



## Article

# Comparing Kriging Estimators Using Weather Station Data and Local Greenhouse Sensors

Pei-Fen Kuo \* , Tzu-En Huang and I Gede Brawiswa Putra 

Geomatics Department, National Cheng Kung University, Tainan 701, Taiwan; neil860515@gmail.com (T.-E.H.); brawiswa@gmail.com (I.G.B.P.)

\* Correspondence: z10608024@email.ncku.edu.tw; Tel.: +886-6-237-0876

**Abstract:** In order to minimize the impacts of climate change on various crops, farmers must learn to monitor environmental conditions accurately and effectively, especially for plants that are particularly sensitive to the weather. On-site sensors and weather stations are two common methods for collecting data and observing weather conditions. Although sensors are capable of collecting accurate weather information on-site, they can be costly and time-consuming to install and maintain. An alternative is to use the online weather stations, which are usually government-owned and free to the public; however, their accuracy is questionable because they are frequently located far from the farmers' greenhouses. Therefore, we compared the accuracy of kriging estimators using the weather station data (collected by the Central Weather Bureau) to local sensors located in the greenhouse. The spatio-temporal kriging method was used to interpolate temperature data. The real value at the central point of the greenhouse was used for comparison. According to our results, the accuracy of the weather station estimator was slightly lower than that of the local sensor estimator. Farmers can obtain accurate estimators of environmental data by using on-site sensors; however, if they are unavailable, using a nearby weather station estimator is also acceptable.

**Keywords:** agriculture; oncidium; weather conditions; sensors; kriging



**Citation:** Kuo, P.-F.; Huang, T.-E.; Putra, I.G.B. Comparing Kriging Estimators Using Weather Station Data and Local Greenhouse Sensors. *Sensors* **2021**, *21*, 1853. <https://doi.org/10.3390/s21051853>

Academic Editors: Kun-chan Lan, Yi-Bing Lin, Teen-Hang Meen, Chi-Yuan Chen and Shih-Lin Wu

Received: 30 January 2021  
Accepted: 2 March 2021  
Published: 6 March 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 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

Weather conditions are the primary factors that have always affected plant growth and agricultural development. However, in recent years, the impacts of climate change have been obvious, such as increased temperatures, highly variable and shifting precipitation patterns, reduced snowpack on mountains, and increased frequency and intensity of extreme weather [1]. Numerous studies have shown that climate change has affected the development of plant crops. For example, in Nepal, the food supply has been affected by increases in CO<sub>2</sub> and temperature, and unstable precipitation [2]. Overall, among mountains (35%), hills (42%), and terai (23%) areas in Nepal, the crops such as rice, wheat, and maize in the Terai area (south of the outer foothills of the Himalayas) have decreased in the past decades. Similarly, in California, these abnormal weather patterns impact on agriculture in many ways, including increased pest and disease pressure and crop yield decline. Due to those reasons, climate change brings a negative impact on the agricultural industry [1]. Therefore, in order to come up with an effective strategy for optimizing crop performances under these abnormal weathers, monitoring these weather conditions is critical.

As previously mentioned, installing on-site sensors and getting data from weather stations are two common methods to collect/observe these weather conditions. Although the sensor could be installed on-site in order to observe accurate weather information, the installation and its management are time-consuming and costly. As an alternative option, the government weather station data could be used in this case, since it is open to the public for free. However, their accuracy is questionable because the locations of these stations are far from the greenhouse. Therefore, in this study two different data sets were

used to compare the accuracy of kriging estimators: First, multiple sensors were installed and dispersed throughout an *Oncidium* greenhouse in order to collect the actual/real local weather data. In addition, this study also utilized the online weather station data from the open data website Central Weather Bureau. The estimators were compared using these two datasets. The data were analyzed over a period of several months in order to determine how seasonal changes affect our study results. If both types of estimators yield similar numbers, farmers with limited budgets can use weather station data to monitor temperature instead of installing their own onsite sensors. Also, researchers will be able to apply historical data from weather stations to represent past weather conditions and use them for time series analysis.

## 2. Literature Review

Environmental factors have a significant impact on agriculture on a global scale. For this reason, it is essential to understand the environmental conditions which affect crop production. In addition, this information will help researchers predict and control dangerous conditions (such as air pollution) [3]. Therefore, determining the accuracy of weather information in a certain location is vital to the health of crops and to the livelihoods of farmers. The most direct way to accomplish this goal is to install the sensors and collect the necessary data. However, due to a limited budget and time, the farmers were unable to install the sensors seamlessly cover the whole study area [4]. An alternative is to use spatial interpolation methods to obtain weather information at a particular place where it can't be attained directly. For example, several traditional interpolation methods have been commonly used to transform weather point data to spatial distribution, such as the inverse distance weighting method (IDW) and the kriging method. The IDW method is relatively straightforward to compute, but the spatial variation of relationships between points cannot be explored and the accuracy of the estimation is limited [5]. In comparison, to estimate the structure of spatial variance, the kriging approach uses variogram analysis and takes into account spatial autocorrelation [6,7]. Recently, some researchers have focused on the accuracy and the performance of these spatial interpolation methods. For example, Oktavia et al. [4] and Zhang et al. [7] used IDW and kriging spatial interpolation for thermal monitoring of a data center and evaluated the accuracy of these two methods based on the Root Mean Squared Error (RMSE) value. Their results showed that the kriging method performed better.

Kriging, which has been widely used for spatial interpolation [8], has the key advantage of incorporating spatial correlation data, leading to a high accuracy rate. Kriging itself is a category of stochastic interpolation methods that include ordinary kriging, simple kriging, co-kriging, universal kriging, regression kriging, and residual kriging [9]. These approaches have been used in a number of researches. For instance, Wu et al. [10] used the residual kriging approach with input variables of longitude, latitude, and elevation to approximate the average monthly temperature in the United States. Liang et al. [11] used the co-kriging method to model the nitrate-nitrogen daily concentration in an agricultural river. Based on the application in the previous study, kriging was a suitable method for our study since our target predictive variable was the temperature, which has spatial variance [12].

Several researchers who have used spatio-temporal kriging to interpolate temperature in the space-time field discovered that it is more accurate than spatial kriging [13]. The additional advantage of spatio-temporal kriging resides in the model's versatility. This method not only interpolates at unobserved spatial locations but also at unobserved instances of time [14]. This reason makes spatio-temporal kriging an appropriate method to cover the gaps of a time series model not only dependent on the time series, but also including all of its spatial neighbors. In sum, the purpose of this study was to compare the spatio-temporal kriging result by using two different datasets in two different seasons.

