sustainability

Article

Predicting Maximum Work Duration for Construction Workers

Ran Yan [1], Wen Yi [2,3,*] and Shuaian Wang [4]

1 Department of Logistics and Maritime Studies, The Hong Kong Polytechnic University, Hung Hom, Hong Kong SAR, China
2 School of Built Environment, College of Sciences, Massey University, Auckland 0632, New Zealand
3 Department of Building and Real Estate, The Hong Kong Polytechnic University, Hung Hom, Hong Kong SAR, China
4 Faculty of Business, The Hong Kong Polytechnic University, Hung Hom, Hong Kong SAR, China
* Correspondence: wenyi0906@gmail.com

Abstract: One of the most common health problems that threaten the transportation infrastructure construction workers in Hong Kong is heat stress. An effective way to reduce this problem is to design a proper work–rest schedule, and the key issue is predicting the maximum working duration given the different conditions of the workers and the surrounding environment, which is the research question of this study. Air temperature, an important input feature, is also determined by the maximum working duration itself, i.e., the input feature is a function of the prediction target. Therefore, the prediction model developed is different from ordinary prediction models and is hard to solve by standard statistical or machine learning models. For the prediction process, a trial-and-error algorithm is proposed to derive a solution based on two theorems that are rigorously proved; there exists a unique solution, and the solution is within a certain range in the prediction model. The proposed model and its solution approach were constructed and validated using simulated data; temperature data were collected from Hong Kong Observatory. The results showed that the mean squared error (MSE), mean absolute percentage error (MAPE), and R^2 of the test set were 0.1378, 0.1123, and 0.8182, respectively, showing that the prediction performance was generally accurate. This study can help construction practitioners and governments to rationally design the work–rest schedules of transportation infrastructure construction workers and thus protect them from the risks brought about by heat stress.

Keywords: transportation infrastructure construction workers; maximum working duration; linear regression; trial-and-error algorithm

Citation: Yan, R.; Yi, W.; Wang, S. Predicting Maximum Work Duration for Construction Workers. *Sustainability* **2022**, *14*, 11096. https://doi.org/10.3390/su141711096

Academic Editor: Tamás Bányai

Received: 4 August 2022
Accepted: 2 September 2022
Published: 5 September 2022

1. Introduction

The Hong Kong construction industry is known for its "fast track" characteristic; its business environment is highly complex and competitive, and severe penalties are imposed on delays in schedules [1,2]. Therefore, construction workers in Hong Kong are prone to high pressure and have to work for long and irregular hours to keep up with the schedule in addition to performing demanding physical tasks characterized by excessive energy expenditure and force demands, highly repetitive actions, and awkward postures [3]. A factor that makes the situation worse is Hong Kong's climate: high temperature and relative humidity, low wind speed, and higher solar radiation levels than many other regions [4]. Consequently, construction workers in Hong Kong are more susceptible to health problems. In recent years, Hong Kong has been planning to be built into a world-class smart city, according to the *Smart City Blueprint for Hong Kong* [5], in which smart mobility and smart mobility infographics, as well as corresponding transportation infrastructure, are to be constructed in the next few years. Therefore, it can be expected that an increasing number of transportation infrastructure construction workers will be involved in this project in Hong Kong, and their working environment and health deserve much attention.

The Construction Industry Council (CIC), which is a statutory coordinating body encompassing all the key sectors in the construction industry of Hong Kong, has made sustained efforts to minimize the risk to site personnel working in hot weather to ensure the safety and health protection of construction workers and enhance working efficiency and productivity [6]. For example, the CIC published the *Guidelines on Site Safety Measures for Working in Hot Weather* in 2013, which aimed to promote the good practice recommendations of the CIC and provide guidance to the construction industry on measures that may be taken to protect construction workers working in hot weather from suffering from heat-related disorders [7]. One of the most common health problems faced by transportation infrastructure construction workers is heat stress [8]. Heat stress can impose heavy risks on transportation infrastructure construction workers' health, increase the incident rate, and reduce the enthusiasm and productivity of the workers [3]. Heat stress can even lead to the lamentable death of construction workers. For example, heat stress caused four deaths of construction workers at a construction site in Hong Kong in July 2011 [9]. In recent years, due to global warming, June was 2022 tied as Earth's warmest June on record, and the highest temperature in the same period was recorded in many places in June and July 2022 around the world [10,11]. Consequently, heat stress is expected to become a more stringent concern, jeopardizing the health and well-being of transportation infrastructure construction workers.

Fortunately, heat stress is preventable by properly designing the work–rest schedules of transportation infrastructure construction workers to improve their comfort, health, and productivity [3]. One critical point of doing so is to estimate the maximum working duration (MWD) of transportation infrastructure construction workers, where MWD refers to the maximum duration of the period (e.g., 1.5 h or 3 h) in which a worker can work continuously considering his/her physical conditions, such as the rating of perceived exertion and heart rate. MWD is related to construction workers' characteristics (age, weight, alcohol and smoking habits, etc.), job nature, and the surrounding environment (temperature, relative humidity, etc.). On the basis of the estimation of the MWD, construction workers' working and rest time can be optimized to reduce the risks brought about by heat stress [8].

A transportation infrastructure construction worker's personal habits and job nature do not change during one day's work, but the surrounding environment changes and is highly related to the MWD itself. Taking temperature as an example, the temperature value might be different at different time points over the MWD, and thus its average value is used as a feature to predict MWD, i.e., MWD is a function of the average temperature. Furthermore, because average temperature is also related to time, it can also be regarded as a function of MWD, indicating that the prediction target can influence the feature value in return in the MWD prediction model.

