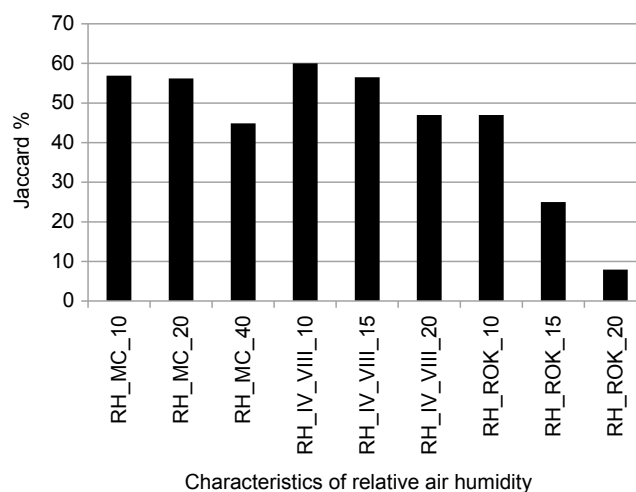

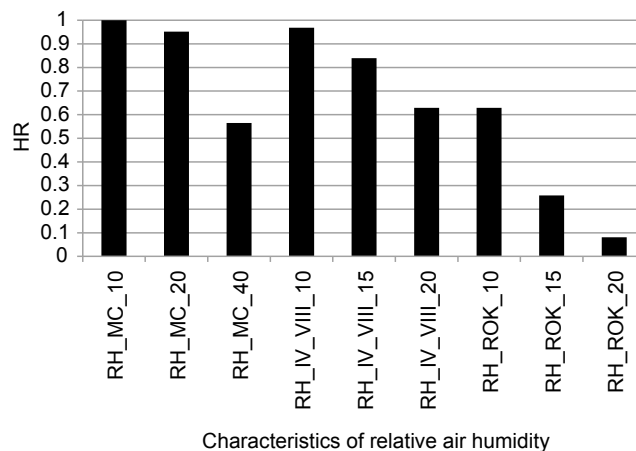**Figure 3.** The Jaccard's index for the experiments discussed

Figure 2. The relative risk for the analyzed experiments

Figure 4. The hit rate (HR) values for selected humidity characteristics and a fire class the new_class 1
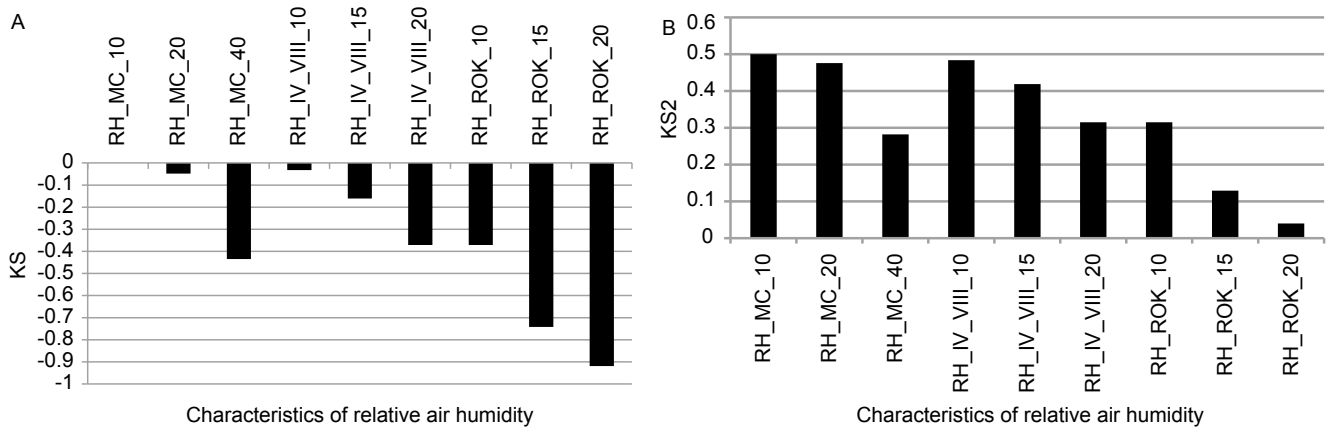
Figure 5. The KS index (A) and the scaled KS2 index (B)

is a good indicator of the occurrence of at least 125 fires in a given subregion (Fig. 4).