### 3. Data and Methods

#### 3.1. Study Area & Study Data

In this study, the temperature data collected from the sensors and the weather stations from the Central Weather Bureau (CWB) were utilized for the spatio-temporal kriging (<http://e-service.cwb.gov.tw> (accessed on 5 March 2021)). There were 24 weather stations, the locations of which are indicated by the blue points in Figure 1. The study area was located in the Xinshe District in Taichung City, indicated by the red point. The distance between the greenhouse and the nearest weather station is approximately four km while the farthest weather station is located 48 km from the greenhouse.

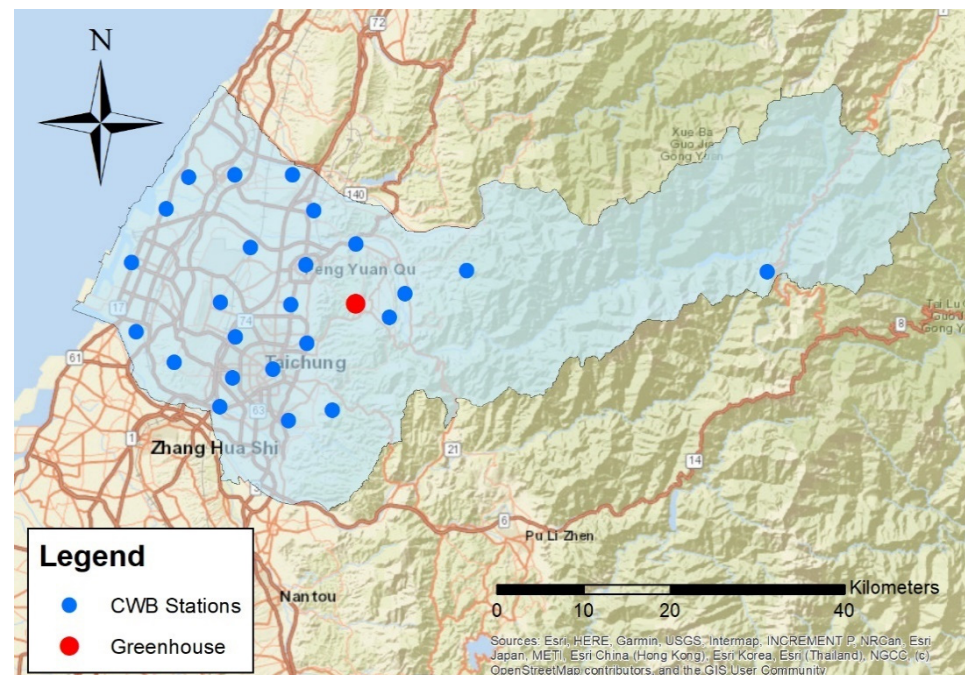


Figure 1. The locations of weather stations.

The weather stations are able to observe several weather parameters such as station pressure, sea surface pressure, temperature, dew point temperature and relative humidity, wind speed, wind direction, visibility, and cloud amount. Table 1 is an example of one-day data obtained from weather stations. The first column is station name, the date column is the record day, the ObsTime column is the hour of record time, StnPres column is the station average air pressure, Temperature is the average temperature, T.max column is the maximum temperature, T.min column is the minimum temperature, RH column is the average relative humidity, WS column is the average wind speed, precp column is the average precipitation, and sunshine is the light intensity. The weather station provides several weather conditions, such as temperature in degrees Celsius, station air pressure in hPa, relative humidity in percentage, wind speed in m/s, precipitation in mm, and light intensity.

However, in this study, only temperature data were used for the analysis. Temperature data from the weather stations were downloaded from the CWB Observation Data Inquiry System online. As previously mentioned, the study periods were from 14–20 June 2020 and 14–20 September 2020. These study periods were chosen by considering the sensor performance in two different seasons (Summer and Autumn) and that the observation days with no missing data fell within the study period. On the other hand, the sensors themselves were installed at the end of March 2020 and became fully functional in May 2020. Therefore, this study only covers two seasons during the study period. The raw data from the website is updated hourly, so it was collected 24 times per day at each station within the fourteen days. The temperature value data was also collected from the 44 sensors

installed in the greenhouse. The sensors measure the temperature every five minutes, so data was collected 288 times by each sensor per day.

**Table 1.** Weather station data.

Station	Date	Obs Time	StnPres	Temperature	T.Max	T.Min	RH	WS	Precp	SunShine
467490	6/14/2020	16	1002.2	26.1	34.1	24.4	85	0.6	32	0.0
467770	6/14/2020	9	1008.9	30.7	32.4	27.1	70	4.3	0.0	1.0
COF000	6/14/2020	4	976.3	24.7	33.0	24.3	100	0.0	0.0	...
COF850	6/14/2020	11	969.5	30.5	32.4	23.0	61	3.2	0.0	...
COF861	6/14/2020	23	786.0	16.1	24.5	13.6	100	0.0	0.0	...
COF930	6/14/2020	01	998.9	28.2	35.1	26.7	94	0.4	0.0	...
COF970	6/14/2020	17	994.9	27.0	33.6	25.2	90	1.6	0.5	...
COF9A0	6/14/2020	4	963.2	24.9	34.8	23.4	85	0.5	0.0	...
COF9I0	6/14/2020	1	988.4	27.1	32.5	25.6	86	1.3	0.0	...
COF9L0	6/14/2020	18	983.6	26.6	33.1	25.9	99	0.5	13.0	...
COF9M0	6/14/2020	7	985.3	28.8	33.0	25.8	67	0.6	0.0	...
COF9N0	6/14/2020	24	1005.9	26.9	35.7	25.7	90	1.4	0.0	...
COF9O0	6/14/2020	23	994.2	27.8	35.8	27.3	81	0.0	0.0	...
COF9Q0	6/14/2020	1	992.8	26.4	33.8	25.0	89	1.3	0.0	...
COF9S0	6/14/2020	2	998.8	27.6	33.9	26.0	91	0.7	0.0	...
COF9T0	6/14/2020	10	1000	32.1	33.7	26.5	61	1.7	0.0	...
COF9U0	6/14/2020	10	1004.6	32.7	34.4	26.6	61	2.0	0.0	...
COF9V0	6/14/2020	6	954.2	25.2	32.6	23.5	82	1.0	0.0	...

The system consists mainly of three main parts: user interface, gateway, and sensor nodes. Figure 2 shows the system structure that illustrates how the environment monitoring system for the greenhouse operates. The sensors, each of which are numbered, are uniformly distributed and deployed throughout the greenhouse to collect environmental information such as temperature, luminosity, and humidity. The data collected by the sensor nodes are transmitted to the gateways via Bluetooth. This study follows the concept of a star network topology where the gateways are deployed in the center of the greenhouse and used to collect the data from sensor nodes, store the data, process, and integrate data [15]. In addition, the gateways are connected to a WiFi network and use a message queuing telemetry transport (MQTT) protocol to transfer data from the gateway to the cloud. The user interface, which is a cloud-based web application, allows users to monitor the environmental data measured by each sensor node in real-time. Figure 3 illustrates the greenhouse with an area of approximately  $5400 \text{ m}^2$  ( $= 180 \text{ m} \times 30 \text{ m}$ ) and the sensor's deployment location.

