

Article

Land Subsidence Prediction Model Based on the Long Short-Term Memory Neural Network Optimized Using the Sparrow Search Algorithm

Peicheng Qiu, Fei Liu and Jiaming Zhang *

Faculty of Civil Engineering and Mechanics, Kunming University of Science and Technology, Kunming 650504, China; qiupeicheng@stu.kust.edu.cn (P.Q.); liufei123@stu.kust.edu.cn (F.L.)

* Correspondence: zhangjiaming@kust.edu.cn

Abstract: Land subsidence is a prevalent geological issue that poses significant challenges to construction projects. Consequently, the accurate prediction of land subsidence has emerged as a focal point of research among scholars and experts. Traditional mathematical models exhibited certain limitations in forecasting the extent of land subsidence. To address this issue, the sparrow search algorithm (SSA) was introduced to optimize the efficacy of the long short-term memory (LSTM) neural network in land subsidence prediction. This prediction model has been successfully applied to the Huanglong Commercial City project in the Guanghua unit of Wenzhou city, Zhejiang province, China, and has been compared with the predictions of other models. Using monitoring location 1 as a reference, the MAE, MSE, and RMSE of the test samples for the LSTM neural network optimized using the SSA are 0.0184, 0.0004, and 0.0207, respectively, demonstrating a commendable predictive performance. This new model provides a fresh strategy for the land subsidence prediction of the project and offers new insights for further research on combined models.

Keywords: sparrow search algorithm (SSA); LSTM; land subsidence prediction; combined models



Citation: Qiu, P.; Liu, F.; Zhang, J. Land Subsidence Prediction Model Based on the Long Short-Term Memory Neural Network Optimized Using the Sparrow Search Algorithm. *Appl. Sci.* **2023**, *13*, 11156. <https://doi.org/10.3390/app132011156>

Academic Editor: Yosoon Choi

Received: 17 September 2023

Revised: 30 September 2023

Accepted: 9 October 2023

Published: 11 October 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The importance of monitoring land subsidence in construction projects cannot be overlooked, as neglecting this aspect could lead to severe consequences [1,2]. Firstly, failing to monitor land subsidence can pose safety hazards at construction sites, such as collapses or landslides, endangering the lives of workers and damaging equipment. Secondly, after the completion of the project, structural issues such as cracks or wall deformations may arise, increasing maintenance costs. Even more seriously, the project party might face legal litigation, bear compensation, and suffer damage to their reputation as a result. Therefore, monitoring land subsidence has always been a crucial step in ensuring construction safety and project quality.

Currently, numerous scholars and researchers have delved deeply into the prediction of land subsidence, and the research approaches can be broadly categorized into two primary types. The first is prediction methods based on independent models. For instance, Zhang et al. [3] analyzed the excessive surface subsidence caused by pit excavation using various machine learning algorithms. Chen et al. [4] employed three neural network methods to determine the model best suited for predicting maximum ground settlement during tunnel construction. Mahmoodzadeh et al. [5] conducted research on the maximum surface subsidence of urban tunnels using seven intelligent methods: long short-term memory (LSTM), deep neural network (DNN), K-nearest neighbor (KNN), Gaussian process regression (GPR), support vector regression (SVR), decision tree (DT), and linear regression (LR). Zhang et al. [6], addressing the issue of ground subsidence during pit construction, proposed a daily land subsidence prediction method based on artificial neural networks. They demonstrated that time-related influencing factors and

previous subsidence monitoring data play a crucial role in predicting daily ground subsidence. Tang et al. [7] integrated four different machine learning methods, including support vector machine (SVM), random forest (RF), back propagation neural network (BPNN), and deep neural network (DNN), to estimate the maximum surface subsidence caused by tunnel excavation. Although independent models had achieved some results in predicting land subsidence, their prediction accuracy still needed improvement. Therefore, many researchers began to explore another prediction strategy, namely, using combined models for prediction. For example, Chen et al. [8] proposed a prediction method for the maximum surface subsidence during shield tunneling, using the BPNN and random forest algorithm (RF) machine learning algorithms, with the optimal hyperparameters determined through the particle swarm optimization (PSO). Moghaddasi et al. [9] introduced a novel hybrid model of artificial neural networks optimized using the imperialist competitive algorithm (ICA-ANN) and precisely predicted the maximum surface subsidence caused by surface structural damage and environmental issues. Yang et al. [10] established a hybrid model by combining three types of meta-heuristic algorithms: ant lion optimizer (ALO), multi-verse optimizer (MVO), and grasshopper optimization algorithm (GOA) to accurately predict the land subsidence caused by underground tunnel construction. Cao et al. [11] used the LSTM to predict land subsidence at typical locations in different subsidence development areas, providing reliable technical support for land subsidence prevention and control.

Existing research clearly showed that ensemble models outperform single models in prediction accuracy, making them particularly favored in land subsidence forecasting. Currently, forecasting efforts are mainly based on existing data related to land subsidence to predict future trends. Given that land subsidence data are often closely associated with time series, selecting an appropriate model is crucial for enhancing forecasting accuracy. The LSTM neural network, as an optimized and improved algorithm, not only addresses the challenge of vanishing gradients during backpropagation but also preserves both long and short-term sequence dependencies. Therefore, LSTM offers a significant advantage when dealing with foundational time series data like land subsidence. However, the prediction accuracy of the LSTM neural network depends on the determination of its parameters, necessitating the optimization of the LSTM neural network [12–14]. To enhance the accuracy of the LSTM model in forecasting land subsidence, this study employs the SSA to optimize the LSTM's learning rate, hidden layer nodes, and regularization coefficient. SSA is a bio-inspired optimization method innovatively introduced by Xue et al. [15] from Donghua University in 2020. It boasts exceptional convergence properties and high precision. This algorithm has showcased its superior performance in model parameter adjustments. Ma et al. [16] developed an advanced LSTM forecasting model to boost the utilization of clean energies like wind power in integrated energy systems, resulting in enhanced economic and environmental benefits. Leng et al. [17] introduced a high-accuracy time series forecasting method for assessing the lifespan of insulated gate bipolar transistors (IGBTs) based on an LSTM network optimized using the sparrow search algorithm. Zu et al. [18] presented a greenhouse environment forecasting model based on an advanced LSTM network, achieving precise predictions of greenhouse environment data. Yu et al. [19] proposed a model based on the sparrow search algorithm and LSTM neural network, demonstrating its significant advantage in precisely predicting soil oxygen content changes after SSA optimization.

Differing from existing research, this study introduces the SSA into the parameter optimization of the LSTM neural network forecasting model, aiming to find the optimal hyperparameter combination. The sparrow search algorithm, by optimizing parameters including the node count in hidden layers and initial learning rates, substantially diminishes the manual parameter tuning workload traditionally seen in LSTM neural network model predictions, thereby improving the precision of land subsidence forecasting. Based on current literature review results, there are still few related research reports available. In this study, the SSA-LSTM model showed good prediction accuracy in the land subsidence

prediction, which provided a new and scientific idea for the monitoring and control of land subsidence in the region.

