Efficiency evaluation of the weighted mean calculation of the forest fire hazard class

R V Kotelnikov* and A N Chugaev
The Branch ‘Center of the Forest Pyrology’, All-Russian Research Institute for Silviculture and Mechanization of Forestry, 42 Krupskoy Street, 660062, Krasnoyarsk, Russia

*E-mail: kotelnikovrv@firescience.ru

Abstract. Nowadays, cost-optimization of aerial patrolling plays a key role in the context of limited aerial forest protection funding. Forest Fire Danger Class is the main indicator that regulates the work of forest fire services. Usually, it’s calculated by the nearest weather station data. Some information systems use the mean of several nearby weather stations to estimate large areas, such as the surveyed area of aerial forest protection. The idea of using the mean weighted index with the weather stations weighting factor is not new. Even though, this idea isn’t widespread due to the calculation complexity and questionable efficiency in practice, this study proposes a scientifically substantiated method of quantitative comparison of two approaches and the direct calculation method of the economic impact when transition to using the mean weighted Forest Fire Danger Class calculation algorithm. The first time such an indicator was used to obtain derivatives of analytical information products. A long-term analysis of forest fire rate showed that the weighted mean of the Forest Fire Danger Class value is 6.7% greater in correlation with the number of forest fires than the usual mean value. The use logarithmic transformation of the forest fire occurrence frequency and population density allows statistical criteria to be reasonably used.

1. Introduction
Taking into account the presence of global warming, the negative effects of forest fires will only continue to increase [1,2]. This leads to the need to find ways of improving the detection and suppression of forest fires.

At the beginning of the twenty-first century, there have been significant breakthroughs in some scientific and technological fields, in particular: remote sensing technologies, unmanned systems, big data analysis and processing, geospatial analysis, and data display technologies (GIS - systems), etc. The ubiquitous digitalization of data opens up fundamentally new possibilities for modeling and forecasting processes in the forest sector in general and in fire protection [3]. Apart from the obvious applications such as detection of forest fires from space and forecasting of forest fire propagation [4], new information technologies also affect organizational issues such as planning. Taking into account these changes, a new round of development was received by the management decision support systems [5].

One of the particular tasks of planning is the adequate estimation of fire danger in forests in order to optimize the costs of aerial patrolling. The regulation of the work of aviation forest fire protection units in terms of the appointment of aircraft departures is based on the fire hazard indicator in forests, depending on weather conditions (hereinafter – Forest Fire Danger Class (FFDC)).
Research in the field of forest fire danger estimation has a long history, and a large number of different indicators exist in the world [6-9]. Studies are periodically carried out to assess which indicators are more effective [10]. In Russia, due to the large areas of forests, there are difficulties with obtaining detailed information about forests, which causes the use of simplified scales in the operational work of forest fire services, based mainly on weather indicators. Almost all scales use indicators based on data from stationary weather stations, which are geographically unevenly distributed. In some cases, the values of meteorological indicators are used, converted to points of the regular grid. At the same time, for management decisions associated with the need to patrol forests located within a certain area (e.g. within the service area of air unit), the initial fire danger indicator needs to be converted accordingly. In Russia, the density of distribution of meteorological stations is very low, which significantly affects the accuracy. Taking into account the geographical distribution of weather stations is the subject of this article.

In the early days of aerial forest protection, FFDC by weather conditions was calculated by employees themselves – usually, radio operators of the air unit, based on the readings of the instruments located directly next to the premises (almost in the center of the surveyed area). It was this data that gradually formed the estimated patrol frequency (the number of required flights depending on the FFDC value). Subsequently, the independent calculation was abandoned in favor of official data generated by certified instruments of the nearest weather station. Even back then, a single FFDC value did not always characterize the fire danger of all forests in a protected area, especially when the area per weather station was significant, or when there were several weather stations in the protected area. Calculation of the mean FFDC value by several weather stations located near forests protected by air units is still used in the Information System for Remote Monitoring of Forest Fires of the Federal Forestry Agency (ISDM-Rosleskhoz) [11,12].

However, this approach has certain disadvantages with the uneven location of weather stations over the territory. Economic factors have led to a reduction in the number of aviation departments and an increase in the area to be inspected which exacerbated the problem.