Bluetooth wireless sensors (shown in Figure 4) from the Xiaomi Corporation were installed throughout the greenhouse (Figure 5) to measure environmental information such as soil moisture and conductivity, as well as light and temperature. This brand was chosen due to its cost-effectiveness compared to other sensors, reliability, and accuracy of up to  $0.5 \text{ }^\circ\text{C}$ . Please see <https://www.mi.com/flowermonitor> (accessed on 5 March 2021) for more details regarding these particular sensors. Moreover, this type is compatible with the Xiaomi Smart Home App, which means that it will be more convenient for farmers. The parameters of these sensors are shown in Table 2. Regarding battery life, in normal circumstances the battery life could exceed one year if the measurement data is stored only once every 20 min to the receiver (the default scan interval is 1200 s). However, in order to ensure the consistency of the observations and prevent missing data, regular maintenance such as battery replacement was performed every three months from the time the sensors were fully operational in May 2020. It must be noted that the sensor data was set to be recorded every five minutes to reduce the missing rate, which resulted in shorter battery life.



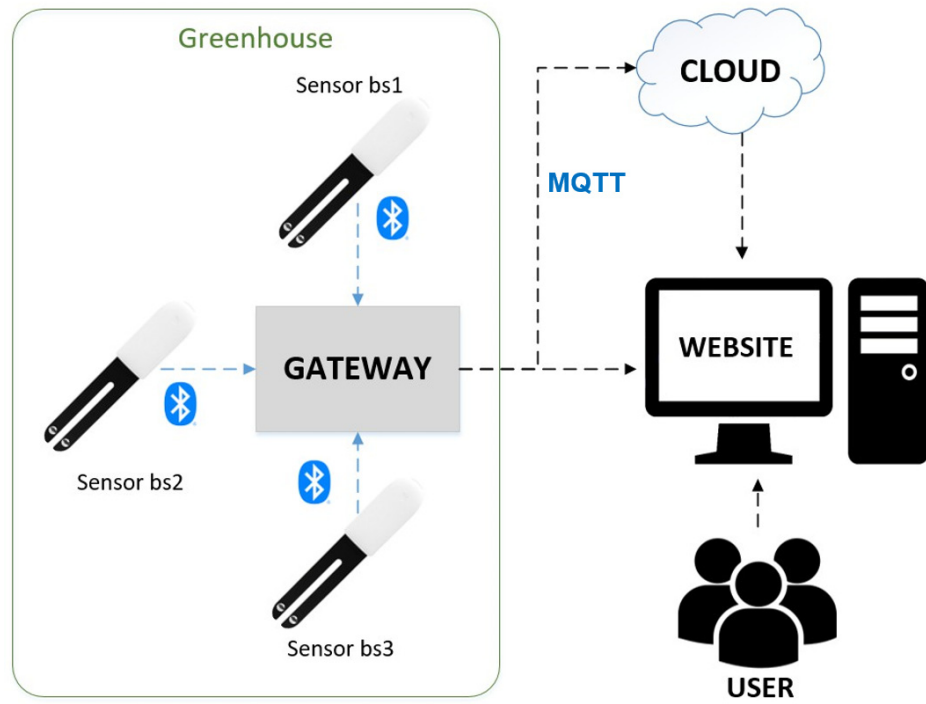


Figure 2. Greenhouse environment monitoring system structure.

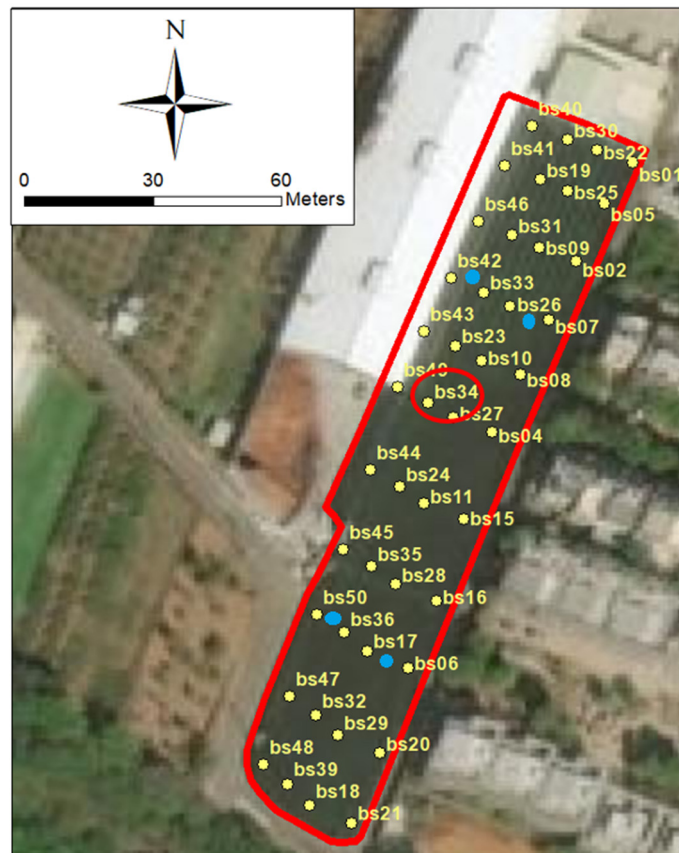


Figure 3. Distribution of sensors and gateways in the Oncidium greenhouse.

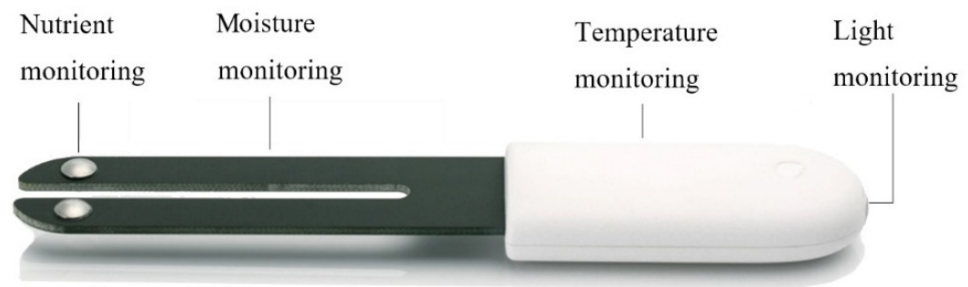


Figure 4. Xiaomi sensor.



Figure 5. Sensor installation placement.

Table 2. Parameters of the Xiaomi sensors.