2. Methods

2.1. Sparrow Search Algorithm (SSA)

The SSA represents a fresh swarm optimization technique influenced by the foraging, anti-predatory, and vigilant behaviors of sparrows. In the algorithm's assumptions, the sparrow population is divided into discoverers, joiners, and scouts. Discoverers are responsible for the group's foraging, while joiners follow the discoverer with the best fitness, monitoring them and seizing opportunities to snatch food. During the foraging process, if scouts sense danger, they will alert the entire flock to quickly fly to a safe area for feeding. The position update for the sparrow acting as the discoverer is shown in Equation (1) [20]:

$$X_{i,j}^{t+1} = \begin{cases} X_{i,j}^t \cdot \exp\left(\frac{-i}{\alpha \cdot \text{iter}_{\max}}\right) & , R_2 < ST \\ X_{i,j}^t + Q \cdot L & , R_2 \geq ST \end{cases} \quad (1)$$

where $X_{i,j}^t$ represents the position of the sparrow in the t th iteration; α is the random number in $(0, 1]$; $\text{iter}_{\max}$ represents the maximum number of iterations; Q is a random number that follows a standard normal distribution; L represents a $1 \times d$ matrix with all elements being 1; R_2 is the alarm value, $R_2 \in [0, 1]$; and as for the alert value (ST), $ST \in (0.5, 1]$. The position update for the sparrow acting as a joiner is presented in Equation (2):

$$X_{i,j}^{t+1} = \begin{cases} Q \cdot \exp\left(\frac{X_{\text{worst}}^t - X_{i,j}^t}{\alpha \cdot \text{iter}_{\max}}\right) & , i > \frac{n}{2} \\ X_p^{t+1} + |X_{i,j}^t - X_p^{t+1}| \cdot A^+ \cdot L & , \text{other} \end{cases} \quad (2)$$

where X_p^{t+1} represents the best position for the $t + 1$ th generation of discoverers; X_{worst}^t is the worst position in the t th generation; n is the total number of sparrows; and A represents a $1 \times d$ matrix with all elements being either 1 or -1 ; $A^+ = A^T(AA^T)^{-1}$. The position update for the sparrow acting as a scout is presented in Equation (3):

$$X_{i,j}^{t+1} = \begin{cases} X_{\text{best}}^t + \beta \cdot |X_{i,j}^t - X_{\text{best}}^t| & , f_i > f_g \\ X_{i,j}^t + K \cdot \frac{|X_{i,j}^t - X_{\text{worst}}^t|}{(f_i - f_w) + \varepsilon} & , f_i = f_g \end{cases} \quad (3)$$

where X_{best}^t denotes the best position of the sparrow within the population in the t th generation; β is a random number from a normal distribution with a mean of 0 and a variance of 1; f_i represents the current fitness value of the i th sparrow; f_g and f_w denote the best and worst fitness values within the current sparrow population, respectively, and K is a random number belonging to $[-1, 1]$.

2.2. LSTM Neural Network

The LSTM network, introduced by Hochreiter and Schmidhuber in 1997 [21], is an optimization of the traditional RNN. Its primary innovation lies in the introduction of gate mechanisms to control long-term states, effectively addressing the long-term dependency issue found in RNNs. This improvement has led to the widespread application of LSTM in various domains, such as speech recognition and image processing, as confirmed by the studies of Vinyals et al. [22] and Soltan et al. [23]. The architecture of the LSTM network predominantly features an input layer, hidden layer, and output layer illustrated in Figure 1.

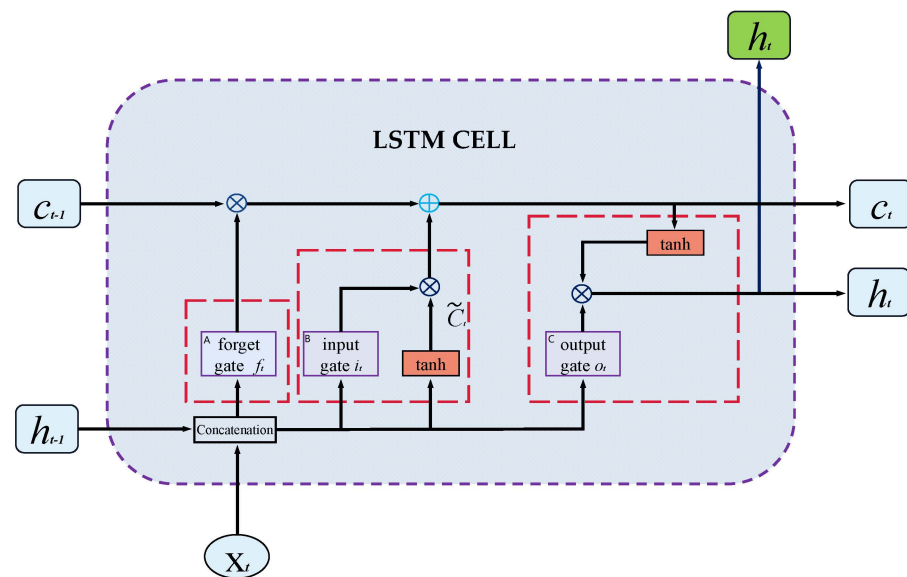


Figure 1. LSTM structural framework.

In Figure 1, A, B, and C represent three nonlinear gate functions termed the forget gate, the input gate, and the output gate, respectively. Unlike traditional RNNs that directly memorize all historical information, LSTM selectively updates and discards historical data through three gates. The algorithm for LSTM is shown in Equations (4)–(9) [24]:

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (4)$$

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (5)$$

$$\tilde{C}_t = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \quad (6)$$

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t \quad (7)$$

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad (8)$$

$$h_t = o_t * \tanh(C_t) \quad (9)$$

In the equations, W_f , W_i , W_c , and W_o are the weight matrices for the input vector x_t . b_f , b_i , b_c , and b_o represent the bias vectors at time t . $\tilde{C}_t$ represents the candidate neuron state at the current time step, encompassing the hidden layer state h_{t-1} from the $t - 1$ time step and the neuron state C_{t-1} from the $t - 1$ time step. C_t is the current neuron state; h_t represents the hidden layer state at the current time step; and σ and $\tanh$ are nonlinear activation functions.

2.3. SSA-LSTM Land Subsidence Prediction Model

The workflow of the SSA-LSTM model consists of two parts: the left side represents the LSTM neural network, while the right side depicts the SSA, as shown in Figure 2. The SSA algorithm optimizes the neural network's learning rate, hidden layer nodes, and regularization coefficient. The global optimal position found by the algorithm is used as the initial structural parameters for the LSTM network.

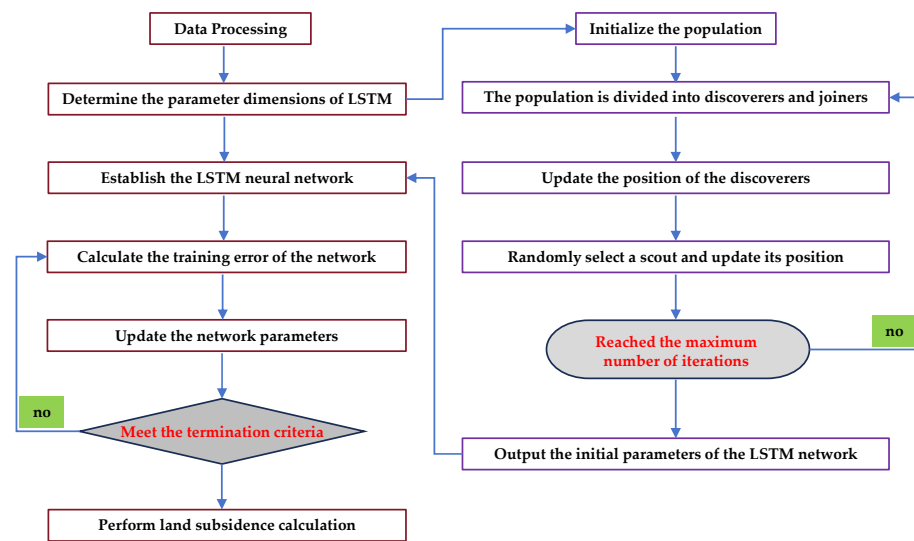


Figure 2. SSA-LSTM neural network flowchart.

