

# Factors Influencing Ground Settlements on Different Geomorphology Units Based on Principal Component Analysis

XiaoBei Yuan<sup>a</sup>, JunCai Li, Peng Zhang, YiLi Dai, and KaiJun Rui

College of Transportation Science&Engineering,Nanjing Tech University, Nanjing 210009, China

**Abstract.** Based on the case Nanjing Metro Line 4, the method Principal Component Analysis (PCA) was used to study the influence of ground settlement by shield tunnel construction on different geomorphology units. Correlation analysis and weighted least square method (WLS) were applied for variables selection and to obtain their relationship with settlement. 5-7 principal components could be used to present the initial 19-21 variables after decreasing the dimensions of data. For the floodplain of Yangtze River, variable parameters that highly linearly dependent on settlement were depth of tunnel, distance between the roof of tunnel and the bottom of soft soil layer, thickness of soft soil, compression modulus of soil that tunnel passed through and speed of the cutter head. For the Qinhuai ancient channel, variable parameters were Poisson's ratio, porosities, moisture content, unit weight, cohesion, internal friction angle, compression modulus of soil that tunnel passed through, advancing speed, earth chamber pressure. For the terrace of Yangtze River, variable parameters were cohesion, porosities, moisture content, Poisson's ratio, compression modulus and unit weight of soil. In addition, for the geomorphology unit with col landform, variable parameters were different. Residuals of regression formula are small, which will have certain reference value in practical engineering.

## 1 Introduction

There are many factors that affect the ground settlement during shield construction. The principal component analysis (PCA) is a statistical analysis method proposed by Pearson and Hotelling [1], which can address the challenge of a large number of factors. It can convert the initial variable to a few independent principal components featured with strong directivity without losing information [2-5]. Wu and Li [6] carried out principal component analysis on 12 selected indexes. The analysis results show that the depth ratio, the disturbance coefficient and the ratio of the width to the depth are the main control factors of the mining subsidence in Yushenfu coal mine. YOUNG et al [7] adopted principal component analysis to analyze the rock spectrum and to further establish the quality evaluation system of rock. Giuliani and Alessandro [8] used principal component analysis and blurred the area between the general statistical technology and quantitative models for specific problems. The statistical mechanics framework of biological system modeling is established. From the perspective of the system, the narrow perspective of reductionism was overcome. Although PCA is widely used in various fields, whereas studies on the settlement analyzed in PCA is very little

The laws of ground settlement under different geomorphic units are various. By preprocessing a large number of data, R language is employed as the carrier to realize data mining and data visualization. Based on the actual data analysis, after analyzing a lot of complicated

data with principal component analysis, the factors that are linearly related to the cumulative value of ground settlement were screened out, and they were used for regression analysis, so as to realize the research on the rule of the development and change of the ground settlement on different geological conditions and establish a relationship between the factors and the ground settlement, which intended to provide reference for practical engineering.

## 2 Principal component analysis of multivariate data

Equation (1) is used for standardized changes:

$$x_k^* = \frac{x_k - \bar{x}_k}{S_k}, \quad k = 1, 2, \dots, m, \quad (1)$$

Where  $x_k^*$  is standardized value;  $x_1, x_2, \dots, x_m$  are the observation values of the sample;  $\bar{x}_k$  and  $S_k$  are average and standard deviation.

The task of principal component analysis is to obtain  $b_{kj}$  ( $k = 1, 2, \dots, m, j = 1, 2, \dots, p, p < m$ ), which is the coefficient of comprehensive variables. According to the  $x_{ik}$  or  $x_k^*$ , the comprehensive variables can be expressed in  $z_j = \sum_k \tilde{b}_{kj} x_k + a_j$ . The cumulative variance  $\lambda_j$  of the comprehensive is equal to  $x_k^*$ s.  $Z_1$  was called the first principal component because of maximum variance, which absorbs the most information of the original variables.  $Z_2$  is second, and so on. Variance contribution rate can be calculated by  $\frac{\lambda_j}{m} \times 100\%$ , If the cumulative

\*Corresponding author: <sup>a</sup> bettyyxb@163.com

contribution rate of the previous principal components is more than 90%, the following principal components can be omitted.

### 3 Geomorphic unit division and data preprocessing

#### 3.1 Geomorphic unit along the subway

As shown in Table 1, there are two types of geomorphology units between Longjiang Railway Station and CaoChang Men Station. The western part of ShiToucheng Road is on floodplain of Yangtze River and the eastern part is on Ancient channel of Qinhuai River. CaoChang Men Station to YunNan Road Station is on the terrace of Yangtze River with subgeomorphology of col landform.