The above characteristic distinguishes the problem of the MWD prediction problem from other prediction models; the prediction targets of the latter have (nearly) no influence on the input features [12–15]. Consequently, standard prediction models based on statistics and machine learning methods cannot be applied directly to capture such mutually restricted relationships in an explicit manner (i.e., represented by a specific equation) [16–21]. To address this issue, this study developed a linear regression model to predict transportation infrastructure construction workers' MWD with the average temperature during this period (which was calculated by collecting and averaging the surrounding temperature values several times) as an input feature, as well as other static input features, including personal habits and job nature. We first showed that the coefficients in the linear regression model could be estimated using the least squares method. Then, to predict the MWD of a worker, we rigorously proved that there must be a unique solution to the linear regression model and showed the range of the solution. Finally, a trial-and-error approach to estimate the functions of the MWD and the average temperature as well as the average temperature and MWD was proposed.

The novelty and contribution of this study are summarized as follows. It developed a linear regression model to predict the MWD of construction workers while considering the interrelationship between the average temperature as an input feature and the MWD as the prediction target. To achieve this, two theorems of the prediction model were first proposed and rigorously proved, and then a trial-and-error method was developed to find the solution to the MWD and the average temperature during the MWD that satisfy the prediction model. The proposed model and its solution are expected to predict the MWD of transportation infrastructure construction workers more accurately and thus help industrial practitioners and governments better design the work–rest schedule of transportation infrastructure construction workers.

2. Literature Review

Several prediction models have been developed in the existing literature to predict various conditions of construction workers. A number of studies have aimed to predict construction workers' physical conditions and heat-related risks. For example, Chan et al. [8] developed a heat stress model to predict construction workers' physiological responses in hot weather based on the wet bulb globe temperature index. Yabuki et al. [4] developed a construction worker heatstroke prevention system, which consisted of a thermal environment prediction system predicting the changes in the thermal environment and a core body temperature prediction system predicting the changes in construction workers' core body temperature. Lazaro and Momayez [22] validated the performance of the heat strain prediction model to predict the core body temperature and water loss of construction workers. Safety is another major issue in construction worker management. Kwon and Kim [23] presented a construction worker accident prediction model based on association rule generation by collecting and analyzing data from environmental sensors. Kim et al. [24] predicted construction workers' inattentive behaviors to hazards by collecting and assessing their biosignal reactivity from a virtual road construction environment using classification models in machine learning. Gao et al. [25] proposed a virtual reality-based system to predict construction workers' safety behavioral tendencies on the basis of personality factors. Another stream of research aims to predict the performance of construction workers. For example, Chih et al. [26] predicted the impact of the supervisor–worker relationship on construction workers' psychological, behavioral, and performance outcomes. Aryal et al. [27] predicted physical fatigue represented by Borg's rating of perceived exertion of construction workers by collecting their skin temperature and heart rate from wearable sensors.

3. Problem Statement

A transportation infrastructure construction worker usually starts work at 7:30 and has a lunch break at or before 13:00 (after a total of 5.5 h) on a working day. However, due to the scorching and humid weather as well as exhaustion, it is highly likely that the MWD of a worker is less than 5.5 h. To predict a certain transportation infrastructure construction worker's MWD considering the surrounding factors, a historical data set on several transportation infrastructure construction workers' MWDs as well as auxiliary data were first collected and then used to construct an MWD prediction model. The auxiliary data, or features, included but were not limited to each worker's age, height (cm), weight (kg), alcohol drinking habits (never, sometimes, or always), smoking habits (never, sometimes, or always), the average surrounding temperature during the working period (°C), and job nature (bar bender and fixer, carpenter, concreter, or plumber).

It was noted that all the features were fixed for a construction worker except for the average surrounding temperature during the working period, as it is correlated with the MWD, which is also the prediction target. Suppose that there are a total of n features collected to predict workers' MWD. The feature vector is denoted by $\mathbf{x} = (x_1, ..., x_n)$ and the prediction function is denoted by $f(\mathbf{x}) = f(x_1, ..., x_n)$. Without the loss of generality, we denote by x_1 the average surrounding temperature during the whole working period. That is, x_1 is a function of $f(\mathbf{x})$ and can be represented by $x_1(f(\mathbf{x}))$. Features x_2 to x_n are fixed for each worker. Then, $f(\mathbf{x})$ can be turned into a more intuitive form: $f(\mathbf{x}) = f(x_1(f(\mathbf{x})), ..., x_n)$. The following two properties hold for functions $f(\mathbf{x})$ and $x_1(f(\mathbf{x}))$:

Property 1. *$f(\mathbf{x})$ is strictly decreasing in $x_1(f(\mathbf{x}))$. This is because, following domain knowledge, given that all other conditions are equal, when the lowest temperature is no less than a certain value (e.g., 28.4 °C, which is the average minimum temperature in July 2022 in Hong Kong [28], a higher average surrounding temperature would reduce transportation infrastructure construction workers' MWD, as a higher temperature would increase a worker's intensity of subjective effort, stress, or discomfort felt during physical activity [8], and thus shorten the maximum period that a worker could work continuously [29].*

Property 2. *$x_1(f(\mathbf{x}))$ is strictly increasing in $f(\mathbf{x})$. This is because the maximum working period of a transportation infrastructure construction worker in the morning is from 7:30 to 13:00. During this period, the temperature increases as time goes by [30], while the lowest temperature should be no less than a certain threshold in the summertime in Hong Kong (e.g., 28.4 °C in July 2022 (Hong Kong Observatory, 2022)). Therefore, the average temperature during a worker's* MWD *$x_1(f(\mathbf{x}))$, which is from 7:30 to 13:00, strictly increases as $f(\mathbf{x})$ increases.*

An illustration of functions $f(\mathbf{x})$ and $x_1(f(\mathbf{x}))$ based on Property 1 and Property 2 is shown in Figure 1. The x-axis represents the average temperature during the MWD, and the y-axis represents the MWD. Function $x_1(f(\mathbf{x}))$ is represented by line AB, and it increases in f as the temperature increases as the working time accumulates. In particular, $A = (T_0, 0)$ shows that the temperature is T_0 when the construction worker starts to work at time 0 (i.e., 7:30). $B = (T, 5.5)$ shows that the average temperature is T at 13:00 when the lunch break starts. Function $f(\mathbf{x})$, which is linearly decreasing with x_1, is represented by Line CD. In particular, point $C = (T_0, h)$ shows that the maximum value of f is h hours (must be larger than 0 h and can be either larger or no larger than 5.5 h), as it is associated with the lowest average temperature. The value of f for point D is negative, as a construction worker cannot continuously work for 5.5 h without a break when the overall average temperature is T.

