Multy-year oscillations investigation of winter temperatures Prediction of integrated temperature difference during the heating period

Ivan Khazheev[*]

[1]Melentiev Energy Systems Institute SB RAS, Department of Applied Mathematics, 664033 Irkutsk, Lermontov st., 130, Russia

Abstract. To estimate possible deviations in fuel consumption for heating based on meteorological observations of previous years, the integrated temperature difference inside and outside the building during the heating season is used. When the heating period is divided into two subperiods relative to the considered date (for example, before and after December 1), the accumulated and residual integral temperature differences are obtained. The assumption about the presence of a statistical relationship between the accumulated and residual integral temperature difference is confirmed. A model for predicting the probability of the expected values of the integral temperature difference for the upcoming heating period is developed. The model is focused on obtaining matrices of conditional probabilities of observations from intervals of dividing the accumulated integral temperature differences into intervals of residual integral temperature differences.

1 Introduction

Due to the long heating period and low winter temperatures, the study of climatic parameters of the heating period is of greater importance for Russia. These are the outdoor temperature, wind speed, air humidity, the intensity of solar radiation. The values of these indicators are used to assess the winter season of the region in question.

Usually the heating period of the region in question is determined by its duration and ambient air temperatures. In coastal areas, an additional cold load can be caused by relative humidity.

The article explores the possibilities of short-term forecasting of the integral temperature difference based on weather data of the first months of the heating period. The assumption of the presence of a statistical relationship between the accumulated and residual integral temperature differences is tested. A model is being developed for predicting the probability of expected values of the integral temperature difference for the upcoming heating period.

Let $\tau = 1,2,...,T$ are the numbers of the studied heating periods, T – the number of heating periods. L_τ denote the duration of the heating period τ. The values $i = 1,2,..., L_\tau$ as the sequence numbers correspond to day heating period τ.

Each heating season starts in the autumn of one calendar year and ends in the spring of the following calendar year. Therefore, each ordinal number of the heating period corresponds to two calendar years. So, the last heating period began in 2019 and ended in 2020.

The article uses the estimated duration of the heating period. To determine it, the following rule is used: the heating period begins if, for five consecutive days, the average daily temperature of atmospheric air is below 8 degrees Celsius. The heating period is considered over if the air temperature is above 8 degrees Celsius for five consecutive days. This formal rule is usually followed by heating systems in settlements.

In solving many problems associated with the heat supply of buildings, an indicator of the integral temperature difference inside and outside the building for the heating period is used. For the considered settlement or district, the indicator of the integral temperature difference is calculated by the formula:

$$B_\tau = \sum_{i=1}(t_n - t_{\tau i}), \quad i=1,2,..., L_\tau, \ \tau=1,2,...,T \quad (1)$$

Here, the value $t_{\tau i}$ is the average daily temperature of the air on the day i of the heating period τ. The indicator B_τ is determined on the basis of meteorological observations.

The value t_n is the specified normative value of the air temperature inside the building. Depending on the purpose of the building, there may be different values of normative temperatures. For residential premises, a temperature of 20 degrees Celsius can be used for the standard, for children's institutions - 24 degrees Celsius, for office premises - 18 degrees Celsius. Whereas in production rooms and warehouses, the normative temperature can be 14 degrees Celsius. In the calculations presented in this article, a standard value of 18 degrees Celsius was used.

The use of the indicator of the integral temperature difference inside and outside the building for the heating

---

[*] Corresponding author: ivan-khazheev@yandex.ru

period (1) is based on the law of thermal conductivity: the loss of thermal energy through the fencing, in the construction of buildings is proportional to the difference in temperature inside and outside the building. Therefore, the ratio of the integral temperature difference in different settlements or in different years for the same settlements reflect the ratio of heat energy consumption for heating and, as a result, fuel consumption for heat for heating.

The integral temperature difference is in practice applicable for calculating the optimal building structures at which heat losses are minimized.

The temperature of the air is not the only factor determining the requirements for the heat supply of buildings. In some cases, other natural factors, such as wind speed and direction, air humidity, the intensity of solar radiation, as well as technical characteristics and features of the operation of buildings, are essential.