<b>Size</b>	<b>120.5 mm × 24.5 mm × 12.5 mm</b>
Wireless Connection	Bluetooth 4.1 BLE
Operating Voltage	3 V
Battery	CR2032 button cell battery

The data collected by the sensors were uploaded to a website and downloaded using a computer, the layout of which is shown in Figure 6. For example, real-time environmental data including temperature, soil moisture, and luminosity is displayed from top to bottom. Each collection sensor is assigned its own color, which makes it easier to determine if a particular sensor is not functioning correctly. The middle panel of the platform is where the management and setting interfaces are located, to which sensors can be added or removed and where the time intervals for recording the environmental components can be set. As previously stated, for this study, the intervals were set to five minutes. The sensor data for a specific date can be downloaded on the left panel. The white section at the bottom is for downloading, and the start and end dates are displayed in the blank section at the upper left corner of the screen.

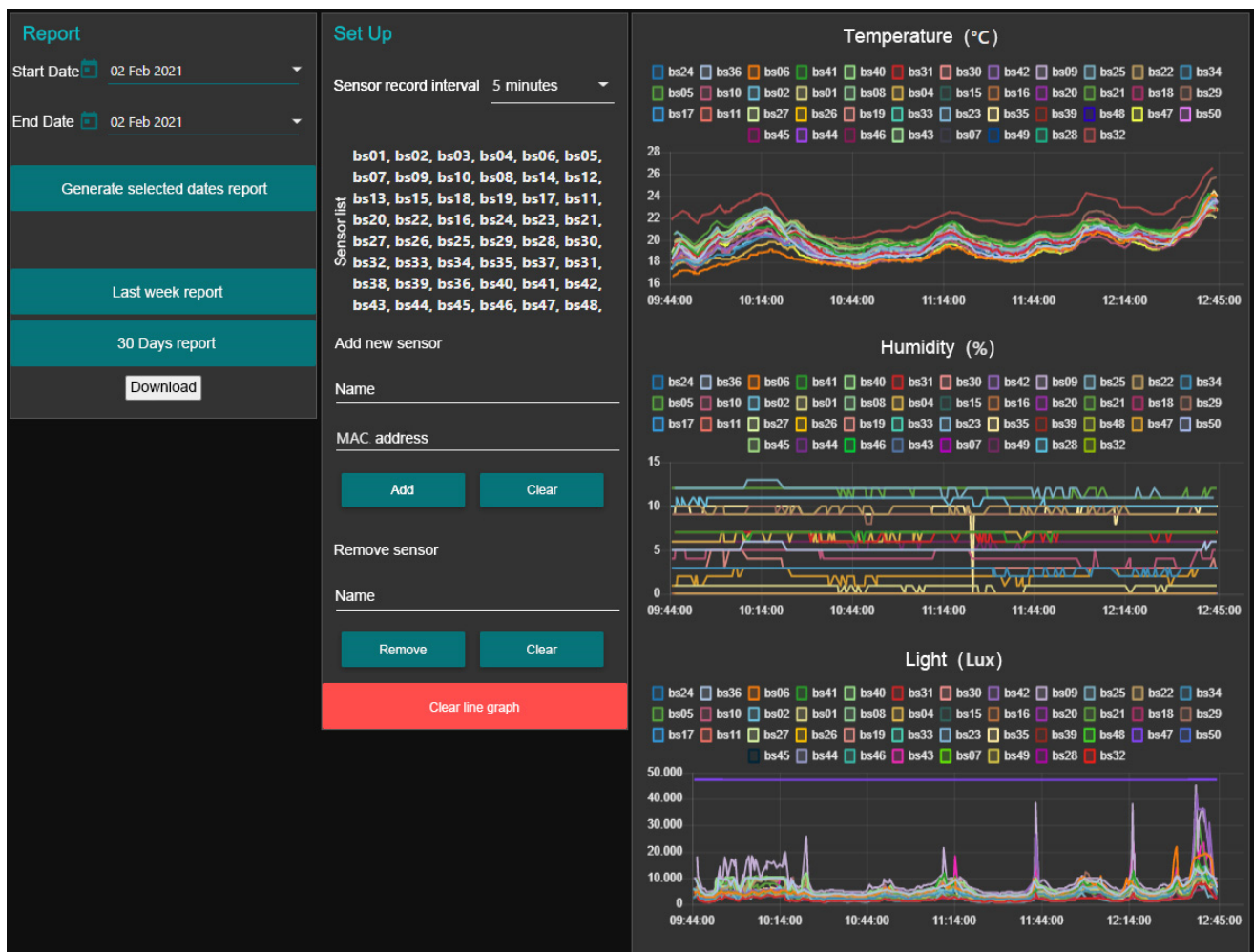


Figure 6. The website that records the sensor data.

The measurements of sensor “bs34” were chosen as the real values at the center of the greenhouse, and the temperature data from the other 43 sensors were selected as the spatio-temporal kriging estimator. The reasons for choosing bs34 as the reference for the actual temperature value are because it is located at the centroid of the greenhouse which is far enough away from factors that might affect the temperature, such as the shadows of nearby buildings, and this sensor had no missing data during the study period. Another estimator was based on the 24 weather stations. Thus, the temperature values measured by sensor “bs34” on June 14th—20th and September 14th—20th were the true values that were used to estimate the error by comparing the results obtained from different input data. The dataset itself was filtered which resulted in fewer than 10% missing observations and outliers. The descriptive statistics of the raw data are shown in Table 3.

Table 3. Statistical descriptions of the raw data for temperatures.

	June 14th—20th			September 14th—20th		
	CWB	Sensor	bs34	CWB	Sensor	bs34
Min.	12.6	20.59	21.12	12.1	20.6	21.19
Q1	26.6	22.76	22.85	26.5	23	22.98
Median	28.6	23.93	23.72	28.3	24.51	24.12
Mean	28.23	26.04	26.05	28.11	26.314	25.86
Q3	30.8	30.00	30.16	30.5	30.31	29.57
Max.	35.7	37.81	34.06	35.7	35.03	32.77
Stdv	3.68	3.73	3.83	3.79	2.85	3.21

Unit: degrees Celsius.

