2017

Study on the methods for predicting the performance of a hybrid solar-assisted ground-source heat pump system

Zhaoyi Zhuang  
*University of Nebraska - Lincoln*, hit6421@126.com

Qiqi Zhu  
*Shandong Jianzhu University*

Jian Song  
*Shandong Jianzhu University*

Xin Zhang  
*Shandong University*

Haorong Li  
*University of Nebraska-Lincoln*, hli3@unl.edu

Follow this and additional works at: [https://digitalcommons.unl.edu/archengfacpub](https://digitalcommons.unl.edu/archengfacpub)

Part of the [Architectural Engineering Commons](https://digitalcommons.unl.edu/archengfacpub), [Construction Engineering Commons](https://digitalcommons.unl.edu/archengfacpub), [Environmental Design Commons](https://digitalcommons.unl.edu/archengfacpub), and the [Other Engineering Commons](https://digitalcommons.unl.edu/archengfacpub)

Zhuang, Zhaoyi; Zhu, Qiqi; Song, Jian; Zhang, Xin; and Li, Haorong, "Study on the methods for predicting the performance of a hybrid solar-assisted ground-source heat pump system" (2017). *Architectural Engineering -- Faculty Publications*. 120.  
[https://digitalcommons.unl.edu/archengfacpub/120](https://digitalcommons.unl.edu/archengfacpub/120)

This Article is brought to you for free and open access by the Architectural Engineering at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Architectural Engineering -- Faculty Publications by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.
10th International Symposium on Heating, Ventilation and Air Conditioning, ISHVAC2017, 19-22 October 2017, Jinan, China

Study on the methods for predicting the performance of a hybrid solar-assisted ground-source heat pump system

Zhaoyi Zhuang\textsuperscript{a,b,c,}, Qiqi Zhu\textsuperscript{a}, Jian Song\textsuperscript{a}, Xin Zhang\textsuperscript{d}, Haorong Li\textsuperscript{b}

\textsuperscript{a}School of Thermal Engineering, Shandong Jianzhu University, Jinan 250101, China
\textsuperscript{b}Durham School of Architectural Engineering & Construction, University of Nebraska–Lincoln, Omaha68182, USA,
\textsuperscript{c}Shandong Zhongrui New Energy Technology Co., Ltd. Jinan250101, China
\textsuperscript{d}School of Information Science and Engineering, Shandong University, Jinan 250100, China

Abstract

It is critical to find suitable setting parameters for designing a hybrid solar-assisted ground-source heat pump system in the practical engineering application, but the heat pump performance is unpredictable after many years of operation. This paper used 2000 sets of performance data collected from solar-assisted GSHP systems that keep operating over 20 years to simulate long term used heat pump with a professional software called GeoStar. Adopted the classification and regression tree (CART) method, the design of solar energy collector areas can be predicted. The multi-linear regression is also utilized to predict average monthly per meter borehole heat exchange. Seasonal factor decomposition and exponential smoothing are used to analyze the average monthly temperature of the circulating fluid, circulating fluid inlet and outlet temperatures of the heat pump after 20 years when we perform the time series prediction. Experimental results demonstrate that CART, multi-linear regression, seasonal factor decomposition and exponential smoothing are promising for practical applications.

© 2017 The Authors. Published by Elsevier Ltd.
Peer-review under responsibility of the scientific committee of the 10th International Symposium on Heating, Ventilation and Air Conditioning.

Keywords: a hybrid solar-assisted ground-source heat pump system; GeoStar software; prediction methods;

1. Introduction

Nowadays, the environmental and natural resources have become key factors in the sustainable development of a country. It is challenging to balance the resource consumption (such as mineral resources) with clean environment.

