Experimental Study of Influence Factor Based on Principal Component Analysis to Frozen Earth Thaw Settlement

Jiaqi Yuan*, Wei Wang* and Weiyan Xia*

School of Prospecting and Surveying Engineering, Changchun Institute Of Technology, Changchun, China

*Corresponding author e-mail: 1012786344@qq.com, *16695496@qq.com, b1587748216@qq.com

Abstract. Thaw settlement property of seasonal frozen earth has great effect to buildings; the grading of thaw settlement can accurately reflect the thaw settlement grades of frozen earth and judge the thaw settlement character of the frozen earth. Taking silty clay and sand of Changchun Area as research object, the article conducts experimental research to moisture content, dry density, plasticity index and initial temperature of freezing; the article adopts principal component analysis method to conduct analysis to the experiment and gets the weighing of all influence factors to thaw settlement property, ie., by calculating the comprehensive scores of all influence factors, we get its incidence to thaw settlement. The result shows that: the primary and secondary relations that affect thaw settlement property factor of silty clay are: moisture content>initial temperature of freezing>plasticity index>dry density; the primary and secondary relations that affect thaw settlement property factor of sand are: moisture content>initial temperature of freezing>silt content>dry density; the comprehensive score ranking got by principal component analysis method can reflect the ranking of thaw settlement property grading to some extent, which provides theoretical foundation for the further deepening of recognition for thaw settlement property.

1. Introduction

The distribution of China frozen earth is very wide; seasonal frozen earth areas and permafrost regions covers the 70% of China’s land total area. However characters like frost heave and thaw settlement property, etc of seasonal frozen earth have vast influence on engineering; conducting engineering design and construction in seasonal frozen earth areas should pay special attention to the influence of seasonal frozen earth to engineering projects as well as buildings and make corresponding precautionary measures.

Many domestic scholars conducts vast researches to thaw settlement property of frozen earth. Xifa Zhang [1], etc have established regression equation between thaw settlement coefficient and moisture content as well as dry bulk density; Chensong Yang [2], etc have found that it will generate thaw settlement when the soil body exceeds initial thaw settlement moisture content; Xiangdong Zhang [3], etc have found thaw settlement value reduces with the increasing of loading for frozen heave; Ping Yang [4], etc have found effect of dry density, moisture content and freezing temperature, etc influence factors to thaw settlement coefficient and have established prediction model of thaw settlement coefficient by...
utilizing BP neural network; Zhihui Zhao [5], etc have simulated four kinds of earth samples to research the effect of soil sample, loading, soak-age as well as temperature, etc factors to frost-heave capacity and thaw settlement; Taking silty clay as in Lanzhou area as example, Yonghu Zhao [6], etc have analyzed the effect of moisture content, density, temperature of cold end and outer loading 4 factors to sensibility of frozen silty clay thaw settlement coefficient by adopting gray correlation degree; Junan Bao [7], etc have established model of gray correlation analysis to subcutaneous layer silt clay in Nanjing area and have analyzed the incidence of all factors to frost heave coefficient and thaw settlement coefficient.

The above research mainly concentrates on the analysis of thaw settlement influence factor as well as thaw settlement capacity prediction aspects of the frozen earth; the research to sensibility of thaw settlement influence factor is scarce to see and related research achievements therefore are scarcer. So the article has conducted experimental research to thaw settlement influence factor sensibility of soil by taking silty clay and sand of Changchun area as research object. By adopting principal component analysis method, the article conducts analysis to all factors that affect thaw settlement property of soil to confirm the primary and secondary relations of frozen earth thaw settlement influence factors, which provides scientific basis and theoretical reference for the better research of frozen earth thaw settlement property as well as its relation between influence factors.

2. Principal Component Analysis

2.1. Fundamental Principle
The principal component analysis utilizes dimensionality reduction idea and transforms multiple indexes into several irrelevant comprehensive indexes on the premise of losing little information, i.e., every principal component is the linear combination of original variable and all principal components are mutually irrelevant, which makes original variable of primary components possess some more excellent properties and must reserve more than 90% information of original variable, which therefore reaches the simplification of system structure and reaches the purpose of grasping problem essential. It is a kind of statistical analysis technique to find out the characteristic quantity of its variance and co-variance matrix in one group variables then to transform multiple variables into a few comprehensive variables by dimensionality reduction.