2.4. Model Evaluation Metrics

In the prediction process, the evaluation of model performance is crucial [25–27]. Common evaluation metrics include: mean square error (MSE), mean absolute percentage error (MAPE), root mean square error (RMSE), and mean absolute error (MAE). Among them, MAE and RMSE are the two most commonly used metrics to measure the accuracy of variables. MAE is the average of absolute errors, while RMSE is the square root of MSE. Its unit aligns with the original data, making it easier to evaluate the model. By substituting the actual values and the model output values into the MAE and RMSE formulas, the larger the value, the greater the prediction error of the model. Based on the aforementioned two indicators, the mean square error (MSE) is additionally selected as an evaluation metric. MSE can display the degree of variation in the evaluation data, thereby reflecting the stability of the model. If the MSE value is smaller, the variation in the residuals is reduced, making the results of the model more stable. The detailed formulas for the three indicators can be found in Equations (10)–(12).

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2 \quad (10)$$

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n |(\hat{y}_i - y_i)|^2} \quad (11)$$

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |\hat{y}_i - y_i| \quad (12)$$

In the equation, $\hat{y}_i$ represents the predicted value for the i th sample; y_i represents the actual value for the i th sample; and n represents the total number of samples.

3. Data Processing and Parameter Determination

3.1. Engineering Overview and Data Acquisition

The Huanglong Commercial City project is located in the central-western part of the Wenzhou Basin, where the geological structure is relatively stable. Based on the drilling data, the site is covered with Quaternary loose sediments with a thickness of 37.60 to 55.00 m. The surface layer consists of miscellaneous fill soil, which is man-made fill. The upper layer, with a thickness of 20.00–28.50 m, is silt-soft soil, which is marine sediment; its era is the Holocene of the Quaternary period (Q4). The middle part consists of fluvial sediments, which are clayey soils; its era corresponds to the Middle Pleistocene stratum (Q3). The lower part consists of slope residual deposits, which are gravelly silty clay

layers; its era corresponds to the Early Pleistocene stratum (Q1). The underlying rock layer is the Upper Jurassic (J₃g) Gaowu Formation of extrusive volcanic rock, with the rock type mainly being tuff. The total land area of the project is 175,562 square meters. The site was originally the Huanglong Commercial City, which is the largest commodity trading market in southern Zhejiang and northern Fujian. The rate of land subsidence in the study area lacks a clear pattern, making it more challenging to accurately predict subsidence. Using the ground subsidence data from monitoring points in the research area between 4 January 2023 and 26 March 2023, this study established a land subsidence prediction model. By combining the prediction with existing ground subsidence data, it is possible to more accurately understand the subsidence trend of the monitoring point and its surrounding areas, offering a valuable reference for subsequent subsidence monitoring and mitigation strategies.

3.2. Processing Sample Data

In the investigated area, observational data from four monitoring points were chosen as the source for subsidence prediction. A subset of the data is seen in Table 1. Taking the data from monitoring point 1 as an example, according to the sample division strategy in Table 2, six consecutive data points were used as input, with the seventh data point serving as the target output, resulting in a total of 77 independent samples. The initial 67 data sets were utilized for training purposes, and the subsequent 10 sets served as test samples. To avoid the instability caused by anomalous sample data, the divided data were normalized. The processed training and testing samples are detailed in Tables 3 and 4, respectively. Although this data partitioning strategy might extend the training duration of the model, it effectively preserves the temporal characteristics of the data, maximizes data utilization, and reduces the risk of overfitting. Due to space limitations, only a portion of the data are displayed in the table.

Table 1. Subsidence value of the monitoring point.

Time	Subsidence Value (mm)			
	1	2	3	4
4 January 2023	−0.87	0.42	−0.08	0.03
5 January 2023	0.14	0.10	0.31	0.26
6 January 2023	−1.09	−0.73	0.03	−0.21
7 January 2023	−0.71	−0.32	−0.92	0.39
8 January 2023	0.33	0.28	−0.53	−0.98
9 January 2023	−0.62	−0.94	−0.88	0.36
10 January 2023	−0.48	−0.37	0.43	−0.83
11 January 2023	−0.88	0.37	−0.59	−0.92
12 January 2023	−0.33	−0.34	−1.03	−0.01
...	...	...	...	...
22 March 2023	−0.15	−0.56	0.03	−0.32
23 March 2023	0.25	−0.66	−0.3	−0.72
24 March 2023	−0.41	−0.02	0.30	0.07
25 March 2023	−0.16	−0.22	−0.34	0.36
26 March 2023	−0.65	−0.41	−0.33	−0.37

Table 2. Sample division strategy.

N Input	M Output
$X_1, X_2, \dots, X_N$	X_{N+1}
$X_2, X_3, \dots, X_{N+1}$	X_{N+2}
...	...
$X_K, X_{K+1}, \dots, X_{N+K-1}$	X_{N+K}

Table 3. Training set sample.

Number	Input Value						Expected Value
1	−0.711	0.618	−1.000	−0.500	0.868	−0.382	−0.197
2	0.618	−1.000	−0.500	0.868	−0.382	−0.197	−0.724
3	−1.000	−0.500	0.868	−0.382	−0.197	−0.724	0.000
4	−0.500	0.868	−0.382	−0.197	−0.724	0.000	0.605
5	0.868	−0.382	−0.197	−0.724	0.000	0.605	−0.737
6	−0.382	−0.197	−0.724	0.000	0.605	−0.737	0.711
7	−0.197	−0.724	0.000	0.605	−0.737	0.711	0.066
8	−0.724	0.000	0.605	−0.737	0.711	0.066	−0.382
9	0.000	0.605	−0.737	0.711	0.066	−0.382	0.539
10	0.605	−0.737	0.711	0.066	−0.382	0.539	−0.395
11	−0.737	0.711	0.066	−0.382	0.539	−0.395	0.816
12	0.711	0.066	−0.382	0.539	−0.395	0.816	−0.566
13	0.066	−0.382	0.539	−0.395	0.816	−0.566	−0.592
14	−0.382	0.539	−0.395	0.816	−0.566	−0.592	0.434
15	0.539	−0.395	0.816	−0.566	−0.592	0.434	0.961
16	−0.395	0.816	−0.566	−0.592	0.434	0.961	0.368
17	0.816	−0.566	−0.592	0.434	0.961	0.368	0.526
18	−0.566	−0.592	0.434	0.961	0.368	0.526	−0.329
19	−0.592	0.434	0.961	0.368	0.526	−0.329	−0.329
20	0.434	0.961	0.368	0.526	−0.329	−0.329	−0.618

Table 4. Test set sample.

