Journal of Marine Science and Technology

Volume 21 | Issue 1 Article 9

RAINFALL PREDICTION USING INNOVATIVE GREY MODEL WITH THE DYNAMIC INDEX

Yong-Huang Lin
Department of Construction Engineering, National Taiwan University of Science and Technology, Taipei, Taiwan, R.O.C.

Chih-Chiang Chiu
Department of Construction Engineering, National Taiwan University of Science and Technology, Taipei, Taiwan, R.O.C.

Yong-Jun Lin Lin
Center for Weather Climate and Disaster Research, National Taiwan University Taipei, Taiwan, R.O.C.

Pin-Chan Lee
Department of Civil Engineering and Hazard Mitigation Design, China University of Technology, Taipei, Taiwan, R.O.C

Recommended Citation
Lin, Yong-Huang; Chiu, Chih-Chiang; Lin, Yong-Jun Lin; and Lee, Pin-Chan (2013) "RAINFALL PREDICTION USING INNOVATIVE GREY MODEL WITH THE DYNAMIC INDEX," *Journal of Marine Science and Technology*: Vol. 21: Iss. 1, Article 9.
DOI: 10.6119/JMST-011-1116-1
Available at: https://jmstt.ntou.edu.tw/journal/vol21/iss1/9

Journal of Marine Science and Technology, Vol. 21, No. 1, pp. 63-75 (2013)
DOI: 10.6119/JMST-011-1116-1

RAINFALL PREDICTION USING INNOVATIVE GREY MODEL WITH THE DYNAMIC INDEX

Yong-Huang Lin[1], Chih-Chiang Chiu[1], Yong-Jun Lin[2], and Pin-Chan Lee[3]

Key words: grey model, Fourier series, exponential smoothing technique, dynamic index.

ABSTRACT

Taiwan's special climate and landforms are affected by summer typhoons, with 78% of its rainfall occurring during the summer and autumn months. The range and the severity of disasters has increased in recent years, thanks in part to climate change, which has caused an unstable rainfall. Accurate rainfall predictions help to forecast rivers' water levels. This study proposes a new rainfall prediction model based on features of the rainfall system.

In order to overcome the drawbacks found in the original grey model, a few corrections have been made to the new model. First, the dynamic index transformation is used to generate an exponentially smooth sequence. When the new model was applied to data from eight different typhoons, the results revealed that the mean peak rainfall error, compared to the original sequences, is close to 0. This technique can also effectively increase the accuracy of maximum rainfall predictions. Next, the grey model's background value was improved through integration. This technique can correct any delay in the peak rainfall as predicted by the conventional model, and make the predicted and actual values closer. Finally, we used the Fourier series and the exponential smoothing technique to correct periodical and random errors. The new model is called the Dynamic index Exponential Fourier Grey Model (DEFGM (1,1)). By examining different indicators, the mean coefficient of efficiency of the DEFGM (1,1) was found to be close to 1, which is indicative of a relatively good overall performance. With this tool, the predictability of rainfall systems during typhoons is more accurate, and disaster prevention measures can be made in advance.

Paper submitted 02/25/11; revised 07/09/11; accepted 11/16/11. Author for correspondence: Pin-Chan Lee (e-mail: alanlee68@cute.edu.tw).
[1] Department of Construction Engineering, National Taiwan University of Science and Technology, Taipei, Taiwan, R.O.C.
[2] Center for Weather Climate and Disaster Research, National Taiwan University Taipei, Taiwan, R.O.C.
[3] Department of Civil Engineering and Hazard Mitigation Design, China University of Technology, Taipei, Taiwan, R.O.C.

I. INTRODUCTION

Climate change has intensified the rainfalls due to stronger typhoons, lengthened their duration, and increased the total amount of rainfall. As a result, the possibility of floods is increasing. Therefore, a new strategy for flood-prevention is needed. A real-time, high-resolution, rainfall prediction model is necessary to support any decision-making.

Generally, short-term rainfall is predicted through the stochastic mathematical model [2, 11]. However, a large amount of explanatory variables as well as a great deal of data is required, and obtaining this is a rather difficult task. The general prediction model must be constructed under such conditions where the explanatory variables are definite values, both random and normally distributed, as well as complying with the statistical hypothesis testing. Also, it is impossible to build a good prediction model without collecting the complete historical and statistical data. Therefore, information obtained by the system often fails to cover the required data completely, and the grey system is the only theory generated to handle such conditions. The grey model can get a satisfactory prediction result, despite having obtained little data. Even with inadequate data and information, the grey model analyzes, predicts and determines the system [6]. This model can be built through generating more than 4 data points [4, 5]. Some studies have proven that the grey model can be built by using only 3 data points [12]. Not only can this model predict equal-interval time sequences but it can also be applied in a non-equal-interval series, or a negative numbered series, etc., given its good applicability. It uses little data and applies the data generation method to reduce the possible influence by each factor. Yang [21] estimated the throughput volume of sea-air transport cargo using the grey prediction model, and Kayacan *et al*. [10] predicted a time sequence using the grey model, and obtained good results in both fit and predictions.

Yu *et al*. [22] presented the rainfall prediction model based on the Grey model. Kang *et al*. [9] predicted the volume of real-time flood discharge by applying the hydrologic grey model in combination with the global search method proposed via Lin [14] based on the Chaos Theory. Bedient *et al*. [1] and Kang *et al*. [8] observed rainfall through a radar rainstorm system. Each method had its own characteristics.

However, the grey forecasting model is constructed, based on the exponential function, which has adverse results in predicting a wavy sequence. The most commonly used grey forecasting model is the GM (1,1). The rainfall intensity using hourly rainfall data is unstable because it is affected by many factors that fluctuate instead of developing at a certain rate. Xiang [20] adopted the dynamic index to convert the river's chemical oxygen demand (COD) to a monotonic incremental sequence, which is applicable to the grey forecasting model. The converted sequence is in line with the system's behaviors, thus solving the fluctuation problem and making prediction more reliable and feasible.

The development coefficient and the grey input of the grey model were estimated by Yu *et al*. [22] using the fuzzy regression technology instead of the least square method and applied it to rainfall predictions for disaster prevention. The basis for the grey model is the exponential function, so using it to predict center-symmetry curves or a random time series is not ideal [12]. In order to obtain desirable results, any model must be built according to the problem's characteristics. Thus different problems should apply different models. Cheng *et al*. [3] extracted the feature signals by combining the grey forecasting model and the wavelet transformation technique, and applied it in the failure prediction of hydraulic and electrical machineries; which improved prediction accuracy. Guo *et al*. [7], Lin *et al*. [13], and Su *et al*. [19] adopted the Fourier series to correct the cyclic residual difference of the grey forecasting model and to improve its prediction accuracy, as well as expand the model's scope of applications. Other studies constructed a high-precision prediction model by combining the grey model and the Fourier series, etc. [13, 15]. They compared the simulated data of this model, the fuzzy prediction model, and the back propagation artificial neural network model with actual data. The results proved that the model had better prediction accuracy than others.

However, the randomness and uncertainty of the wavy sequence does not change with the exponential law. It is possible to fully use the information from the data series to greatly reduce the randomness, if the single exponential smoothing of the time sequence is introduced into the grey prediction model. Therefore, the original data series is reconstructed through exponential smoothing to improve calculations of the grey model's background value. As a result, the original sequence is transformed into an exponential series with stronger regularity [16, 18].