At the same time, the indicator of the integral temperature difference is one of the most important characteristics of the degree of climate severity during the heating period for the region under consideration.

2 Cumulative and residual integral temperature differences inside and outside the building

Divide the heating period into two sub-periods: before and after the date of consideration (for example, before and after December 1). Then the integral temperature difference B_τ can be decomposed into two indicators. There are cumulative integral temperature difference inside and outside the building:

$$B_{\tau\ before} = \sum_{j=1}(\underline{t}-t_{\tau j}), \qquad j=1,2,...,L_\tau', \qquad (2)$$

and residual integral temperature difference inside and outside the building:

$$B_{\tau\ after} = B_\tau - B_{\tau\ before}. \qquad (3)$$

Is it possible to predict the indicator of the integral temperature difference for the heating period on the basis of meteorological observations of the beginning and the first half of the heating period? Is there a relationship between the accumulated and residual integral temperature differences inside and outside the building?

Let's check the assumption that if the first half of the heating year was cold (warm), then the probability of a colder (warm) remaining part of the heating period increases.

To test the assumption, a method is used that is based on counting and estimating the frequency of distribution of observations over the quadrants of the coordinate plane.

The sets of observations (accumulated and residual integral temperature differences) for all considered heating periods are divided into two subsets according to two criteria. Firstly, the accumulated integral temperature difference of a given heating period relative to the arithmetic mean value for the entire long-term period is more or less. Secondly, more or less the residual integral temperature difference of the mean long-term value of this indicator. In the unlikely cases of coincidence in terms of the indicator under consideration with its average annual value, this heating period is half taken into account in one subset, and half in the other.

The result is a partition of the set into four subsets. The distributions of heating periods by the accumulated and residual integral temperature differences in Irkutsk are shown in Figure 1. The index of the cumulative temperature difference accumulated by a given date is considered along the abscissa axis, and the residual integral temperature difference along the ordinate.

The resulting quadrants of the division of the coordinate plane can be interpreted as follows:

- I quadrant describes the distribution of heating periods corresponding to the "cold winter" (large values of accumulated and residual integral temperature differences);

- III quadrant describes the distribution of heating periods corresponding to "warm winter" (small values of accumulated and residual integral differences);

- II and IV quadrants describe situations of "asynchronous fuel consumption", until a certain moment winter is cold (warm), then it becomes warm (cold).

on December 1

on January 1
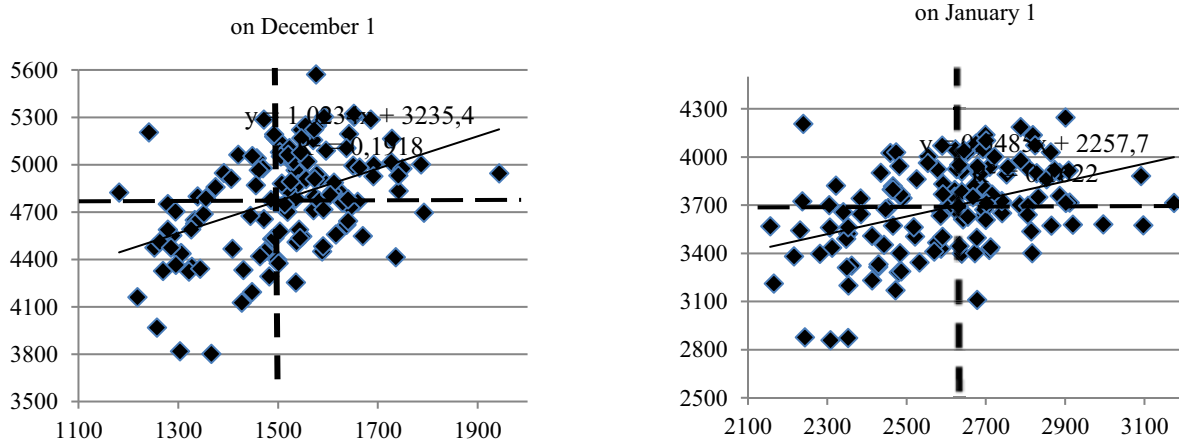
on February 1
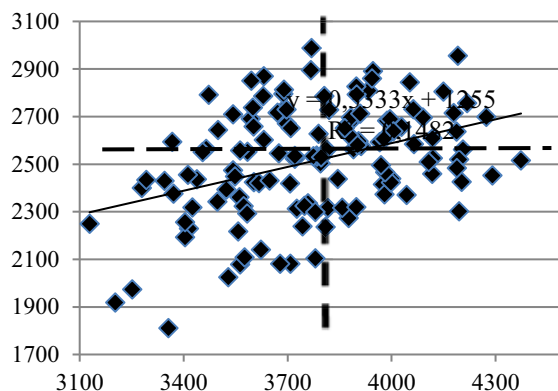
Fig. 1. Distribution of heating periods by accumulated and residual integral temperature differences, Irkutsk