2.2. Calculation Steps
Taking sensibility of thaw settlement as example, assuming there are i groups data a j PCs influence factors, we can get data matrix:

\[
X = \begin{bmatrix}
X_{i1} & X_{i2} & \cdots & X_{ij} \\
X_{21} & X_{22} & \cdots & X_{2j} \\
\vdots & \vdots & \ddots & \vdots \\
X_{ni} & X_{ni} & \cdots & X_{nj}
\end{bmatrix}
\] (1)

In formula 1: expresses the jth factors of the i group data. Conduct the Z-SCORE standard treatment according to formula 2:

\[
ZX = \frac{X - \mu}{\delta}
\] (2)

In the formula: \(X\)-Original data; \(\mu\)-Overall average value; \(\delta\)-Overall standard deviation. We can obtain standard data matrix:

\[
ZX = \begin{bmatrix}
ZX_{i1} & ZX_{i2} & \cdots & ZX_{ij} \\
ZX_{21} & ZX_{22} & \cdots & ZX_{2j} \\
\vdots & \vdots & \ddots & \vdots \\
ZX_{ni} & ZX_{ni} & \cdots & ZX_{nj}
\end{bmatrix}
\] (3)

Confirm related coefficient matrix R:
Determine characteristic value, finally we can get corresponding characteristic vector. According to characteristic equation of related coefficient matrix $R$: $|R-\lambda E|=0$.

We get characteristic value ($\lambda_1, \lambda_2, \ldots, \lambda_m$) as well as the corresponding characteristic vector ($\mu_1, \mu_2, \ldots, \mu_m$). When the comparable bigger characteristic value of the front $m$, $\lambda_1>\lambda_2>\ldots>\lambda_m>1$, it is the corresponding variance of the front $m$ PCs principal components; the corresponding unit characteristic $\mu_m$ of $\lambda_m$ is the coefficient about original variable of principal component $F_m$, then the $m$th principal component $F_m$ of the original variable is

$$F_m = \mu_m X \tag{5}$$

Variance rate of contribution of principal components is calculated according to formula 6, which is used for reflecting size of information amount, $a_i$ is:

$$a_i = \frac{\lambda_i}{\sum_{i=1}^{m} \lambda_i} \tag{6}$$

The finally chosen principal component quantity, ie., $F_1, F_2, \ldots, F_m$, the confirming of $m$ in $F_m$ is confirmed by the variance accumulated contribution rate $G(m)$ in formula 7:

$$G(m) = \frac{\sum_{i=1}^{m} \lambda_i}{\sum_{k=1}^{m} \lambda_k} \tag{7}$$

When the accumulated contribution rate is bigger than 70%, it is thought to be able to reflect the information of original variable enough; the corresponding $m$ is the extracted front $m$th principal components.

Principal component coefficient $\eta_i$ is to reflect the correlative degree between principal component $F_i$ and original variable, which is calculated according to formula 8:

$$\eta_{ij} = \sqrt{\lambda_i Y_{ij}} \tag{8}$$

In the formula: $\lambda_i$ - Characteristic value of component $I$; $Y_{ij}$ - the loading coefficient of $j$th influence factor in component $I$; Finally calculate comprehensive scores of component $i$ according to formula 9:

$$F_i = \eta_{i1} X_1 + \eta_{i2} X_2 + \ldots + \eta_{ij} X_j \tag{9}$$

3. Experiment of Frozen Earth Thaw settlement

3.1. General Situation of the Experiment

This experiment of thaw settlement adopts YDRS type frozen earth melting compression test to conduct the experiment; the experiment instrument is mainly made up of bigeminy counter-force, sample canister, air pressure control cabinet, air pump, industrial personal computer, data collection system of melting
and compression, controller of melting and compression, circulating water flume of low temperature and constant temperature, etc.; the experiment adopts co-cycle circulatory bath to simulate the heat transmission for the upper part of the soil body, which is shown as figure 1. Put the testing sample with inner diameter: $\phi 79.8$mm and height: 40.0mm into the canister. In order to simulate the soil sample in frozen status, the article adopts low temperature and constant temperature groove and selects liquid medium of absolute ethyl alcohol to conduct refrigeration. The article add the loading of 0–200Pa to every soil sample, which therefore provides experimental data for principal component analysis method.