This study mainly adopts the dynamic index transformation model to transform the original data sequence into an exponential smoothing function sequence. In other words, it converts the rainfall system with an unstable development into a stable system. The background value of the grey model is integrated to improve the accuracy of predictions. This combined with modification of the exponential smoothing and the Fourier series to correct its residual is called an EFGM (1,1) to predict rainfall. The rainfall prediction accuracy of the GM (1,1), DGM (1,1), EFGM (1,1) and DEFGM (1,1) is also evaluated, based on different indicators. The results indicate that the DEFGM (1,1) is the most accurate indicator when compared with others [16]. Finally, the study's conclusions and recommendations are proposed.

II. FUNDAMENTAL CONCEPTS OF GREY THEORY

This section introduces the concepts of the grey prediction model and the dynamic index, which form the basis of construction of the prediction model.

1. Grey Forecasting Model

Owing to the variations of the internal and external environments, the system development is usually irregular. Therefore, Deng [6] recommended adopting a technique of accumulation generating operation (AGO) to reveal the regular pattern hidden in the system development. The grey forecasting model is the GM (1,1), which indicates that one variable is employed in the model with the first order differential equation being adopted to match the data generated by the AGO. The generating function of the grey systems can be expressed in the following equations:

$$X^{(0)}(k) - aZ^{(1)}(k) = b \tag{1}$$

$$X^{(1)}(k) = \sum_{m=1}^{k} X^{(0)}(m) \tag{2}$$

$Z^{(1)}(k) = \alpha X^{(1)}(k) + (1 - \alpha)X^{(1)}(k - 1)]$ is the background value, where α is often set as its representative value of 0.5. However, to effectively resolve the prediction problem, value α needs to be adjusted according to the series features. Parameter a is called the developing coefficient, and b is the grey input.

The predicting series of $X^{(0)}$ is sorted as, $X^{(0)} = (X^{(0)}(1), X^{(0)}(2), \ldots, X^{(0)}(j), \ldots, X^{(0)}(n))$, where $X^{(0)}(j)$ is the datum for the j-th time and n is the total number of modeling data. This equation estimates the value of $X^{(0)}(n + i)$, where i is a positive number. The calculation steps are shown below:

Step 1: The general form of $X^{(0)} = (X^{(0)}(1), X^{(0)}(2), \ldots, X^{(0)}(j), \ldots, X^{(0)}(n))$ is represented as:

$$X^{(1)}(k) = \sum_{j=1}^{k} X^{(0)}(j) \tag{3}$$

Step 2: Set the first order ordinary differential equation of $X^{(1)}$ as:

$$\frac{dX^{(1)}(t)}{dt} + aX^{(1)}(t) = b \tag{4}$$

Step 3: Use the least square method to get a and b

$$\begin{bmatrix} a \\ b \end{bmatrix} = (U^T U)^{-1} U^T V_n \tag{5}$$

where

$$U = \begin{bmatrix} -0.5(X^{(1)}(1) + X^{(1)}(2)) & 1 \\ -0.5(X^{(1)}(2) + X^{(1)}(3)) & 1 \\ \dots & \dots \\ -0.5(X^{(1)}(n-1) + X^{(1)}(n)) & 1 \end{bmatrix} \tag{6}$$

$$V_n = (X^{(0)}(2), X^{(0)}(3), \dots, X^{(0)}(n))^T \tag{7}$$

Step 4: By using the initial condition of $X^{(1)}(1)$, a and b are put into the grey differential equation to get the specific solution shown below:

$$\hat{X}^{(1)}(k+1) = (X^{(0)}(1) - \frac{b}{a})e^{-ak} + \frac{b}{a} \tag{8}$$

where, $\hat{X}^{(1)}(k+1)$ is the predicted value of $X^{(1)}(k+1)$. Therefore, the original sequence can be expressed in the following equation:

$$\hat{X}^{(0)}(k+1) = \hat{X}^{(1)}(k+1) - \hat{X}^{(1)}(k) \tag{9}$$

As the rainfall records are all positive numbers ($X^{(0)}(j) \geq 0$, for $j = 1, 2, \dots, n, \dots$), they do not defy the basic principle of the GM (1,1) model. In accordance with this logic, only a small amount of data is needed to construct the GM (1,1), and there are only two parameters to be estimated in Eq. (5). In view of this, the GM (1,1) model is often used to predict short-term flows, and rainfall caused by typhoons.

2. Dynamic Index

Whether or not the original sequence of the grey forecasting model meets the exponential is an important fact that affects the grey model's predictive accuracy. The AGO is certainly one of the methods that can be applied but, fundamentally, it cannot improve the smoothness of the series. Specifically, the bigger restoration error controls the scope of the application. Xiang [20] pointed out that the dynamic index transformation was a better method. It not only enabled the generated series to have a better smoothness but also made no amplification of the restoration error. This study adopts Xiang's dynamic index conversion formula in order to improve the smoothness of the original data and reduce the grey model's prediction error. It is called the DGM (1,1), and is expressed below:

$$D(x(t)) = \frac{x(t)}{\alpha} t^{1/n} \beta^{(-1)^{\gamma} \xi} \tag{10}$$

α, n, β, γ, and ξ are the assigned dynamic index transfer parameters, where $\alpha > 0$, $n > 0$, $\beta > 0$, $\gamma > 0$, and $\xi > 0$. After investigating the values of γ and α in Eq. (10), Xiang [20] found that these two parameters are affected by the trend of the original sequence.

When $x(2)/x(1) > 1$ and $x(3)/x(2) < 1$, this is called the grey left-wobbly sequence, with the following restriction:

$$\gamma = t - 1 \tag{11}$$

$$\alpha = (\beta - 1)^{1/2} \tag{12}$$

when $x(2)/x(1) < 1$ and $x(3)/x(2) > 1$, this is called the grey right-wobbly sequence, with the following restriction:

$$\gamma = t \tag{13}$$

$$\alpha = \beta + 1 \tag{14}$$

And, the values of β, n, and ξ are calculated by Eqs. (15), (16), and (17):

$$\beta = 10\xi \tag{15}$$

$$\frac{2x(m)}{x(1)} \leq n \leq \frac{3x(m-1)}{x(2)} \tag{16}$$

$$\sqrt[m]{\frac{x(m)}{x(1)}} - 0.9 \leq \xi \leq \sqrt[m-2]{\frac{3x(m-1)}{x(2)}} - 0.9 \tag{17}$$

To meet conditions $n > 0$, $\xi > 0$, in Eqs. (16) and (17), $m \geq 4$. The grey prediction model is the only one that needs more than 4 data points for model construction [4]. Thus, if m is set as the minimum value within the given constraint, it may extend the prediction period.

III. INTEGRATED INNOVATIVE GREY MODEL WITH DYNAMIC INDEX

The GM (1,1) improves the prediction accuracy by adjusting the value of α and correcting its residual by using the exponential smoothing and the Fourier series called EFGM (1,1) or this study's innovative grey model. The innovative grey model with the pre-processing of the dynamic index is called the DEFGM (1,1). The flowchart of the study's analysis is presented in Fig. 1.