### 3.2. Spatio-Temporal Kriging

Kriging is a common method for obtaining unbiased estimators in a particular location for spatial interpolation. Ordinary kriging is the most commonly used kriging method in different subjects. The basic concept of the calculation of it is shown as Equation (1), where  $Z(x)$  represents the estimator at the point  $x$ ,  $\omega_i$  represents the weight of each sample point, and  $n$  means the number of the sample point.

$$Z(x) = \sum_{i=1}^n \omega_i Z(x_i) \quad (1)$$

To determine whether the estimation is unbiased, the weight sum ( $\omega_i$ ) should equal to 1, as shown in Equation (2).

$$\sum_{i=1}^n \omega_i = 1 \quad (2)$$

Ordinary kriging interpolation considers the 2-D distance between sample points, also called known/observed points. The coordinates and the  $Z$ -values of the known points are used to calculate the semi-variogram for obtaining the weight, which is used to predict the unknown point. The calculation of the variogram can be written as Equation (3), where  $\hat{\gamma}(h)$  represents the semi-variogram,  $N(h)$  is the number of the pairs of sample point which separated by distance  $h$  (Euclidean Distance),  $Z(s_i)$  and  $Z(s_j)$  is the estimator at the coordinate  $s_i$  and  $s_j$ . Then, the semi-variogram is fitted and can be used to calculate the weight for getting an unknown value of a certain location.

$$\hat{\gamma}(h) = \frac{1}{2N(h)} \sum_{N(h)} \{ [Z(s_i) - Z(s_j)] \}^2 \quad (3)$$

The illustration of the semi-variogram is shown in Figure 7. The x-axis in Figure 7 represents the distance  $h$ , the y-axis is the semi-variance, and it shows that the model levels out at a certain distance. The range is the distance where the model first flattens, the value at which the semi-variogram model attains the range (the value on the y-axis) is called the sill, the nugget is the intercept on the y-axis when the distance equal to 0, and the partial sill is the sill minus the nugget. According to McBratney and Webster [16], the semi-variogram value will be 0 when the separation distance is also equal to zero. However, at an extremely small distance of separation, the semi-variogram usually shows a nugget effect higher than 0. The nugget effect could be attributed to spatial variance or measurement errors at distances shorter than the sample interval. Measurement error occurs due to an error underlying in the measurement instrument. Natural phenomena could still vary spatially across a variety of scales. Microscale variations smaller than the sample interval will appear as part of the nugget effect.

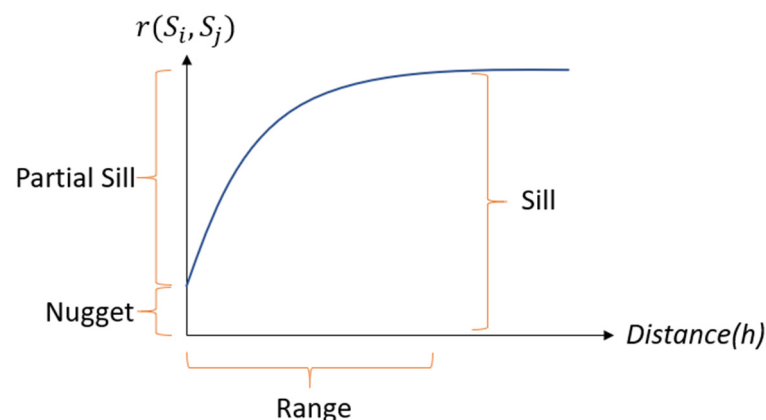


Figure 7. The illustration of semi-variogram.



In this study, a method known as spatio-temporal kriging was used to spatially interpolate the temperature data over time. This method takes time into account during the kriging process by adapting the covariance function. Unlike the standard 2-D kriging technique, spatial-temporal kriging not only considers the coordinates but also the element of time [14]. Specifically, for each point  $s_i$ , there is a time  $t_i$  associated with it to calculate the variance between this point and another point of interest with a spatial separation of  $h$  as well as their temporal separation  $u$ . Thus, the spatio-temporal variogram can be calculated using Equation (4).

$$\hat{\gamma}(h, u) = \frac{1}{2N(h, u)} \sum_{N(h, u)} \{ [Z(s_i, t_i) - Z(s_j, t_j)] \}^2 \quad (4)$$

After utilizing Equation (4), the weight of each point with a known value can be calculated using the spatio-temporal variogram and the unknown value of a certain location can be determined via interpolation.

For this study, the “sp”, “gstat”, “spacetime” packages for use with software “Rstudio” were selected to process the spatial-temporal kriging. First, the spatio-temporal variogram was constructed based on the sample data collected for predictions, which is similar to a normal 2-D kriging experiment. The simple sum metric model was chosen as the most effective for fitting the model to our variogram. It combines the spatial, temporal and joint nugget effects to restrict the spatial, temporal, and joint variograms into a nugget-free model. Thus, a single spatio-temporal nugget and the variogram can be written as Equation (5) [14].

$$\gamma(h, u) = \text{nug} \cdot 1_{h>0 \vee u>0} + \gamma_s(h) + \gamma_t(u) + \gamma_{\text{joint}} \left( \sqrt{h^2 + (k \cdot u)^2} \right) \quad (5)$$

where  $\gamma(h, u)$  represents the spatio-temporal variogram,  $\gamma_s(h)$  symbolizes the spatial variogram,  $\gamma_t(u)$  is the temporal variogram,  $\gamma_{\text{joint}} \left( \sqrt{h^2 + (k \cdot u)^2} \right)$  represents the joint variogram, and  $k$  denotes the spatio-temporal anisotropy scaling parameter. The selected variogram fitting model was used to calculate the weight of the location with a known value for prediction, which allowed for the interpolation of that particular location.

The ten-fold cross-validation technique was utilized to evaluate the kriging method’s performance [17]. This method is one of the cross-validation methods that is commonly used in order to assess the quality of kriging prediction. The sensor data was divided into groups of ten with approximately the same sensor numbers. Then, one data group was set aside as the testing dataset while the remaining nine were utilized as the training datasets. The kriging test was conducted using the training dataset to predict the values of the test dataset. This process was repeated until predictions were determined for all groups. This is primarily used to assess how accurate the kriging model would be in reality. In order to test the prediction accuracy of the kriging, the R-square and RMSE measurements were determined by comparing the prediction values to the observations.

### 3.3. Data Preprocessing

The weather station data was measured automatically per hour, and the sensor data was calculated every five minutes. As mentioned previously, the goal of the study was to compare the accuracy of the spatio-temporal kriging results using these two types of data. Therefore, all the sensor data was converted into the same format for comparison. In order to convert the sensor data into a per-hour format, the average hourly temperature data of each sensor was calculated. With regard to the missing data, the null value was replaced with each individual average.

### 3.4. Comparison

To compare the accuracy of the spatio-temporal kriging predictions using weather station data and sensor data, in this study the Root Mean Square Error (RMSE) and Mean Absolute Error (MAE) values were compared. Then, the real and predicted temperatures from the weather station and sensor data could also be projected and analyzed as to whether they were affected by the different seasons. Equation (6) was used to calculate the RMSE, and Equation (7) calculates the MAE.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (6)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |(y_i - \hat{y}_i)| \quad (7)$$

where  $n$  represents the number of samples,  $y_i$  is the true value (the temperature value measured by sensor bs34), and  $\hat{y}_i$  symbolizes the predicted value (the interpolation result of spatio-temporal kriging based on the weather station and sensor data). The smaller the value, the more accurate the predicted result.