Figure 1. YDRS type froze earth melting compression experimental instrument.

Thaw settlement coefficient $a_0$ of the soil is calculated according to formula 10; melting compression coefficient $a$ is calculated according to formula 11:

$$a_0 = \frac{\Delta h_0}{h_0}$$ (10)

$$a = \frac{e_i - e_{i+1}}{P_{i+1} - P_i}$$ (11)

In this formula: $a_0$—frozen earth thaw settlement coefficient, %; $\Delta h_0$—deflection of frozen earth melting, mm; $h_0$—initial height of frozen earth testing sample, mm; $a$—melting and compression coefficient within one pressure range, MPa$^{-1}$; $P_{i+1}$, $P_i$—grading pressure value, kPa; $e_i$, $e_{i+1}$—The corresponding void ratio with grading pressure.

3.2. Analysis of Experimental Result

3.2.1. Silty Clay. Conduct regression analysis of thaw settlement coefficient $a_0$ got by the experiment respectively together with moisture content $w$, dry density $\gamma_d$, plasticity index $I_p$ as well as initial frozen temperature $T_0$ and draw regression curve, which is shown as figure 2–5.

Figure 2. Relation curve between silty clay thaw settlement coefficient and moisture content.
From the data regression analysis we can know that: thaw settlement coefficient of silty clay presents linear variation relation with moisture content, dry density, plasticity index as well as initial frozen temperature. It increases with the increasing of moisture content; it reduces with the increase of dry density; it increases with the increase of initial frozen temperature; it reduces with the increase of plasticity index.

**Table 1.** Grading of silty clay thaw settlement property.

| Silty Clay Serial number | Thaw settlement coefficient $a_0$ | Grading of thaw settlement |
|--------------------------|----------------------------------|-----------------------------|
| 1                        | 4.6                              | 2                           |
| 2                        | 0.8                              | 3                           |
| 3                        | 0.8                              | 3                           |
| 4                        | 9.8                              | 3                           |
| 5                        | 23.6                             | 4                           |
3.2.2. Sand. Respectively conduct regression analysis to the thaw settlement coefficient got by the experiment together with moisture content $w$, dry density $\gamma_d$, initial frozen temperature $T_0$ and silt content $C$ and draw regression curve, which is show as figure 6~9.

![Figure 6](image1.png)  
**Figure 6.** Relation curve between sand thaw settlement coefficient and moisture content.

![Figure 7](image2.png)  
**Figure 7.** Relation curve between sand thaw settlement coefficient and dry density.

![Figure 8](image3.png)  
**Figure 8.** Relation curve between sand thaw settlement coefficient and initial frozen temperature.

![Figure 9](image4.png)  
**Figure 9.** Relation curve between sand thaw settlement coefficient and silt content.
Table 2. Grading of sand thaw settlement property.

| Sand Serial number | Thaw settlement coefficient $a_0$ | Grading of thaw settlement property |
|--------------------|-----------------------------------|------------------------------------|
| 1                  | 28.6                              | 4                                  |
| 2                  | 8.6                               | 3                                  |
| 3                  | 30.7                              | 4                                  |
| 4                  | 2.3                               | 4                                  |
| 5                  | 2.9                               | 3                                  |
| 6                  | 22                                | 4                                  |
| 7                  | 9                                 | 4                                  |
| 8                  | 15.6                              | 4                                  |
| 9                  | 13.4                              | 4                                  |

From the data regression analysis we can know that: thaw settlement coefficient of sand presents linear variation relation with moisture content, dry density, initial frozen temperature as well as silt content. It increases with the increasing of moisture content; it reduces with the increase of dry density; it increases with the increase of initial frozen temperature; it increases with the increase of silt content.