1. Improve the Background Value of Grey Model

It's known from Eq. (8) that the forecasting accuracy depends on parameters a and b, and their solution depends on the structural form of the background value of GM (1,1), $z^{(1)}(k)$. Therefore, $z^{(1)}(k)$ becomes the key factor that directly influences the accuracy and adaptability of the GM (1,1) model. Having $Z^{(1)}(k) = \alpha X^{(1)}(k) + (1 - \alpha)X^{(1)}(k-1)$ as the smoothing

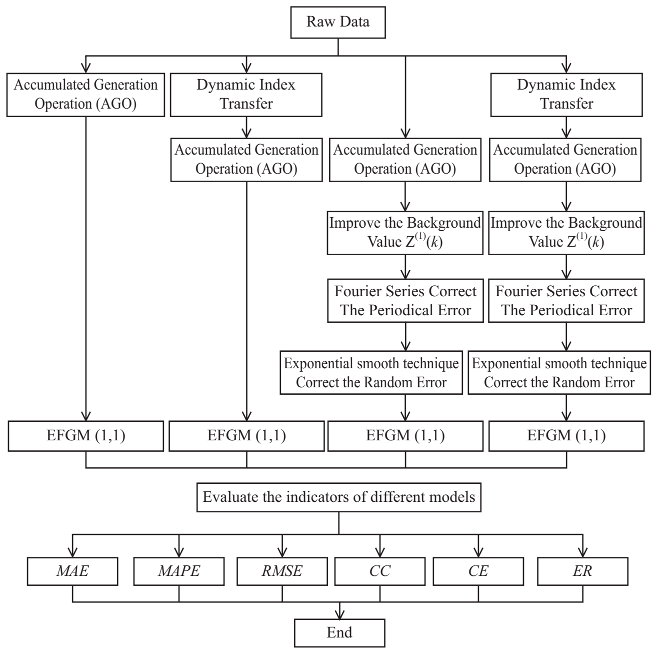

Fig. 1. Flow Diagram of Analysis.

α is the key parameter of the background value. When the data sequence changes greatly, the longer time lag error will cause a larger error.

In order to increase the accuracy of the GM (1,1), literature shows that α is an important parameter in the accuracy of the GM (1,1). For this reason, this study improves the background value by integrating as:

$$z^{(1)}(k) = \alpha x^{(1)}(k) + (1-\alpha)x^{(1)}(k-1) \approx \int_{k-1}^{k} x^{(1)}(t)dt \quad (18)$$

In order to avoid problems resulting from a continuous adjustment of value α, this study deals with the background value of Eq. (18) directly, instead of estimating α to improve the precision of the GM (1,1).

The grey differential equation is a kind of exponential function. In this study, $x^{(1)}(t)$ is represented as

$$x^{(1)}(t) = \lambda e^{\omega t} + \theta \quad (19)$$

It is substituted in Eq. (18), and t is substituted in Eq. (19) by k, k-1, and k-2, respectively. Finally, values ω and θ are substituted to get the improved background value as shown below:

$$\begin{aligned} z^{(1)}(k) &= \int_{k-1}^{k} (\lambda e^{\omega t} + \theta) = \frac{x^{(0)}(k)}{\omega} + \theta \\ &= \frac{x^{(0)}(k)}{\ln \dfrac{x^{(0)}(k)}{x^{(0)}(k-1)}} + x^{(1)}(k) - \frac{(x^{(0)}(k))^2}{x^{(0)}(k) - x^{(0)}(k-1)} \end{aligned} \quad (20)$$

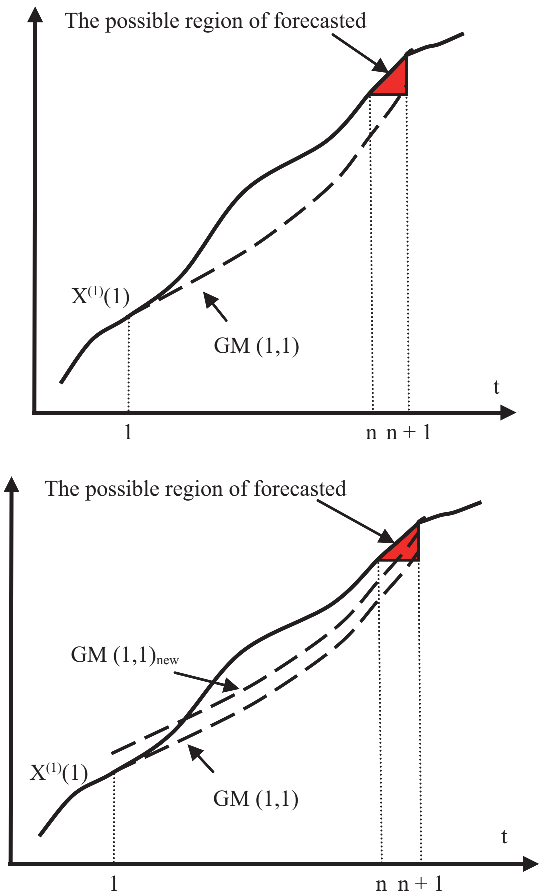

Fig. 2. Conceptaual Diagram of GM (1,1) and GM $(1,1)_{new}$.

The new grey prediction model GM $(1,1)_{new}$ is based on a minimum deviation of the model. Data can be obtained by using the above calculations. Detailed illustrations of the operations are shown in Fig. 2.

2. Error Correction Model

System characteristics include periodicity, randomness, and tendency. In order to get the tendency of the series and the context of development of the system effectively, this study, apart from improving the background value by integrating, has improved the accuracy by correcting the model's periodical errors through the Fourier series and correcting the residual, random errors through an exponential smoothing method. When using the GM $(1,1)_{new}$ to calculate the one-step predicted value $\hat{X}^{(0)}(n+1)$ of $X^{(0)}$, the first residual error series $\delta'^{(0)}$ is:

$$\delta'^{(0)} = (\delta'^{(0)}(1), \delta'^{(0)}(2), ..., \delta'^{(0)}(n))$$

where $$\delta'^{(0)}(k) = X^{(0)}(k) - \hat{X}^{(0)}(k) \quad (21)$$

1) Periodic Correction Model

This study adopts the Fourier series to extract the periodical feature hidden in the first series $\delta'^{(0)}$ of the residual sequences. Utilizing the Fourier series, this residual sequence is shown below:

$$\hat{\delta}^{(0)}(k) = \frac{1}{2}a_0 + \sum_{i=1}^{k_a}(a_i \cos(\frac{i2\pi}{T}k) + b_i \sin(\frac{i2\pi}{T}k))$$

for $\quad k = 2, 3, \ldots, n \quad (22)$

where $T = (n - 1)$ is the interval length of the finite residual error series. Additionally, the integral part of $k_a = [((n-1)/2) - 1]$ is the lower limit of the reasonable expanded degrees for the finite series. Parameter a_0 is estimated by the least square method, and a_0 and b_i (for i = 1, 2, …, n) are expanded as follows:

$$C = (\Gamma^T \Gamma)^{-1} \Gamma^T \delta^{(0)} \tag{23}$$

where $C = [a_0, a_1, b_1, a_2, b_2, \ldots, a_{k_a}, b_{k_a}]^T$ and

$$\Gamma = \begin{bmatrix} 1/2 & \cos(\frac{2\pi\times 1}{T}2) & \sin(\frac{2\pi\times 1}{T}2) & \ldots\ldots\ldots & \cos(\frac{2\pi\times k_a}{T}2) & \sin(\frac{2\pi\times k_a}{T}2) \\ 1/2 & \cos(\frac{2\pi\times 1}{T}3) & \sin(\frac{2\pi\times 1}{T}3) & \ldots\ldots\ldots & \cos(\frac{2\pi\times k_a}{T}3) & \sin(\frac{2\pi\times k_a}{T}3) \\ \ldots & \ldots & \ldots & \ldots\ldots\ldots & \ldots & \ldots \\ 1/2 & \cos(\frac{2\pi\times 1}{T}n) & \sin(\frac{2\pi\times 1}{T}n) & \ldots\ldots\ldots & \cos(\frac{2\pi\times k_a}{T}n) & \sin(\frac{2\pi\times k_a}{T}n) \end{bmatrix} \tag{24}$$

By substituting the Eq. (24) result for Eq. (23), it can be observed that simulated sequence $\hat{\delta}'^{(0)}$ of $\delta^{(0)}$ is

$$\hat{\delta}'^{(0)} = (\hat{\delta}'^{(0)}(2), \hat{\delta}'^{(0)}(3), \ldots, \hat{\delta}'^{(0)}(n)) \tag{25}$$

The one-step estimated periodicity value $\hat{\delta}^{(0)}(n+1)$ in the original data, which can be obtained by substituting $k = (n + 1)$ in Eq. (22).