## 4. Results

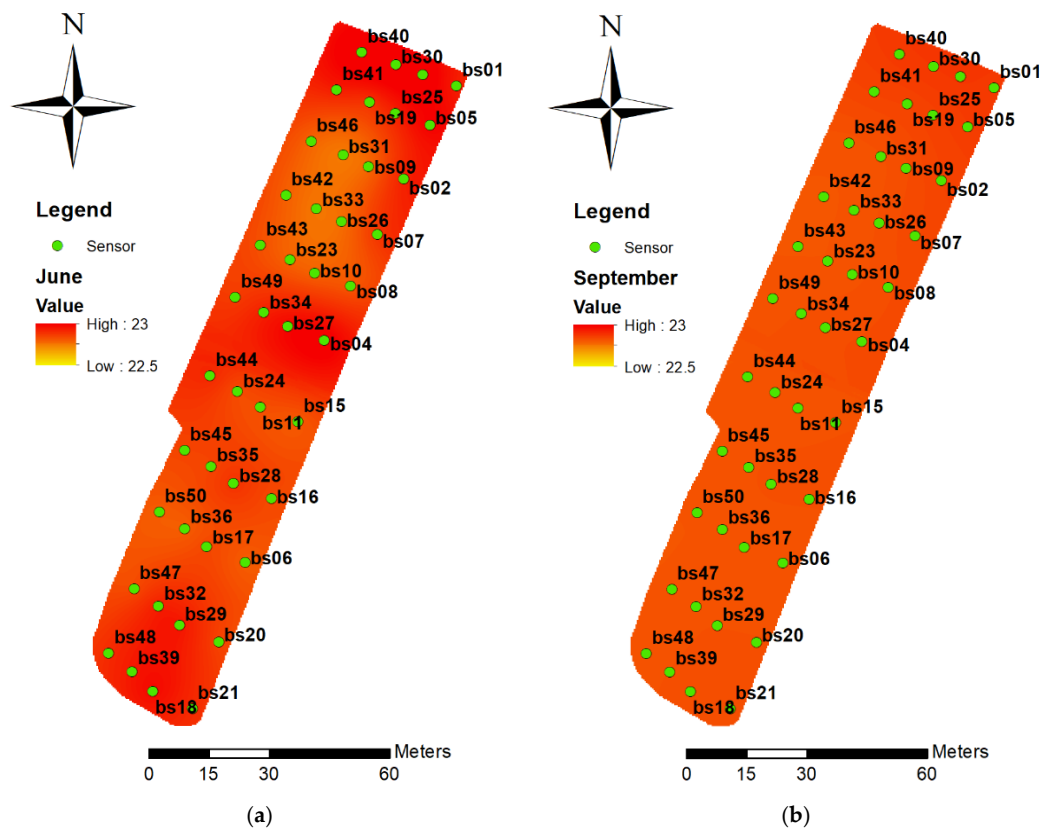
### 4.1. Cross-Validation Results

The kriging results illustrate the temperature variations in the greenhouse that were recorded using sensors, as shown in Figure 8. The average temperatures on July 14th from 12:00 a.m.—00:55 a.m. (Figure 8a) and on September 14th from 12:00 a.m.—00:55 a.m. (Figure 8b) were utilized as the sample, which as shown in the legend, and measured in degrees Celsius. As seen in Figure 7, the average temperatures in July, especially in the northern, middle, and southern sections of the greenhouse, were slightly higher than those in September, which was expected. Furthermore, there were fewer variances in September than in June, largely because there were more missing data on September 14th than on June 14th. Since in this study the missing data were filled in with the averages of all sensor measurements, the temperature distribution, shown in Figure 8b, was less than that in Figure 8a.

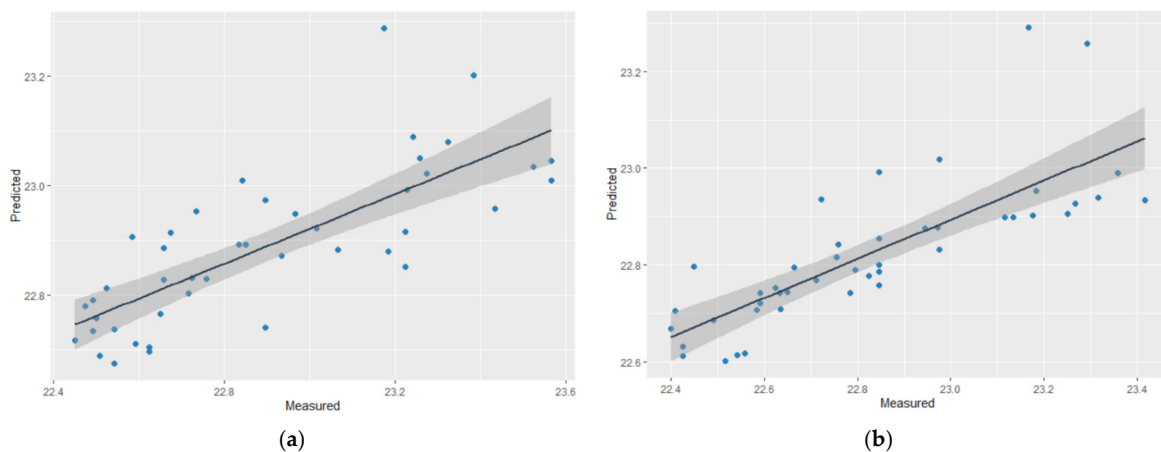
Validation is crucial for evaluating a model's robustness and performance. For this study, the R-square and RMSE values from the results of the k-fold cross-validation were utilized as the parameters to evaluate the kriging from the sensor observations [17]. Figure 9 shows scatterplots of the observed and predicted temperatures (based on 95% prediction intervals). The kriging consistently performed well in both seasons with R-squares of 0.59 and 0.62 (RMSE 0.3, and 0.19) in summer (Figure 9a) and autumn, respectively (Figure 9b).

### 4.2. Comparison Result

To estimate the temperature at the center of the greenhouse (sensor # bs34) on June 14th—20th and September 14th—20th, the temperature data from the nearby weather stations and 43 greenhouse sensors were used to determine the spatio-temporal kriging respectively. The temperature value, measured by sensor bs34, was set as the true value and the difference between the predicted and the true value was the error. To compare the model performance, the RMSE and the MAE were calculated, as shown in Table 4. The temperature value that was measured by sensor bs34 was designated as  $y_i$  in Equations (6) and (7) and the kriging results from the sensors and the CWB weather station were defined as  $\hat{y}_i$ . The RMSE was calculated by dividing the sum of the square from the difference between each bs34 observation and the kriging result with the number of sample sizes within one week (4032 sample for CWB sensor data; 7392 sample for sensor data). In addition, the results from adding the absolute value to the subtracted true value and the kriging predicted value were divided by the number of sample sizes from a week of data were calculated in order to obtain the MAE value.



**Figure 8.** The kriging results for the greenhouse temperatures (a) June 14th from 12:00 a.m.—00:55 a.m. (summer); (b) September 14th from 12:00 a.m.—00:55 a.m. (autumn).