In all 9 variants, the FAR variant was 1. This confirmed the tendency of the reconstruction model to overestimate the number of fires.

The KS index did not have a positive value and the scaled KS2 was not greater than 0.5, which indicates that information on low humidity is useful in research on the number of forest fires, but it is not sufficient to correctly predict the threat (Fig. 5).

AD. The reconstruction of fire classes using fuzzy analogues

Figures 6–8 present the percentage of distribution in the subregions of the number of underestimated, accurate and overestimated reconstructions using the six-class and two-class approaches.

Experiment B

The values of the obtained correlation coefficient for the monthly average and the minimum monthly relative humidity (RH) as well as the percentage of low humidity days in the year ($^{\%}RH_{40}$) are included in Table 3. The correlation coefficient reached the highest, positive values for $^{\%}RH_{40}$; it was negative for the remaining parameters; and the highest absolute values were reached in the case of the annual mean with the monthly averages $^{SR}RH_{SR\_MC}$ (the absolute value of the correlation coefficient was below 0.5 for only three stations) and average for the warm half-year with the minimum monthly $^{SR4-9}RH_{MIN\_MC}$ (the absolute value of the correlation coefficient was below 0.5 only for Sulejów). In the case of the latter characteristic, the absolute value of the correlation coefficient was more than 0.7 for five subregions. Apart

from two cases (for Sulejów and Kozienice, and the characteristic $^{SR6-8}RH_{MIN\_MC}$, where the values are very close to zero), the correlation coefficient has negative values, suggesting the expected inverse linear relationship between the number of fires and the given humidity characteristics. Excluding September data from calculating the average minimum monthly relative humidity caused a reduction in the correlation coefficient. By restricting the calculations only to the summer months (June, July and August), this effect was even more visible. This reduction in the absolute value of the correlation coefficient suggests the important role of the neighbouring spring and autumn months for the summer included in the warm half-year.

The strongest correlations occurred for the variable describing the percentage of days in the year with humidity below 40% (Table 3), as in three subregions – Siedlce, Warszawa-Okęcie and Legionowo – the correlation coefficient was more than 0.9 and the level of significance of this coefficient was less than 0.001. For six subregions, it ranged from 0.7 to 0.95 and was statistically significant at $p = 0.01$; for the subregion of the city of Warsaw, it was 0.69 and statistically significant at $p = 0.05$.

The total of all the subregions were analysed (for a sample size of N = 110). The correlation coefficient for the number of fires per year and the percentage of days of the year with low humidity for the whole sample was 0.45, indicating the occurrence of a moderate correlation, in other words, a significant linear relationship between the variables.

The level of significance of the above correlation coefficient is less than 0.01. Calculations were also conducted for the categorical variables described for the experiment B methodology (classes defined on the basis of quantiles). For such processed variables, a correlation was confirmed only in the case of the

$^{SR4-9}RH_{MIN\_MC}$ characteristic. For this characteristic, a statistical significance below 0.05 was obtained for five subregions: Siedlce, Skierniewice, Warszawa-Okęcie, Sulejów and Kozienice.