2) Random Correction Model

The exponential smoothing technique only needs a few data points, and the correction randomness residual can be handled by it. However, the 2nd residual series needs to be shifted to a non-negative series and the shifted back after the prediction process is completed [13]. An exponential smoothing technique is a method that predicts and analyzes a time sequence trend. It is generally used in the prediction of short and mid-term trends of economic behaviours.

With regard to system randomness, this study adopts the exponential smoothing for the 2nd residual correction. The 2nd residual sequence $\delta''^{(0)}$ is represented as:

$$\delta''^{(0)} = (\delta''^{(0)}(2), \delta''^{(0)}(3), \ldots, \delta''^{(0)}(n))^T$$

where $\quad \delta''^{(0)}(k) = \delta'^{(0)}(k) - \hat{\delta}'^{(0)}(k) \quad (26)$

By applying the exponential smoothing technique to extract the random characteristics hidden in the 2nd residual sequence $\delta''^{(0)}$, the following formula is the best fit for the 2nd residual sequence $\delta''^{(0)}$:

$$\hat{\delta}''^{(0)}(k) = \varphi\delta''^{(0)}(k-1) + (1-\varphi)\hat{\delta}''^{(0)}(k-1),\ k = 2, 3, \ldots, n \tag{27}$$

where φ is the smoothing coefficient, and 0 < φ < 1. Eq. (27) means that the predicted 2nd residual value $\hat{\delta}''^{(0)}(k)$ for point k is composed of the actual 2nd residual value $\hat{\delta}''^{(0)}(k-1)$, and the predicted 2nd residual value $\hat{\delta}''^{(0)}(k-1)$ for point (k-1). Wherein $\delta''^{(0)}(1)$ has no predicted value, while the predicted value of $\delta''^{(0)}(2)$ is $\delta''^{(0)}(1)$, namely $\hat{\delta}''^{(0)}(2) = \hat{\delta}''^{(0)}(1)$.

For an exponential smoothing technique, smoothing coefficient φ has an impact on the simulation and prediction accuracy of the 2nd residual series. Generally speaking, smaller smoothing coefficient values can be adopted in case of a stable 2nd residual sequence while larger smoothing coefficient values can be applied in a heavily fluctuated 2nd residual sequence. Similarly, an optimized smoothing coefficient is difficult to obtain in case of stronger subjective judgement factors. Therefore, this study chooses a smoothing coefficient value that has a minimum deviation with the 2nd residual sequence through an optimum technique, i.e. solving the following formula's minimum value.

$$SSE = \sum_{k=2}^{n}\delta'''^{(0)}(k)^2 = \sum_{k=2}^{n}(\delta''^{(0)}(k) - \hat{\delta}''^{(0)}(k))^2 \tag{28}$$

where $\hat{\delta}''^{(0)}(k)$ can be further derived as:

$$\hat{\delta}''^{(0)}(k) = \sum_{\tau=1}^{k}(\phi(1-\phi)^{\tau-1}\delta''^{(0)}(k-\tau) - (1-\phi)^{k-1}\delta''^{(0)}(1)) \tag{29}$$

Thus, the optimization equation is:

$$\begin{cases} \min \quad SSE = \sum_{k=2}^{n}[\delta''^{(0)}(k) - \sum_{\tau=1}^{k}(\phi(1-\phi)^{\tau-1}\delta''^{(0)}(k-\tau) - (1-\phi)^{k-1}\delta''^{(0)}(1))]^2 \\ s.t. \quad 0 < \varphi < 1 \end{cases} \tag{30}$$

After getting the optimum ϕ value, the one-step predicted value of the 2nd residual series $\delta''^{(0)}$ is derived as follows:

$$\hat{\delta}''^{(0)}(n+1) = \phi\delta''^{(0)}(n) + (1-\phi)\hat{\delta}''^{(0)}(n) \tag{31}$$

Finally, through a combination of previously corrected values, the one-step predicted value $\overline{X}^{(0)}(n+1)$ of original sequence $X^{(0)}$, can be derived as follows:

$$\overline{X}^{(0)}(n+1) = \hat{X}^{(0)}(n+1) + \hat{\delta}'^{(0)}(n+1) + \hat{\delta}''^{(0)}(n+1) \tag{32}$$

3. Innovative Grey Model through Integration of the Dynamic Index

A new series is obtained through transforming original sequence $x^{(1)}(k) = \sum_{j=1}^{k} x^{(0)}(j)$ and by using dynamic index formula $D(x(t)) = \frac{x(t)}{\alpha} t^{1/n} \beta^{(-1)^\gamma \xi}$:

$$D^{(1)}(k) = \sum_{j=1}^{k} D^{(0)}(j) \tag{33}$$

Then, the following equation is obtained through the EFGM (1,1):

$$\bar{D}^{(0)}(n+1) = \hat{D}^{(0)}(n+1) + \hat{\varepsilon}'^{(0)}(n+1) + \hat{\varepsilon}''^{(0)}(n+1) \tag{34}$$

Eq. (34) is substituted by the reductive formula of Eq. (10) to get the predicted value of the original sequence shown below:

$$\bar{x}^{(0)}(t) = \frac{\alpha \bar{D}^{(0)}(t)}{t^{1/n} \beta^{(-1)^\gamma \xi}} \tag{35}$$

Using the dynamic index to smooth the original sequence allows the series to have the monotonic series with the grey exponential law. By using the EFGM (1,1), Eq. (35) is adopted to get the prediction series, which has better prediction accuracy. This model is known as the DEFGM (1,1).

IV. CASE STUDY

1. Background

There are over 7 million people in northern Taiwan's Taipei Basin. It has a flat centre surrounded by high mountains and hills, along with low-lying areas below sea level. The Danshui River is the main river system, crossing the area to form a kettle shape. This kettle-shaped basin and the winding Danshui River, causes damage to lives and properties during the flood season, with the low-lying areas bearing the brunt since the floodwaters cannot be discharged successfully. It is easily affected by tidal backwaters if a flood and huge tide coincide. Part of the population in this basin is concentrated on the western side, which is located in the Dahan River Watershed, upstream of the Danshui River. The elevation of this watershed is between 0.5 m and 15 m. Due to heavy water flows along the dike, river blockages caused by mud-flows and rocks, or sediment accumulation in the river, and poor drainage systems, a large body of water can accumulates upstream, especially during heavy rainstorms. Furthermore, abnormal weather conditions, due to global climate change, have led to an increase in the amount of annual rainfall in Taiwan. While the number of rainy days has decreased, rainfall is more concentrated, and there are inadequate preventive measures to withstand dangerous flooding. Obviously, the accuracy of predicting rainfall and flood peaks is extremely important for flood prevention. Therefore, this study used rainfall records from the Dabao rain-gauge station in the Dahan River Watershed on the Danshui River as raw data. The layout of the Danshui River showing the location of the Dabao rain-gauge station is shown in Fig. 3.