Number	Input Value						Expected Value
1	−0.382	−0.421	0.539	0.921	0.921	0.921	0.763
2	−0.421	0.539	0.921	0.921	0.921	0.763	0.158
3	0.539	0.921	0.921	0.921	0.763	0.158	−0.013
4	0.921	0.921	0.921	0.763	0.158	−0.013	0.934
5	0.921	0.921	0.763	0.158	−0.013	0.934	0.618
6	0.921	0.763	0.158	−0.013	0.934	0.618	0.237
7	0.763	0.158	−0.013	0.934	0.618	0.237	0.763
8	0.158	−0.013	0.934	0.618	0.237	0.763	−0.105
9	−0.013	0.934	0.618	0.237	0.763	−0.105	0.224
10	0.934	0.618	0.237	0.763	−0.105	0.224	−0.421

3.3. Determination of SSA-LSTM Model Parameters

The sparrow search algorithm is a collective intelligence optimization method unveiled in 2020. Since its inception, it has received widespread attention in the academic community. However, in practical applications, parameter selection for the algorithm often relies on experience, and the optimal parameters might vary for different models. The key parameters of the sparrow search algorithm include population size, number of iterations, and proportion of discoverers, among others [28–30]. Inappropriate parameter settings might result in decreased prediction performance or excessively long training times. To circumvent these issues, the potential range of parameters is initially determined based on the existing data. Subsequently, numerical experiments are employed to precisely ascertain the optimal value for each parameter.

In the application of the SSA, the selected population size is typically between 30 and 100. To further investigate how population size affects the algorithm's efficacy, experiments were conducted every 10 units. Under the premise of fixed iteration times and other parameters, the fitness values and prediction durations under different population sizes were analyzed. As shown in Figure 3, with the increase in the population size, the fitness reaches its optimum when the population size is 30. Afterward, the optimal fitness shows an initial increase followed by a decrease, indicating that the performance of the sparrow search algorithm is not entirely positively correlated with the population size. Moreover, with the growth of the population size, the computation time also shows a significant

increasing trend. Therefore, considering both time and performance, 30 is determined to be the optimal population size.

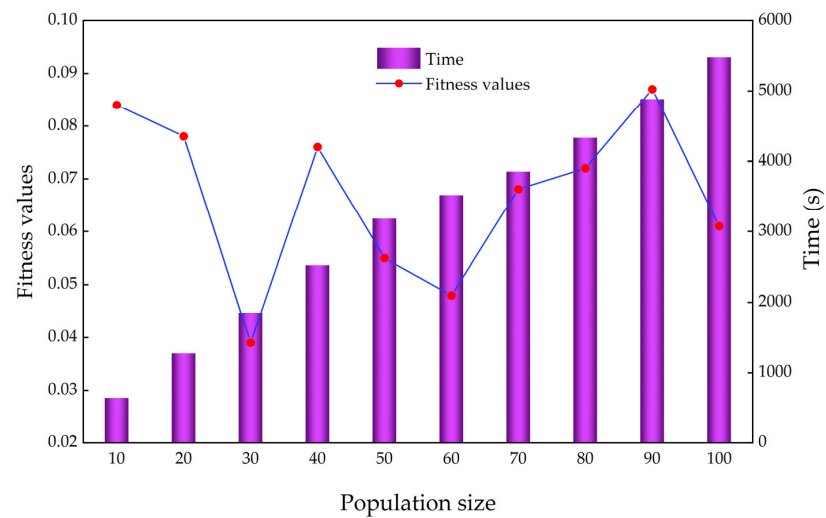


Figure 3. Determination of the sparrow population.

On the other hand, if the number of iterations is set too high, it might lead to overfitting, where the algorithm performs exceptionally well on the training data but poorly on new, unseen data. Furthermore, excessive iterations could also lead to a waste of computational resources and time. Therefore, it is crucial to find an optimal number of iterations that provides a balance between accuracy and computational efficiency. Conversely, excessive iteration might lead the algorithm to over-iterate before reaching the optimal solution, which is not only counterproductive but also increases computational time. Therefore, selecting an appropriate number of iterations is crucial for ensuring both prediction efficiency and accuracy. Figure 4 illustrates the fitness values under different hidden layers of LSTM. The results indicate that under the given input data dimension and varying numbers of hidden layers, the SSA algorithm could converge to a stable state within 25 iterations. Thus, the maximum number of iterations for the SSA algorithm was set to 25. As for the alert value (ST), its range typically lies between 0.5 and 1.0. Figure 5 indicates that when the alert value was 0.7, the fitness achieved its optimum. The number of producers (PDs) ranged from [0, 1]. Figure 6 illustrates the fitness values at different PD levels, and the results indicate that the fitness was optimal when the PD was set to 0.4.

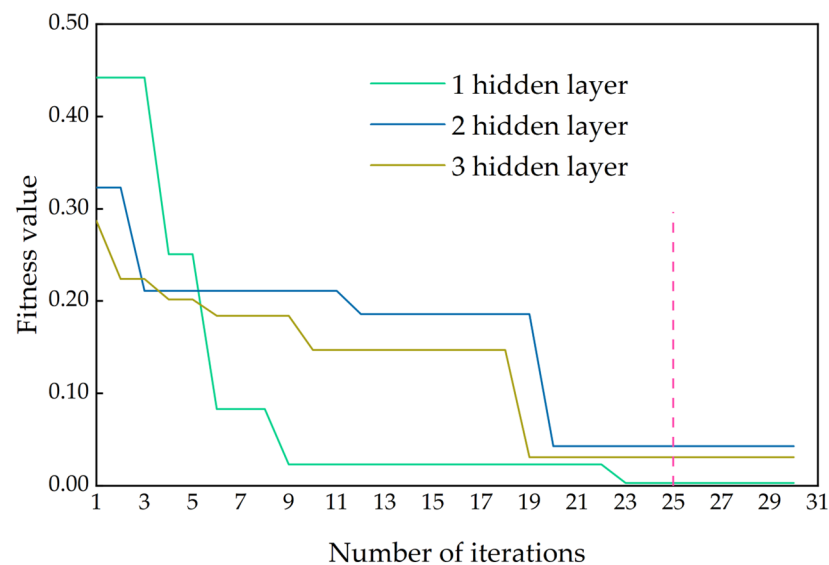


Figure 4. Determination of the number of iterations.

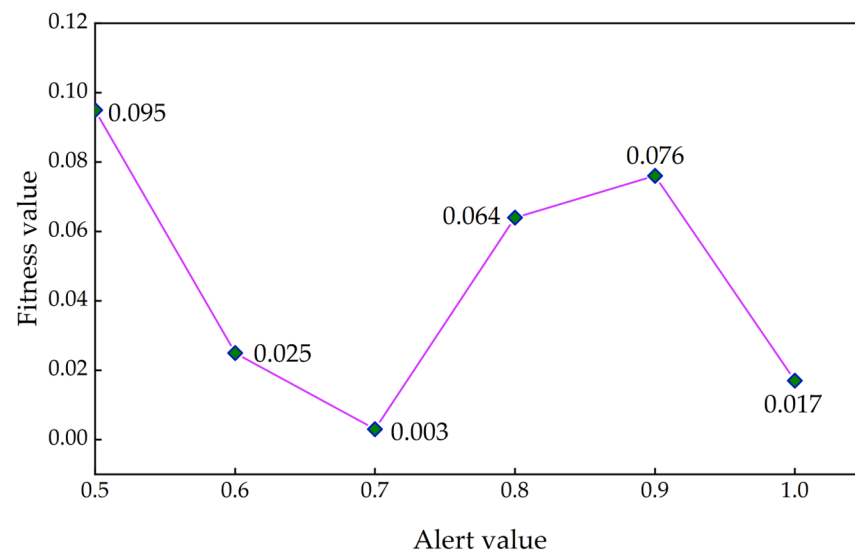


Figure 5. Determination of the number of alert values.

