Use of the multi-factor regression model for temperature control in structurally isolated rooms of the railway stations

V S Tynchenko¹, I V Markevich¹, A R Ogol¹, O V Baryshnikova¹, D V Rogova¹ and O N Kotova²

¹ Reshetnev Siberian State University of Science and Technology, 31, Krasnoyarsky Rabochy Av., Krasnoyarsk, 660037, Russian Federation
² Kemerovo State University, 6, Krasnaya St., Kemerovo, 650000, Russian Federation

E-mail: vadimond@mail.ru

Abstract. The article is devoted to the analysis and construction of a multivariate regression model. The mode of operation of generalized heat points is considered. The positive and negative aspects of using heat points to maintain the temperature at the desired level have been identified. The analysis of the multivariate model is carried out. The input data for the generated model has been defined. The factors that affect the temperature in structurally isolated rooms have been identified. A multifactorial regression model of temperature control in heated rooms has been built, with the help of which it is possible to determine how much heat needs to be supplied to a particular structurally isolated room, which will allow more efficient use of resources to maintain the temperature at the required level.

1. Introduction

Maintaining the temperature in various rooms at a certain level is one of the most important tasks [1 - 3]. For example, heat enters the apartments after the average daily temperature is kept at +8 and below degrees for five days in a row. But the final decision on the beginning of the heating season is made by the executive authorities [4, 5].

Starting up a central heating system is a gradual process. After opening the valves at CHPPs, boiler houses and other heat sources, heat enters the rooms for several days, and it takes time to obtain the necessary heat supply parameters.

The living quarters use both local and central heating. The heating mode in isolated rooms of the railway stations is most often not chosen by residents. It depends on if the heaters installed by the builders support energy from a common heat source and do not have the ability to be regulated. But, at the same time, the supply of heat to rooms and premises often occurs in strange modes, which forces the owners of apartments and premises to change radiators, install adjusting devices on them and form additional heating circuits.

Based on this, the purpose of this article is to develop a multifactorial model for temperature control in isolated rooms of the railway stations during the heating season, which will lead to a decrease in resource use. To achieve this goal, it is necessary to solve the following tasks:

• To identify factors influencing the temperature in living quarters.
• To develop a multifactor model of temperature control in isolated rooms of the railway stations during the heating season.
2. Materials and methods
In the modern world, in practice, used in construction, central water heating systems are used. A fairly large part of the housing stock is heated by central networks. Central heating provides a more uniform indoor climate and shows great demand and efficiency, despite the active introduction and use of such sources to maintain a favorable microclimate at home [6]. Central heating is most often used in railway stations. The uninterrupted operation and efficiency of centralized networks at home depends on high-quality assembly and proper operation of all constituent elements.

The main difference between central heating and local heating is that there is no pollution by combustion products and fuel, more convenient and reliable control. The delivery of the coolant is carried out through the main pipelines. This is a complex ramified engineering system that provides heat energy to many objects and is dispersed over a large area.

For the modern practice of multi-storey construction, central heating is a necessary element of the sanitary engineering of buildings.

The central heating system has the following advantages:

- Use of inexpensive fuels.
- Regular monitoring of performance and technical condition by special services.
- Use of environmentally friendly equipment.
- Ease of use.
- Ensuring a more uniform thermal regime in the premises during the day.
- Lack of pollution in the premises by combustion products.
- More convenient, reliable and safe management.
- No labor costs by the population for heating the premises.

Among the disadvantages of such a heating scheme for the railway stations, the following should be noted:

- Operation of the system according to a strict seasonal schedule.
- Impossibility of individual regulation of the temperature of heating devices.
- Frequent pressure drops in the system.
- Significant heat loss during transportation and heating.
- High cost of equipment and its installation.

Due to the above disadvantages, it was decided to develop a multifactor model for temperature control in isolated rooms of the railway stations.

A decision was made to conduct a regression analysis in order to establish the degree of influence of independent variables on the dependent variable. To identify dependencies, a generally accepted linear regression equation will be used [7 - 9].

In the regression equation, Y means the variable, the influence of factors on which we are trying to study. In our case, this is the number of buyers. The x-value is the various factors that affect the variable.