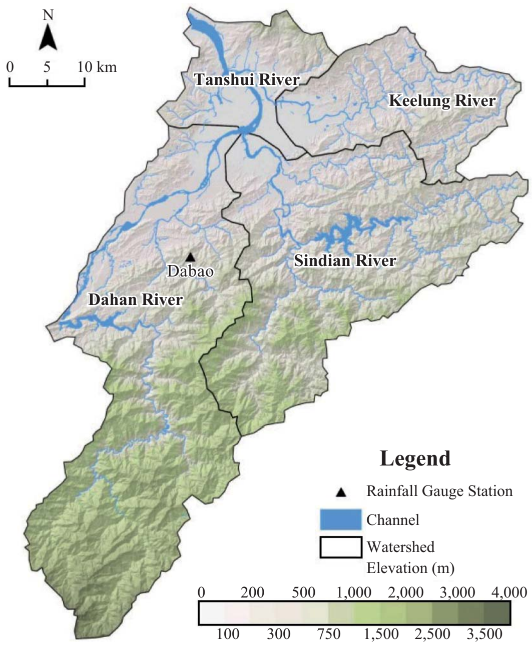

Fig. 3. Layout of Tanshui River and Dabao Rainfall Guage Station.

2. Case Analysis

This study first constructed a model based on the data obtained from the original sequence, which was then converted by the dynamic index technique, and the converted sequence was calculated by using the GM (1,1). The constructed model was called the DGM (1,1). To generate the model, there is a need to set parameters: α, n, β, γ, and ξ. Using records from the Dabao rain gauge, the amount of rain from Typhoon KROSA was taken as an example. As shown in Fig. 4, the first three data points display a "$\vee$" shape, and this sequence is the grey right-wobbly sequence of $x(2)/x(1) < 1$, $x(3)/x(2) > 1$. Therefore, $\gamma = t$, $\alpha = \beta + 1$, and $\beta = 10\xi$.

This sequence meets the restrictions of $\alpha > 0$, $n > 0$, $\beta > 0$, $\gamma > 0$, $\xi > 0$, and $m \geq 4$. Since the foundation of the GM (1,1) is the exponential function, the predicting of centre-symmetry curves or random time series is not ideally done by the GM (1,1), and causes an undesirable effect in predicting the sequences with wavy patterns [12]. Similarly, this dynamic index is a function of time $t^{1/n}$ and $\beta^{(-1)^\gamma \xi}$. When $m = 4$, $1.89 \leq n \leq 10.5$, $0.09 \leq \xi \leq 2.34$ can be obtained for later parameter calibration. For $n = 10$, $\xi = 1.2$, and $n = 2$, $\xi = 0.2$ in Table 1, it can be observed that the waving equation is larger than the sequence of the actual value. By calibration from conditions

Table 1. Parameters of Dynamic Index of Typhoon KROSA.

m	n	α	β	ξ	$\sqrt[m]{\frac{x(m)}{x(1)}}-0.9\le\xi\le$	$\sqrt[m-2]{\frac{3x(m-1)}{x(2)}}-0.9$	$\frac{2x(m)}{x(1)}\le n\le$	$\frac{3x(m-1)}{x(2)}$
4	10	3	2	0.2	0.09	2.34	1.89	10.5

Table 2. Parameters of Dynamic Index of Typhoon SINLAKU.

m	n	α	β	ξ	$\sqrt[m]{\frac{x(m)}{x(1)}}-0.9\le\xi\le$	$\sqrt[m-2]{\frac{3x(m-1)}{x(2)}}-0.9$	$\frac{2x(m)}{x(1)}\le n\le$	$\frac{3x(m-1)}{x(2)}$
4	-	-	-	-	-0.9	-0.21	0	0.47
5	-	-	-	-	-0.9	-0.9	0	0
6	-	-	-	-	-0.9	-0.9	0	0
7	-	-	-	-	0.1	-0.9	2	0
8	-	-	-	-	0.32	-0.16	10	0.16
9	-	-	-	-	0.36	0.07	16	0.79
10	1.2	0.71	1.5	0.15	-0.9	0.13	0	1.26

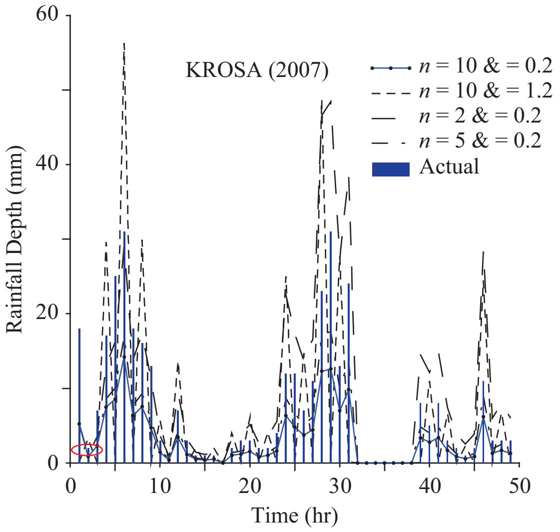

Fig. 4. Oringial Sequence and the Dynamic Index Converted Sequence of different parameters of Typhoon KROSA (2007).

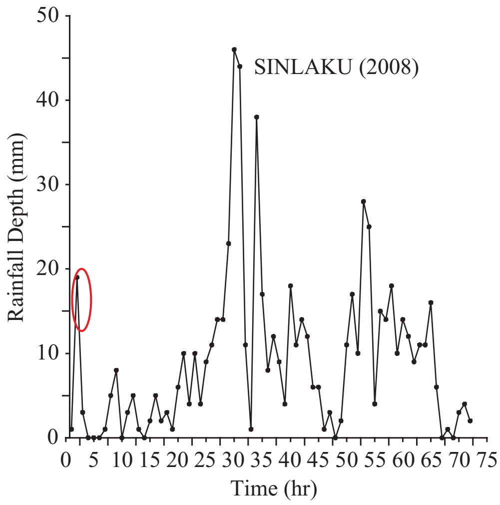

Fig. 5. Rainfall Trend of Typhoon SINLAKU (2008).

$n = 5$, $\xi = 0.2$, it can be found that the waving equation has decreased significantly. In order to lower the influence of $t^{1/n}$ which may reduce the wavy equation of this sequence, n needs to be set to a larger value. To reduce interference by ξ, and ensure the smooth and stable development of this sequence ξ needs to be set to a smaller value. Thus, this study adopts the $n = 10$, $\xi = 0.2$ to get $\beta = 2$, $\alpha = 3$, $\gamma = t$, as shown in Table 1. Finally, the dynamic index conversion function can be obtained as $D(x(t)) = \frac{x(t)}{3}t^{1/10}2^{(-1)^t 0.2}$.

Taking the amount of rainfall from typhoon SINLAKU as another example, the first three data points shown in Fig. 5 display a "∧" shape, which is the grey left-wobbly sequence of $x(2)/x(1) > 1$ and $x(3)/x(2) < 1$. The conditions are $\gamma = t - 1$, $\alpha = (\beta - 1)^{1/2}$, $\beta = 10\xi$, $\alpha > 0$, $n > 0$, $\beta > 0$, $\gamma > 0$, $\xi > 0$, and $m \ge 4$.

When m = 4~8, it does not meet the condition of $\xi > 0$. When $m = 9$, it gets the contradictory conditions of $n \ge 16$ and $n \ge 0.79$. Therefore $0 \le n \le 1.26$, $-0.9 \le \xi \le 0.13$ was selected when $m = 10$ for parameter calibration. This study adopts $n = 1.2$, $\xi = 0.15$ to get $\beta = 1.5$, $\alpha = 0.71$, $\gamma = t - 1$, shown in Table 2. Finally, a dynamic index conversion function of $D(x(t)) = \frac{x(t)}{0.71}t^{1/1.2}1.5^{(-1)^{t-1}0.15}$ can be obtained.