4. Analysis for Sensibility of Thaw Settlement Influence Factor

4.1. Analysis for Sensibility of Silty Clay Thaw Settlement Influence Factor

Table 3. Standard data of silty clay.

| Z moisture content w | Z dry density $\gamma_d$ | Z plasticity index $I_p$ | Z initial frozen temperature $T_0$ |
|----------------------|-------------------------|--------------------------|-----------------------------------|
| -0.72                | 1.04                    | 1.69                     | 0.04                              |
| -0.06                | 0.69                    | -0.37                    | -0.02                             |
| -0.50                | 0.11                    | -0.37                    | -1.30                             |
| -0.45                | -0.31                   | -0.02                    | -0.22                             |
| 1.74                 | -1.53                   | -0.94                    | 1.51                              |

Table 4. Common factor variance.

|                          | Initial | Extract |
|--------------------------|---------|---------|
| Zscore(moisture content w) | 1.000   | 0.951   |
| Zscore(dry density $\gamma_d$) | 1.000   | 0.855   |
| Zscore(plasticity index $I_p$) | 1.000   | 0.564   |
| Zscore(initial frozen temperature $T_0$) | 1.000   | 0.598   |

Table 5. Explanation for total variance.

| Component | Initial characteristic value | Extract loading quadratic sum |
|-----------|------------------------------|--------------------------------|
|           | Total | Variance percentage | Accumulation% | Total | Variance percentage | Accumulation% |
| 1         | 2.968 | 74.197              | 74.197        | 2.968 | 74.197              | 74.197        |
| 2         | 0.822 | 20.541              | 94.738        |
| 3         | 0.179 | 4.464               | 99.203        |
| 4         | 0.032 | 0.797               | 100.000       |

Table 6. Component Matrix$^a$.

|                          | Component 1 |
|--------------------------|-------------|
| Zscore(moisture content w) | 0.975       |
| Zscore(dry density $\gamma_d$) | -0.925     |
| Zscore(plasticity index $I_p$) | -0.751     |
| Zscore(initial frozen temperature $T_0$) | 0.773     |
Conduct standard calculation to the data got by the experiment according to formula 2 and conduct dimension reduction process to it, as is shown by table 3. Get related coefficient matrix $R$ to the sample after treatment via SPSS:

$$
R = \begin{pmatrix}
1.000 & -0.841 & -0.657 & 0.837 \\
-0.841 & 1.000 & 0.725 & -0.588 \\
-0.657 & 0.725 & 1.000 & -0.216 \\
0.837 & -0.588 & -0.216 & 1.000
\end{pmatrix}
$$

From table 4 we can see that there are four factors that extract most information from moisture content, however the information that the plasticity index loses is the most. From table 5 we can know that among the four factors component 1 can explain 74.197% variance therefore we put forward to regard component 1 as the main component.

The value in table 6 is the correlative coefficient between common factor and original variable; the more the absolute value is, it means the relation is more intimate. The common factor has negative correlation with moisture content and initial frozen temperature and has positive correlation with dry density and plasticity index; therefore it is thought that common factor 1 can be regarded as comprehensive factor. The value in table 6 is loading coefficient and transform the loading coefficient as principal coefficient according to formula 8, as is shown by table 7. Finally, conduct comprehensive score calculation by utilizing the formula in table 6; the comprehensive score achieved by component 1 is shown as table 8:

**Table 7. Coefficient of principal component.**

| Zscore(moisture content $w$) | Zscore(dry density $\gamma_d$) | Zscore(plasticity index $I_p$) | Zscore(initial frozen temperature $T_0$) |
|-----------------------------|-------------------------------|-------------------------------|----------------------------------|
| 0.57                        | -0.54                         | -0.44                         | 0.45                             |

$$F_1 = 0.57 \times \text{Zscore (moisture content w)} - 0.54 \times \text{Zscore (dry density $\gamma_d$)} - 0.44 \times \text{Zscore (plasticity index $I_p$)} + 0.45 \times \text{Zscore (initial frozen temperature $T_0$)}$$

**Table 8. Comprehensive score of component 1.**

| $F_1$  | Grading of thaw settlement property |
|-------|-------------------------------------|
| -1.70 | 2                                   |
| -0.26 | 3                                   |
| -0.77 | 3                                   |
| -0.18 | 3                                   |
| 2.91  | 4                                   |

By the comprehensive analysis of scores for five group’s data and grading of thaw settlement we can know that: the component 1 got by principal component analysis method can conduct grading to thaw settlement property. Among which the primary and secondary relation that affects silty clay thaw settlement property factor is: moisture content $>$ initial frozen temperature $>$ plasticity index $>$ dry density.
4.2. Analysis for sensitivity of sand thaw settlement influence factor