To assess the tightness of the relationship between the accumulated and residual integral temperature differences, an indicator of the synchronicity of the deviations of the integral temperature differences for the past and forthcoming heating periods is introduced:

$$k = (n_1 + n_3) / (n_2 + n_4), \qquad (4)$$

where n_i, $i = 1, 2, 3, 4$ is the frequency of distribution of heating periods along the quadrants of the coordinate plane.

Table 1. Indicator of synchronicity of deviations of accumulated and residual integral temperature differences in cities of Russia for the period 1900-2019.

City	December 1	January 1	February 1
Irkutsk	1,95	2,10	1,60
Novosibirsk	1,08	1,20	1,08
Moscow	1,17	1,41	1,95

The data in Table 1, as in the following, are ranked in descending order of the integral temperature difference inside and outside the building during the heating period.

Table 1 shows the results of calculating the synchronicity index (4) for the selected observation points and three dates of the heating period. For all observation points, the synchronicity index is greater than one.

To identify the relationship between the accumulated and residual integral temperature differences, we will use the second method - the construction of a regression relationship between the values under consideration.

Table 2. The values of the indicators of the tightness of the relationship between the accumulated and residual integral temperature differences for the period 1900-2019.

City	December 1	January 1	February 1
Paired correlation coefficients			
Irkutsk	0,4	0,4	0,4
Novosibirsk	0,3	0,3	0,3
Moscow	0,1	0,2	0,4
Linear regression slope coefficients			
Irkutsk	1,0	0,5	0,3
Novosibirsk	0,7	0,4	0,2
Moscow	0,3	0,4	0,3
Determination coefficients			
Irkutsk	0,2	0,2	0,1
Novosibirsk	0,1	0,1	0,1
Moscow	0,0	0,1	0,1

Table 2 shows the values of indicators of paired correlation, linear regression slope, determination coefficients. The positive values of the slope and correlation coefficients indicate a stable positive statistical relationship between the deviations of the accumulated and residual integral temperature difference inside and outside the building. At the same time, the coefficient of determination shows that this linear relationship cannot be considered significant.

3 Model for predicting the probability of the expected integral temperature difference for the upcoming heating period

The set of values of the accumulated integral temperature difference for the heating period $B_{\tau \, before}$ is divided into n intervals with the same discrete step:

$$\delta = (max \, B_{\tau \, before} - min \, B_{\tau \, before})/n. \qquad (5)$$

Similarly, the residual integral temperature difference is divided into m intervals with the same discrete step.

The vectors of numbers of the partition intervals are introduced for the accumulated integral temperature difference:

$$N_{before} = \{1, 2, ..., i, ..., n\}, \qquad (6)$$

for the residual integral temperature difference:

$$N_{after} = \{1, 2, ..., j, ..., m\}, \qquad (7)$$

Next, a matrix of distribution of the accumulated and residual integral temperature differences over the intervals is formed (to simplify the calculations, $n = m = 3$ is taken).

Table 3. Distribution matrix of accumulated and residual integral temperature differences over intervals $i, j = 1,2,3$.