*Corresponding author. Tel: +86 18906442878
E-mail address: hit6421@126.com

1877-7058 © 2017 The Authors. Published by Elsevier Ltd.
Peer-review under responsibility of the scientific committee of the 10th International Symposium on Heating, Ventilation and Air Conditioning.
10.1016/j.proeng.2017.10.049

Open access per Scopus (Elsevier)
A vision to solve this challenge is to employ green, pollution-free alternative energy source. This paper approaches this vision by a ground-source heat pump (GSHP) [1]. The GSHP system performance degrades if installed in heating-dominated buildings in severe cold climate areas, where more heat are extracted from the ground than rejected into the ground. This will cause a lower temperature in the intake water due to the annual released heat. A solution is to increase the spacing between boreholes, or to use strip type and block layout with a high level of initial investment to eliminate this thermal imbalance[2]. Chen and Yang designed and numerically simulated a solar assisted GCHP system in northern China[3], which showed that this optimal design reduced the borehole length of 3.9m/m² and the designed system efficiency could be 3.55 with 36% annual space heating solar fraction and 75% annual domestic hot water solar fraction. Verma et al [4] studied the solar energy storage and space heating by using solar-assisted GCHP system for Indian climatic conditions. Man and Yang [5] simulated hybrid GSHP system with heat transfer process of principal components model, and combined nocturnal cooling radiator works with hybrid GSHP system[6]. You et al [7] proposed multiple functions, such as heat compensation, direct domestic hot water (DHW) as well as direct space heating, to support heat compensation unit with thermosyphon unit for the purpose of energy-saving of hybrid GCHP system in cold regions requiring considerable air-conditioning and DHW supply[8]. In our previous work [9], we have used Partial Least Squares Regression (PLSR), Support Vector Regression (SVR) and M5 Model Tree to predict the heat transfer performance for the GCHP system. However, there has been little work on predicting the performance of solar-assisted ground-source heat pump systems. To address the problem, 2000 groups of simulation data of solar-assisted GSHP systems operating for over 20 years were created using simulation with a professional software GeoStar [10]. Then, the classification and regression tree (CART), multi-linear regression, and time series analysis are used to predict the performance of solar-assisted ground-coupled heat pump system.

2. Data collation analysis

This paper employs 2000 sets of performance data simulated from a professional software called GeoStar[10]. Collector area $Y_1$, average monthly per meter borehole heat exchange $Y_2$, average monthly temperature of the circulating fluid $Y_3$, circulating fluid inlet temperature of the heat pump $Y_4$ and circulating fluid outlet temperature of the heat pump $Y_5$, $Y_6$, $Y_7$, $Y_8$, and $Y_9$ are all time series data for 240 months (20 years). The system working condition parameters includes drilling vertical spacing $X_1$, drilling column spacing $X_2$, drilling radius $X_3$, drilling geometry arrangement $X_4$, drilling nominal external diameter $X_5$, U-tube spacing $X_6$, thermal conductivity coefficient $X_7$, drilling depth $X_8$, number of drilling $X_9$, ground temperature $X_{10}$, ground thermal conductivity $X_{11}$, circulating fluid parameter $X_{12}$, collector type $X_{13}$, collector efficiency $X_{14}$, collector installation angle $X_{15}$, heat loss efficiency $X_{16}$ and surface albedo $X_{17}$. $X_1$, $X_2$, $X_3$, $X_4$, $X_5$, $X_6$, $X_7$, $X_8$, and $X_9$ are catalog, and their introduction can be found in Ref [9]. $X_{10}$ is also catalog, and Table 1 provides the specific collector type $X_{10}$ of different materials in this study. Table 2 provides descriptive statistics of these 17 system working condition parameters.