**Table 1.** Geological unit

Section	Geological unit	Mileage
Longjiang Railway Station to ShiToucheng Road	floodplain of Yangtze River	CK11+238-CK12+371
ShiToucheng Road to CaoChang Men Station	ancient channel of Qinhuai River	CK12+371-CK12+563
CaoChang Men Station to YunNan Road Station	terrace of Yangtze River (subgeomorphology of col landform)	CK12781-CK14+416

#### 3.2 Data preprocessing

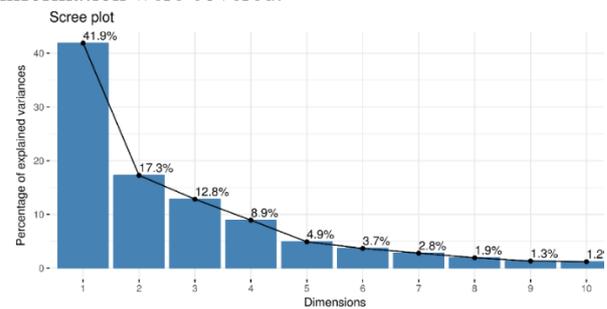
By consulting the literature [9-13], the variables were designated as depth of tunnel; groundwater level; moisture content, porosities, compression modulus, internal friction angle, cohesion, Poisson's ratio, unit weight, advancing speed of shield machine, speed of the cutter head, earth chamber pressure. The variables related to soil parameters were calculated by weighting the soil layers above tunnel floor and soil that tunnel passed through respectively. For the floodplain of Yangtze River, the additional variables of " thickness of soft soil" and "distance between the roof of tunnel and the bottom of soft soil layer" were taken into account because it has a thick layer of soft soil.

### 4 Principal component analysis of different geomorphic units

#### 4.1 Floodplain of Yangtze River

The principal components were obtained by solving the equation and the eigenvalues. Contribution rate and cumulative contribution were obtained by calculating the correlation coefficient matrix. As shown in Figure.1, the cumulative contribution rate of the first 7 principal

components is 92.27%, which means 92.3% of the total information were covered.



**Figure 1.** Principal component lithotripsy of Yangtze River floodplain.

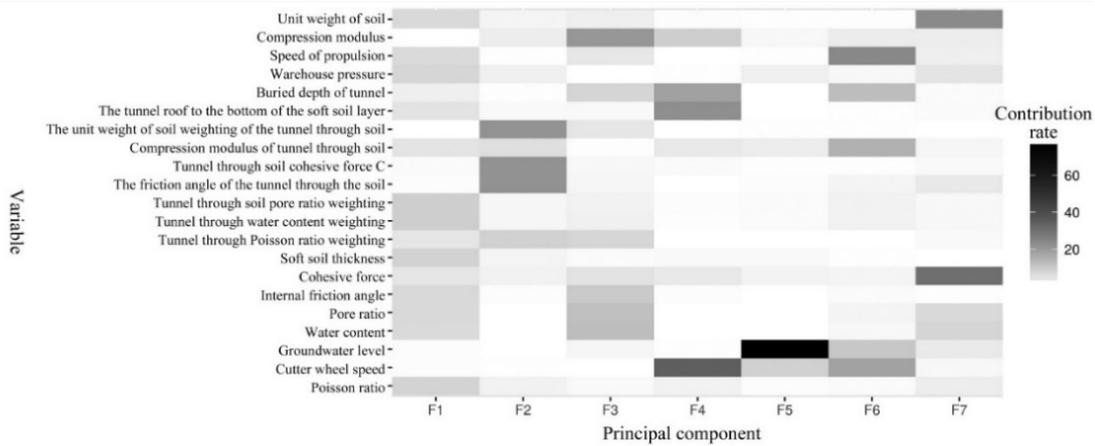
The load of factors is shown in Table 2,  $X_1, X_2, \dots, X_{21}$  are initial variables.

Fig 2 shows the contribution of the variables to the principal component, the horizontal axis is the first 7 principal components, and the longitudinal axis is the variables. In a rectangle, the deeper the color (close to black) is, the greater the contribution rate to the corresponding principal component is. For example, groundwater level has the largest contribution rate that is over 60% to the 5<sup>th</sup> principal component. Variables that have a larger contribution rate to the 2<sup>nd</sup> principal component are unit weight of soil that tunnel passed through; cohesion and internal friction angle of soil.

The 1<sup>st</sup> principal component are thickness of soft soil, Poisson's ratio, porosities and moisture of soil that tunnel passed through, earth chamber pressure. The 2<sup>nd</sup> principal component mainly points to unit weight; Poisson's ratio; internal friction angle and cohesion of soil that tunnel passed through. The 3<sup>rd</sup> principal component mainly points to internal friction angle; porosities, moisture content; compression modulus. The 4<sup>th</sup> principal component mainly refers to depth of tunnel, the distance between the tunnel roof and the bottom of the soft soil, thickness of soft soil; the compression modulus of soil that tunnel passed through, and speed of cutter head. The 5<sup>th</sup> principal component mainly points to the groundwater level. The 6<sup>th</sup> principal component mainly points to depth of tunnel, groundwater level, speed of the cutter head, compression modulus of soil that tunnel passed through, the 7<sup>th</sup> principal component mainly points to cohesion and unit weight.