To build a model, you must specify the following input data:

- Purpose of the isolated rooms of the railway stations (residential / non-residential).
- Number of storeys.
- Floor height.
- Height of the window.
- Outside air temperature text equal to the average temperature of the coldest five-day period.
- Average temperature of the heating season (tht).
- Duration of the heating season (zht).
- Maximum of the average wind speeds in terms of points for January V.
- Heat carrier source.
- Heat carrier parameters: in the heating network and in the heating system.
- Differential pressure at the subscriber input.
For calculations, we will use the following parameters: railway station, 5 floors, floor height - 3 meters, window height - 1.5 meters, text = -31 ºC, tht = -4 ºC, zht = 221 days, V = 5.5 m / s, heat carrier source - city heating networks, heat carrier parameters: in the heating network - 130-70 ºC, in the heating system - 95-70 ºC, pressure drop at the subscriber input = 60 kPA.

### 3. Experimental study and discussion

The problem under study can be solved using a multivariate regression model, which will allow you to build and test the relationship between one resulting variable and several independent factors that influence it [10, 11].

In general terms, the multivariate regression equation can be written as follows:

\[
Y = f(\overline{X}) = f(\overline{x}) + e, \quad (1)
\]

where \( (\overline{X}) = X_1, X_2, \ldots X_n \) – independent explanatory variables; \( e \) - random variable, an additional residual term that reflects the influence of random errors, measurement features and actions that affect the resulting variable, and other explanatory variables not included in the equation. It is also called resentment or remnant. This random component can be considered a random forecast error \( Y \) for a given value \( X \).

**Table 1. Data for regression.**

| Apartment | X0  | X1  | X2  | X3  | X4  | Y   |
|-----------|-----|-----|-----|-----|-----|-----|
| 101       | 1020| 20  | 22  | 18.91| 1   | 1834|
| 102       | 490 | 20  | 20  | 20.13| 1   | 1316|
| 103       | 430 | 10  | 20  | 16.47| 1   | 1106|
| 104       | 470 | 20  | 20  | 18.3 | 1   | 1221|
| 105       | 490 | 20  | 20  | 20.13| 1   | 1316|
| 106       | 490 | 20  | 20  | 20.13| 1   | 1316|
| 107       | 470 | 20  | 20  | 18.3 | 1   | 1221|
| 108       | 430 | 10  | 20  | 16.47| 1   | 1106|
| 109       | 490 | 20  | 20  | 20.13| 1   | 1316|
| 110       | 1070| 70  | 20  | 18.91| 1   | 1884|
| 111       | 1070| 70  | 20  | 18.91| 1   | 1884|
| 112       | 490 | 20  | 20  | 20.13| 1   | 1316|
| 113       | 430 | 10  | 20  | 16.47| 1   | 1106|
| 114       | 470 | 20  | 20  | 18.3 | 1   | 1221|
| 115       | 470 | 20  | 20  | 18.3 | 1   | 1221|
| 116       | 490 | 20  | 20  | 20.13| 1   | 1316|
| 117       | 490 | 20  | 20  | 20.13| 1   | 1316|
| 118       | 470 | 20  | 20  | 18.3 | 1   | 1221|
| 119       | 470 | 20  | 20  | 18.3 | 1   | 1221|
| 120       | 430 | 10  | 20  | 16.47| 1   | 1106|
| 121       | 490 | 20  | 20  | 20.13| 1   | 1316|
| 122       | 1020| 20  | 20  | 18.91| 1   | 1834|
| 201       | 880 | 20  | 20  | 18.91| 2   | 1694|
| 202       | 350 | 20  | 20  | 20.13| 2   | 1176|
| 203       | 310 | 10  | 20  | 16.47| 2   | 986 |
| 204       | 350 | 20  | 20  | 18.3 | 2   | 1101|
| 205       | 340 | 20  | 20  | 20.13| 2   | 1166|
| 206       | 350 | 20  | 20  | 20.13| 2   | 1176|
| 207       | 350 | 20  | 20  | 18.3 | 2   | 1101|
| 208       | 310 | 10  | 20  | 16.47| 2   | 986 |
| 209       | 350 | 20  | 20  | 20.13| 2   | 1176|
| 210       | 930 | 70  | 20  | 18.91| 2   | 1744|
| 211       | 620 | 70  | 20  | 18.91| 2   | 1434|
| 212       | 350 | 20  | 20  | 20.13| 2   | 1176|
The value of the required load on heating devices in each apartment is taken as the dependent variable $Y$.