Table 9. Standard data of sand.

| Z moisture content w | Z dry density $\gamma_d$ | Z initial frozen temperature $T_0$ | Z silt content C |
|----------------------|-------------------------|-----------------------------------|-----------------|
| 1.86                 | -1.39                   | 0.74                              | 1.23            |
| -1.32                | 0.10                    | -0.29                             | 1.77            |
| 0.61                 | -1.06                   | -0.27                             | 0.32            |
| -0.22                | -0.07                   | -0.22                             | -0.82           |
| -1.30                | 1.51                    | -0.34                             | 0.39            |
| 0.45                 | -0.73                   | 1.53                              | -1.21           |
| 0.51                 | -0.24                   | 0.68                              | -0.72           |
| -0.44                | 0.58                    | -2.07                             | -0.50           |
| -0.15                | 1.30                    | 0.23                              | -0.45           |

Table 10. Variance of common factor.

| Component | Initial characteristic value | Extract quadratic sum of the loading |
|-----------|------------------------------|--------------------------------------|
|           | Total                        | Variance percentage | Accumulation% | Total          | Variance percentage | Accumulation% |
| 1         | 2.156                        | 53.897                 | 53.897        | 2.156          | 53.897                 | 53.897        |
| 2         | 1.068                        | 26.709                 | 80.606        | 1.068          | 26.709                 | 80.606        |
| 3         | 0.587                        | 14.666                 | 95.272        | 1.068          | 26.709                 | 80.606        |
| 4         | 0.189                        | 4.728                  | 100.000       | 1.068          | 26.709                 | 80.606        |

Table 12. Component Matrix $^a$.

| Component | Component 1 | Component 2 |
|-----------|-------------|-------------|
| Zscore(moisture content w) | 0.915       | -0.010      |
| Zscore(dry density $\gamma_d$) | -0.882      | -0.273      |
| Zscore(initial frozen temperature $T_0$) | 0.732       | -0.228      |
| Zscore(silt content C) | -0.067      | 0.971       |

According to the previous method, we get related coefficient matrix $R$ to the sample data after treatment with the aid of SPSS:

$$R = \begin{pmatrix}
1.000 & -0.777 & 0.498 & -0.102 \\
-0.777 & 1.000 & -0.433 & -0.129 \\
0.498 & -0.433 & 1.000 & -0.132 \\
-0.102 & -0.129 & -0.132 & 1.000
\end{pmatrix}$$

Among which, from table 10 we can see that the information extracted from silt content is the most; however the information that the initial frozen temperature loses is the most. From table 11 we can know that among the four factors, component 1 can explain 53.897% variance; however accumulation of component 1 and component 2 can explain 80.606% variance, therefore we put forward to regard component 1 and component 2 as the main components.
The common factor 1 in table 12 has the positive relation with dry density and silt content; it has negative relation with moisture content and initial frozen temperature; it has negative relation with silt content; therefore it is thought that common factor 1 can be regarded as the comprehensive factor; however common factor 2 can be regarded as intensity factor. Principal component coefficient is calculated according to previous method and is reflected by table 13. Finally, conduct comprehensive score calculation by utilizing the formula in table 12 and we can get nine groups data scores, as is show as below table:

**Table 13. Coefficient of principal component.**

|                      | Coefficient of principal component 1 | Coefficient of principal component 2 |
|----------------------|--------------------------------------|--------------------------------------|
| Zscore (moisture content w) | 0.62                                 | -0.01                                |
| Zscore (dry density γd)       | -0.60                                | -0.26                                |
| Zscore (initial frozen temperature T0) | 0.50                                  | -0.22                                |
| Zscore (silt content C)       | -0.05                                | 0.94                                 |

F1=0.62×Zscore (moisture content w)+0.60×Zscore (dry density γd)+0.50×Zscore (initial frozen temperature T0)-0.05×Zscore (silt content C)

F2=-0.01×Zscore (moisture content w)-0.26×Zscore (dry density γd)-0.22×Zscore (initial frozen temperature T0)+0.94×Zscore (silt content C)