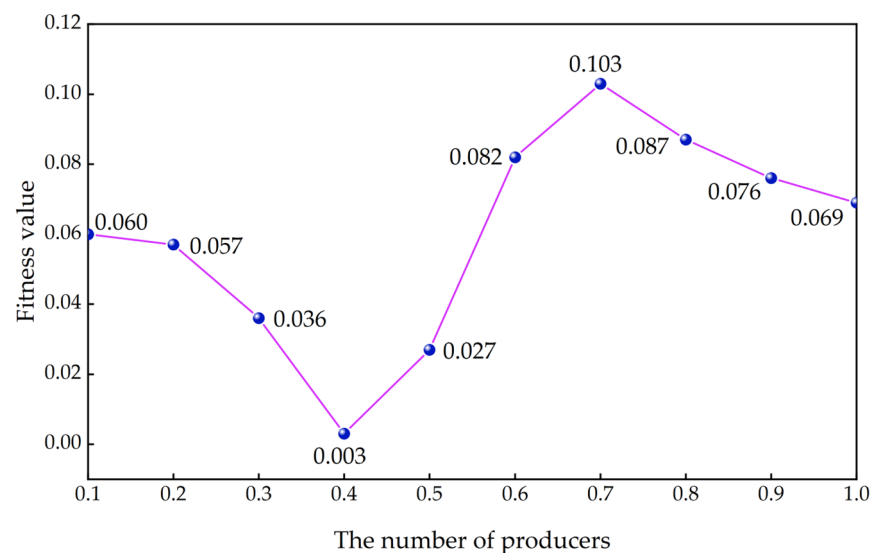


Figure 6. Determination of the number of PDs.

In the LSTM neural network architecture, the number of nodes in the hidden layer is a crucial parameter. Increasing the number of nodes in the hidden layer can enhance the predictive capability of the model and reduce errors. However, too many nodes in the hidden layer may lead to the problem of “overfitting” during training. Although there are many studies discussing how to determine the count of nodes within the hidden layer, there is currently no universally effective method. Some proposed calculation methods might only be suitable for large sample data. In real applications, specific projects might necessitate distinct hidden layer configurations. To ensure the network’s performance and generalization capacity, the number of nodes selected should be as minimal as possible, while still meeting the accuracy criteria. As illustrated in Figure 7, when the number of nodes in the hidden layer ranged from 1 to 10, the prediction accuracy of the LSTM network fluctuated, with the mean absolute error first increasing and then decreasing. When the number ranged from 10 to 20, the mean absolute error began to rise, which could be attributed to overfitting of the network. Although the mean absolute error decreased in the 20 to 30 range, to avoid overfitting, the chosen number of nodes in the hidden layer was set to 10.

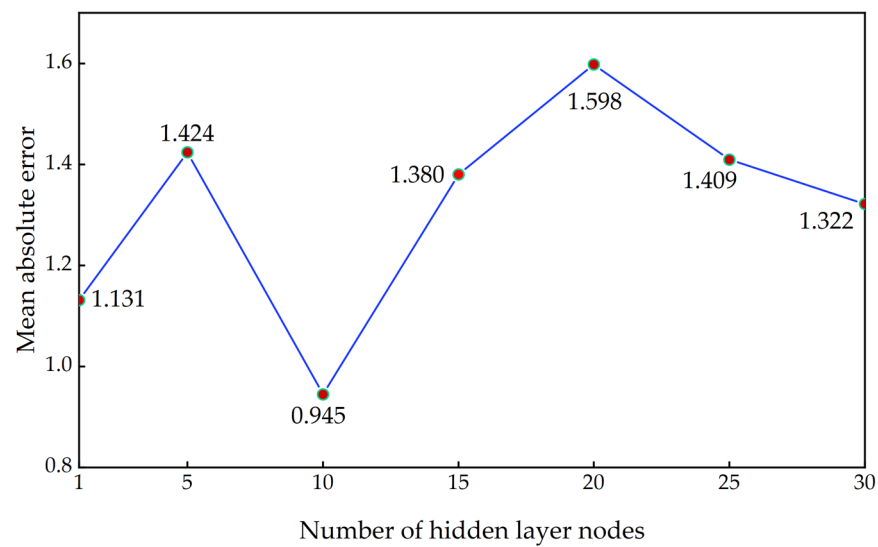


Figure 7. Determination of the number of hidden layer nodes.

After a series of numerical experiments, the following parameter configuration was determined for the sparrow search algorithm in the ground subsidence prediction model: a population size of 30, a maximum iteration number of 25, an alert threshold of 0.7, a producer ratio of 0.4, and a vigilance threshold (SD) typically set at 0.2. Simultaneously, the number of nodes in the hidden layer of the LSTM neural network was set to 10.

4. Results and Discussion

4.1. Prediction and Analysis of Land Subsidence

After applying the SSA-LSTM neural network to the training and testing samples of monitoring point 1, the output of the network was obtained. Using de-normalization and comparison with the actual values, the prediction residuals were calculated. As shown in Table 5, the model outputs the predicted settlement values, actual values, and their prediction residuals for the training samples. Based on Table 5 and Equations (10)–(12), the MAE, MSE, and RMSE for the LSTM neural network training samples were calculated to be 0.3650, 0.1456, and 0.3816, respectively. The corresponding values for the SSA-LSTM neural network were 0.0216, 0.0008, and 0.0278, respectively. By comparing the evaluation metrics of the two models, it is evident that the SSA-LSTM has significantly improved in terms of prediction accuracy, resulting in more stable predictions. Some of the model's output for test sample settlement predictions, actual values, and prediction residuals are presented in Table 6.

Table 5. The results of the training set prediction at monitoring point 1.

Number	True Value	Subsidence Value (mm)			
		Predicted Value 1	Predicted Value 2	Residual Value 1	Residual Value 2
1	−0.48	−0.13	−0.43	−0.35	−0.05
2	−0.88	−0.70	−0.86	−0.18	−0.02
3	−0.33	−0.03	−0.35	−0.30	0.02
4	0.13	0.60	0.13	−0.47	0.00
5	−0.89	−0.64	−0.81	−0.25	−0.08
6	0.21	0.64	0.16	−0.43	0.05
7	−0.28	0.03	−0.31	−0.31	0.03
8	−0.62	−0.35	−0.60	−0.27	−0.02
9	0.08	0.53	0.07	−0.45	0.01
10	−0.63	−0.38	−0.62	−0.25	−0.01

Note: Value 1 represents LSTM prediction, while Value 2 represents SSA-LSTM prediction.

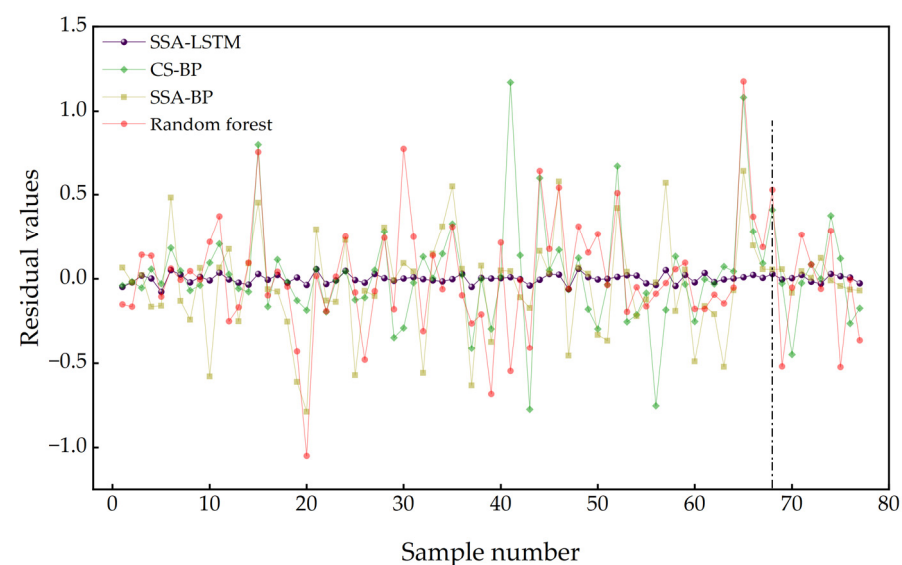
Table 6. The results of test set prediction at monitoring point 1.