**Table 2.** Factor loading of Yangtze River floodplain.

	F1	F2	F3	F4	F5	F6	F7
X <sub>1</sub>	0.168156	0.05788	-0.27672	0.429066	-0.02972	0.353165	-0.09641
X <sub>2</sub>	0.224664	0.108536	-0.10672	0.474041	0.017651	-0.04568	0.090025
X <sub>3</sub>	-0.28682	-0.13638	0.089483	-0.09677	0.108784	0.050464	0.009904
X <sub>4</sub>	0.063379	0.023522	-0.12734	0.059665	-0.87669	-0.32438	0.192387
X <sub>5</sub>	0.213858	0.154269	-0.22819	-0.20279	0.143029	0.1352	0.554036
X <sub>6</sub>	0.263841	0.07459	0.315191	0.074658	-0.02502	-0.08451	-0.03708
X <sub>7</sub>	-0.26159	0.013909	-0.34444	0.016769	0.008889	-0.13375	-0.26525
X <sub>8</sub>	-0.25682	0.016355	-0.35092	0.01899	0.009529	-0.11532	-0.27703
X <sub>9</sub>	-0.28443	-0.15396	-0.10886	0.151241	0.069772	0.085685	0.177635
X <sub>10</sub>	0.003959	0.178797	-0.45619	-0.30256	0.119299	-0.1869	0.175042
X <sub>11</sub>	0.261407	0.135912	0.174112	-0.07734	-0.05022	-0.0605	-0.48528
X <sub>12</sub>	-0.30268	0.126892	0.153483	0.07515	-0.08849	0.157654	0.113294
X <sub>13</sub>	-0.30348	0.123946	0.159686	0.050596	-0.09189	0.155085	0.11391
X <sub>14</sub>	0.031209	0.46564	0.215997	-0.0064	0.050646	-0.06088	-0.01113
X <sub>15</sub>	-0.21472	0.292011	0.2735	-0.03611	0.023631	-0.03071	0.106195
X <sub>16</sub>	-0.21693	0.241501	0.037835	0.200764	-0.18145	0.385597	0.123285
X <sub>17</sub>	-0.08228	0.469215	-0.11834	-0.08616	-0.04585	0.060705	-0.09696
X <sub>18</sub>	-0.07604	0.469454	-0.12962	-0.01177	0.067649	-0.13436	-0.19796
X <sub>19</sub>	-0.25732	-0.03785	0.203653	-0.02344	0.052084	-0.49218	0.167238
X <sub>20</sub>	0.067198	-0.04081	0.003934	-0.59187	-0.29433	0.428446	-0.11837
X <sub>21</sub>	0.27359	0.161941	0.009651	-0.04739	0.164826	-0.11444	0.222219

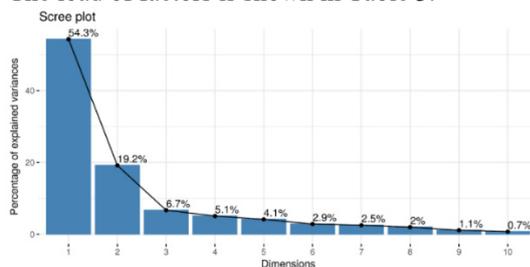


**Figure 2.** The contribution diagram of the Yangtze River flood plain.

#### 4.2 Ancient channel of Qinhuai River

As shown in Figure 3, The first 6 principal components contain 92.3% of the total information of the 19 original variables.

The top 6 principal components that have the highest contribution rate to the comprehensive index are screened out. The load of factors is shown in Table 3.

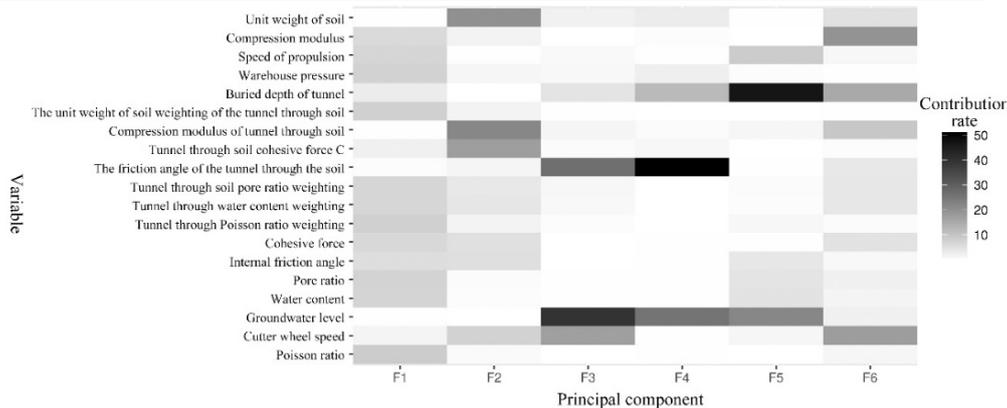


**Figure 3.** The principal component lithotripsy of Ancient channel of Qinhuai River.

From Figure 4, the 1<sup>st</sup> principal component contain wider information, and the contribution value of each variable is comparative average. The 2<sup>nd</sup> principal component mainly links to unit weight, compression modulus and cohesion of soil that tunnel passed through. The 3<sup>rd</sup> principal component links to groundwater level, internal friction angle of soil that tunnel passed through; speed of cutter head. The 4<sup>th</sup> principal component mainly links to depth of tunnel, groundwater level, internal friction angle of soil that tunnel passed through. The 5<sup>th</sup> principal component mainly links to depth of tunnel and groundwater level. The 6<sup>th</sup> principal component mainly links to depth of tunnel, compression modulus and speed of cutter head.