<table>
<thead>
<tr>
<th>Nominal value</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>metallic glass vacuum tube</td>
</tr>
<tr>
<td>2</td>
<td>flat-plate collector</td>
</tr>
<tr>
<td>3</td>
<td>glass tube collector with vacuum</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Parameters</th>
<th>$X_1$</th>
<th>$X_2$</th>
<th>$X_3$</th>
<th>$X_4$</th>
<th>$X_5$</th>
<th>$X_6$</th>
<th>$X_7$</th>
<th>$X_8$</th>
<th>$X_9$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average</td>
<td>3.998</td>
<td>4.000</td>
<td>0.065</td>
<td>2.88</td>
<td>36.75</td>
<td>2.50</td>
<td>2.034</td>
<td>96.16</td>
<td>59.07</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>0.708</td>
<td>0.707</td>
<td>0.007</td>
<td>1.453</td>
<td>9.313</td>
<td>1.118</td>
<td>0.455</td>
<td>13.317</td>
<td>8.210</td>
</tr>
<tr>
<td>Minimum</td>
<td>2.5</td>
<td>3</td>
<td>0.055</td>
<td>1</td>
<td>25</td>
<td>1</td>
<td>1.3</td>
<td>80</td>
<td>46</td>
</tr>
<tr>
<td>Maximum</td>
<td>5</td>
<td>5</td>
<td>0.075</td>
<td>5</td>
<td>50</td>
<td>4</td>
<td>2.8</td>
<td>120</td>
<td>70</td>
</tr>
<tr>
<td>Range</td>
<td>2.5</td>
<td>2</td>
<td>0.02</td>
<td>4</td>
<td>25</td>
<td>3</td>
<td>1.5</td>
<td>40</td>
<td>24</td>
</tr>
</tbody>
</table>
3. Simulation and calculation

3.1 Performance Prediction on \( Y \)

Generally speaking, when establishing and training the prediction model, addition of unnecessary features will reduce the prediction performance. Moreover, a wide difference between the value ranges of different features will also affect the regression prediction. Therefore, we mainly adopt feature selection and normalization to preprocess the original data.

3.1.1 Feature Selection

Feature selection is a procedure that investigates all the features to eliminate irrelevant ones, and the standards for investigation vary. Now, data used in this paper are mainly studied in two aspects: variance and correlation.

Variance:

\[
S^2 = \frac{1}{N} \sum_{i=1}^{N} \left( X_i - \bar{X} \right)^2
\]

(1)

where \( \bar{X} \) is mean of a set \( \{x_i, i=1,\cdots,N\} \), and \( N \) is the size of the set.

Pearson correlation coefficient:

\[
r = \frac{1}{N-1} \sum_{i=1}^{N} \left( \frac{X_i - \bar{X}}{S_X} \right) \left( \frac{Y_i - \bar{Y}}{S_Y} \right)
\]

(2)

where \( S_X \) and \( S_Y \) are standard deviations of \( \{x_i, i=1,\cdots,N\} \) and \( \{y_i, i=1,\cdots,N\} \), and \( \bar{Y} \) is mean of a set \( \{y_i, i=1,\cdots,N\} \). Pearson Correlation Coefficient ranging in value from -1 through +1, +1 means that there is a positive linear correlation between two variables, -1 means that there is a linear negative correlation, and 0 means irrelevance.

Firstly, the variances of 17 variables are calculated and the threshold, as shown in Fig.1, was set to 0.1. Meanwhile, any feature whose variance is lower than 0.1 among the 2000 samples will be eliminated for it cannot show sufficient differences and has a little help in establishing the model. Feature \( X_3 \) and \( X_{17} \) are eliminated during the regression prediction of \( Y'_1 \) for smaller variances.

Single variable selection then is processed after eliminating features with variance lower than the threshold, where Pearson correlation coefficient is used in the measurement of correlation. Consequently, the Pearson correlation coefficient of predicted value \( Y'_1 \) and the 17 eigenvalues were calculated, as shown in Fig.2, and needed features were selected according to their absolute value.

3.1.2 Normalization

Firstly, the two features with smaller variance are eliminated and then the dimensionality of features is reduced to 8 (remaining \( X_1, X_3, X_5, X_9, X_{12}, X_{14}, X_{15}, X_{16} \)) from 15 by calculating and comparing the Pearson correlation coefficient. Dimensionality reduction has not only reduced the workload in an establishment of the prediction model but also can improve its performance. Then, the data is normalized by obtaining the average and standard deviation.
of the remaining eight feature variables, therefore avoiding harmful effects may be caused by a wide difference between value ranges.