Number	True Value	Subsidence Value (mm)		Residual Value 1	Residual Value 2
		Predicted Value 1	Predicted Value 2		
1	0.25	0.73	0.22	−0.48	0.03
2	−0.21	0.16	−0.21	−0.37	0.00
3	−0.34	−0.02	−0.34	−0.32	0.00
4	0.38	0.90	0.36	−0.52	0.02
5	0.14	0.64	0.16	−0.50	−0.02
6	−0.15	0.27	−0.12	−0.42	−0.03
7	0.25	0.73	0.22	−0.48	0.03
8	−0.41	−0.13	−0.43	−0.28	0.02
9	−0.16	0.21	−0.17	−0.37	0.01
10	−0.65	−0.38	−0.62	−0.27	−0.03

Note: Value 1 represents LSTM prediction, while Value 2 represents SSA-LSTM prediction.

Based on Table 6 and Equations (10)–(12), the MAE, MSE, and RMSE values of the LSTM neural network on the test samples were 0.4020, 0.1693, and 0.4115, respectively. In contrast, the MAE, MSE, and RMSE values of the test samples from the SSA-optimized LSTM neural network were 0.0184, 0.0004, and 0.0207, respectively. Comparing the prediction evaluation indicators of the two models on the test samples, it was evident that the SSA-LSTM model notably surpassed the original LSTM model regarding forecast precision and robustness. Therefore, based on these evaluation results, the SSA-LSTM model is more suitable as a prediction model for land subsidence.

When the sample data were input into other commonly used models, as shown in Figure 8, the residual values of the SSA-LSTM model were distributed around 0 with minimal fluctuations. The residual values of the other three models had shown larger fluctuations, indicating poor prediction accuracy. The SSA-LSTM model had the highest and most consistent prediction accuracy, further demonstrating its suitability for predicting ground settlement.

**Figure 8.** Comparison of prediction effect of different models.

4.2. Generalization Capability of the SSA-LSTM Model

To delve deeper into the model's generalization potential, data from monitoring points 2, 3, and 4 were, respectively, input into the SSA-LSTM model for prediction and compared with the prediction results from monitoring point 1. The corresponding prediction results are shown in Figures 9–12. By combining the three evaluation indexes MAE, MSE, and

RMSE, the predictive accuracy of the model was assessed. Detailed comparisons can be seen in Table 7. From Figures 9–12, it was clearly observed that the prediction error ranges for monitoring points 1 to 4 were, respectively, $(-0.0753$ to $0.0593)$, $(-0.0685$ to $0.1237)$, $(-0.0524$ to $0.0579)$, and $(-0.0876$ to $0.0774)$. These findings indicate that the fluctuations in prediction errors were relatively minor, attesting to the model's stable performance. Based on the data from Table 7, the SSA-LSTM model excelled in its predictions for monitoring point 4, while it also performed commendably on the other three monitoring points. This further validated the high accuracy, stability, and superior generalization capability of the SSA-LSTM model in predicting ground settlement, aligning with the conclusions drawn in Section 2.4. Hence, with its robust and accurate predictive qualities, this model can provide strong decision-making support for early warning and mitigation of land subsidence.

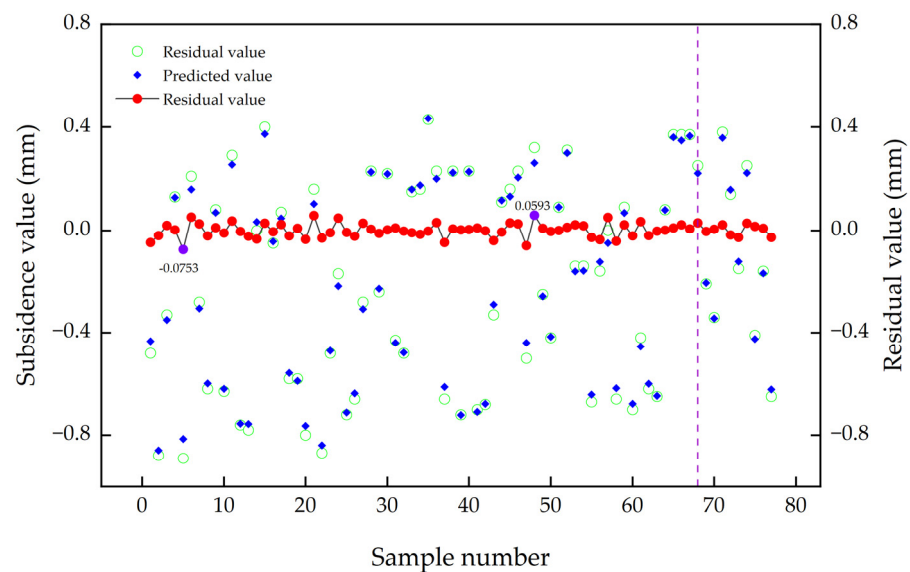


Figure 9. Prediction results at monitoring point 1.

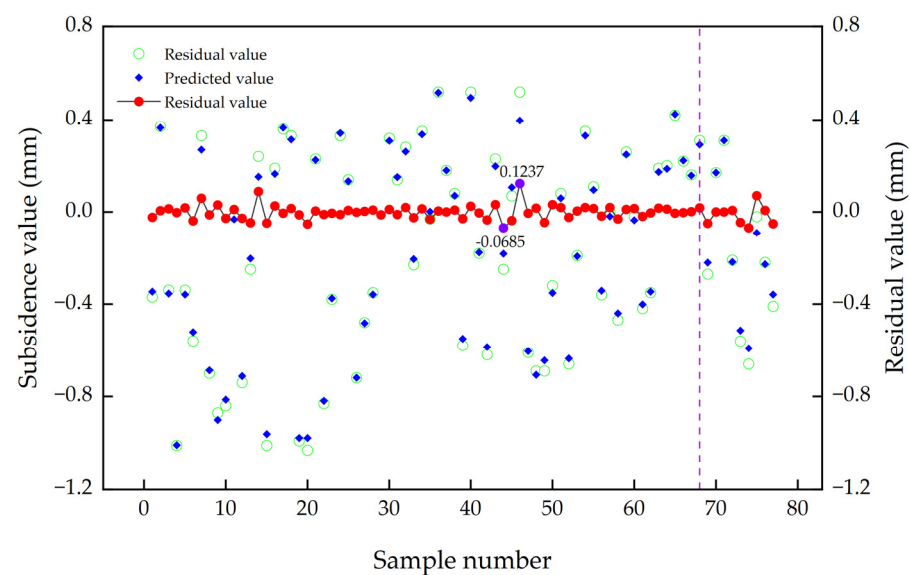


Figure 10. Prediction results at monitoring point 2.