Forest legislation strictly regulates aerial patrolling of forests depending on FFDC [13]. Thus, for example, at an FFDC value of II, aerial patrols have to be carried out once every two days, while at an FFDC value of III the aerial patrols are carried out twice in one day. Taking into account the significant cost of aircraft leasing (for example, in the Republic of Sakha (Yakutia) the mean rate of An-2 aircraft is 172 thousand rubles per hour) [14], and the requirements for the adequacy of FFDC value increase significantly.

One way to improve the quality of management decisions on assigning patrol flights on a given day is to use a weighted mean of the fire danger class obtained by summing the products of the FFDC for each weather station by the weight factor of each weather station (1):

\[
\text{FFDC} = \sum_{i=0}^{n} ffdc_i \cdot r_i
\]

where \( ffdc \) – weighted mean fire risk class of air unit; \( n \) – the number of weather stations affecting the value of indicators in forest district; \( ffdc \) – fire risk class according to weather station; \( i \) – serial number of weather station; \( r_i \) – weight factor of the i-th weather station.

The proportion of the area of influence of each weather station (Thiessen polygon, i.e. points equidistant from the weather stations) can be used as a weight factor. However, this approach is not used in practice due to the complexity of calculations and questionable efficiency.

Within the framework of this work, it is supposed to statistically prove that the weighted mean FFDC value more accurately reflects the fire hazard of forests than the conventional mean value.

In addition, using the approved values of the estimated frequency of forest patrols, it is possible to numerically estimate financial savings for aerial forest patrols in the event of a transition to a new approach.
2. Methods

To choose one of the two options for calculating FFDC, which to a greater extent characterizes the fire hazard of forests, we estimate the degree of relationship between them and the frequency of forest fires.

The main problem with this method is the large variation in the values examined associated with a large number of the factors unaccounted for. Due to the fact that simple averaging of indicators for a long period loses its meaning, because except for compensation of deviations of unnecessary factors, there also occurs averaging of the FFDC value itself. So, it is proposed to take the fire hazardous season as the calculation period, but instead of the mean FFDC, use the intensity of the fire hazardous season \( H \) (2):

\[
H = \frac{d_{3,4,5}}{D}
\]

where \( D \) – duration of the fire season; \( d_{3,4,5} \) – number of days with FFDC >2.5.

The intensity of the fire season is understood as the number of days from FFDC III to FFDC V of the total number of days in the fire season.

Given that the main cause of forest fires is the human factor, it is advisable to adjust the calculation to account for the influence of population density. This can be done using algorithms to calculate partial correlation factors (estimating the degree of weather impact on fires with a fixed population size).

Another problem is the shape of distribution of random variables. To use the correlation factor as a criterion for the relationship between indicators, the samples being compared must have a distribution that is close to normal.

Considering that the frequency of forest fires and population density are distributed by lognormal law, it is proposed to use hyperbolic logarithms of these values in the calculation.

For this study, we select data by aerial forest protection service areas in the Republic of Sakha (Yakutia), and we take the fire season as the interval under consideration.

The cost of forest patrols is directly proportional to the number of flights to patrol the area. The number of flights is determined taking into account the regulations [13] and is calculated by formula (3):

\[
V = 0.5 \cdot n_2 + 1 \cdot n_3 + 2 \cdot n_4 + 2 \cdot n_5
\]

where \( V \) – the number of aerial patrol flights; \( n_2, n_3, n_4, n_5 \) – number of days when the fire danger class depending on weather conditions was respectively 2, 3, 4 and 5; multipliers 0.5, 1, 2 and 2 – normative frequency of patrolling for appropriate fire danger classes according to weather conditions [13].

To assess the impact of the new methodology on costs, find the ratio of the number of flights calculated by the weighted mean algorithm to the number of aerial patrol flights calculated by the mean, then multiply by 100 (to express as a percentage) and subtract 100 (to show in ‘deviation’ format) (4).

\[
dZ = \frac{V_w}{V_{\text{mean}}} \cdot 100 - 100
\]

where \( dZ \) – deviation of the number of aerial patrol flights at transition to weighted mean FFDC calculation (percentage); \( V_{\text{mean}} \) – number of aerial patrol flights, which is calculated based on the number of days with the weighted mean of the fire danger class; \( V_w \) – number of aerial patrol flights, which is calculated based on the number of days with the weighted mean of the fire danger class; multipliers 0.5, 1, 2 and 2 – normative aerial patrol frequency for corresponding fire danger classes according to weather conditions [13].