By comparing the GM (1,1) with the DGM (1,1), it shows that the GM (1,1) prediction is affected by the previous data. It is somewhat concentrated, and unable to display the extreme value of the flood peak, plus there is a prediction delay. After converting the dynamic index into an exponential function, calculation of the prediction value approximates the actual value, and the fitting level becomes relatively high. However, it is still unable to improve the prediction value for the extreme value, as shown in Figs. 6(a)-6(h).

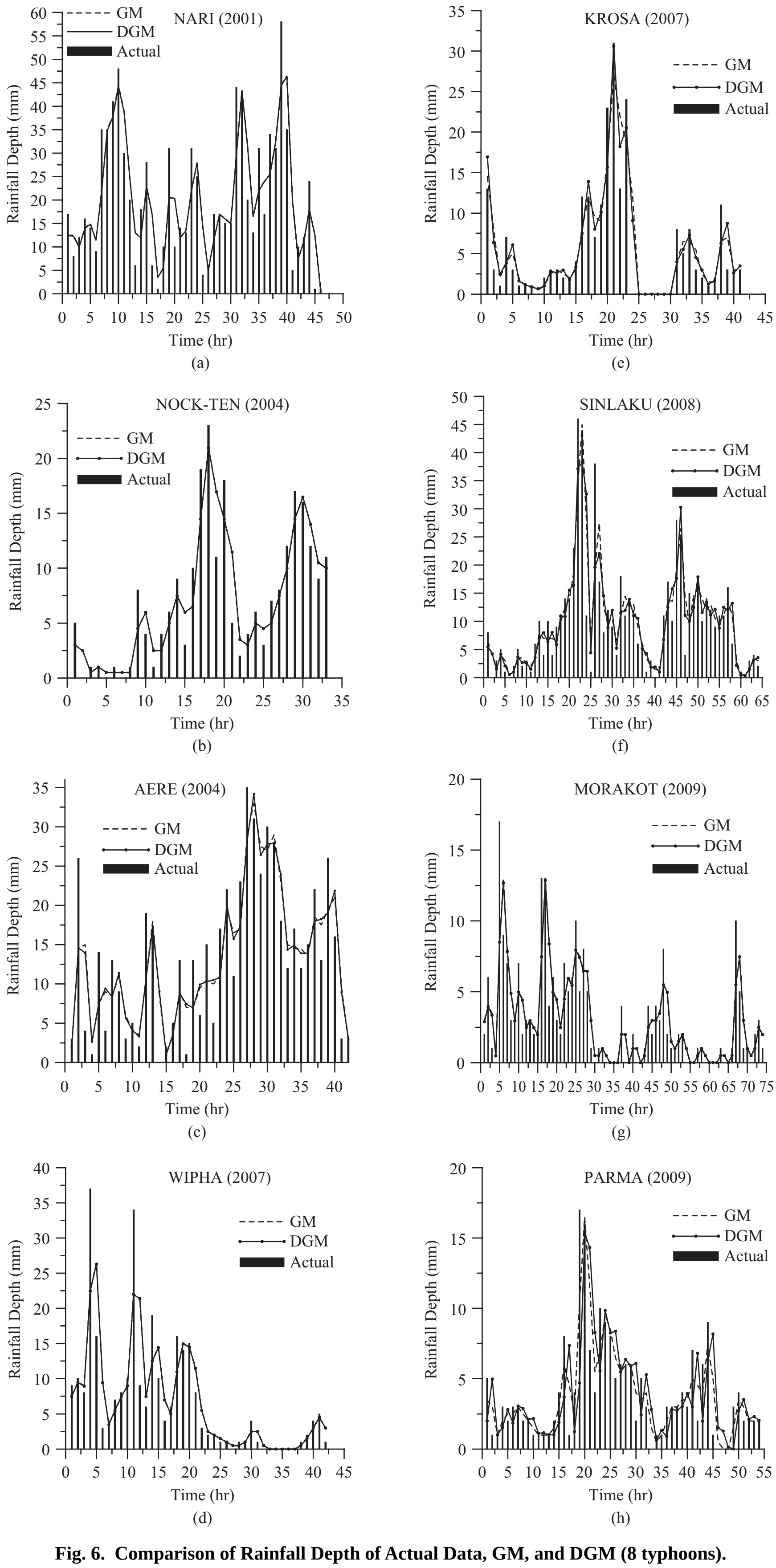

Fig. 6. Comparison of Rainfall Depth of Actual Data, GM, and DGM (8 typhoons).

Since the prediction value of the GM (1,1) is somewhat concentrated, the extreme value cannot be displayed, or there is a time delay for the prediction. For this reason, the EFGM (1,1) improves the background value of the grey model by integration, while combining it with the Fourier series and exponential smoothing for error correction. This would finally achieve better accuracy, and reflect the actual situation [16]. This study converts the exponential function sequence by a dynamic index to generate prediction model DEFGM (1,1), which can obtain a higher prediction accuracy than the EFGM (1,1), which is shown in Figs. 7 and 8.

As shown in Figs. 7(a)-7(h), based on the rainfalls produced by the eight typhoons, it can be seen that, the EFGM (1,1) and DEFGM (1,1) have significantly improved the time delay problems with the typhoons' extreme values, when compared to the GM (1,1). The Peak rainfall simulated by the DEFGM (1,1) is closer to the actual value.

As shown in Fig. 8(a)-8(h), based on the accumulative rainfall prediction for the eight typhoons, it can be seen that the DEFGM (1,1) is more accurate than either the EFGM (1,1) or the GM (1,1) when dealing with the extreme value's turning point. This proves that after applying the dynamic index conversion, DEFGM (1,1) has relatively improved the accuracy of prediction of the extreme value for the event. It can be observed that the DEFGM (1,1) and EFGM (1,1) perform better than the GM (1,1). This proves that periodical and random residual correction can significantly solve the problem of the delayed peak value commonly seen in the GM (1,1).

3. Statistics for Indicators of Different Models

Based on eight other typhoons, this study analyzes the following four models: GM (1,1), DGM (1,1), EFGM (1,1), and DEFGM (1,1). Six indicators have been used to evaluate the performance of each model. Among these, the EFGM (1,1) and DEFGM (1,1) can get better prediction performance. The good performance by the DEFGM (1,1) is then verified by the mean value of indicators from the eight typhoons.

The mean absolute error (*MAE*) is a kind of valid error measurement used in sequences with the same unit. The mean absolute percentage (*MAPE*) is a good evaluation criterion often used to compare prediction performances. The root mean square error (*RMSE*) can present the discrete degree between the actual and predicted values. The closer the correlation coefficient (*CC*) is to 1, the better the prediction, with *MCC* as the mean correlation coefficient. The closer the coefficient of efficiency (*CE*) is to 1, the more the prediction matches the actual situation. *MCE* is the mean coefficient of efficiency. Then, the error of peak rainfall (ER_P) evaluates the errors between the predicted maximum rainfall and the actual maximum rainfall. MER_P is the mean error of peak rainfalls, and the indicators are stated as follows:

$$MAE = \frac{1}{n}\sum_{k=1}^{n}\left|X(k)-\hat{X}(k)\right| \tag{36}$$

$$MAPE = \left(\frac{1}{n}\sum_{k=1}^{n}\left|\frac{(\hat{X}(k)-X(k))}{X(k)}\right|\right)\times 100\% \tag{37}$$

$$RMSE = \sqrt{\frac{\sum_{k=1}^{n}(X(k)-\hat{X}(k))^2}{n}} \tag{38}$$

$$CC = \frac{\sum_{t=1}^{n}(X(k)-\bar{X}(k))(\hat{X}(k)-\bar{\hat{X}}(k))}{\sqrt{\sum_{t=1}^{n}(X(k)-\bar{X}(k))^2\sum_{t=1}^{n}(\hat{X}(k)-\bar{\hat{X}}(k))^2}} \tag{39}$$

$$CE = 1-\frac{\sum_{t=1}^{n}(X(k)-\hat{X}(k))^2}{\sum_{t=1}^{n}(X(k)-\bar{X}(k))^2} \tag{40}$$

$$ER_P = \left|\frac{(\hat{X}(k)_P - X(k)_P}{X(k)_P}\right| \tag{41}$$

where $X(k)$ is the actual value for time point k, then $\hat{X}(k)$ is the predicted value for it. The equation $\bar{\hat{X}}(k)$ is the mean value for n predicted values, $X(k)_P$ is the actual maximum rainfall, while $\hat{X}(k)_P$ is the predicted maximum rainfall, and n is the number of all predicted points. This can be shown by $MCE = \frac{1}{n}\sum_{k=1}^{N} CE_k$ where CE_k is event k's CE value, and $MER_P = \frac{1}{N}\sum_{k=1}^{N} ER_{Pk}$ where ER_{Pk} is k's ER_P value.