The obtained results indicated that in the case of combining class 1 (less than 72 fires) with class 3 (from 146 to 299 fires) and class 1 with class 4 (more than 299 fires),
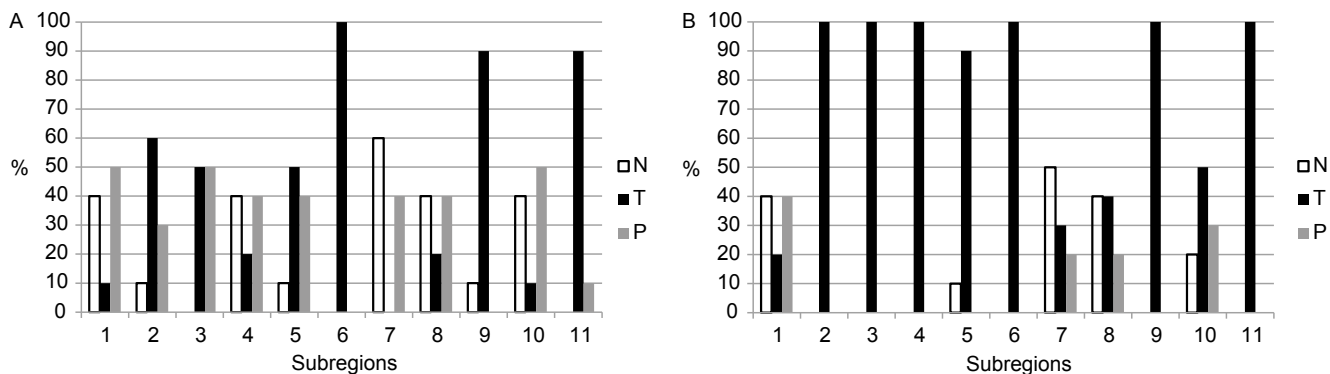

Figure 6. Percentage of underestimations (N), accurate reconstructions (T) and overestimations (P) of fire classes determined on the basis of information on monthly relative humidity deficiencies. A for 6 classes, B for cases: class > 3 / class < 4. Numbers of subregions as in Table 1.
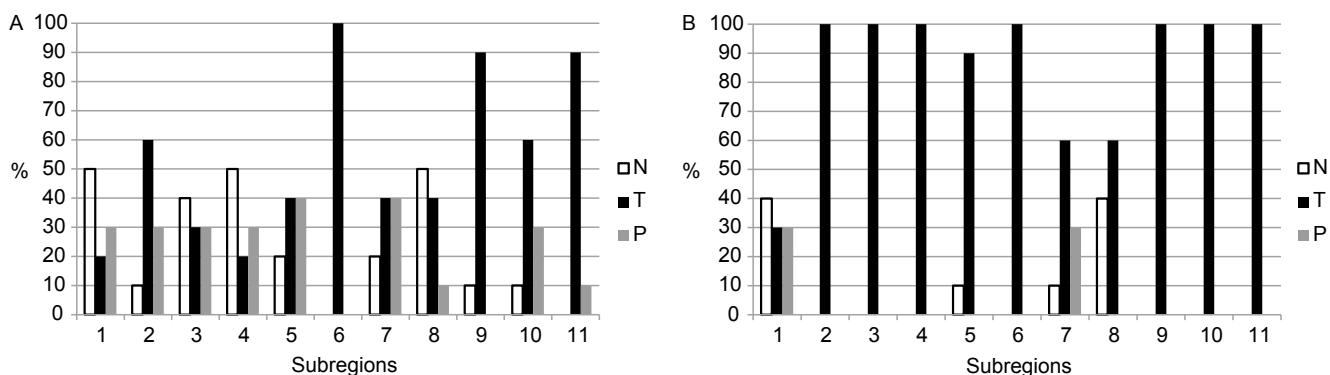

Figure 7. Percentage of underestimations (N), accurate reconstructions (T) and overestimations (P) of fire classes determined on the basis of information on annual relative humidity deficiencies. A for 6 classes, B for cases: class >3 / class <4. Numbers of subregions as in Table 1.
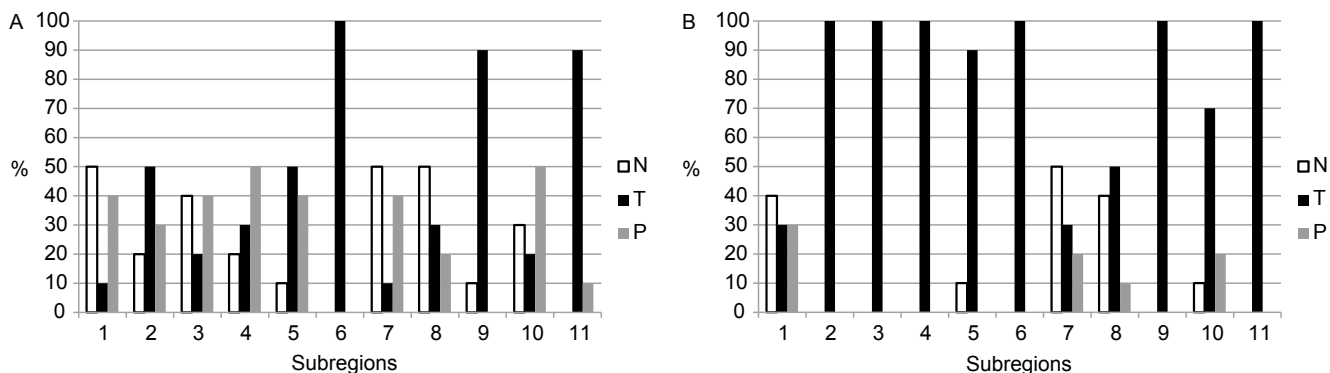