Taking into account the above-stated, the new algorithm more adequately estimates current fire danger in forests depending on weather conditions. Hence, the situation where there is no need to carry out aerial patrols under the new calculation and patrols under the old calculation is necessary, can be interpreted as a waste of resources.
In some cases, the opposite situation is possible, when a departure is required according to the new algorithm, but flights are not performed according to the old algorithm. Based on the argument that the old algorithm is less representative of fire hazards, and that the risk of missing a fire is higher than the savings from a canceled flight, such a situation cannot be considered economically reasonable. Thus, one way of financially justifying the need to switch to the new calculation algorithm could be to consider the amount of money spent on flights when there is no need to fly under the new algorithm or in simpler terms, the proportion of extra flights. To calculate this indicator, we divide the difference of assigned aerial patrol flights according to the new method and the old method calculated for days when FFDC according to the new method is less than the old method by the total number of aerial patrol flights, which is calculated based on the number of days with a mean fire danger class value (5):

$$d = \frac{V_{\text{mean}} - V_{w}}{V_{\text{mean}}} \quad (5)$$

where $d$ – proportion of aerial patrol flights that were not required; $V_{\text{mean}}$ – total number of aerial patrol flights, which is calculated based on the number of days with a mean fire danger class value; $V_{\text{mean}}'$ – number of aerial patrol flights, which is calculated based on the number of days with mean fire danger class only for days for which the mean FFDC is higher than the weighted mean; $V_{w}$ – number of aerial patrol flights, which is calculated based on the number of days with a weighted mean fire danger class value; multipliers 0.5, 1, 2 and 2 – standard patrol frequency for respective fire danger classes based on weather conditions [13].

Thus, the proposed calculation method implies the following algorithm:

- calculate the frequency of forest fires;
- calculate the FFDC value for each day for each weather station;
- calculate the mathematical FFDC mean for each day for each weather station;
- calculate the fire period tension by the mean FFDC value for each air unit (for each year);
- build the impact zones of weather stations and calculate their weighting factors;
- calculate the weighted mean FFDC value for each day for each air unit;
- calculate the tension of the fire period based on the weighted mean FFDC values (for each year);
- calculate the partial correlation between the hyperbolic logarithm of forest frequency and the mean FFDC (choosing the logarithm of population density as the controlled variable);
- calculate the partial correlation between the hyperbolic logarithm of forest frequency and the weighted mean FFDC (choosing the logarithm of population density as the controlled variable);
- compare the value of the partial correlation factor;
- estimate possible changes in the costs of aerial patrols;
- calculate the proportion of aerial patrol flights with the assignment more rational in case of transition to the new method.

The Republic of Sakha (Yakutia) is chosen as an example to estimate savings for an entity of the Russian Federation.

3. Results and discussion

Using the above method, we obtained samples of values for the years 2004 -2020 for the air units of the Republic of Sakha (Yakutia).

Sample 1 (FFDC value by day for each air unit), contains 107,673 observations and is structured as follows:

- year;
- air unit;
- date;
- mean FFDC
- weighted mean FFDC.
Sample 2 (fire season tension value by year for each air unit) contains 393 observations and is structured as follows:
- year;
- air unit;
- frequency of forest fire occurrences;
- hyperbolic logarithm of the frequency of forest fire occurrences (lnN);
- fire season tension by mean FFDC (Hawg);
- fire season tension by weighted mean FFDC fire dangerous season (Hw);
- population density (within air unit boundaries);
- hyperbolic logarithm for population density (lnNas).

Sample 3 (number of days with mean FFDC) contains 10322 observations and is structured as follows:
- year;
- air unit;
- number of days with mean FFDC=2;
- number of days with mean FFDC=3;
- number of days with mean FFDC=4;
- number of days with mean FFDC=4;
- number of days with mean FFDC=5.

Sample 4 (number of days with each FFDC for the case where the mean is greater than the weighted mean) contains 6870 observations and is structured as follows:
- year;
- air unit;
- number of days with mean FFDC=2;
- number of days with mean FFDC=3;
- number of days with mean FFDC=4;
- number of days with mean FFDC=5;
- number of days with weighted mean FFDC=2;
- number of days with weighted mean FFDC=3;
- number of days with weighted mean FFDC=4;
- number of days with weighted mean FFDC=5.

To clear the data from random outliers, let us construct a box plot for each indicator (figure 1).

**Figure 1.** Box plot of indicators: Hw – weighted mean tension of fire season; Havg – mean tension of fire season; lnN – frequency of forest fires; lnNas – logarithm of population density.
The presented box plot partially confirms the closeness of distributions to normal, giving an opportunity to exclude outliers from the calculation.