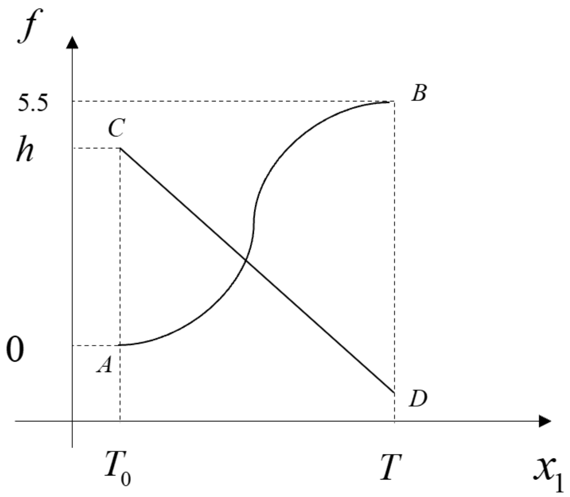

Figure 1. An illustration of functions f and x_1.

4. Development of a Prediction Model for MWD

We assume that $f(\mathbf{x})$ has a linear relation with $\mathbf{x}$, shown as follows:

$$\begin{aligned} f(\mathbf{x}) &= f(x_1(f(\mathbf{x})), \ldots, x_n) \\ &= b_0 + b_1 x_1(f(\mathbf{x})) + b_2 x_2 + \ldots + b_n x_n, \end{aligned} \quad (1)$$

where b_0 to b_n are coefficients that need to be learned from data. Given a data set consisting of historical working records (including values of all features and the MWD as the target) of construction workers, $x_1(f(\mathbf{x}))$ is a fixed value and $f(\mathbf{x})$ (i.e., MWD) is known. Therefore, $x_1(f(\mathbf{x}))$ can be treated as a normal feature in the model training (i.e., parameter estimation) process. Then, the coefficients of the MWD prediction model in Equation (1) can be estimated using standard algorithms for coefficient estimation in the linear regression, such as the least squares method [31], and the values of b_0 to b_n can be obtained.

The prediction process using the MWD prediction model in Equation (1) is different from that using a normal linear regression model as well as other machine learning models, as the predicted MWD, i.e., $f(\mathbf{x})$, can influence the average temperature during this period, i.e., $x_1(f(\mathbf{x}))$, as an input feature. In return, the function $x_1(f(\mathbf{x}))$ also determines the value of $f(\mathbf{x})$. It should also be mentioned that, as the surrounding temperature information is constantly collected for a construction worker whose MWD needs to be predicted, the value of $x_1(f(\mathbf{x}))$ can be calculated when a specific $f(\mathbf{x})$ is given. In other words, both functions are mutually restricted from each other, as they have opposite monotonicity properties to each other. Meanwhile, for a given worker whose MWD needs to be predicted, all the other features of this worker, i.e., x_2 to x_n, are known. Therefore, we can simplify the function $x_1(f(\mathbf{x}))$ to $x_1(f(x_1))$ or x_1 and the function $f(\mathbf{x})$ to $f(x_1)$ or f. The solution to the MWD prediction problem shown in Equation (1) is denoted by $(x_1^*(f^*(x_1^*)), f^*(x_1^*))$ or (x_1^*, f^*) for short. The following two theorems hold.

Theorem 1. *There exists a unique solution.*

Proof. As shown in Figure 1, there must exist a solution to the MWD prediction model in Equation (1) considering the coordinates and relations of the four points determining the two functions. Figure 1 also shows that, due to the monotonicity property of the two functions, a unique solution exists, which can also be mathematically proved by contradiction. Suppose there are two solutions, $(x_1^{\#}, f^{\#})$ and $(x_1^{\&}, f^{\&})$. Note that we must have $x_1^{\#} \neq x_1^{\&}$, because otherwise $f^{\#} = f^{\&}$ and thus they are the same solution. Without the loss of generality, assume $x_1^{\#} < x_1^{\&}$. As $x_1(f(\mathbf{x}))$ is strictly increasing in $f(\mathbf{x})$, we can have $f^{\#} < f^{\&}$, where the relationship between these two solutions is shown in Figure 2. □

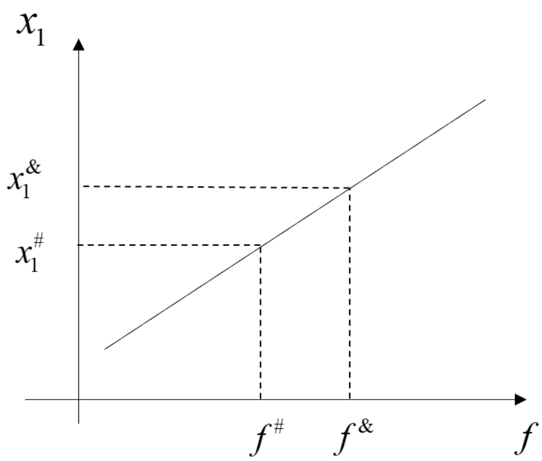

Figure 2. The relationship between $f^{\#}$ and $f^{\&}$ given $x_1^{\#} < x_1^{\&}$.

Then, if we have $f^{\#} < f^{\&}$, and given the fact that f is strictly decreasing in x_1, the relationship between $x_1^{\#}$ and $x_1^{\&}$, which are denoted by $x_1^{\#\prime}$ and $x_1^{\&\prime}$ in this case, respectively, is shown in Figure 3. Figure 3 shows that $x_1^{\#\prime} > x_1^{\&\prime}$ given $f^{\#} < f^{\&}$, which is contradictory to the assumption $x_1^{\#} < x_1^{\&}$.