	Intervals of accumulated integral temperature difference N_{before}		
Intervals of residual integral temperature difference N_{after}	№	1	2	3
	1	γ_{11}	γ_{12}	γ_{13}
	2	γ_{21}	γ_{22}	γ_{23}
	3	γ_{31}	γ_{32}	γ_{33}
$\sum_{j=1} \gamma_{ij}$		γ_1	γ_2	γ_3

Here γ_{ij}, $i, j = 1,2,3$ is the number of observations from the i interval of accumulated integral differences in the j interval of the residual integral temperature differences; γ_i, $i = 1,2,3$ is the total number of observations in the i interval of accumulated integral differences:

$$\gamma_i = \sum_{j=1} \gamma_{ij}, \quad i,j = 1,2,3. \tag{8}$$

To obtain the matrix of conditional probabilities of the "transition" of observation from the i interval of accumulated integral differences to the j interval of residual integral temperature differences, use the formula

$$p_{ij} = \gamma_{ij}/\gamma_i, \quad i,j = 1,2,3. \tag{9}$$

Table 4. Matrix of conditional probabilities of "transition" of observation from the i interval of accumulated integral differences to the j interval of residual integral temperature differences

p_{11}	p_{12}	p_{13}
p_{21}	p_{22}	p_{23}
p_{31}	p_{32}	p_{33}

Where p_{ij}, $i, j = 1,2,3$ is the conditional probability of the observation "transition" from the i interval of accumulated integral differences to the j interval of residual integral temperature differences.

The sum of the conditional probabilities of the matrix over the columns is equal to one. This effect is achieved due to the previously adopted method of obtaining conditional probabilities:

$$\sum_{i=1} p_{ij} = 1, \quad i,j = 1,2,3. \tag{10}$$

For the matrix of conditional probabilities, the following relation is fulfilled (shown in Table 4).

Table 5. Matrix of conditional probabilities and distribution frequencies of cumulative and residual integral temperature differences

p_{11}	p_{12}	p_{13}		$p_{before\ 1}$		$p_{after\ 1}$
p_{21}	p_{22}	p_{23}	×	$p_{before\ 2}$	=	$p_{after\ 2}$
p_{31}	p_{32}	p_{33}		$p_{before\ 3}$		$p_{after\ 3}$

Here $p_{before\ i}$, $i = 1,2,3$ is the distribution frequency of the accumulated integral difference over the partition intervals N_{before}; $p_{after\ j}$, $j = 1,2,3$ is the distribution frequency of the residual integral difference over the partition intervals N_{after}.

The splitting intervals of the accumulated integral temperature difference can be interpreted as follows. Interval 1 characterizes the past part of the heating period

as a relatively "warm winter". Interval 2 is like "average winter", the realized part of the heating year is within the range of average annual indicators. Interval 3 is "cold winter".

The partitioning intervals of the residual integral temperature difference can be interpreted as follows. Interval 1 characterizes the coming part of the heating season as a relatively "warm winter". Interval 2 is like "average winter". And interval 3 is "cold winter".

Consider the matrices of conditional probabilities for the three cities of Irkutsk, Novosibirsk, Moscow by states for December 1, January 1, February 1.

Table 6. Conditional probability matrices for selected cities on December 1, January 1, and February 1

| on December 1			on January 1			on February 1		
Irkutsk								
0,3	0,0	0,0	0,3	0,0	0,0	0,2	0,1	0,0
0,6	**0,6**	**0,6**	**0,5**	**0,6**	**0,5**	**0,7**	**0,5**	**0,5**
0,1	0,4	**0,4**	0,2	**0,4**	**0,5**	0,1	**0,4**	**0,5**
Novosibirsk								
0,4	0,3	0,2	**0,4**	0,2	0,0	0,3	0,2	0,0
0,6	**0,7**	**0,5**	**0,6**	**0,8**	**0,9**	**0,6**	**0,7**	0,3
0,0	0,0	0,3	0,0	0,1	0,1	0,1	0,1	**0,7**
Moscow								
0,5	0,3	0,3	**0,4**	0,2	0,0	**0,4**	0,2	0,0
0,3	**0,5**	**0,6**	**0,5**	0,3	**1,0**	**0,6**	**0,5**	**0,6**
0,2	0,2	0,1	0,2	**0,4**	0,0	0,0	0,3	**0,4**

The conditional probabilities shown in Table 6 can be commented on as follows. Suppose for Irkutsk, based on the temperature data for December 1, we observe a "warm winter", then with a probability of 0.3 the upcoming part of the heating period will be warm, with a probability of 0.6 the rest of the winter is expected within the mean annual values, the winter will be cold with a probability of 0.1.