After the normalization of

\[
\tilde{x} = \frac{x - \bar{x}}{S_x}
\]

the average of variables will reach around 0, and standard deviation approaches 1. CART is used to make regression prediction of \( Y_1 \); the average error is 3.031, the error rate was 0.97%, the RMSE is 4.044. The red curve shown in Fig.3 is the actual value of \( Y_1 \), and the green curve is the predictive value \( Y_1 \).

3.2 Performance Prediction on \( Y_2 \)

Using 12 months as a period \( Y_2 \), the multiple linear regression and least square method are adopted. The prediction expression of \( Y_2 \) for using the former 1600 groups of data as a training set and remaining 400 ones as a testing set is:

\[
Y_2^l = -15.972 + 0.002X_1 + 0.001X_2 + 0.359X_3 + 0.009X_4 + 0.002X_5 - 0.005X_6 + 0.003X_7 + 0.001X_8 + 0.002X_9 + 0.002X_{10} + 0.02X_{11} - 0.253X_{12}
\]

The result is \( F=754.284 \), significance SIG=0.000. So it is judged at a probability of 99.9% that arguments \( X_1, \ldots, X_{12} \) all have a significant effect on the dependent variable \( Y_2 \) and the RMSE is 0.006, quite ideal. Then each of results of 12 months can be obtained in turn, and the relationship between average monthly per meter borehole heat exchange of all 12 months and 17 inputs are formulated as

\[
Y_2 = AX + b
\]

where

\[
Y_2 = \begin{bmatrix} Y_1^l \\ Y_2^l \\ \vdots \\ Y_{12}^l \end{bmatrix}, \quad X = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_{12} \end{bmatrix},
\]

\[
\beta = [-15.972, -10.942, -5.161, 6.281, 9.025, 8.461, 6.793, 7.060, 7.254, 4.010, 6.242, 14.701]
\]