It can be observed in Table 3 that all the *MAE* values of the DEFGM (1,1) for all eight typhoons are smaller than 1, which indicates the best performing model. The FEGM (1,1) comes in second, which indicates that the grey model can correct periodical and random errors through the Fourier series and exponential smoothing technology. The relative accuracy of the DEFGM (1,1) is improved after the original sequence is smoothed by a dynamic index conversion. Where Typhoon NOCK-TEN had the least number of rainfall data points, it also revealed that the GM (1,1) and DGM (1,1) had better prediction accuracy for shorter periods of time. For events with larger data points, such as Typhoons SINLAKU, MORAKOT, and PARMA, the indicators for their EFGM (1,1) and DEFGM (1,1) show better results than those of both the GM (1,1) and DGM (1,1). Thus, it can be said that periodical and random residual correction of sequences with large data points can have more accurate predictions. The DEFGM (1,1) has the smallest values for *MAPE* and *RMSE*, followed by the FEGM (1,1), which shows that the accuracy can be

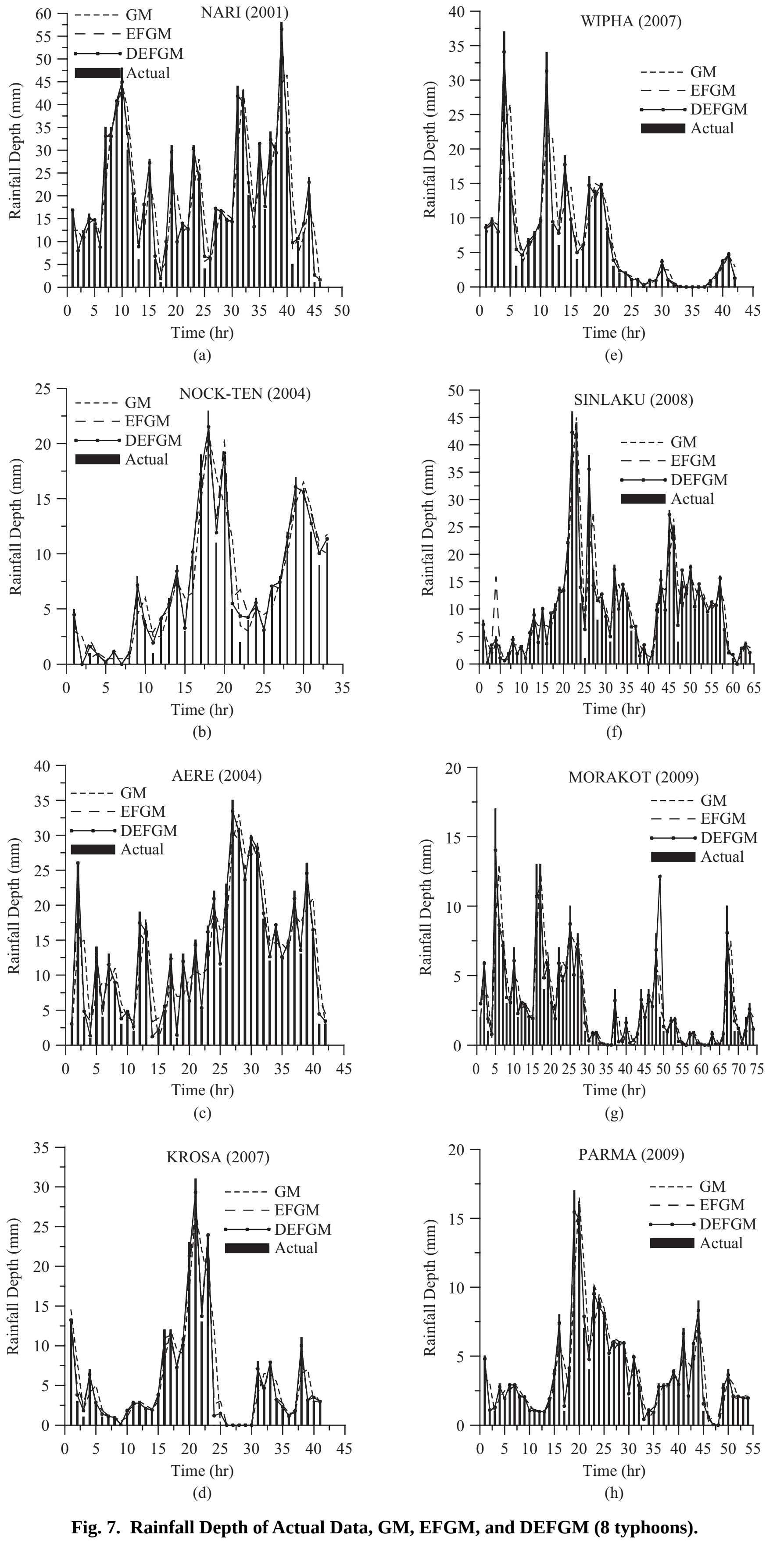

Fig. 7. Rainfall Depth of Actual Data, GM, EFGM, and DEFGM (8 typhoons).