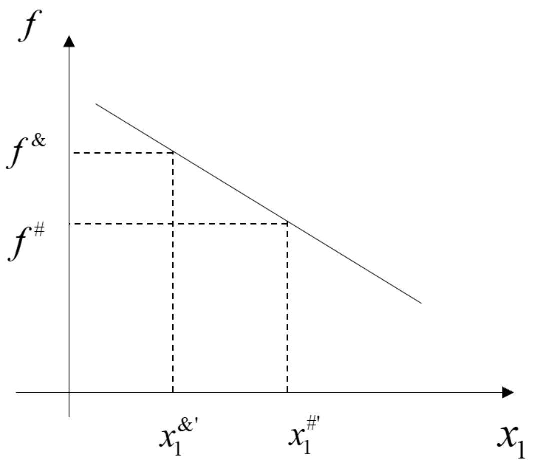

Figure 3. The relationship between $x_1^{\#}$and$x_1^{\&}$ given $f^{\#} < f^{\&}$.

Therefore, it can be concluded that there is a unique solution to the MWD prediction model in Equation (1). □

Theorem 2. *For an $x_1^{\#}$ and the calculated $f^{\#} = f(x_1^{\#})$ (note that $(x_1^{\#}, f^{\#})$ might not be a solution), we have $x_1^* \in [\min(x_1^{\#}, x_1(f^{\#})), \max(x_1^{\#}, x_1(f^{\#}))]$ and $f^* \in [\min(f^{\#}, f(x_1(f_1^{\#})), \max(f^{\#}, f(x_1(f_1^{\#}))]$, where (x_1^*, f^*) is the unique solution of the MWD prediction problem shown in Equation (1).*

Proof. Theorem 2 can be proven by contradiction. Assume $x_1^* < \min(x_1^{\#}, x_1(f^{\#}))$. As f is strictly decreasing in x_1, we can have $f^* > \max(f^{\#}, f(x_1(f^{\#})))$. However, as x_1 is strictly increasing in f, we have $x_1^{*\prime} > \max(x_1^{\#}, x_1(f^{\#}))$, which is contradictory to the assumption. Assume that $x_1^* > \max(x_1^{\#}, x_1(f^{\#}))$; we can have $f^* < \min(f^{\#}, f(x_1(f^{\#})))$, as f is strictly decreasing in x_1, and thus $x_1^{*\prime} < \min(x_1^{\#}, x_1(f^{\#}))$, as x_1 is strictly increasing in f, which is also contradictory to the assumption. A similar situation also applies to f^* when $f^* < \min(f^{\#}, f(x_1(f^{\#})))$ and $f^* > \max(f^{\#}, f(x_1(f^{\#})))$. □

Theorem 2 shows that in solution (x_1^*, f^*) to the MWD prediction model, x_1^* and f^* restrict each other by their mutual monotonicity relationships, as Equation (1), which involves both functions, needs to be satisfied. Therefore, the values of both functions are within certain range that is determined by its original value and the value calculated given the value of the other function.

On the basis of Theorems 1 and 2, (x_1^*, f^*) in Equation (1) can be solved by the trial-and-error method shown in Algorithm 1.

Algorithm 1 Trial-and-error method to find (x_1^*, f^*)

Initialize a sufficiently small f_0 (i.e., smaller than f^*). Define $\hat{f}_0 = f_0$ and $\hat{f}_{-1} = f_0$;
Set the tolerance gap $t = 0.01$ and the current gap $g = \inf$;
$k = 0$;
while $g > t$:
$k = k + 1$;
$f_k = f(\overline{x}(\hat{f}_{k-1}))$; // predict the initial value of f using the final predicted value in the last round
if $f_{k-1} \geq \hat{f}_{k-1}$:
$f_k = \max(f_k, \hat{f}_{k-2})$;
else:
$f_k = \min(f_k, \hat{f}_{k-2})$;
$\hat{f}_k = \frac{\hat{f}_{k-1} + f_k}{2}$;
$g = \left|\hat{f}_k - \hat{f}_{k-1}\right|$;
Return $f^* = \hat{f}_k$.

$\overline{x}(\hat{f}_{k-1})$ is a function to calculate the average temperature given the working duration $\hat{f}_{k-1}$ of a specific day. Algorithm 1 guarantees that the current gap is no larger than half of the gap in the last round. In particular, f_k is the initial predicted value, and $\hat{f}_k$ is the final predicted value for f in this round. f_k is used to control the range of $\hat{f}_k$ by considering the updating direction of f, i.e., the two possible relationships between $\hat{f}_{k-2}$ and $\hat{f}_{k-1}$. To be more specific, when $f_{k-1} > \hat{f}_{k-1}$, and recall that $\hat{f}_{k-1} = \frac{\hat{f}_{k-2}+f_{k-1}}{2}$, we have $\hat{f}_{k-2} < \hat{f}_{k-1}$, which means that the value of f increases in the last round of updating. Further considering that function x_1 is linearly increasing in f and f is linearly decreasing in x_1, we can expect that $f_{k-1} > f_k$. Therefore, there are three possible positions of f_k, as shown in Figures 4–6.