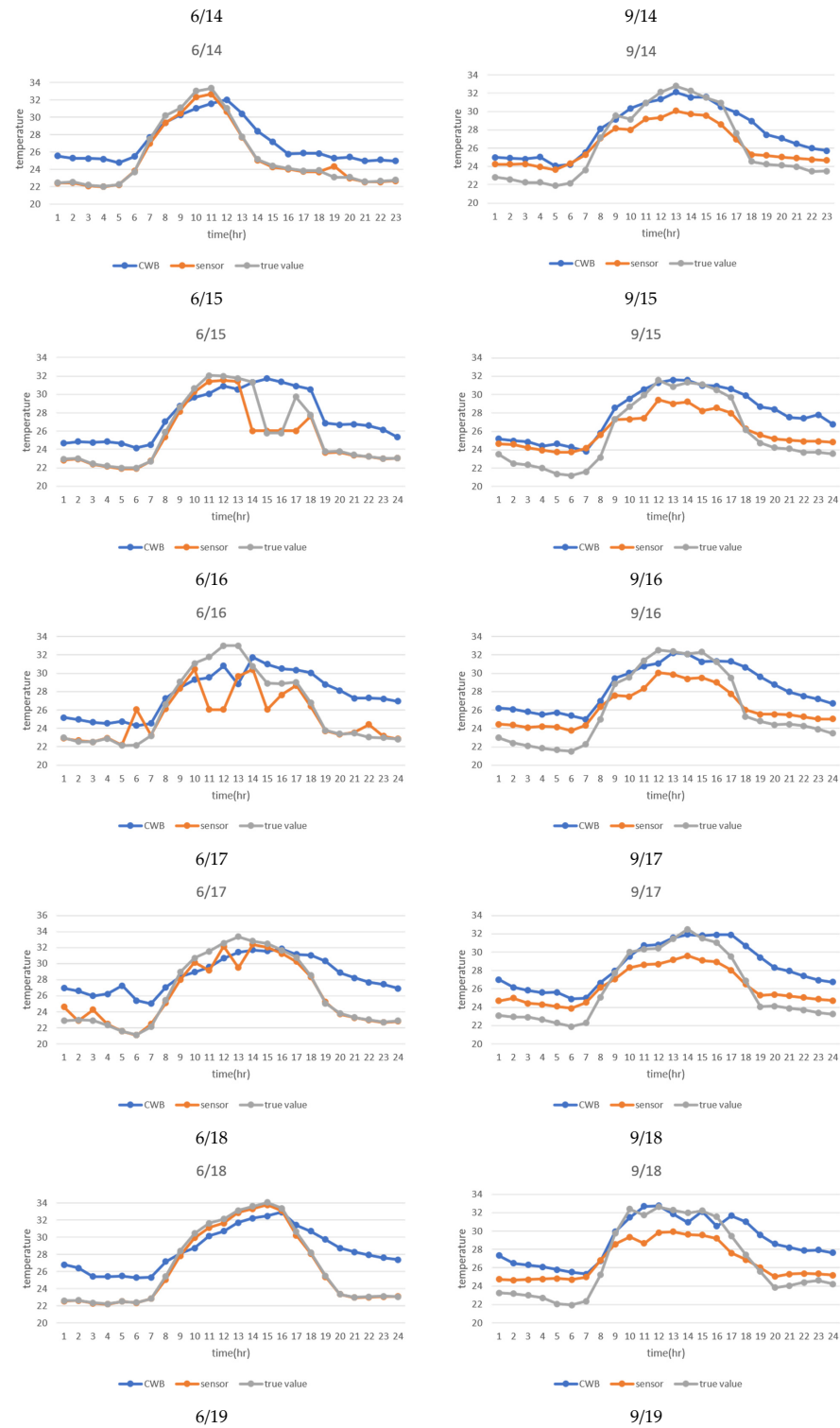


**Figure 9.** Sample of scatterplots of the observed and predicted greenhouse temperatures (a) June 14th from 12:00 a.m.—00:55 a.m. (summer); (b) September 14th from 12:00 a.m.—00:55 a.m. (autumn)

In addition, the true value and the two estimators were plotted on the graphs below to compare the differences in the results between June and September. Figure 9 shows the line chart of the true and the predicted values based on the weather station and sensor data. The blue line is the predicted value from the weather station data, the orange line is the predicted value from the sensor data, and the gray line shows the observed value from sensor bs34 (real value). The x-axis is the time, and the y-axis is the temperature. As shown in Figure 10, the sensor data prediction on June 14th—20th was closer to the true value than the one on September 14th—20th. During this week, the weather station data prediction was a bit higher than that from the sensor data.

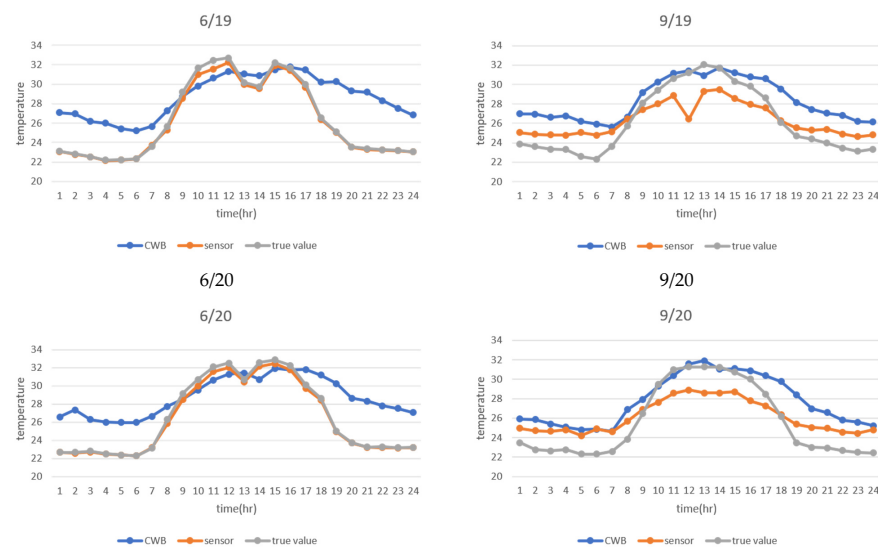
**Table 4.** RMSE and MAE (hourly temperatures).

Date	RMSE		MAE	
	CWB	Sensors	CWB	Sensors
June 14th—20th	3.01	1.10	2.63	0.47
September 14th—20th	2.66	1.87	2.26	1.72



**Figure 10.** Cont.





**Figure 10.** Line chart of the predicted and true values.

## 5. Conclusions

According to Table 4, the RMSE and the MAE from the sensor data predictions were all lower than from the weather station data. Thus, the spatio-temporal kriging sensor data results were more accurate than the weather station data. Moreover, the RMSE and MAE values of the sensor data prediction in September were obviously higher than in June. The possible reason for this is that the missing rate (percentage of the data missing in dataset) of the sensor data in September was higher than in June. These results indicate that fewer sample points used for the spatio-temporal kriging will obtain less accurate results, as seen in September.