To confirm the normality of distributing the indicators studied, let us construct the distribution of the logarithm of fire forest frequency, the logarithm of population density, mean fire season tension, and weighted mean fire season tension (figure 2, respectively a, b, c, d) and calculate the numerical values of the most common normality criteria (table 1), using the Statistica software package.

![Figure 2](image_url)

**Table 1.** Numerical values of main criteria for distribution normality.

| Indicator                        | Kolmogorov Smirnov, $d$ | $p$     | Shapiro-Wilk, $W$ | $p$     | Lilliefors, $p$ |
|----------------------------------|------------------------|---------|-------------------|---------|----------------|
| Logarithm of fire forest frequency | 0.04                   | >0.2    | 0.99              | 0.15    | >0.2           |
| Mean fire season tension         | 0.03                   | >0.2    | 0.99              | 0.11    | >0.2           |
| Weighted mean fire season tension | 0.04                   | >0.2    | 0.99              | 0.08    | >0.2           |
| Logarithm of population density  | 0.16                   | <0.1    | 0.95              | 0.00    | <0.1           |

The presented criterion values for most of the indicators do not give reason to reject the hypothesis of distribution normality. The only exception is the logarithm of the population density. Given the fact that linear regression is not used in the problem in question, and the indicator in question is only needed to reduce the impact of population density on the result of comparing different algorithms, the deviation from normality can be neglected. This means that this deviation does not have a strong impact on the conclusions obtained.
To calculate the individual correlation values (in the case of logarithm of population density as a control variable), we will use the Statistica statistical package. After removing the outliers line by line, a sample of 300 observations remains. All results are significant at the level of p<0.05.

**Table 2.** Results of calculating partial correlation of indicators with the logarithm of fire forest frequency (when the logarithm of population density is controlled).

| Indicator                        | Mean  | Standard deviation | r(X,Y) | r²    | t     | Degrees of freedom | p      | Number of observations |
|----------------------------------|-------|--------------------|--------|-------|-------|--------------------|--------|------------------------|
| Mean fire season tension         | 0.32  | 0.15               | 0.48   | 0.23  | 9.46  | 297                | 0      | 300                    |
| Weighted mean fire season tension| 0.34  | 0.15               | 0.51   | 0.26  | 10.34 | 297                | 0      | 300                    |

As can be seen from table 2 the factor of partial (i.e. calculated when excluding impact of logarithm of population density), correlation r for relation between logarithm of forest fires frequency and weighted mean value of tension of fire danger season is 6.9% higher than analogous index for usual mean tension value.

Let us substitute the values in formula (4) and obtain an estimate of the change in the frequency of aviation patrols when switching to the calculation according to the weighted mean indicator of the intensity of the fire hazardous season 0.47.

Thus, when switching to a new calculation algorithm, the costs will practically not increase (an increase of less than 0.5%) due to multidirectional changes. Using formula (4), we determine that the percentage of unreasonable flights that will be used more efficiently when transitioning to the new algorithm is about 4%.

Since the tension of the fire season is calculated based on the forest fire danger class depending on the weather conditions, the findings on the greater relationship between the tension of the fire season, calculated by the weighted mean FFDC and the frequency of forest fires can also be attributed to the weighted mean FFDC itself.

The results obtained statistically significantly prove the hypothesis that the weighted mean value of forest fire danger class depending on the weather conditions more adequately estimates the potential forest fire rate than the mean value. Despite the fact that the result is based on data for one region, the multi-year nature of the data used excludes the random nature of the relationship, which suggests that the conclusion will be approximately the same throughout the Russian Federation. Nevertheless, a similar calculation for all regions in the future has been planned.

The results obtained can be used as a scientific justification for the need to include in the ISDM-Rosleskhoz modernization plan work on changing the algorithms for calculating the fire hazard class in the context of structural units of forest fire units carrying out work on aviation protection of forests.

According to preliminary expert estimates, the cost of revising ISDM-Rosleskhoz for the implementation of the new approach will require only 7 man-shifts.

4. **Conclusion**

Thus, to estimate adequacy of mean and weighted mean FFDC values, it is suggested to use individual correlation factors of each indicator with the number of forest fire occurrences. The indicators of the frequency of occurrence of forest fires and population density need to be transformed by logarithms.

Due to the need to compare long periods (during the whole fire season) it is necessary to use in the calculation not the FFDC value itself, but the indicator calculated on its basis - the intensity of the fire season (the share of days with FFDC III or higher).