**Table 14. Comprehensive score of competent 1 and component 2.**

| F1  | F2  | Grading of thaw settlement property |
|-----|-----|-------------------------------------|
| 2.29| 1.33| 4                                   |
| -1.11| 1.71| 3                                   |
| 0.86| 0.63| 4                                   |
| -0.16| -0.71| 4                                  |
| -1.90| 0.06| 3                                   |
| 1.55| -1.29| 4                                  |
| 0.84| -0.77| 4                                  |
| -1.63| -0.17| 4                                  |
| -0.74| -0.81| 4                                  |

By the comprehensive analysis between scores of nine groups data and grading of thaw settlement property, we can know that the comprehensive factor score of component 1 got by principal component analysis method can conduct grading to thaw settlement; however as the intensity factor, component 2 can not conduct explanation to thaw settlement property. Among which, primary and secondary relation that affects sand thaw settlement property factor is moisture content>initial frozen temperature>silt content>dry density.

5. Conclusion

By the experimental research analysis for thaw settlement influence factor of silty clay and sand in Changchun area and by adopting principal component analysis method to analyze the sensitivity of all influence factors to thaw settlement, we get main conclusions as below:

1. The principal component analysis method utilizes idea of dimensionality reduction to transform multiple indexes into several irrelevant comprehensive indexes on the premise of losing very little information. The article utilizes principal component analysis method to transform the influence factor dimensionality reduction into the comprehensive score of component 1, ie., comprehensive factor.

2. Thaw settlement coefficient of silty clay all presents linear variation relation with moisture content, dry density, plasticity index and initial frozen temperature, which increases with the increase of moisture content; it reduces with the increase of dry density; it reduces with the increase of plasticity index; thaw settlement coefficient of sand all presents linear variation relation with moisture content,
dry density, initial frozen temperature and silty content; it increases with the increase of moisture content; it reduces with the increase of dry density and it increases with the increase of initial frozen temperature; it increases with the increase of silt content.

(3) The primary and secondary relation that affects silty clay thaw settlement property factor is moisture content>initial frozen temperature>plasticity index>dry density; the primary and secondary relation that affects sand thaw settlement property factor is moisture content>initial frozen content>silt content>dry density; The comprehensive factor score ranking got by principal component analysis method can reflect the ranking of thaw settlement property grading to some extent, which provides theoretical basis for the further deepening of recognition to thaw settlement property.

Acknowledgements
This work was financially supported by Students’ innovation and entrepreneurship training program of Jilin province (201911437039).

References
[1] Xifa Zhang, Ji Chen, Dongqing Zhang. Application of thawing settlement coefficient to the research on the roadbed frost damage of freeway in seasonal frost region [J]. Journal of Glaciology and Geocryology, 2002 (5): 634-638
[2] Chensong Yang, Ping He, Guodong Cheng, et al. Testing study on influence of freezing and thawing on dry density and water content of soil [J]. Chinese Journal of Rock Mechanics and Engineering, 2003 (S2): 2695-2699
[3] Xiangdong Zhang, Shuang Zhang, Fu Yi. Experimental study on freeze-thaw characteristics of aeolian soil in western Liaoning province [J]. Rock and Soil Mechanics, 2005, 26 (S2): 79-82
[4] Yang P, Ke J, Wang J G, et al. Numerical simulation of frost heave with coupled water freezing, temperature and stress fields in tunnel excavation [J]. Computer and Geotechnics, 2006, 33 (6-7): 330-340
[5] Zhihui Zhao. Shuangguo Yang. Ruiqin Hao. Shaoqin Niu. Huanhuan Yang. Xing Pang. Simulating Research to Frost Heave Thaw Settlement Feature Experiment and Value of Different Soil Sample [J]. Science and Technology and Engineering. 2019, 19 (21): 245-252.
[6] Yonghu Zhao. Weijun Mi. Xiaopeng Wu. Xueyun Miao. Analysis to Sensibility of Silty Clay Thaw Settlement Coefficient Influence Factors Basing on Gray Relational Analysis [J]. Construction Technique. 2016, 45 (S1): 353-355.
[7] Junan Bao. Ping Yang. Cheng Chen. Research of Frost Heave Thaw Settlement Influence Factor Basing on Gray Relational Analysis [J]. Subgrade Engineering. 2012 (01): 21-23.