**Table 3.** Factor loading of Ancient channel of Qinhuai River.

	F1	F2	F3	F4	F5	F6
X <sub>1</sub>	-0.18239	-0.01914	0.216852	0.353186	0.684026	-0.39567
X <sub>2</sub>	-0.04255	0.00409	-0.63074	-0.5081	0.471712	0.16467
X <sub>3</sub>	-0.26363	0.234968	0.009656	0.046876	-0.04093	0.226829
X <sub>4</sub>	-0.24638	-0.23826	0.025221	-0.0483	-0.20776	0.10883
X <sub>5</sub>	0.278991	-0.07986	-0.04689	0.056996	0.2208	0.161335
X <sub>6</sub>	0.278368	-0.07385	-0.06398	0.060341	0.226249	0.139787
X <sub>7</sub>	0.301462	0.089434	-0.0271	0.048722	0.010793	-0.1275
X <sub>8</sub>	-0.25489	0.145648	-0.00136	0.073443	-0.02049	0.446075
X <sub>9</sub>	-0.03694	-0.44924	0.155472	-0.18477	0.055388	0.232462
X <sub>10</sub>	0.270194	0.203508	0.121043	-0.00327	0.088284	0.206493
X <sub>11</sub>	0.268913	0.211092	0.108938	-0.00633	0.074516	0.204789
X <sub>12</sub>	-0.29026	-0.14595	-0.06123	-0.00756	0.004024	0.047039
X <sub>13</sub>	0.288891	0.144936	-0.07142	-0.00359	-0.11608	-0.08175
X <sub>14</sub>	-0.04034	0.472056	0.13198	0.082619	0.127806	0.316894
X <sub>15</sub>	-0.1645	0.421016	0.08337	-0.12042	0.021125	-0.0775
X <sub>16</sub>	-0.06965	0.13075	0.518225	-0.7131	0.065001	-0.20283
X <sub>17</sub>	0.273432	0.052303	-0.1018	-0.06263	-0.30273	-0.11911
X <sub>18</sub>	-0.14221	0.283311	-0.41584	-0.05024	-0.12902	-0.42348
X <sub>19</sub>	-0.28395	0.12771	-0.10785	0.16542	-0.07126	0.060164

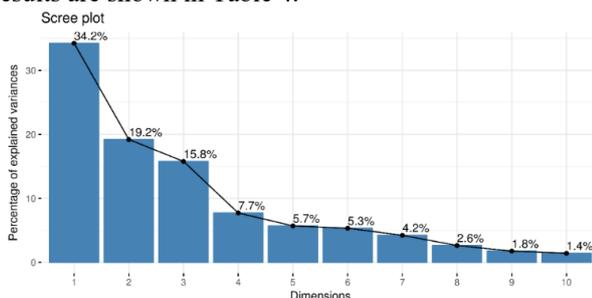


**Figure 4.** The contribution diagram of Ancient channel of Qinhuai River.

### 4.3 Terrace of Yangtze River

As show in Figure 5, the first 7 principal components included 92.2% of the total information of the 16 initial variables.

The top 7 principal components are screened out, and the results are shown in Table 4.

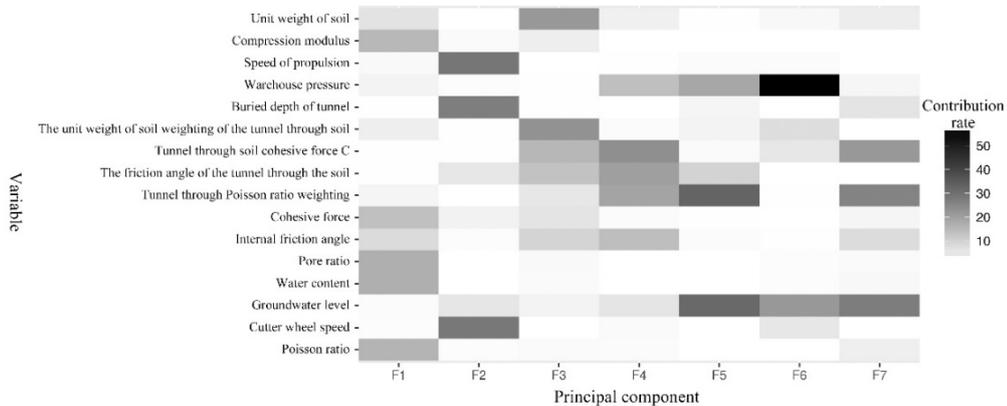


**Figure 5.** The principal component lithotripsy of Terrace of Yangtze River

As shown in Figure 6, the 1<sup>st</sup> principal component mainly refers to cohesion, porosities, moisture content, Poisson's ratio, compression modulus and unit weight of soil. The 2<sup>st</sup> principal component mainly refers to depth of tunnel, advancing speed and speed of cutter head. The 3<sup>st</sup> principal component mainly refers to unit weight, cohesion, and internal friction angle of soil that tunnel passed through. The 4<sup>th</sup> principal component mainly refers to internal friction angle, Poisson's ratio, cohesion, and internal friction angle of soil that tunnel passed through, earth chamber pressure. The 5<sup>th</sup> principal component mainly refers to underground water level, Poisson's ratio of soil that tunnel passed through, earth chamber pressure. The 6<sup>th</sup> principal component refers to groundwater level; earth chamber pressure. The 7<sup>th</sup> principal component mainly refers to groundwater level, Poisson's ratio and cohesion of soil that tunnel passed through.