Figure 8. Percentage of underestimations (N), accurate reconstructions (T) and overestimations (P) of fire classes determined on the basis of information on relative humidity deficiencies in the months of IV–VIII. Fig. 8A for 6 classes, Fig. 8B for the case, class > 3 / class < 4. Numbers of subregions as in Table 1.

Table 3. Values of the correlation coefficient between the annual relative humidity characteristics and the number of fires in the subregion, 2007–2016. Significance level for correlation coefficient: * $p<0.05$; ** $p<0.01$; *** $p<0.001$.

Station	$^{SR}RH_{SR\_MC}$	^{MIN}RH	$^{SR4-9}RH_{MIN\_MC}$	$^{SR4-8}RH_{MIN\_MC}$	$^{SR6-8}RH_{MIN\_MC}$	$^{\%}RH_{40}$
Płock	-0.66*	-0.40	-0.67*	-0.61	-0.49	0.76*
Siedlce	-0.60	-0.58	-0.71*	-0.67*	-0.48	0.90***
Skierniewice	-0.72*	-0.46	-0.70*	-0.57	-0.29	0.70*
Warszawa-Okęcie	-0.80**	-0.57	-0.78**	-0.67*	-0.62	0.93***
Warszawa-Bielany	-0.63*	-0.47	-0.53	-0.49	-0.38	0.69*
Legionowo	-0.93**	-0.31	-0.80**	-0.72*	-0.64*	0.95***
Wieluń	-0.65*	-0.75*	-0.77**	-0.73*	-0.31	0.75*
Puczniew	-0.20	-0.38	-0.50	-0.42	-0.26	0.48
Łódź	-0.41	-0.67*	-0.65*	-0.59	-0.32	0.40
Sulejów	-0.58	0.23	-0.40	-0.24	0.04	0.33
Kozienice	-0.28	-0.54	-0.68*	-0.25	0.09	0.55

statistically significant changes occur in the distributions of the percentage of days with low humidity. The results of the Mann–Whitney test showed a statistically significant difference only between the distributions of the number of fires in the first group ($^{\%}RH_{40} \leq Q^U25$) and each of the other groups. For the $Q^U25 < {}^{\%}RH_{40} \leq Q^U50$ group, the level of significance was $p = 0.025$; for the $Q^U50 < {}^{\%}RH_{40} \leq Q^U75$ group, it was $p = 0.039$, and for the last group $^{\%}RH_{40} > Q^U75$, it was $p = 0.006$. For the remaining sets of groups, the p value of the tests ranged from 0.3 to 0.8. An illustration of the relationship described above is the distributions of one variable in classes divided according to the second variable (Figs. 9 and 10).

The tests performed in this case showed a statistically significant difference between the distribution of the number of fires in the first class (percentage of days with low humidity below 7.7%) and the distribution in each of the other three classes of the percentage of days with low humidity. The calculations suggest a statistical association between the percentage of days with low humidity and the annual number of fires.