Therefore, the transition to the new algorithm will not increase the costs (less than 0.5% increase) due to the multidirectional changes. Using formula (4), we determine that the percentage of unreasonable flights that will be used more efficiently when transition to the new algorithm is about 4%.
The use of a new approach in planning aviation patrols will increase the adequacy of management decisions and allow more efficient spending of funds allocated to regions for monitoring fire hazards in forests and forest fires.

At the same time, a rough estimation (using one region as an example) shows a 4% increase in efficiency with an increase in costs of no more than 0.5%.

Due to the mathematical substantiation obtained in the course of the research, the weighted average values of the intensity of the fire hazardous season have been actively used by the specialists of the Center for Forest Pyrology since 2019 as the main indicator characterizing weather factors in assessing the effectiveness of the activity of forest fire formations.

References

[1] Jia G et al. 2019 SPM2 Land–climate interactions SPM2 Land–Climate Interactions Climate Change and Land: an IPCC Special Report on Climate Change, Desertification, Land Degradation, Sustainable Land Management, Food Security, and Greenhouse Gas Fluxes in Terrestrial Ecosystems (Intergovernmental Panel on Climate Change) chapter 2 pp 131-248

[2] Feurdean A et al. 2020 Recent fire regime in the southern boreal forests of western Siberia is unprecedented in the last five millennia. Quaternary Sci. Rev. 244 106495 doi: 10.1016/j.quascirev.2020.106495

[3] Kotelnikov R V, Verkhovets S V and Ageev A A 2019 Krasnoyarsk cluster of forest science – a new stage of development. Forestry Information 3 156 doi: 10.24419/LHI.2304-3083.2019.3.13 [In Russian]

[4] Taylor S W, Woolford D G, Dean C B and Martell D L 2013 Wildfire prediction to inform fire management: statistical science challenges. Statistical Science 28(4) 586 doi: 10.1214/13-STS451

[5] Kotelnikov R V, Korshunov N A and Giryaev N M 2017 Objectives of decision making in protecting forests from fires. Main priorities on development of informational support. Sib. J. For. Sci. 5 18 doi: 10.15372/SJFS20170502 [In Russian]

[6] Ivanov V A, Goroshko A A, Baksheeva E O, Golovina A N and Morozov A S 2020 Regional scales of fire danger according to weather conditions for the forests of the Amur region. Coniferous Forests of the Boreal Zone 38(1-2) 34 [In Russian]

[7] Srock A F, Charney J J, Potter B E and Goodrick S L 2018 The Hot-Dry-windy index: a new fire weather index. Atmosphere 9(7) 279 doi: 10.3390/atmos9070279

[8] Jain P, Tye M R, Paimazumder D and Flannigan M 2020 Downscaling fire weather extremes from historical and projected climate models. Climatic Change. 163(1) 189 doi: 10.1007/s10584-020-02865-5

[9] Pezzatti G B, De Angelis A, Bekar I, Ricotta C, Bajocco S and Conedera M 2020 Complementing daily fire-danger assessment using a novel metric based on burnt area ranking. Agr. Forest Meteorol. 295 1 doi: 10.1016/j.agrformet.2020.108172

[10] Ziel R H, Bieniek P A, Bhatt U S, Strader H, Rupp T S and York A A 2020 Comparison of fire weather indices with modis fire days for the natural regions of Alaska. Forests. 11(5) 516 doi: 10.3390/f11050516

[11] Kotelnikov R V, Lupyan E A, Bartalev S A, Ershov D V 2020 Space monitoring of forest fires: history of the creation and development of ISDM-Rosleskhoz. Contemp. Probl. Ecol. 13(7) 795 doi: 10.1134/S1995425520070045

[12] Kovaliev N A, Lupian E A, Balashov I V, Bartalev S A, Beerst M A, Ershov D V, Krivosheev N P and Mazurov A A 2020 ISDM-Rosleskhoz: 15 years of operation and evolution. Sovr. Probl. DZZ Kosm. 17(7) 283 doi: 10.21046/2070-7401-2020-17-7-283-291

[13] Order of the Ministry of Natural Resources of Russia of 15.11.2016, available at: http://www.consultant.ru/document/cons_doc_LAW_214709 [In Russian]

[14] Official Website of the Unified Information System in the Field of Procurement, available at: https://zakupki.gov.ru/epz/main/public/siteMap.html [In Russian]