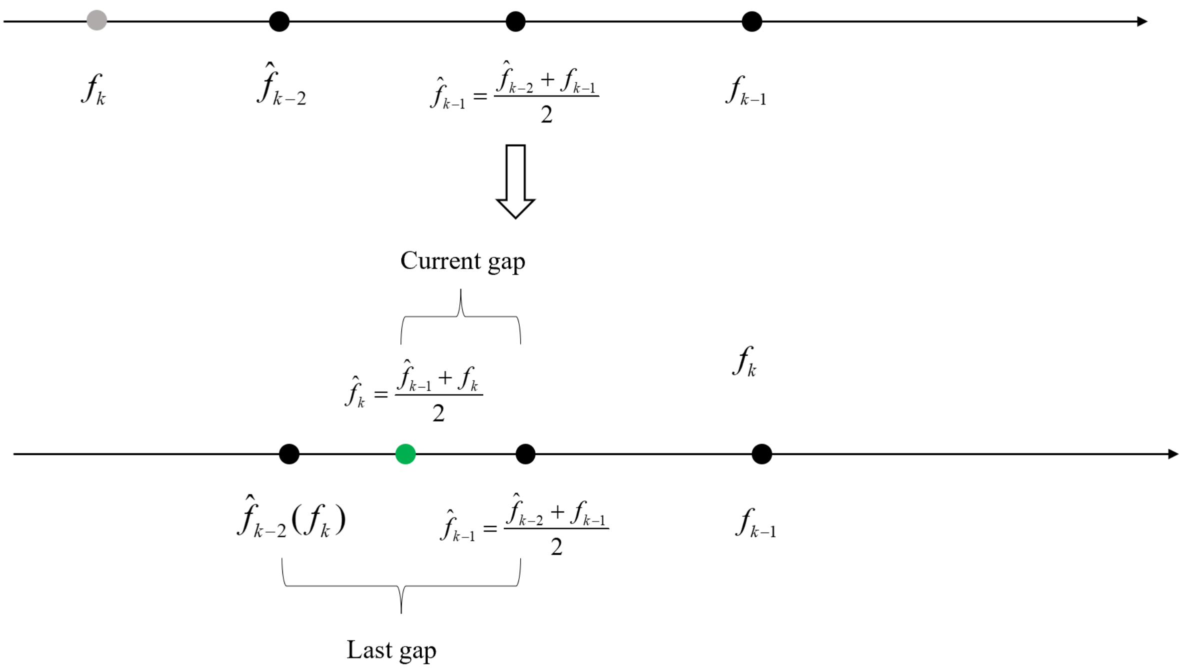

Figure 4. The situation when $f_k \leq \hat{f}_{k-2}$.

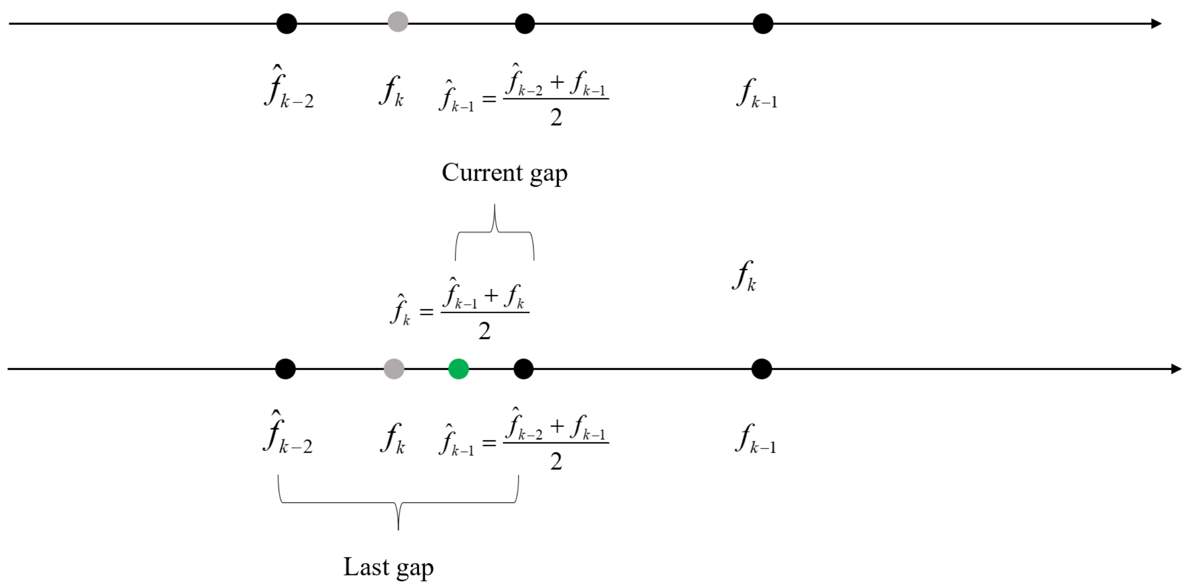

Figure 5. The situation when $\hat{f}_{k-2} < f_k \leq \hat{f}_{k-1}$.

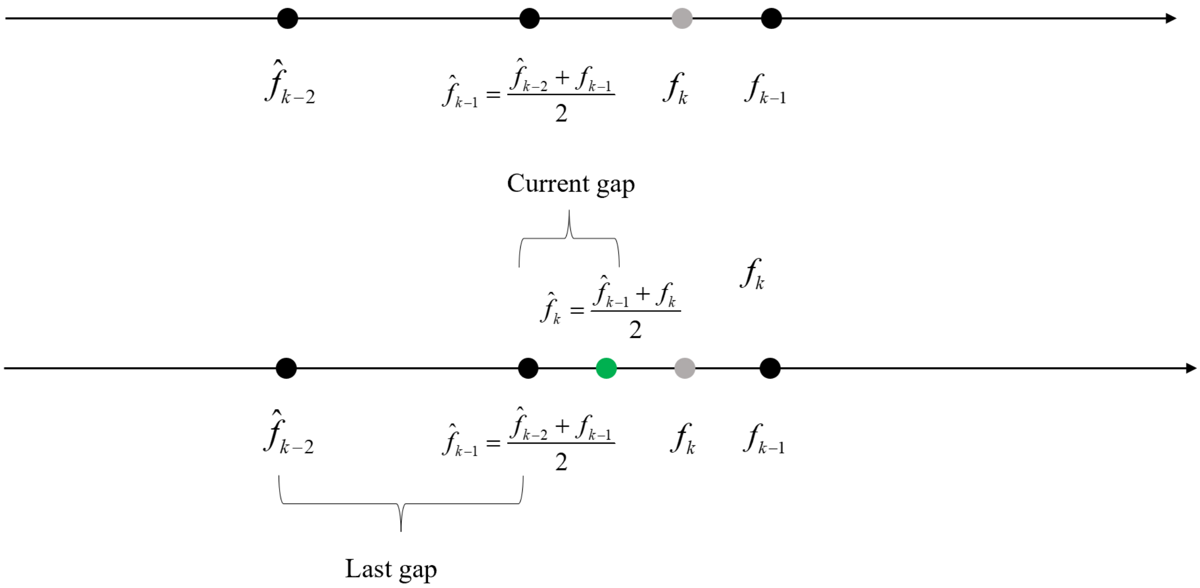

Figure 6. The situation when $\hat{f}_{k-1} < f_k < f_{k-1}$.