\[
A = \begin{bmatrix} 0.002 & 0.001 & 0.519 & 0.000 & 4.475 \times 10^{-5} & 0.534 \times 10^{-5} & 0.002 & -5.005 & 0.001 & -3.607 \times 10^{-5} & 0.001 & -3.657 \times 10^{-6} & -7.087 \times 10^{-5} & 0.020 & 0.000 & -0.253 \\
0.002 & 0.000 & 0.302 & 0.006 & 3.608 \times 10^{-5} & 1.521 \times 10^{-5} & 0.001 & -0.004 & 0.001 & 1.065 \times 10^{-5} & 0.001 & -1.518 \times 10^{-5} & -2.411 \times 10^{-5} & 0.015 & 0.000 & -1.181 \\
0.001 & 0.001 & 0.175 & 0.005 & 3.609 \times 10^{-5} & -1.638 \times 10^{-5} & 0.001 & -0.002 & 0.000 & 3.402 \times 10^{-5} & 0.001 & -0.002 & 0.000 & 6.313 \times 10^{-6} & 0.007 & 0.000 & -0.079 \\
0.000 & 0.000 & 0.000 & -0.006 & 2.125 \times 10^{-5} & -2.740 \times 10^{-5} & 0.000 & -0.002 & 0.000 & 0.000 & 0.000 & -0.001 & -0.001 & -0.001 & 0.000 & -0.001 \\
0.002 & 0.000 & -0.229 & -0.002 & 3.400 \times 10^{-5} & -7.490 \times 10^{-5} & 0.000 & -0.003 & 0.000 & -0.001 & -0.001 & -0.001 & 0.000 & -0.002 & 3.614 \times 10^{-5} & 0.191 \\
-0.002 & -0.001 & -0.347 & -0.005 & 1.448 \times 10^{-5} & -3.903 \times 10^{-5} & 0.000 & -0.002 & -0.001 & 0.000 & 0.000 & -0.001 & -0.001 & -0.001 & -0.001 & 0.000 & -0.001 \\
-0.003 & -0.001 & -0.347 & -0.005 & 1.448 \times 10^{-5} & -3.903 \times 10^{-5} & 0.000 & -0.002 & -0.001 & 0.000 & 0.000 & -0.001 & -0.001 & -0.001 & -0.001 & -0.001 & 0.123 \\
-0.002 & -0.001 & -0.340 & -0.005 & 2.207 \times 10^{-5} & -1.804 \times 10^{-5} & 0.000 & -0.002 & 0.001 & 0.000 & -0.002 & 0.000 & -0.002 & 0.000 & -0.002 & 0.000 & -0.144 \\
0.000 & 2.937 \times 10^{-5} & -9.040 & -0.004 & 3.117 \times 10^{-5} & -3.088 \times 10^{-5} & 0.000 & -0.002 & -0.002 & 0.000 & 0.000 & 0.000 & 0.000 & 7.349 \times 10^{-5} & -0.001 & 0.000 & 0.000 & 0.079 \\
0.002 & 0.001 & 0.274 & -0.002 & 3.902 \times 10^{-5} & 0.000 & 0.000 & 0.000 & 9.900 \times 10^{-6} & -1.099 \times 10^{-5} & 0.002 & 1.672 \times 10^{-5} & 0.000 & -5.414 \times 10^{-6} & -0.001 & 0.000 & -0.018 \\
0.003 & 0.001 & 0.382 & 0.004 & 4.566 \times 10^{-5} & 0.000 & 0.000 & -0.002 & 0.000 & -7.228 \times 10^{-5} & 0.002 & 0.591 & 0.000 & -1.737 \times 10^{-5} & -0.001 & 0.000 & -0.120 \\
0.003 & 0.001 & 0.384 & 0.008 & 4.560 \times 10^{-5} & 0.000 & 0.000 & -0.002 & -0.004 & 0.000 & -5.707 \times 10^{-5} & 0.001 & 0.000 & -2.506 \times 10^{-5} & -7.775 \times 10^{-5} & 0.022 & 0.000 & -0.265 \end{bmatrix}
\]
3.3 Performance Prediction on $Y_i$

3.3.1 Seasonal factor decomposition
As can be seen from Section 2, $Y_i$ shows a cyclical downtrend. We can analyze this from the seasonal factor decomposition, and Fig 4 shows the seasonal factor decomposition results. The effect of seasonal factor item on time series is fixed and remains unchanged every year. The error term is a random value around 0. Due to that the appearance of error is random, we mainly explore the trend circulation item in time series. As we can see from Fig 4 (3), the trend circulation item shows a fluctuant downtrend.

3.3.2 Time series analysis of $Y_i$
We use exponential smoothing method to predict the result after 20 years. The result respectively figured by (1) Simple non-seasonal model (Single exponential smoothing), (2) Simple seasonal model (Second exponential smoothing method) and (3) Holt-winters additive model (Triple exponential smoothing) show in Fig.5.

It can be seen that if we use a smoothing method, it is not able to predict the time series, and the predicted value is a straight line. Table 3 lists $R^2$, root-mean-square error (RMSE), Mean Absolute error percentage (MAPE), maximum absolute error percentage(MaxAPE), mean absolute error (MAE), maximum absolute error (MaxAE) and Standardized BIC.

<table>
<thead>
<tr>
<th></th>
<th>$R^2$</th>
<th>RMSE</th>
<th>MAPE</th>
<th>MaxAPE</th>
<th>MAE</th>
<th>MaxAE</th>
<th>Standardized BIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Second Exponential Smoothing Method</td>
<td>0.993</td>
<td>0.031</td>
<td>0.237</td>
<td>0.689</td>
<td>0.026</td>
<td>0.074</td>
<td>-6.973</td>
</tr>
<tr>
<td>Holt-winters Additive Model</td>
<td>0.994</td>
<td>0.030</td>
<td>0.226</td>
<td>0.634</td>
<td>0.025</td>
<td>0.070</td>
<td>-6.973</td>
</tr>
</tbody>
</table>

Simple seasonal model (Second exponential smoothing method) predicts the time series successfully, and the value of $R^2$ is 0.993. Holt-winters additive model (Triple exponential smoothing) also predict the time series successfully, and we can see the value of $R^2$ is 0.994. As is evident from the above methods, the Holt-winters additive model gets the maximum $R^2$ and the best prediction.