From Table 6, we see that as the heating period ends, the distribution of observations in the conditional probability matrices tends to the main diagonal. The distributions of observations over Novosibirsk and Moscow are indicative, while "medium winters" and "cold winters" are typical for Irkutsk.

4 Conditional entropy as an estimate of the model forecasting efficiency

Entropy is used to assess the efficiency of predicting the integral temperature difference inside and outside the building for the heating period using conditional probability matrices.

Entropy can be viewed as a "measure of uncertainty" when an event occurs.

on January 1

on December 1

on November 1

I

on October1

II

III

...
...
...
...
...
...
...
...
...
...
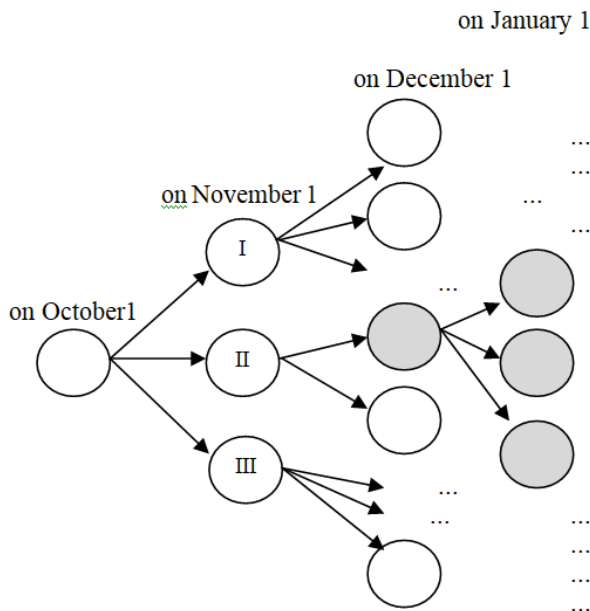
Fig. 2. Hierarchical graph future scenarios

Indeed, choosing a certain scenario for the implementation of the remaining part of the heating period using the matrices of conditional probabilities of the development scenario, we reduce the uncertainty of the future, thereby reducing the value of entropy (Figure 2).

Let's introduce the indicators:

1) the general entropy $H(J)$ is the entropy of the frequencies of the distribution of the residual integral temperature difference over the partition intervals N_{after}:

$$H(J)=-\sum_{j=1}p_{after\,j}\cdot log\,p_{after\,j}\,, \qquad (10)$$

where $p_{after\,j}$ is the distribution frequency of the residual integral difference over the partition intervals N_{after}.

The general entropy $H(J)$ shows the uncertainty of the future in case of refusal to choose the scenario for the implementation of the current winter.

2) The partial entropy $H(J_i)$ is the entropy of the column vector of the matrix of conditional probabilities:

$$H(J_i)=-\sum_{i=1}p_{ij}\cdot log\,p_{ij}\,, \qquad (11)$$

where p_{ij} is the conditional probability of the observation "transition" from the i interval of accumulated integral differences N_{before} to the j interval of residual integral temperature differences N_{after}.

The partial entropy $H(J_i)$ shows the uncertainty associated with the choice of the i-th scenario of future development.

In most cases, the partial entropy is less than the general one, since when choosing a scenario, we have less uncertainty of the future than when refusing to choose (Figure 2).

However, exceptions are possible, when the private entropy is greater than the general one: the future uncertainty arising when choosing a certain development option is greater than uncertainty - without choosing a scenario.

3) The weighted average entropy:

$$weighed\ H(J)=\sum_{i=1}p_{before\,i}\,H(J_i), \qquad (12)$$

where $p_{before\,i}$ is the distribution frequency of the accumulated integral difference over the partition intervals N_{before}; $H(J_i)$ is the private entropy of the i-th scenario of future development.

The weighted average entropy reflects the weighted average future uncertainty with the probabilities of the current winter.