5. Summary

Meteorological factors have a large impact on the occurrence and spread of fires in forests. Wiler and Wcisło (2013) named, amongst others, wind that can move fire from areas outside the forest, precipitation, cloudiness, atmospheric pressure, and heat exchange in the atmosphere. Szczygieł et al. (2009) emphasised the importance of temperature,
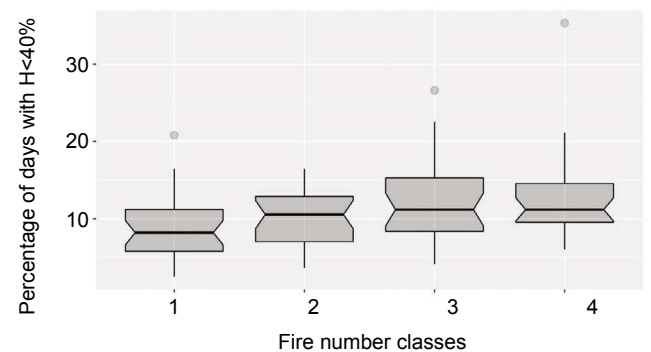
Figure 9. Distribution of the percentage of the number of days with low humidity $^{\%}RH_{40}$ in the classes of the number of fires
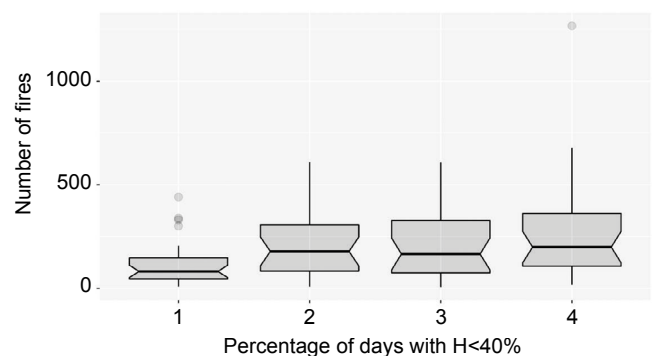
Figure 10. The dependence of the number of fires on classes by percentage of days with low humidity $^{\%}RH_{40}$

relative air humidity, wind speed and precipitation. In the case of short- and medium-term weather forecasts (e.g. The European Centre for Medium-Range Weather Forecasts – ECMWF), there are models that, based on the above factors, can determine the threat of forest fires in a given area several dozen hours ahead of time. From the point of view of managing crisis situations, forecasts for the next month or even the season may be useful. Climate researchers are interested in both the impact of climate change on the occurrence of forest fires and the way in which forest fires affect the climate. The development of such a long-term forecasting model requires the construction of an appropriate forecasting algorithm and the selection of predictors, that is, the manner of providing input information on the state of the atmosphere. The advancement of work on monthly, quarterly and seasonal weather forecasts concerning temperature, total precipitation and the number of days with precipitation is optimistic. We hope that the applied model will also work for long-term air humidity forecasts. The relative humidity research presented in this article confirms that the proposed characteristics contain much information about the threat of forest fires.

The cases noted in experiment A (Fig. 3), when the Jaccard index exceeded 50%, mean that low relative air humidity not only accompanies the occurrence of forest fires but is also inseparably connected with it. The obtained values of the Hanssen and Kuipers score KS index (no positive values) and the scaled Hanssen and Kuipers score KS2 index (not exceeding the threshold of 0.5) (Fig. 5a and b, respectively) mean that the list of predictors needs to be expanded. Therefore, in further studies, temperature and precipitation characteristics need to be added. It would be interesting to check the role of the air temperature predictor not only at a height of 2 m but also at ground level (because of the influence of the litter conditions on the emergence and spread of fires in forests). Long-term predictions of cloudiness and wind seem to be too prone to error at the moment; however, it is possible that this topic will also be addressed in further research.

Reconstruction of a fire class using fuzzy analogues seems to be a good direction of work on a model for a long-term forecast. It is clear that when using two fire classes, the accuracy of the reconstruction is much higher than with the use of six classes. It should be remembered that the limited length of the data series (only 10 years) had a significant impact on the obtained results. This is followed by the lack of reports on many situations (e.g. the occurrence of a fire class in the eastern Warsaw subregion other than class 6) and does not contain all possible fire and weather situations. Thus, the model may not be sufficient to make correct inferences with greater variation at the start. However, the selection of the best time interval for the predictor (month, year, IV–VIII period) depends on the choice of the subregion (Figs. 6–8).