**Table 4.** Factor loading of Terrace of Yangtze River.

	F1	F2	F3	F4	F5	F6	F7
X <sub>1</sub>	-0.0703	0.520004	-0.05002	0.019839	-0.14495	0.00739	-0.2351
X <sub>2</sub>	0.086356	-0.22505	-0.15466	-0.22858	-0.56578	-0.46089	-0.52322
X <sub>3</sub>	-0.35837	-0.15839	-0.23361	0.08732	0.036171	0.001939	-0.14014
X <sub>4</sub>	-0.26963	-0.08134	-0.29498	0.366161	0.084871	0.067133	-0.26597
X <sub>5</sub>	0.406204	0.005427	0.107717	-0.00708	0.014733	0.081365	-0.09815
X <sub>6</sub>	0.407105	0.00931	0.104433	0.006814	0.017227	0.072142	-0.1177
X <sub>7</sub>	0.392076	0.076536	-0.10438	0.091314	0.013917	0.029162	-0.18161
X <sub>8</sub>	-0.38066	-0.10067	-0.18504	0.001945	-0.00444	-0.00786	0.001756
X <sub>9</sub>	-0.23542	-0.02162	0.461378	-0.17288	-0.06476	-0.11824	0.188149
X <sub>10</sub>	-0.18315	0.036953	0.477177	-0.08381	-0.15504	-0.26381	0.02056
X <sub>11</sub>	0.142041	0.025292	-0.22086	0.433081	-0.57588	0.056643	0.508991
X <sub>12</sub>	0.04344	-0.03991	-0.38342	-0.48131	-0.09937	-0.22009	0.460149
X <sub>13</sub>	0.059936	0.21983	-0.35213	-0.44694	0.299165	0.014526	0.001811
X <sub>14</sub>	-0.10541	0.537149	-0.01925	0.031961	-0.09372	-0.08839	-0.01675
X <sub>15</sub>	0.079744	-0.53406	0.028758	-0.09535	0.021497	0.221722	-0.00568
X <sub>16</sub>	0.15817	-0.07069	-0.05293	0.362244	0.425618	-0.76018	0.131052

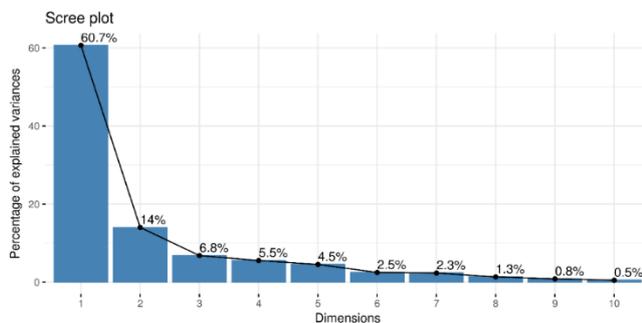


**Figure 6.** The contribution diagram of Terrace of Yangtze River.

#### 4.4 Col landform

From Figure 7, in the condition of the subgeomorphology of col landform, the cumulative contribution rate of first 5 principal components is 91.5%, which means it contains 91.5% of the total information.

The Factor loading of the top 5 principal components are extracted as the Table 5.

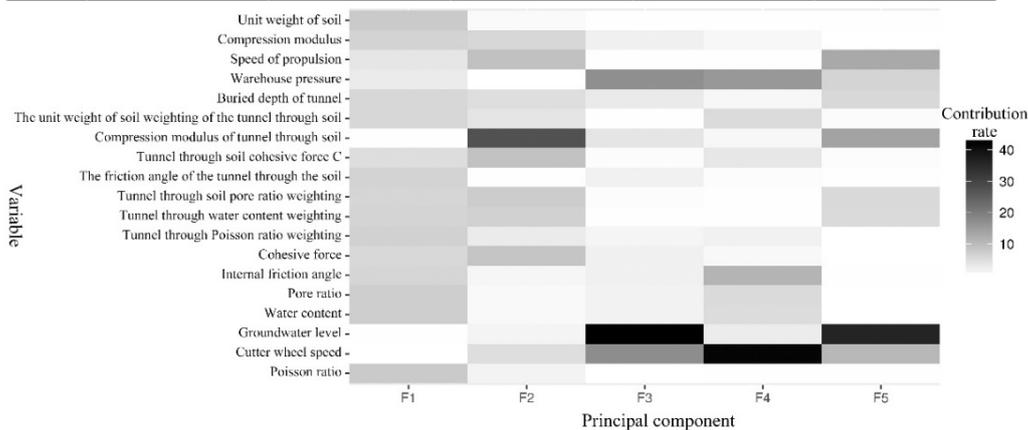


**Figure 7.** The principal component lithotripsy of col landform.

As shown in Figure 8, 1<sup>st</sup> principal component refers to many variables, the 2<sup>nd</sup> principal component points to the compression modulus, cohesion of soil that tunnel passed through; advancing speed, The 3<sup>rd</sup> principal component points to groundwater level, speed of cutter head, earth chamber pressure. The 4<sup>th</sup> principal component points to the internal friction angle, speed of cutter head, and earth chamber pressure. The 5<sup>th</sup> principal component points to groundwater level, compression modulus of soil that tunnel passed through, speed of advancing and cutter head.