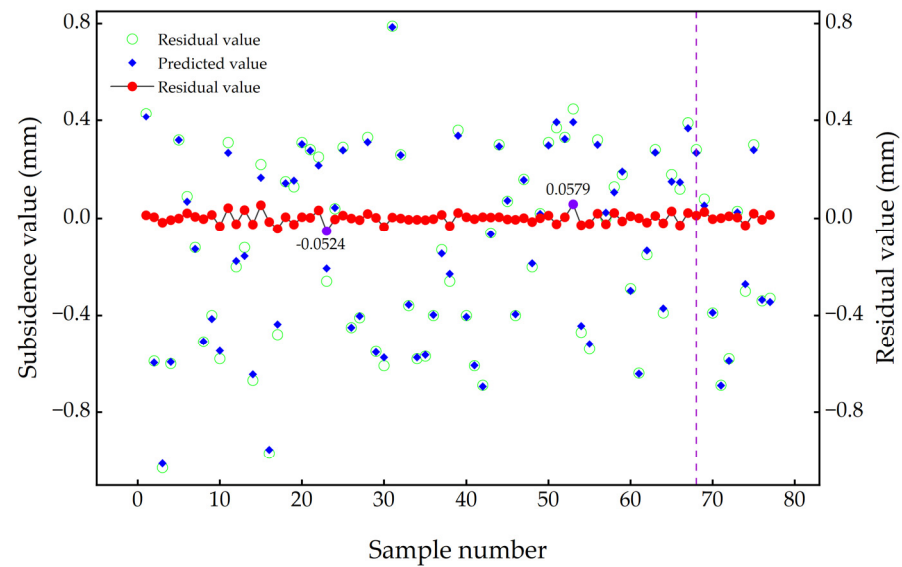


Figure 11. Prediction results at monitoring point 3.

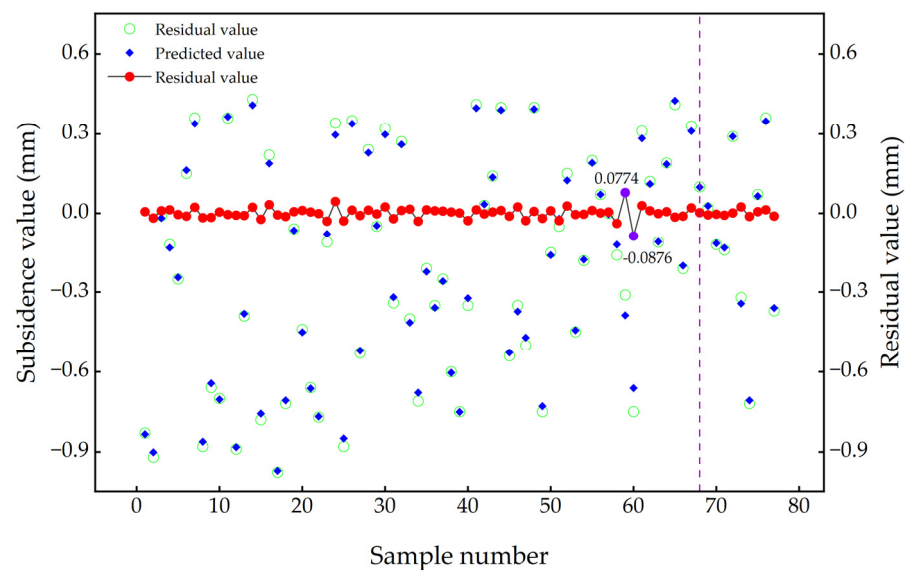


Figure 12. Prediction results at monitoring point 4.

Table 7. Evaluation indicators for prediction results of different monitoring points.

Point	MAE	MSE	RMSE
1	0.0212	0.0007	0.0269
2	0.0229	0.0010	0.0319
3	0.0155	0.0004	0.0204
4	0.0149	0.0004	0.0210

5. Conclusions

(1) The study employed numerical experimental strategies using variables such as the population size of the SSA, the iteration count, and the proportion of discoverers. A numerical relationship between each indicator and fitness was established, which was subsequently used to determine the optimal values for the model parameters. The results revealed that when the swarm count was 30, the iteration count was 25, the alert value was 0.7, the PD was 0.4, and the SD was 0.2 the SSA-LSTM ground settlement prediction model demonstrated optimal performance. The optimal parameters obtained through numerical

experiments can optimize the combined model more effectively, enabling it to be better applied in practical engineering applications.

(2) The experiments further indicate that combining multiple models often achieves better performance than any single model. For instance, ensemble models can effectively reduce the variance in the model, thereby decreasing the risk of overfitting. Ensemble models can combine the advantages of various models, performing better than a singular model. A single model might overly focus on a specific aspect of the data and neglect others, while ensemble models can achieve a better balance. Different models might make distinct mistakes. When combined, their errors tend to cancel each other out, resulting in a lower overall error rate. The outputs of ensemble models are generally more stable.

(3) From the evaluation of the predicted values at the monitoring points and the depiction of the prediction outcomes, it was clear that the SSA-LSTM model significantly improved prediction accuracy. Furthermore, the adaptive features of the SSA-LSTM model were fully utilized, demonstrating its strong generalization capability. This provided a novel method for predicting ground settlement in areas affected by multiple factors. In addition, compared with other combined models, this model simplified the operational process and enhanced its practicality in ground settlement prediction. However, when applying the SSA-LSTM model, the segmentation strategy for input and output data still needs further exploration and optimization.

Author Contributions: Conceptualization, P.Q., F.L. and J.Z.; methodology, J.Z.; software, F.L.; validation, P.Q. and F.L.; formal analysis, P.Q.; investigation, P.Q. and J.Z.; resources, F.L.; writing—original draft preparation, P.Q. and J.Z.; writing—review and editing, P.Q. and J.Z.; visualization, F.L.; supervision, F.L.; project administration, F.L.; funding acquisition, J.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Nature Science Foundation of China, grant number 41807258.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: All data included in this study are available upon request to the corresponding author.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Wang, W.; Xu, K.; Wang, H.; Li, L. Application of real-time monitoring technology to foundation settlement of sloping breakwaters in construction. *Chin. J. Geotech. Eng.* **2017**, *39*, 85–90. (In Chinese with English Abstract) [[CrossRef](#)]
2. Zhou, Z.; Heng, C.; Zhang, J.; Sun, X. Surface Settlement Characteristics of An Ultra-shallow-buried Rectangular Section Underpass. *J. Basic Sci. Eng.* **2020**, *28*, 900–912. (In Chinese with English Abstract) [[CrossRef](#)]
3. Zhang, R.; Li, Y.; Goh, A.T.C.; Zhang, W.; Chen, Z. Analysis of ground surface settlement in anisotropic clays using extreme gradient boosting and random forest regression models. *J. Rock Mech. Geotech. Eng.* **2021**, *13*, 1478–1484. [[CrossRef](#)]
4. Chen, R.-P.; Zhang, P.; Kang, X.; Zhong, Z.-Q.; Liu, Y.; Wu, H.-N. Prediction of maximum surface settlement caused by earth pressure balance (EPB) shield tunneling with ANN methods. *Soils Found.* **2019**, *59*, 284–295. [[CrossRef](#)]
5. Mahmoodzadeh, A.; Mohammadi, M.; Daraei, A.; Ali, H.F.H.; Al-Salihi, N.K.; Omer, R.M.D. Forecasting maximum surface settlement caused by urban tunneling. *Autom. Constr.* **2020**, *120*, 103375. [[CrossRef](#)]
6. Zhang, Z.; Xu, R.; Wu, X.; Wang, J. ANN-Based Dynamic Prediction of Daily Ground Settlement of Foundation Pit Considering Time-Dependent Influence Factors. *Appl. Sci.* **2022**, *12*, 6324. [[CrossRef](#)]
7. Tang, L.; Na, S. Comparison of machine learning methods for ground settlement prediction with different tunneling datasets. *J. Rock Mech. Geotech. Eng.* **2021**, *13*, 1274–1289. [[CrossRef](#)]
8. Chen, R.P.; Dai, T.; Zhang, P.; Wu, H.N. Prediction Method of Tunneling-induced Ground Settlement Using Machine Learning Algorithms. *J. Hunan Univ.* **2021**, *48*, 111–118. (In Chinese with English Abstract).
9. Moghaddasi, M.R.; Noorian-Bidgoli, M. ICA-ANN, ANN and multiple regression models for prediction of surface settlement caused by tunneling. *Tunn. Undergr. Space Technol.* **2018**, *79*, 197–209. [[CrossRef](#)]
10. Yang, P.; Yong, W.; Li, C.; Peng, K.; Wei, W.; Qiu, Y.; Zhou, J. Hybrid Random Forest-Based Models for Earth Pressure Balance Tunneling-Induced Ground Settlement Prediction. *Appl. Sci.* **2023**, *13*, 2574. [[CrossRef](#)]