Explanatory variables:
- $X_0$ is the value of the main heat losses.
- $X_1$ - value of additional heat losses.
- $X_2$ - required temperature in isolated rooms.
- $X_3$ - room area.
- $X_4$ - floor.

Table 1 presents some of the input data for constructing the regression.
Tables 2 - 4 show the performed regression calculations.

### Table 2. Regression statistics.

|                |            |            |            |            |            |
|----------------|------------|------------|------------|------------|------------|
| Multiple R     | 1.00       | R- square  | 0.999      |            |            |
| Normalized R-  |            | square     | 1.00       |            |            |
| Standard error | 5.88       |            |            |            |            |
| Observations   | 110        |            |            |            |            |

### Table 3. Analysis of variance.

|            | df | SS        | MS        | $F$            | Significance $F$ |
|------------|----|-----------|-----------|----------------|------------------|
| Regression | 5  | 7087520,10| 1417504,02| 40983,8012     | 1.3725E-169      |
| Balance    | 104| 3597,04   | 34,58693384|               |                  |
| Total      | 109| 7091117,14|           |                |                  |

### Table 4. Regression coefficients.

|                | Odds  | Standard error | t- statistics | P- Value | Lower 95% | Upper 95% |
|----------------|-------|----------------|---------------|----------|-----------|-----------|
| $Y$- crossing  | -50.07| 47.34          | -1.06         | 0.29     | -143.95   | 43.81     |
| Variable $X_1$ | 1.05  | 0.00           | 299.27        | 0.00     | 1.04      | 1.05      |
| Variable $X_2$ | 1.31  | 0.05           | 27.70         | 0.00     | 1.22      | 1.40      |
| Variable $X_3$ | 2.55  | 2.36           | 1.08          | 0.28     | -2.13     | 7.23      |
| Variable $X_4$ | 39.90 | 0.44           | 89.81         | 0.00     | 39.02     | 40.78     |
| Variable $X_5$ | -0.65 | 0.40           | -1.62         | 0.11     | -1.44     | 0.14      |

One of the main indicators of regression is the R-squared, which indicates the quality of the model [12]. In our case, this ratio is 0.999 or about 99%. This is an acceptable level of quality. Dependency less than 0.5 is bad.

The $Y$-intersection shows what will be $Y$ if all factors in the model are equal to 0. In our case, this value is -50.07.
- 1.05 - coefficient that shows the importance of the influence of the factor $X_0$ on $Y$, that is, the main heat losses within this model affect the heat load with a weight of only 1.05.
- 1.31 - coefficient that shows that additional heat loss within the given model affects the heat load with a weight of 1.31.
- 2.55 - coefficient that shows that the required temperature in the apartment within the given model affects the heat load with a weight of 2.55.
- 39.9 is a coefficient that shows that the area of an apartment within a given model affects the heat load with a weight of 39.9, i.e. is the most significant factor in the model.
- -0.65 is a coefficient that shows that a floor within a given model affects the heat load with a weight of -0.65. A negative value of the coefficient means that an inverse relationship is established between the factor and the dependent variable, i.e., the higher the floor, the lower the value of the heat load on electrical appliances.

Using these coefficients, you can build a regression model and it will look like this:
\[ Y = 1.05 \times X_0 + 1.31 \times X_1 + 2.55 \times X_2 + 39.9 \times X_3 - 0.65 \times X_4 - 50.07. \] (2)

Using a regression model, you can make predictions for this, you need to substitute the resulting coefficients instead of X into the model.

Table 5 shows a fragment of the forecast calculation.