**Table 5.** Factor loading of col landform.

	F1	F2	F3	F4	F5
X <sub>1</sub>	-0.2436	0.220523	-0.17788	0.112415	-0.24308
X <sub>2</sub>	-0.02955	-0.12739	-0.6533	-0.16681	0.604079
X <sub>3</sub>	-0.24242	-0.29676	0.150519	-0.10224	0.018964
X <sub>4</sub>	-0.25237	-0.1158	0.147124	-0.33825	-0.05686
X <sub>5</sub>	0.272032	0.091728	-0.13727	0.234425	0.015321
X <sub>6</sub>	0.272202	0.097823	-0.13856	0.229768	0.017222
X <sub>7</sub>	0.282124	0.131511	-0.01142	0.032914	-0.05277
X <sub>8</sub>	-0.25735	-0.24369	0.149525	-0.11056	0.013858
X <sub>9</sub>	-0.28132	0.088587	-0.03307	0.052427	-0.02962
X <sub>10</sub>	0.250154	-0.27588	-0.0648	-0.05348	-0.23993
X <sub>11</sub>	0.254158	-0.26337	-0.05936	-0.0551	-0.23405
X <sub>12</sub>	-0.24503	0.197857	0.0396	0.233189	0.082117
X <sub>13</sub>	0.262529	0.173354	0.123179	-0.1443	-0.01062
X <sub>14</sub>	0.060301	-0.52683	-0.19597	0.108634	-0.37444
X <sub>15</sub>	-0.2267	-0.30466	-0.06424	0.190772	0.065391
X <sub>16</sub>	-0.25586	0.055827	-0.14428	0.063462	-0.0301
X <sub>17</sub>	-0.19467	0.306346	-0.03002	0.054752	-0.36329
X <sub>18</sub>	0.025871	-0.2207	0.417855	0.645678	0.328199
X <sub>19</sub>	0.170523	0.020906	0.415102	-0.39705	0.252479



**Figure 8.** The contribution diagram of col landform.

### 5 Influence factors of ground settlement

The correlation analysis between principal components and ground settlement was carried out with the R Programming Language. The significance was described by Pearson coefficient. Variables that has the greatest linear correlation with ground settlement on each geomorphic unit were screened out.

#### 5.1 Floodplain of the Yangtze River

As shown in Table 6, in the perspective of statistical significance, the 1<sup>st</sup> principal component and 4<sup>th</sup> principal component are highly correlated with the settlement. The 2<sup>nd</sup> principal component and 3<sup>rd</sup> principal component are significantly correlated with the settlement. In the perspective of the correlation, the 4<sup>th</sup> principal component is highly linear correlation with settlement.

Considering the factor loading of each principal component, the most important factors are given in Table 7.

**Table 6.** Correlation coefficient and Saliency.

Principal component	Correlation coefficient	Statistical significance	Lower bound of confidence interval	Upper bound of confidence region	Confidence degree
F1	-0.24285817	0.000358703**	-0.365641572	-0.111766741	0.95

F2	0.160988451	0.01900217*	0.026821304	0.289458209	0.95
F3	0.158467607	0.020983212*	0.024236146	0.287086335	0.95
F4	0.546134852	7.02443E-18**	0.444065117	0.634211503	0.95
F5	-0.098300319	0.15379078	-0.230002716	0.036937939	0.95
F6	-0.03645762	0.597590906	-0.170369641	0.098776616	0.95
F7	0.018309298	0.790983666	-0.116727671	0.152681592	0.95

**Table 7.** Impact factor table.

Principal component	Variables with high contribution rate
F1**	thickness of soft soil; Poisson's ratio; porosities of soil that tunnel passed through; moisture content of soil that tunnel passed through; earth chamber pressure
F2*	unit weight of soil that tunnel passed through; porosities of soil that tunnel passed through; cohesion of soil that tunnel passed through; internal friction angle of soil that tunnel passed through
F3*	Internal friction angle, porosities, compression modulus
F4**	depth of tunnel; distance between tunnel roof and soft soil bottom; thickness of soft soil; compression modulus of soil that tunnel passed through; speed of cutter head

### 5.2 Ancient channel of Qinhuai River

As shown in Table 8, in the perspective of statistical significance, the 1<sup>st</sup> principal component and 2<sup>nd</sup> principal

component are significantly related to the settlement. In the perspective of correlation coefficient, the 1<sup>st</sup> principal component is highly linearly related to settlement. The 2<sup>nd</sup> one is lower linearly.

The linear correlation factors of Qinhuai ancient channel were given in Table 9.

**Table 8.** Correlation analysis results.

Principal component	Correlation coefficient	Statistical significance	Lower bound of confidence interval	Upper bound of confidence interval	Confidence degree
F1	-0.673709001	1.5381E-07**	-0.803948566	-0.481803013	0.95
F2	-0.335596212	0.019707733**	-0.565781118	-0.056885836	0.95
F3	0.203706881	0.164914445	-0.08536899	0.461150121	0.95
F4	0.179483299	0.222214632	-0.110275149	0.441121912	0.95
F5	0.106752157	0.47018841	-0.182931209	0.379379518	0.95
F6	0.024323904	0.869650082	-0.261618912	0.306341395	0.95

**Table 9.** Impact factor table.

Principal component	High contribution rate variable
F1**	Cohesion, internal friction angle; porosities, moisture content; Poisson's ratio; compression modulus; advancing speed; earth chamber pressure; porosities of soil that tunnel passed through; moisture of soil that tunnel passed through; unit weight of soil that tunnel passed through; soil Poisson's ratio of soil that tunnel passed through
F2**	unit weight; compression modulus of soil that tunnel passed through; cohesion of soil that tunnel passed through

### 5.3 Terrace of Yangtze River

As shown in Table 10, it is found that the 2<sup>nd</sup> and 5<sup>th</sup> principal components are closely related to the settlement, and the 1<sup>st</sup> principal component is significantly related to

the settlement. According to the correlation coefficient, the 2<sup>nd</sup> principal component is high linearly related to the settlement, and the 5<sup>th</sup> is linearly at a relatively low level .