The percentage of uncertainty of the future eliminated by choosing the i-scenario is calculated.

$$\delta H_i=[H(J_i)-H(J)]\,/\,H(J)\cdot100\%,\ i=1,2,3, \qquad (13)$$

where $H(J)$ is the general entropy, $H(J_i)$ is the partial entropy.

Uncertainties of the future, eliminated by predicting the integral temperature difference using conditional probability matrices, are calculated:

$$weighed\ \delta H_i=\sum_{i=1}p_{before\,i}\,\delta H_i\,, \qquad (14)$$

Numerical calculations of the partial entropy, the general entropy and the weighted average entropy for the city of Irkutsk on December 1 are presented in Table 7.

Table 7. Partial entropy, general entropy and weighted average entropy of Irkutsk on December 1

Conditional probability matrix			The distribution frequency	
			of the accumulated integral difference	of the residual integral difference
0,3	0,0	0,0	0,3	0,1
0,6	0,6	0,6	0,7	0,6
0,1	0,4	0,4	0,1	0,3

Partial entropy for each scenario i			Weighted average entropy	General entropy
0,8	0,8	0,7	0,8	0,9

δH_i for each scenario i		
-9%	-10%	-29%

From Table 7 we see that the partial entropy does not exceed the general entropy, i.e. the choice of any scenario for the implementation of the current winter for Irkutsk on December 1 reduces the uncertainty than in the case of refusal to choose the scenario. As expected, the weighted average is also less than the total entropy.

From Table 7 it follows that the choice of the future scenario allows reducing the uncertainty to an average of 16%, in some cases, this effect can be achieved by 30%. Knowledge of possible scenarios for the implementation of the current winter can significantly reduce future uncertainty.

Conclusions

The possibility of predicting the integral temperature difference inside and outside the building for the heating period on the basis of meteorological data for the first months of the heating year has been investigated.

An assumption is proved about the statistical relationship between the accumulated and residual integral temperature differences.

A model has been developed for predicting the probability of expected consumption of heat energy and fuel for heating for the entire heating period according to meteorological data of the beginning of the heating period. The model is focused on the construction of matrices of conditional probabilities of "transition" of observation from the interval of partitioning the accumulated integral temperature differences into the interval of residual integral temperature differences.

The conditional probability matrices allow to estimate the probabilities of the scenarios for the rest of the heating period. Matrices can be used when adjusting fuel supply programs.

Entropy is used to assess the efficiency of predicting the integral temperature difference inside and outside the building for the heating period using conditional probability matrices. On average, the choice of future scenarios can reduce entropy by up to 16%.

References

1. M. Velikanov, A. Nekrasov, A. Reznikovskij, Proceedings of Academy of Sciences. Power Engeneering and Transport, **2**, 151-157 (1988).
2. V. Zorkaltcev. *Long-term temperature variations and impact on the economy and energy* (Geo, Novosibirsk, 2017).
3. V. Zorkaltcev. *Fluctuations in fuel demand for heating in the economic regions of the USSR* (ESI SB RAS, Irkutsk, 1998).
4. V. Zorkaltcev, I. Mokryi, I. Khazheev. Long-term variations in winter temperatures in the regions of Russia and their impact on the reliability of power supply. Energy policy, 4, 81-91 (2017).
5. I. Khazheev. How does climate affect the economy? Eco, 7, 147-162 (2015).
6. I. Khazheev. How do you know when you meet? About predicting winter by the date of its beginning. Eco, 4, 124-133 (2017).
7. I. Khazheev. Research of fluctuations in the demand for energy resources for heating based on long-term meteorological data. System research in power engineering, 45, 144-151 (2015).
8. L. Khrilev. On the influence of the climatic factor on the perspective installation of the fuel and energy balance. Teploenergetika, 2, 16-20 (1966).
9. L. Khrilev. *District heating and fuel and energy complex.* (Nauka, Novosibirsk, 1979).
10. Yu. Rudenko. Methodological issues in the study of the reliability of large power systems. Proceedings of Academy of Sciences of the USSR. Energy and transport, 1, 7-24 (1976).
11. Russian Hydrometeorological Portal All-Russian Research Institute of Hydrometeorological Information. URL: http://aisori.meteo.ru/ClimateR.