In Figure 3, $f_k \leq \hat{f}_{k-2}$. To guarantee that the gap in the current round k is no more than half of the gap in the last round $k-1$, we need to require that f_k is no smaller than $\hat{f}_{k-2}$, and thus the value of f_k is constrained to $f_k = \max(f_k, \hat{f}_{k-2}) = \hat{f}_{k-2}$, as shown by the bottom of Figure 3. Then, the predicted value of f in this round, i.e., $\hat{f}_k$, is shown by the green dot, and the gap in the current round is equal to half of the gap in the last round.

In Figure 4, $\hat{f}_{k-2} < f_k \leq \hat{f}_{k-1}$, and the current gap is smaller than half of that in the last round. If in the next round of solution updating is needed, the "if" condition in Algorithm 1 holds. Similarly, in Figure 5, $\hat{f}_{k-1} < f_k < f_{k-1}$, and the current gap is also smaller than half of that in the last round. If in the next round of solution updating is needed, the "else" condition in Algorithm 1 holds.

By contrast, when $\hat{f}_{k-1} < f_{k-1}$, the situations are the exact opposite of the situations in which $\hat{f}_{k-1} > f_{k-1}$, where the current gap equals half of the gap in the last round when $f_k \geq \hat{f}_{k-2}$ and is less than half of the gap in the last round when $f_{k-1} < f_k < \hat{f}_{k-2}$. After finding f^*, the corresponding x_1^* can be calculated, and thus the solution to construction workers' MWD prediction model (x_1^*, f^*) can be obtained.

5. Case Study

A case study based on simulated data is presented in this section. A construction worker's age, height, weight, alcohol drinking habits, smoking habits, and job nature are assumed. We required that the difference between the construction worker's height (cm) and weight (kg) was between 910 and 110. The average temperature was calculated according to daily weather information in Hong Kong in August 2022 according to the Hong Kong Observatory (https://www.hko.gov.hk/en/cis/dailyExtract.htm?y=2022&m=08, accessed on 1 September 2022). Specifically, the minimum and maximum temperatures on each day in August 2022 in Hong Kong were found, and the hourly temperatures from 7:30 to 13:00 were assumed on the basis of the minimum and maximum temperature values. We assumed that the temperature was identical for one hour, i.e., 7:30 to 8:29 had the same temperature, 8:30 to 9:29 had the same temperature, etc., and the hourly temperature was between the minimum and maximum temperature values and increased as time went by. Then, the average temperature during the MWD (in minutes) could be calculated accordingly. We required that the MWD of each worker satisfied Equation (2):

$$0.2\% \times \max(1, (t-25)) \times MWD + \varepsilon = 100\% \tag{2}$$

We set the base temperature to 25 °C, which is a temperature that makes people feel comfortable. t is the temperature, which can be different in different periods within the MWD. We further assumed that if the surrounding temperature was no more than 25 °C, working for one minute would increase the fatigue degree of a construction worker by 0.2%. Otherwise, working for one minute would increase the fatigue degree by $0.2\% \times (t - 25)$. When the fatigue degree reached 100%, the MWD of the construction worker was reached. ε is a random term following the normal distribution with a mean of zero and variance of 0.001.

Based on the above assumptions, we generate a data set consisting of 310 records by using one day's temperature information in August 2022 in Hong Kong to generate 10 records. The descriptive statistics of the data set formulated are shown in Tables 1 and 2.

Table 1. Descriptive statistics of continuous variables in the data set.

Feature	Way of Generation	Min Value	Mean Value	Max Value	Standard Deviation
Age	Assumption: a random integer between 28 to 65	28	48.24	65	10.58
Height (cm)	Assumption: a random integer between 155 to 185	155	170.03	185	9.06
Weight (kg)	Assumption: a random integer between 55 to 85	55	69.99	85	8.20
Average temperature	Calculated from Hong Kong Observatory considering the daily highest and lowest temperatures in August 2022	25.84	27.80	29.95	0.94
MWD	Calculated by Equation (2)	1.68	5.50	3.18	0.97

Table 2. Descriptive statistics of categorical variables in the data set.

Feature	Way of Generation	Distribution
Alcohol drinking habits *	Assumption: a random integer in set {0,1,2}	0: 104, 1: 96, 2: 110
Smoking habits **	Assumption: a random integer in set {0,1,2}	0: 94, 1: 112, 2: 104
Job nature ***	Assumption: a random variable with any of the following binary features to be 1 and the other three to be 0: is_bar_bender_and_fixer, is_carpenter, is_concretor, and is_plumber	is_bar_bender_and_fixer = 1: 74, is_carpenter = 1: 88, is_concretor = 1: 79, is_plumber = 1: 69

Note *: 0 = none, 1 = sometimes, 2 = often; **: 0 = none, 1 = sometimes, 2 = often; ***: this feature has four values: bar bender and fixer, carpenter, concretor, and plumber. It is discretized into four binary features: is_bar_bender_and_fixer, is_carpenter, is_concretor, and is_plumber.