Each factor that has higher correlation with settlement under the unit of terrace of Yangtze River Terrace was given in Table 11.

**Table 10.** Correlation analysis results.

Principal component	Correlation coefficient	Statistical significance	Lower bound of confidence interval	Upper bound of confidence interval	Confidence degree
F1	-0.136925783	0.013202871*	-0.241793546	-0.028896331	0.95
F2	-0.59453417	1.22929E-32**	-0.660408099	-0.51957927	0.95

F3	0.069221571	0.211867018	-0.039533822	0.176356132	0.95
F4	-0.025547751	0.645308353	-0.133636047	0.083141208	0.95
F5	-0.20539295	0.000184111**	-0.307012336	-0.099142936	0.95
F6	-0.016065935	0.772253303	-0.124307915	0.092553924	0.95
F7	-0.043588823	0.432114816	-0.151331974	0.065177896	0.95

**Table 11.** Impact factor table.

Principal component	High contribution rate variable
F1*	cohesion; porosities; soil moisture, Poisson's ratio, compression modulus; unit weight
F2**	depth of tunnel; speed of advancing and cutter head
F5**	groundwater level; Poisson's ratio of soil that tunnel passed through; earth chamber pressure

### 5.4 Col landform

From Table 12, in the perspective of statistical significance, the 1<sup>st</sup>, 3<sup>rd</sup>, 4<sup>th</sup>, and 5<sup>th</sup> principal component

are closely related to the settlement. In the perspective of the correlation, 2<sup>nd</sup> and 4<sup>th</sup> are linearly related to settlement at a relatively low level.

As shown in Table 13, Extremely significant factors was given.

**Table 12.** Correlation analysis results.

Principal component	Correlation coefficient	Statistical significance	Lower bound of confidence interval	Upper bound of confidence interval	Confidence degree
F1	-0.180369946	0.002405323**	-0.291235626	-0.064723613	0.95
F2	0.331952063	1.1825E-08**	0.223625893	0.432177566	0.95
F3	0.052604035	0.379682181	-0.064807232	0.168578736	0.95
F4	-0.32555593	2.32502E-08**	-0.42632773	-0.21680246	0.95
F5	-0.248083303	2.59892E-05**	-0.354796344	-0.134989531	0.95

**Table 13.** Impact factor table.

Principal component	High contribution rate variable
F1**	depth of tunnel; cohesion; internal friction angle; porosities; moisture content; Poisson's ratio; compression modulus; unit weight of soil that tunnel passed through; moisture of soil that tunnel passed through; Poisson's ratio of soil that tunnel passed through; cohesion of soil that tunnel passed through, internal friction angle of soil that tunnel passed through
F2**	compression modulus of soil that tunnel passed through; cohesion of soil that tunnel passed through; advancing speed
F4**	internal friction angle; speed of cutter head; earth chamber pressure
F5**	groundwater level; compression modulus of soil that tunnel passed through; advancing speed of shield machine; speed of cutter head

## 6 Predicted research of ground settlement caused by shield method

Due to the deviation of measured data, the method weighted least squares (WLS) was applied for regression analysis with elimination of heteroscedasticity. It is evasive to carry on regression analysis with Ancient channel of Qinhuai River because of the limit of data.

### 6.1 Floodplain of Yangtze River

For the floodplain of Yangtze River, the 4<sup>th</sup> principal component is linearly related to the settlement. 5 variables of the 4<sup>th</sup> principal component were screened out to get regression equation, they are X<sub>1</sub> (depth of tunnel), X<sub>2</sub>

(distance from tunnel roof to bottom of soft soil), X<sub>3</sub> (cohesion), X<sub>4</sub> (Compression modulus), X<sub>5</sub> (speed of cutter head). Equation (2) shows the result:

$$Y = 2.983157X_1 + 2.1389587X_2 + 1.248048X_3 - 1.39273X_4 - 89.2393X_5 - 20.3879 \quad (2)$$

As shown in Figure 9, Except for a few abnormal points, the residual values of most points are evenly distributed near the 0 level. The fitting results are basically normal distribution, which is consistent with the theory.

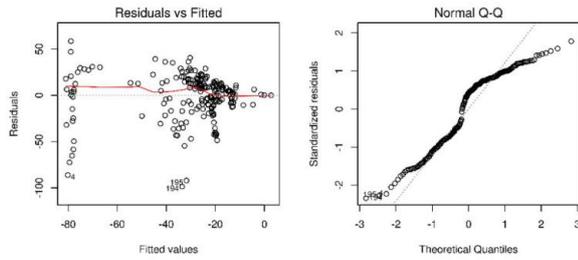


Figure 9. The residuals map of floodplain of Yangtze River.