3.4 Performance Prediction on $Y_4$
3.4.1 Seasonal factor decomposition

As can be seen from Section 2, $Y_4$ basically changes circularly and periodically, and the specific trend can be judged by seasonal decomposition.

From Fig. 6, after seasonal decomposition, the effect of Seasonal factor item on time series is fixed and remains unchanged every year. The error term is a random value around 0. Fig. 6 (3) shows that the trend circulation item also displays a fluctuant declining trend.

3.4.2 Time Series Analysis of $Y_4$

The prediction of $Y_4$ is performed by the Winters additive model in exponential smoothing method, and its result is shown in Fig. 7.

It can be seen from Fig. 7 that this Winters additive model has successfully predicted the time series $Y_4$, the Table 4 provides all the same statistical indexes as Table 3, and the value of R2 is 1, quite ideal.

3.5 Performance Prediction on $Y_5$

3.5.1 Seasonal Factor Decomposition

The results of seasonal factor decomposition for $Y_5$ are shown in Fig. 8. The results are similar with that of $Y_4$. 
3.4.1 Seasonal factor decomposition

As can be seen from Section 2, \( Y \) basically changes circularly and periodically, and the specific trend can be judged by seasonal decomposition.

(1) Stochastic error item
(2) Seasonal factor item
(3) Trend circulation item

Fig. 6. Results of seasonal factor decomposition of \( Y \)

From Fig. 6, after seasonal decomposition, the effect of Seasonal factor item on time series is fixed and remains unchanged every year. The error term is a random value around 0. Fig. 6 (3) shows that the trend circulation item also displays a fluctuant declining trend.

3.4.2 Time Series Analysis of \( Y \)

The prediction of \( Y \) is performed by the Winters additive model in exponential smoothing method, and its result is shown in Fig. 7, and all statistical indexes of the prediction are listed in Table 5.

Fig. 7 Predictive results of by Time series analysis

<table>
<thead>
<tr>
<th>Table 5. Various statistical indexes</th>
</tr>
</thead>
<tbody>
<tr>
<td>R(^2)</td>
</tr>
<tr>
<td>--------</td>
</tr>
<tr>
<td>Holt-winters Additive Model</td>
</tr>
</tbody>
</table>

4. Conclusion

In this paper, we have shown that the collector area has been triumphantly predicted by Feature Selection based on variance and Pearson correlation coefficient together with CART with the average error of 3.031, the error rate of 0.97% and the RMSE of 4.044. The average monthly per meter borehole heat exchange has been predicted by multi-linear regression, and the relationship between average monthly per meter borehole heat exchange of all 12 months and 17 inputs are formulated as a matrix product form. The average monthly temperature of the circulating fluid, circulating fluid inlet temperature of the heat pump, circulating fluid outlet temperature of the heat pump has been successfully analyzed by seasonal factor decomposition, and their performances have been estimated by exponential smoothing. Experimental results demonstrate that the CART, multi-linear regression, seasonal factor decomposition and exponential smoothing are promising for practical applications in predicting performances of the solar-assisted GCHP systems.

Acknowledgements

This work was supported by National Natural Science Foundation of China (Grant No. 51708339), the China Postdoctoral Science Foundation Funded Project (Grant No.2017M612303), Shandong University of Science and Technology Plan Projects (Grant No. J15LG03), Shandong Co-Innovation Center of Green Team Construction Funds (Grant No. LSXT201519) and the China Scholarship Council.
Reference