| Apartment | X0  | X1  | X2  | X3  | X4  | Y    |
|-----------|-----|-----|-----|-----|-----|------|
| 101       | 1020| 20  | 22  | 18.91| 1   | 1832 |
| 102       | 490 | 20  | 20  | 20.13| 1   | 1321 |
| 103       | 430 | 10  | 20  | 16.47| 1   | 1110 |
| 104       | 470 | 20  | 20  | 18.3 | 1   | 1228 |
| 105       | 490 | 20  | 20  | 20.13| 1   | 1321 |
| 106       | 490 | 20  | 20  | 20.13| 1   | 1321 |
| 107       | 470 | 20  | 20  | 18.3 | 1   | 1228 |
| 108       | 430 | 10  | 20  | 16.47| 1   | 1110 |
| 109       | 490 | 20  | 20  | 20.13| 1   | 1321 |
| 110       | 1070| 70  | 20  | 18.91| 1   | 1893 |
| 111       | 1070| 70  | 20  | 18.91| 1   | 1893 |
| 112       | 490 | 20  | 20  | 20.13| 1   | 1321 |
| 113       | 430 | 10  | 20  | 16.47| 1   | 1110 |
| 114       | 470 | 20  | 20  | 18.3 | 1   | 1228 |
| 115       | 470 | 20  | 20  | 18.3 | 1   | 1228 |
| 116       | 490 | 20  | 20  | 20.13| 1   | 1321 |
| 117       | 490 | 20  | 20  | 20.13| 1   | 1321 |
| 118       | 470 | 20  | 20  | 18.3 | 1   | 1228 |
| 119       | 470 | 20  | 20  | 18.3 | 1   | 1228 |
| 120       | 430 | 10  | 20  | 16.47| 1   | 1110 |
| 121       | 490 | 20  | 20  | 20.13| 1   | 1321 |
| 122       | 1020| 20  | 20  | 18.91| 1   | 1827 |
| 201       | 880 | 20  | 20  | 18.91| 2   | 1680 |
| 202       | 350 | 20  | 20  | 20.13| 2   | 1174 |

Looking at the table, you can see that the discrepancy between the actual data and the forecast is not significant.

4. Conclusion
As a result of this work, the positive and negative aspects of using heat points to maintain the temperature at the desired level have been identified. The analysis of the multivariate model was carried out, while the input data for the generated model has been defined. The factors that affect the temperature in structurally isolated rooms have been identified.

A multifactorial regression model of temperature control in heated rooms has been built, with the help of which it is possible to determine how much heat needs to be supplied to a particular structurally isolated room. It allows one to use the resources more efficiently to maintain the temperature in isolated rooms of the railway station at the required level.

References
[1] Ferdyn-Grygierek J 2014 Indoor environment quality in the museum building and its effect on heating and cooling demand Energy and Buildings 85 32-44
[2] Sodha M S, Sharma A K, Singh S P, Bansal N K and Kumar A 1985 Evaluation of an earth—air tunnel system for cooling/heating of a hospital complex Building and Environment 20(2) 115-22
[3] Stryjakowska-Sekulska M, Piotraszewska-Pajak A, Szyszka A, Nowicki M and Filipiak M 2007 Microbiological quality of indoor air in university rooms Polish Journal of Environmental Studies 16(4) 623

[4] Wan G H, Chung F F and Tang C S 2011 Long-term surveillance of air quality in medical center operating rooms American journal of infection control 39(4) 302-8

[5] Ng S F, Oo C S, Loh K H, Lim P Y, Chan Y H and Ong B C 2003 A comparative study of three warming interventions to determine the most effective in maintaining perioperative normothermia Anesthesia & Analgesia 96(1) 171-6

[6] Moriske H J, Drews M, Ebert G, Menk G, Scheller C, Schöndube M and Konieczny L 1996 Indoor air pollution by different heating systems: coal burning, open fireplace and central heating Toxicology letters 88(1-3) 349-54

[7] Bukhtoyarov V V, Tynchenko V S, Petrovsky E A, Doksharin S G and Kukartsev V V 2019 Research of methods for design of regression models of oil and gas refinery technological units IOP Conference Series: Materials Science and Engineering 537(4) 042078

[8] Chamberlain G 1982 Multivariate regression models for panel data Journal of econometrics 18(1) 5-46

[9] Boyko A A, Kukartsev V V, Tynchenko V S, Korpacheva L N, Dzhioeva N N, Rozhkova A V and Aponasenko S V 2020 Using linear regression with the least squares method to determine the parameters of the Solow model Journal of Physics: Conference Series 1582(1) 012016

[10] Eerola T, Lartillot O and Toiviainen P 2009 Prediction of Multidimensional Emotional Ratings in Music from Audio Using Multivariate Regression Models In: ISMIR pp 621-6

[11] Azen R and Budescu D V 2006 Comparing predictors in multivariate regression models: An extension of dominance analysis Journal of Educational and Behavioral Statistics 31(2) 157-80

[12] Santiago C B, Guo J Y and Sigman M S 2018 Predictive and mechanistic multivariate linear regression models for reaction development Chemical science 9(9) 2398-412