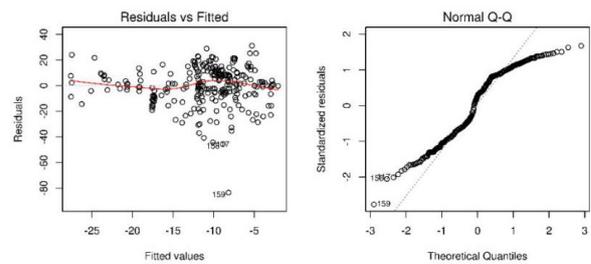


Figure 11. Variance diagram of col landform.

## 6.2 Terrace of Yangtze River

For the terrace of Yangtze River, the 2<sup>nd</sup> principal component is highly linearly related to settlement, and the 5<sup>th</sup> principal component is low linear correlation. Analysis applies 6 variables for regression. The settlement can be estimated by equation (3)

$$Y = -1.7786X_1 + 4.428968X_2 + 123.906X_3 - 0.20031X_4 - 42.7107X_5 - 8.909101X_6 - 119.888 \quad (3)$$

Where  $X_1$  is depth of tunnel;  $X_2$  is advancing speed;  $X_3$  is speed of cutter head;  $X_4$  is groundwater level;  $X_5$  is Poisson's ratio of soil that tunnel passed through;  $X_6$  is earth chamber pressure.

As shown in Figure 10, the regression effect of the first grade terrace of the Yangtze River is ideal.

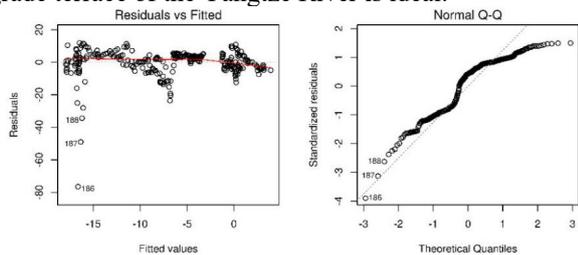


Figure 10. Variance diagram of Terrace of Yangtze River.

## 6.3. Col landform

According to the regression coefficient, the settlement of Subgeomorphology of the depressions can be estimated by equation (4):

$$Y = -2.00531X_1 + 0.662063X_2 - 1.41525X_3 + 0.326726X_4 - 0.906938X_5 - 1.95293X_6 - 6.79388 \quad (4)$$

Where  $X_1$  is compression modulus of soil that tunnel passed through;  $X_2$  is cohesion of soil that tunnel passed through;  $X_3$  is advancing speed;  $X_4$  is internal friction angle;  $X_5$  is speed of cutter head;  $X_6$  is earth chamber pressure.

As shown in Figure 11, The amount of measured data is less, the regression effect is lower than the other unit. But most of the regression results are evenly distributed around the 0-horizontal line. So, the regression result still has practical reference significance.

## 7 Conclusion

The influence factors of ground settlement caused by shield tunnel construction is investigated by employing the method principal components analysis. The dimension of numerous variables was significantly reduced by introducing PCA. It is found that factors which has highly influence on ground settlement are different on each geomorphology unit. The factors that were screened out by correlation analysis are coincide with the engineering experience. In addition, the residuals of regression equation formed by selected factors are acceptable.

## References

1. Hotelling H. Analysis of a complex of statistical variables into principal components. J. Journal of Educational Psychology. **24**,41-441 (1933)
2. Aflalo Y, Kimmel R. Regularized principal component analysis. J. Annals of Mathematics B. **38(1)**, 1-12 (2017)
3. Han L, Wu Z, Zeng K, et al. Online Multilinear Principal Component Analysis J. Neurocomputing. (2017).
4. Bandalos D L, Boehm Kaufman M R. Four common misconceptions in exploratory factor analysis. J. 61-87 (2008).
5. Aguilera A M, Ocaña F A, Valderrama M J. Forecasting with unequally spaced data by a functional principal component approach. J. Test. **8(1)**, 233-253 (1999)
6. Wu Chaoyang, Li Ning. Comprehensive Evaluation Of Ming Subsidence Damage In Yushenfu Mining Area Based On Principal Component Analysis C. National annual meeting of engineering geology. (2011)
7. YOUNG R P, COFFEY J R, HILL J J. The application of spectral analysis to rock quality evaluation for mapping purposes J. Bulletin of Engineering and Geology and the Environment. **19(1)**,268-274 (1979)
8. Giuliani A. The application of principal component analysis to drug discovery and biomedical data J. Drug Discovery Today. **22(7)**, 1069 (2017)
9. Saha P, Roy N, Mukherjee D, et al. Application of Principal Component Analysis for Outlier Detection

- in Heterogeneous Traffic Data. *J. Procedia Computer Science*. **83**,107-114 (2016)
10. Feng H. Analysis of the influence of subway tunnel under construction on the settlement of the upper subgrade. *J. Journal of Jiamusi Vocational Institute*. (2017)
  11. Suwansawat S, Einstein H H. Artificial neural networks for predicting the maximum surface settlement caused by EPB shield tunneling. *J. Tunnelling & Underground Space Technology*. **21(2)**, 133-150 (2006)
  12. Liu J, Qi T, Wu Z. Analysis of ground movement due to metro station driven with enlarging shield tunnels under building and its parameter sensitivity analysis. *J. Tunnelling & Underground Space Technology*. **28(3)**, 287-296 (2012)
  13. Kasper T, Meschke G. A numerical study of the effect of soil and grout material properties and cover depth in shield tunneling. *J. Computers & Geotechnics*. **33(4-5)**, 234-247 (2006)