11. Cao, X.Y.; Zhu, L.; Gong, H.L.; Guo, L.; Wei, Y.J.; Guo, T.; Chen, B.B.; Wang, H.G.; Li, H.J. Simulation of Ground Settlement in the Eastern Beijing Plain Using AM-LSTM Network. *Natl. Remote Sens. Bull.* **2022**, *26*, 1302–1314. (In Chinese) [[CrossRef](#)]
12. Dai, Y.; Dai, W.; Yu, W.; Bai, D. Determination of Landslide Displacement Warning Thresholds by Applying DBA-LSTM and Numerical Simulation Algorithms. *Appl. Sci.* **2022**, *12*, 6690. [[CrossRef](#)]
13. Kim, D.; Han, H.; Wang, W.; Kang, Y.; Lee, H.; Kim, H. Application of Deep Learning Models and Network Method for Comprehensive Air-Quality Index Prediction. *Appl. Sci.* **2022**, *12*, 6699. [[CrossRef](#)]
14. Zhu, Z.; Song, X.; Zhang, R.; Li, G.; Han, L.; Hu, X.; Li, D.; Yang, D.; Qin, F. A Hybrid Neural Network Model for Predicting Bottomhole Pressure in Managed Pressure Drilling. *Appl. Sci.* **2022**, *12*, 6728. [[CrossRef](#)]
15. Xue, J.; Shen, B. A novel swarm intelligence optimization approach: Sparrow search algorithm. *Syst. Sci. Control Eng.* **2020**, *8*, 22–34. [[CrossRef](#)]
16. Ma, Z.X.; Zhang, L.Q.; Qiu, Z.J.; Wang, Z.P.; Liu, G.C.; Wang, X. Two-stage Optimal Scheduling of the Park Integrated Energy System Based on CEEMD-SSA-LSTM. *High Volt. Eng.* **2022**, *49*, 1430–1440. (In Chinese with English Abstract) [[CrossRef](#)]
17. Leng, L.Y.; Fu, J.Z.; Ning, B. Research on IGBT Sequentially Prediction Based on SSA-LSTM Mode. *Semicond. Technol.* **2023**, *48*, 66–72. (In Chinese with English Abstract) [[CrossRef](#)]
18. Zu, L.L.; Liu, P.Z.; Zhao, Y.P.; Li, T.H.; Li, H. Solar Greenhouse Environment Prediction Model Based on SSA-LSTM. *Trans. Chin. Soc. Agric. Mach.* **2022**, *54*, 351–358. (In Chinese with English Abstract) [[CrossRef](#)]
19. Yu, Z.Z.; Zou, H.F.; Yu, D.S.; Wang, C.; Liu, T.X.; Zhang, X.Y. SSA-STM-based Model for Predicting Soil Oxygen Content in Maize. *Trans. Chin. Soc. Agric. Mach.* **2022**, *53*, 360–368+411. (In Chinese with English Abstract) [[CrossRef](#)]
20. Chen, Y.; Liu, Z.Y.; Xiao, C.Y.; Zhao, X.L.; Li, K.; Pang, L.L.; Shi, Y.X.; Li, S.H. Overlapping Peak Analysis of Soil Heavy Metal X-Ray Fluorescence Spectra Based on Sparrow Search Algorithm. *Spectrosc. Spectr. Anal.* **2021**, *41*, 2175–2180. (In Chinese with English Abstract)
21. Hochreiter, S.; Schmidhuber, J. Long Short-Term Memory. *Neural Comput.* **1997**, *9*, 1735–1780. [[CrossRef](#)] [[PubMed](#)]
22. Vinyals, O.; Toshev, A.; Bengio, S.; Erhan, D. Show and tell: A neural image caption generator. In Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, USA, 7–12 June 2015. [[CrossRef](#)]
23. Soltau, H.; Liao, H.; Sak, H. Neural Speech Recognizer: Acoustic-to-Word LSTM Model for Large Vocabulary Speech Recognition. *arXiv* **2016**, arXiv:1610.09975. [[CrossRef](#)]
24. Greff, K.; Srivastava, R.K.; Koutnik, J.; Steunebrink, B.R.; Schmidhuber, J. LSTM: A Search Space Odyssey. *IEEE Trans. Neural Netw. Learn. Syst.* **2017**, *28*, 2222–2232. [[CrossRef](#)]
25. Hu, Z.R.; Zhao, W.F.; Song, Y.X.; Wang, F.; Lin, Y.M. Prediction Spatial Distribution of Soil Organic Matter Based on Improved BP Neural Network with Optimized Sparrow Search Algorithm. *Environ. Sci.* **2023**, *15*. (In Chinese with English Abstract) [[CrossRef](#)]
26. Wu, J.; Huang, Z.; Qiao, H.; Wei, B.; Zhao, Y.; Li, J.; Zhao, J. Prediction about residual stress and microhardness of material subjected to multiple overlap laser shock processing using artificial neural network. *J. Cent. South Univ.* **2022**, *29*, 3346–3360. [[CrossRef](#)]
27. Li, Q.L.; Liu, Y.M.; Zhang, L.; Zhou, M.; Liang, J.; Wang, Y.; Sun, L. A Neural Networks Based Method for Suspended Sediment Concentration Retrieval from GF-5 Hyperspectral Images. *J. Infrared Millim. Waves* **2022**, *41*, 323–336. (In Chinese with English Abstract) [[CrossRef](#)]
28. Zhao, H.Y.; Huang, J.; Wang, J.D.; Wen, H.D.; Li, X. Improved SSA-VMD Algorithm and Its Application in Fault Diagnosis of Reciprocating Compressors. *Lubr. Eng.* **2022**, *47*, 147–152. (In Chinese with English Abstract) [[CrossRef](#)]
29. Li, D.D.; Wu, Y.X.; Zhu, C.C.; Li, Z.K. Parameter Identification Method for J-A Hysteresis Model Based on the Sparrow Search and Genetic Algorithm. *High Volt. Eng.* **2021**, *48*, 4181–4188. (In Chinese with English Abstract) [[CrossRef](#)]
30. Ma, X.Y.; Zhang, H.F.; Wang, Z.; Zhang, C.; Dong, Z. Prediction of transformer bushing temperature based on SSA-BPNN-ARIMA residual correction. *Foreign Electron. Meas. Technol.* **2022**, *41*, 151–156. (In Chinese with English Abstract) [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.