On the basis of the results obtained in experiment B, we can conclude that amongst the humidity characteristics considered there, the annual mean and average mean for the warm half-year from the monthly minimum air humidity, as well as the percentage of low humidity in the year can play a significant role in the construction of the forecasting model.

As mentioned earlier, this study is a pilot, with the results allowing us to prepare the research apparatus for further analyses. The results obtained show that combining relative air humidity deficiencies with the number of fires and taking into account these dependencies in further work on improving long-term weather forecasts is of great importance. The cooperation of specialists in the field of weather forecasting and climate change researchers with persons working on fire risks in Poland can bring many benefits. The effects of such cooperation will help to improve the methods of preventing the effects of phenomena that threaten the life of people and the environment.

Further work on this issue should concern
• Obtaining information on the number of fires before 2007,
• Expanding the list of meteorological measurement points to include forest stations,
• Providing information about fires at the seasonal or monthly scale,
• Broadening the list of potential predictors to include temperature and precipitation characteristics,
• Adapting and verifying various models of monthly, three-month and seasonal forecasts for fire protection needs.

Conflict of interest

The authors declare the lack of potential conflicts of interest.

Acknowledgments and source of funding

References

Bauer D.F. 1972. Constructing confidence sets using rank statistics. *Journal of the American Statistical Association* 67: 687–690. DOI 10.2307/2284469.

Dyrekcji Generalnej Lasów Państwowych. 2012. Instrukcja ochrony przeciwpożarowej lasów. Centrum Informacyjne Lasów Państwowych, 108 s. ISBN 978-83-61633-59-4

Forest Fire in Europe 2009. Report no 10. Publications Office of the European Union, Luxembourg, 88 s. DOI 10.2788/74089.

Hollander M.I., Wolfe D.A. 1973. Nonparametric Statistical Methods. John Wiley & Sons, New York, 848 s. ISBN 978-0-470-38737-5.

Klasyfikacja NUTS. 2017. http://stat.gov.pl/statystyka-regionalna/ jednostki-terytorialne/klasyfikacja-nuts/ [5.01.2017].

KSIPL 2017. Krajowy System Informacji o Pożarach Lasów. http://bazapozarow.ibles.pl/ibl_ppoz/faces/index.jsp [15.01. 2017].

Kwiatkowski M., Szczygieł R., Piwnicki J. 2010. Opracowanie nowej metody prognozowania zagrożenia pożarowego lasu. Instytut Badawczy Leśnictwa, Sękocin Stary, 90 s.

Kwiatkowski M., Szczygieł R. 2015. Klimatyczne aspekty ochrony przeciwpożarowej lasu. Konferencja pt. „Klimatyczne uwarunkowania życia lasu", Rogów, 16-17 czerwca 2015 r. http://cepl.sggw.pl/konferencja/klimat/abstract/Klimatyczne_ aspekty_ochrony_przeciwpozarowej_lasu.htm [5.02.2017].

Matteucci M. 2017. A Tutorial on Clustering Algorithms. https://home.deib.polimi.it/matteucc/Clustering/tutorial_html/cme-ans.html [7.03.2017].

World Meteorological Organization 2002. Standardised Verification System (SVS) for Long-Range Forecasts (LRF). New attachment II-9. *Manual on the GDPS* 485(1).

Prędecka A. 2011. Ogień w lesie a przyroda. Dyrekcja Generalna Lasów Państwowych, Warszawa, 166 s.

R Core Team 2017. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/ [11.01.2017].

Real R., Vargas H.M. 1996. The Probabilistic Basis of Jaccard's Index of Similarity. *Systematic Biology* 45(3): 380–385. DOI 10.1093/sysbio/45.3.380

Szczygieł R., Ubysz B., Piwnicki J. 2009. Kryteria oceny ryzyka zagrożenia pożarowego lasu. *Technika i Bezpieczeństwo Pożarowe* 3: 83–132.

Szczygieł R. 2012. Wielkoobszarowe pożary lasów w Polsce. *Bezpieczeństwo i Technika Pożarnicza* 1: 67–78.

Wiler K., Wcisło P. 2013. Ochrona lasów przed pożarami. CILP, Warszawa, 6–88. ISBN 978-83-61633-95.

Authors' contribution