As shown in Figure 10, the trend of the predicted value from the sensor data and true value were similar on June 14th–20th. The RMSE and MAE were also similar. Moreover, the prediction based on sensor data in June was more accurate than the one in September. On June 14th–20th, the prediction based on weather station data was a bit higher than that from the true value and sensor data, except from 9 a.m. to 12 a.m. As indicated in all figures, most of the predictions from sensor data were lower than those from weather station data. The possible reason is that several fans are installed in the greenhouse and only operated during summer. When the results in summer (June) and autumn (September) were compared, there were no obvious differences between them. Moreover, there was a more significant difference between that from the sensor estimator and the true value at noon (11 a.m. to 12 a.m.) than from the CWB estimators on 6/17 and 9/20. The possible reason is that the daily average was used to fill the missing value at and these two days have more missing values than usual.

According to our results, the local-sensor kriging method performed better than the weather station data. It must be noted that the distance between the known and the unknown point will affect the accuracy of the spatio-temporal kriging result. The sample size and original value of the data integrity will also affect the accuracy of the result. Thus, future scholars may wish to focus on defining the optimal sample size of sensors for the kriging method and analyze the difference between the kriging estimators. Another limitation is the high consumption rate of the sensor battery. When some sensors are out of battery, the missing value increases and then reduces the accuracy of the kriging result. The researchers also suggest future work could include long-term observation of seasonal and yearly temperature.

**Author Contributions:** Conceptualization, P.-F.K.; methodology, T.-E.H.; validation, T.-E.H. and I.G.B.P.; formal analysis, T.-E.H. and I.G.B.P.; data curation, T.-E.H.; writing—original draft preparation, T.-E.H. and I.G.B.P.; writing—review and editing, P.-F.K., T.-E.H. and I.G.B.P.; visualization,

T.-E.H.; supervision, P.-F.K.; project administration, P.-F.K. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the Ministry of Science and Technology Taiwan, grant number MOST 108-2321-B-055-002 and MOST 109-2124-M-006-002.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Not applicable.

**Acknowledgments:** This study has been supported by the Ministry of Science and Technology Taiwan under research on wisdom technology in promoting the industry efficiency and establishing the expert system of oncidium.

**Conflicts of Interest:** The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

## References

1. Pathak, T.B.; Maskey, M.L.; Dahlberg, J.A.; Kearns, F.; Bali, K.M.; Zaccaria, D. Climate change trends and impacts on California Agriculture: A detailed review. *Agronomy* **2018**, *8*, 1–27. [[CrossRef](#)]
2. Malla, G. Climate change and its impact on Nepalese agriculture. *J. Agric. Environ.* **2008**, *9*, 62–71. [[CrossRef](#)]
3. van Zoest, V.; Osei, F.B.; Hoek, G.; Stein, A. Spatio-temporal regression kriging for modelling urban NO<sub>2</sub> concentrations. *Int. J. Geogr. Inf. Sci.* **2020**, *34*, 851–865. [[CrossRef](#)]
4. Oktavia, E.; Mustika, I.W. Inverse distance weighting and kriging spatial interpolation for data center thermal monitoring. In Proceedings of the 2016 1st International Conference on Information Technology, Information Systems and Electrical Engineering (ICITISEE), Yogyakarta, Indonesia, 23–24 August 2016; pp. 69–74.
5. Lu, G.Y.; Wong, D.W. An adaptive inverse-distance weighting spatial interpolation technique. *Comput. Geosci.* **2008**, *34*, 1044–1055. [[CrossRef](#)]
6. Matheron, G. Principles of geostatistics. *Econ. Geol.* **1963**, *58*, 1246–1266. [[CrossRef](#)]
7. Zhang, J.; Li, X.; Yang, R.; Liu, Q.; Zhao, L.; Dou, B. An extended kriging method to interpolate near-surface soil moisture data measured by wireless sensor networks. *Sensors* **2017**, *17*, 1390. [[CrossRef](#)] [[PubMed](#)]
8. Li, C.; Lu, Z.; Ma, T.; Zhu, X. A simple kriging method incorporating multiscale measurements in geochemical survey. *J. Geochem. Explor.* **2009**, *101*, 147–154. [[CrossRef](#)]
9. Oliver, M.A.; Webster, R. Kriging: A method of interpolation for geographical information systems. *Int. J. Geogr. Inf. Syst.* **1990**, *4*, 313–332. [[CrossRef](#)]
10. Wu, T.; Li, Y. Spatial interpolation of temperature in the United States using residual kriging. *Appl. Geogr.* **2013**, *44*, 112–120. [[CrossRef](#)]
11. Liang, X.; Schilling, K.; Zhang, Y.-K.; Jones, C. Co-Kriging Estimation of Nitrate-Nitrogen Loads in an Agricultural River. *Water Resour. Manag.* **2016**, *30*, 1771–1784. [[CrossRef](#)]
12. Appelhans, T.; Mwangomo, E.; Hardy, D.R.; Hemp, A.; Nauss, T. Evaluating machine learning approaches for the interpolation of monthly air temperature at Mt. Kilimanjaro, Tanzania. *Spat. Stat.* **2015**, *14*, 91–113. [[CrossRef](#)]
13. Sha, L. Geostatistical space-time modeling for temperature estimation. In Proceedings of the 2012 First International Conference on Agro-Geoinformatics (Agro-Geoinformatics), Shanghai, China, 2–4 August 2012; pp. 1–5.
14. Pebesma, E.; Heuvelink, G. Spatio-temporal interpolation using gstat. *RFID J.* **2016**, *8*, 204–218.
15. Pahuja, R.; Verma, H.K.; Uddin, M. A wireless sensor network for greenhouse climate control. *IEEE Pervasive Comput.* **2013**, *12*, 49–58. [[CrossRef](#)]
16. Mcbratney, A.B.; Webster, R. Choosing functions for semi-variograms of soil properties and fitting them to sampling estimates. *J. Soil Sci.* **1986**, *37*, 617–639. [[CrossRef](#)]
17. Mercer, L.D.; Szpiro, A.A.; Sheppard, L.; Lindström, J.; Adar, S.D.; Allen, R.W.; Avol, E.L.; Oron, A.P.; Larson, T.; Liu, L.J.; et al. Comparing universal kriging and land-use regression for predicting concentrations of gaseous oxides of nitrogen (NO<sub>x</sub>) for the Multi-Ethnic Study of Atherosclerosis and Air Pollution (MESA Air). *Atmos. Environ.* **2011**, *45*, 4412–4420. [[CrossRef](#)] [[PubMed](#)]