The whole data set was randomly divided into a training set containing 80% of the records (i.e., 248 records) and a test set containing 20% of the records (i.e., 62 records). The coefficients in Equation (1) were estimated by the least squares method, and the equation took the following form:

$$\begin{aligned} f(\mathbf{x}) &= f(x_1(f(\mathbf{x})), \ldots, x_n) \\ &= 5.2174 + 0.0094x_1(f(\mathbf{x})) + (-0.0704)x_2 + 0.1365x_3 + (-0.0836)x_4 + (-0.0069)x_5 \\ &+ (-0.0090)x_6 + 0.0085x_7 + 0.0394x_8 + (-0.0389)x_9 + (-4.2270)x_{10}. \end{aligned} \tag{3}$$

Then, model performance was evaluated on the test set using Algorithm 1. The mean squared error (MSE), mean absolute percentage error (MAPE), and R^2 of the test set were 0.1378, 0.1123, and 0.8182, respectively, showing that the prediction performance was generally accurate.

6. Conclusions and Future Research

To build Hong Kong into a world-class smart city, a large number of transportation infrastructure construction workers are expected to become involved in the related projects, and their working environment and health should be paid close attention to. Due to the scorching and humid climate and the "fast track" nature of construction projects in Hong Kong, these transportation infrastructure construction workers are prone to heat stress, which imposes a heavy risk to their health and working status. The potential risks of heat stress can be reduced by designing reasonable work–rest schedules for the workers. This study aimed to deal with one critical point in work–rest schedule design: the prediction of the MWD of transportation infrastructure construction workers considering their personal habits, job nature, and the surrounding environmental conditions. The surrounding temperature is a critical feature for MWD prediction and varies as time goes by on a working day. Therefore, the average temperature during the whole working period was considered as a feature to predict the MWD. In return, the MWD also influences the value of the average temperature as an input feature. This characteristic of the MWD prediction problem distinguishes it from ordinary prediction models in which the target to be predicted has (nearly) no influence on the features. Therefore, standard statistical and machine learning methods cannot be directly applied.

To deal with this issue, this study first developed a linear regression model for MWD prediction with the average temperature as one of the inputs. The coefficients of the linear regression model were estimated by ordinal least squares in the model construction stage. In the prediction stage, a trial-and-error algorithm was proposed to find the function forms of the MWD and the average temperature and of the average temperature and the MWD as the model solution based on two properties of the model: there is a unique solution to the model whose range is between an upper bound and a lower bound. The proposed model and its solution approach can help construction practitioners and the governments to better predict construction workers' working status and thus protect them from the risks brought about by heat stress.

There are some limitations to this study. First, relative humidity, an important influence factor on the transportation infrastructure construction workers' MWD, was not considered in this study. It should also be mentioned that relative humidity has a similar characteristic to the average temperature, as it is highly related to the MWD while it is also influenced by the surrounding temperature [32]. In addition, the data used in the case study are not real data but simulated ones, which may not be able to reveal the performance of the proposed model and its solution accurately and comprehensively. Moreover, linear regression is a relatively elementary model to address prediction tasks. More sophisticated and accurate machine learning models for regression can be developed for MWD prediction.

Therefore, there are several directions for future research. First, real data collected from construction workers in Hong Kong can be used to construct the model and evaluate model performance. Then, the model and situation of construction workers in Hong Kong can be compared with those in similar cities, such as Singapore, to validate the model's relevance and generate more useful managerial insights. In addition, more features that could influence construction workers' MWD can be used to predict the MWD. For example, the lowest and highest temperatures of one day could be incorporated as normal features (i.e., the target is not a function of them) for MWD prediction, as they can also influence construction workers' MWD because a day's lowest temperature usually occurs before and during sunrise (i.e., just before the start of the work), and the highest temperature usually occurs at noon (i.e., around the lunch break). In addition, relative humidity, which is also a function of the prediction target, i.e., MWD, and is highly dependent on the surrounding temperature, can be considered as a feature together with the average temperature. In this case, the model training process would not be influenced much and would be similar to the model proposed in this study. However, the test process could be much more complex, as the average humidity and temperature and the MWD are interrelated. One viable way to predict the MWD of a specific construction worker would be to first calculate the specific function forms of the relative humidity and the MWD and of the temperature and the MWD, then input the specific function forms into the prediction model (and there is only one unknown parameter in the prediction model), and finally calculate the corresponding MWD. Another viable way is similar to the solution proposed in this study: the properties of the model would be derived first, and the existence, uniqueness, and range of the solutions would then be figured out. Finally, proper algorithms, such as the trial-and-error method, could be developed to find the values of the MWD together with the average humidity and temperature over this period.

Author Contributions: Conceptualization, W.Y.; methodology, R.Y., W.Y. and S.W.; investigation, R.Y.; writing—original draft preparation, R.Y.; writing—review and editing, W.Y. and S.W. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: Data will be available upon request.

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

References

1. Hong Kong Housing Authority. 2022. Available online: https://www.housingauthority.gov.hk/mini-site/site-safety/common/resources/handbook/201603/PG_res-en.pdf (accessed on 29 July 2022).
2. Yi, W.; Chan, A.P.C. Optimal work pattern for construction workers in hot weather: A case study in Hong Kong. *J. Comput. Civ. Eng.* **2015**, *29*, 1–11. [CrossRef]
3. Yi, W.; Chan, A.P.C. Optimizing work–rest schedule for construction rebar workers in hot and humid environment. *Build. Environ.* **2013**, *61*, 104–113. [CrossRef]
4. Yabuki, N.; Onoue, T.; Fukuda, T.; Yoshida, S. A heatstroke prediction and prevention system for outdoor construction workers. *Vis. Eng.* **2013**, *1*, 1–16. [CrossRef]
5. Smart City Bluepoint for Hong Kong. "Vision and Mission of HK Smart City Blueprint". 2017. Available online: https://www.smartcity.gov.hk/vision-and-mission.html (accessed on 1 August 2022).
6. CIC. About CIC. 2022. Available online: https://www.cic.hk/eng/index.html (accessed on 28 August 2022).
7. CIC. Guidelines on Site Safety Measures for Working in Hot Weather. 2013. Available online: https://www.cic.hk/cic_data/pdf/about_cic/publications/eng/V10_6_e_V00_Guidelines%20on%20Site%20Safety%20Measures%20for%20Working%20in%20Hot%20Weather.pdf (accessed on 28 August 2022).
8. Chan, A.P.; Yam, M.C.; Chung, J.W.; Yi, W. Developing a heat stress model for construction workers. *J. Facil. Manag.* **2012**, *10*, 59–74. [CrossRef]
9. AM730. Four Construction Workers Died of Heatstroke, and the Labor Union Urged Legislation to Increase Rest Time [Chinese Version]. 2011. Available online: http://archive.am730.com.hk/article-66052 (accessed on 30 July 2022).
10. Chinanews. Many Countries Around the World Recorded Record High Temperatures in June and Continued to Set New Records after Entering July [Chinese Version]. 2022. Available online: https://www.chinanews.com.cn/gj/2022/07-07/9797843.shtml (accessed on 30 July 2022).
11. Yale Climate Connections. NASA: June 2022 Tied as Earth's Warmest June on Record. 2022. Available online: https://yaleclimateconnections.org/2022/07/nasa-june-2022-tied-as-earths-warmest-june-on-record/ (accessed on 31 July 2022).
12. Wang, S.; Yan, R.; Qu, X. Development of a non-parametric classifier: Effective identification, algorithm, and applications in port state control for maritime transportation. *Transp. Res. Part B Methodol.* **2019**, *128*, 129–157. [CrossRef]
13. Yan, R.; Wang, S.; Du, Y. Development of a two-stage ship fuel consumption prediction and reduction model for a dry bulk ship. *Transp. Res. Part E Logist. Transp. Rev.* **2020**, *138*, 101930. [CrossRef]
14. Yan, R.; Wang, S.; Cao, J.; Sun, D. Shipping domain knowledge informed prediction and optimization in port state control. *Transp. Res. Part B Methodol.* **2021**, *149*, 52–78. [CrossRef]
15. Yan, R.; Wang, S.; Zhen, L.; Laporte, G. Emerging approaches applied to maritime transport research: Past and future. *Commun. Transp. Res.* **2021**, *1*, 100011. [CrossRef]

16. Wang, S.; Yan, R. A global method from predictive to prescriptive analytics considering prediction error for "Predict, then optimize" with an example of low-carbon logistics. *Clean. Logist. Supply Chain* **2022**, *4*, 100062. [CrossRef]
17. Zhang, J.; Long, D.Z.; Wang, R.; Xie, C. Impact of penalty cost on customers' booking decisions. *Prod. Oper. Manage* **2021**, *30*, 1603–1614. [CrossRef]
18. Long, D.Z.; Qi, J. Distributionally robust discrete optimization with entropic value-at-risk. *Oper. Res. Lett.* **2014**, *42*, 532–538. [CrossRef]
19. Yi, W.; Wu, S.; Zhen, L.; Chawynski, G. Bi-level programming subsidy design for promoting sustainable prefabricated product logistics. *Clean. Logist. Supply Chain* **2021**, *2021*, 100005. [CrossRef]
20. Yi, W.; Zhen, L.; Jin, Y. Stackelberg game analysis of government subsidy on sustainable off-site construction and low-carbon logistics. *Clean. Logist. Supply Chain* **2021**, *2021*, 100013. [CrossRef]
21. Wu, L.; Adulyasak, Y.; Cordeau, J.-F.; Wang, S. Vessel service planning in seaports. *Oper. Res.* **2022**, *70*, 2032–2053. [CrossRef]
22. Lazaro, P.; Momayez, M. Validation of the predicted heat strain model in hot underground mines. *Min. Metall. Explor.* **2019**, *36*, 1213–1219. [CrossRef]
23. Kwon, J.H.; Kim, E.J. Accident prediction model using environmental sensors for industrial internet of things. *Sens. Mater.* **2019**, *31*, 579–586. [CrossRef]
24. Kim, N.; Kim, J.; Ahn, C.R. Predicting workers' inattentiveness to struck-by hazards by monitoring biosignals during a construction task: A virtual reality experiment. *Adv. Eng. Inform.* **2021**, *49*, 101359. [CrossRef]
25. Gao, Y.; González, V.A.; Yiu, T.W.; Cabrera-Guerrero, G.; Li, N.; Baghouz, A.; Rahouti, A. Immersive virtual reality as an empirical research tool: Exploring the capability of a machine learning model for predicting construction workers' safety behaviour. *Virtual Real.* **2022**, *26*, 361–383. [CrossRef]
26. Chih, Y.Y.; Kiazad, K.; Cheng, D.; Lajom, J.A.L.; Restubog, S.L.D. Feeling positive and productive: Role of supervisor–worker relationship in predicting construction workers' performance in the Philippines. *J. Constr. Eng. Manag.* **2017**, *143*, 04017049. [CrossRef]
27. Aryal, A.; Ghahramani, A.; Becerik-Gerber, B. Monitoring fatigue in construction workers using physiological measurements. *Autom. Constr.* **2017**, *82*, 154165. [CrossRef]
28. Hong Kong Observatory. The Weather of July 2022. 2022. Available online: https://www.hko.gov.hk/en/wxinfo/pastwx/mws2022/mws202207.htm (accessed on 2 August 2022).
29. Chan, A.P.; Yi, W.; Chan, D.W.; Wong, D.P. Using the thermal work limit as an environmental determinant of heat stress for construction workers. *J. Manag. Eng.* **2013**, *29*, 414–423. [CrossRef]
30. Caltech Teaching Notes. Air Temperature, GEOL 1350: Introduction to Meteorology. 2022. Available online: http://web.gps.caltech.edu/~{}xun/course/GEOL1350/Lecture5.pdf (accessed on 31 July 2022).
31. Plackett, R.L. A historical note on the method of least squares. *Biometrika* **1949**, *36*, 458–460. [CrossRef]
32. Jones, R.A.C. Future scenarios for plant virus pathogens as climate change progresses. *Adv. Virus Res.* **2016**, *95*, 87–147.