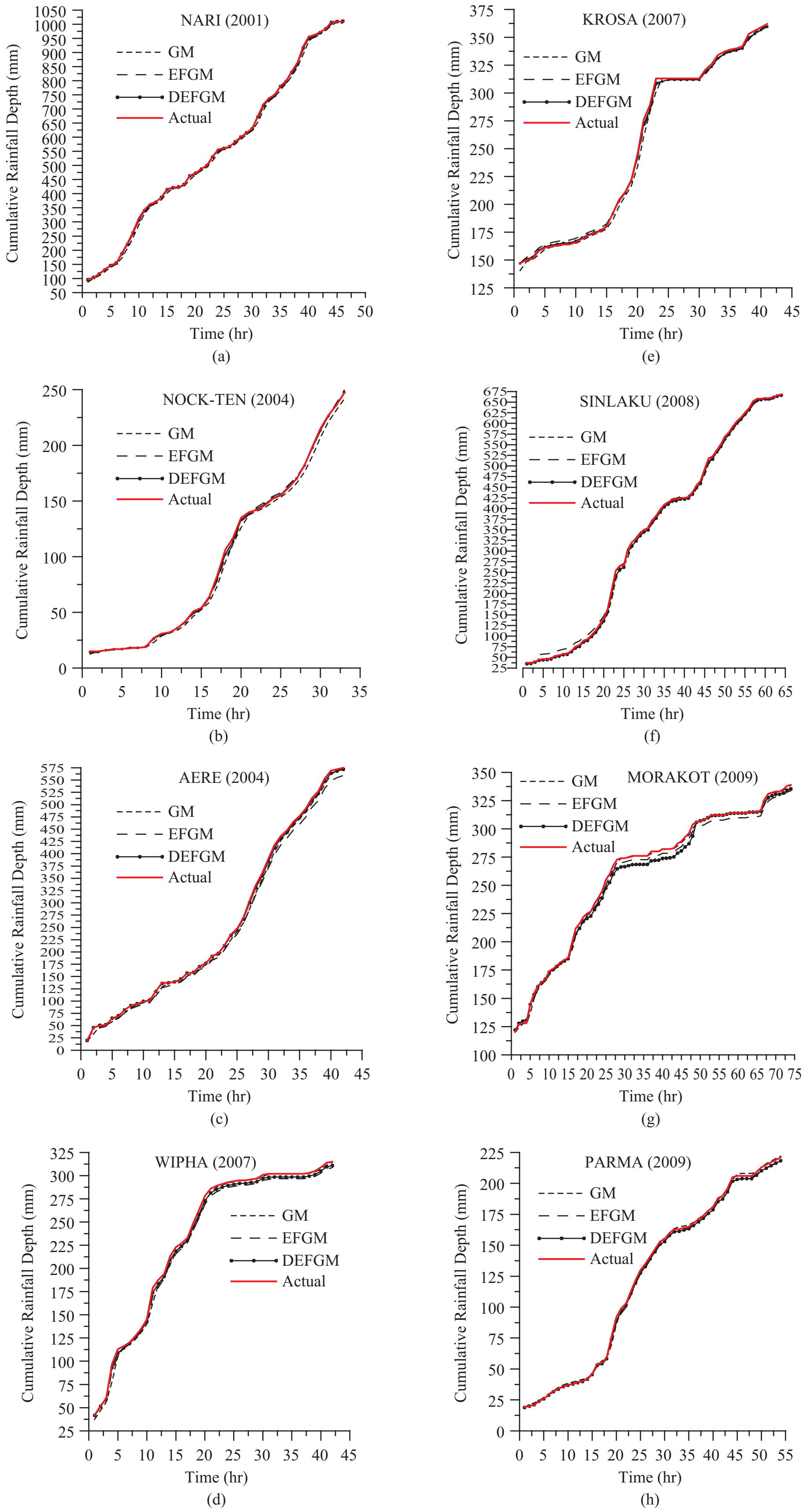

Fig. 8. Cumulative Rainfall Depth of Actual Data, GM, EFGM, and DEFGM (8 typhoons).

Table 3. Statistics of GM, DGM, EFGM, and DEFGM Models.

	GM (1,1)						DGM (1,1)					
	MAE	*MAPE*	*RMSE*	*CC*	*CE*	ER_P	*MAE*	*MAPE*	*RMSE*	*CC*	*CE*	ER_P
NARI	2.84	30.21	4.55	0.86	0.86	0.20	2.65	27.24	4.37	0.86	0.88	0.20
AERE	3.19	58.18	4.46	0.87	0.80	0.06	2.26	39.16	3.73	0.88	0.82	0.02
NOCK-TEN	1.40	33.39	2.08	0.93	0.87	0.09	1.02	21.97	1.78	0.93	0.87	0.09
WIPHA	2.5	31.36	4.43	0.83	0.70	0.28	1.70	19.90	3.68	0.85	0.73	0.29
KROSA	2.64	42.35	4.92	0.82	0.68	0.10	1.41	17.84	2.79	0.92	0.87	0.06
SINLAKU	2.84	47.11	5.08	0.87	0.80	0.02	2.21	29.31	4.61	0.86	0.78	0.17
MORAKOT	2.46	44.99	4.17	0.89	0.79	0.18	1.70	29.78	3.43	0.90	0.81	0.22
PARMA	0.87	29.89	1.57	0.88	0.83	0.03	1.22	41.44	2.52	0.58	0.43	0.10

	EFGM (1,1)						DEFGM (1,1)					
	MAE	*MAPE*	*RMSE*	*CC*	*CE*	ER_P	*MAE*	*MAPE*	*RMSE*	*CC*	*CE*	ER_P
NARI	0.40	3.94	0.76	1.00	1.00	0.05	0.45	5.04	0.96	1.00	0.99	0.03
AERE	1.00	16.55	1.60	0.99	0.97	0.13	0.45	6.89	0.68	1.00	1.00	0.04
NOCK-TEN	0.70	19.43	1.25	0.97	0.95	0.13	0.35	9.62	0.62	0.99	0.99	0.06
WIPHA	0.88	13.12	1.70	0.99	0.96	0.16	0.44	6.40	0.85	1.00	0.99	0.08
KROSA	0.60	11.11	1.14	0.99	0.98	0.00	0.31	5.85	0.55	1.00	1.00	0.00
SINLAKU	1.31	26.86	2.66	0.98	0.94	0.12	0.74	19.26	1.41	0.99	0.98	0.08
MORAKOT	0.46	7.68	0.82	0.99	0.99	0.00	0.39	7.51	0.74	0.98	0.99	0.00
PARMA	0.25	6.87	0.50	0.99	0.98	0.10	0.19	4.34	0.37	1.00	0.99	0.09

Table 4. Statistics of Average Value of Indicators of Different Model.

	GM (1,1)	DGM (1,1)	EFGM (1,1)	DEFGM (1,1)
MCC	0.868	0.848	0.987	**0.994**
MCE	0.791	0.772	0.973	**0.991**
MER_P	0.119	0.144	0.086	**0.048**
MMAE	2.340	1.770	0.699	**0.414**
MMAPE	39.685	28.330	13.196	**8.112**
MRMSE	3.909	3.365	1.304	**0.773**

Note: Bold indicates the best accuracy.

greatly improved after periodical and random residual correction of the sequence. Moreover, the values for *MAPE* and *RMSE* obtained by the DEFGM (1,1) and EFGM (1,1) are smaller than that of the GM (1,1), which means that modeling through sequence smoothing by the dynamic index conversion can help increase prediction accuracy. Furthermore, the *CC* values for the DEFGM (1,1) and FEGM (1,1) are closer to 1, indicating that these two models have the better performance. Their *CE*'s are close to 1, which indicates a better overall performance than for the GM (1,1). Finally, ER_P is considered the most important indicator for an evacuation announcement. The ER_R value of the DEFGM (1,1) is closest to 0 and that means that there is no time lag for predicting the time of the peak rainfall.

In order to prevent the prediction accuracy judgment from being affected by features of individual typhoons, this study evaluated those models with the mean values of *CC*, *CE*, ER_P, *MAE*, *MAPE*, and *RMSE* for eight typhoons. As shown in Table 4, the *MCC* and *MCE* of the DEFGM (1,1) is 0.994 and 0.991, respectively, and are found to be closest to 1, indicating its superiority. The FEGM (1,1) comes in second, whose *MCC* and *MCE* are 0.987 and 0.973, respectively. It means that the overall performance of the DEFGM (1,1) is better than that of the FEGM (1,1). When viewing the MER_P value, the error of peak rainfall for the DEFGM (1,1) is smaller than 0.05. Moreover, the *MMAE*, *MMAPE*, and *MRMSE* of the EFGM (1,1) and DEFGM (1,1) have the least deviation between the actual and predicted values for The DEFGM (1,1) with a *MMAPE* smaller than 10. Obviously, modeling through a dynamic index conversion sequence and correcting the periodical and random residual can greatly increase prediction accuracy.

V. CONCLUSION

It has been observed that the GM (1,1) cannot obtain the desired results for an unstable system with wavy changes. This study proposed a new Grey Model: the DEFGM (1,1) and its process is summarized as follows. First, the original rainfall system is transformed into a stable one by using the dynamic index transformation, which then improves the background value of the grey model. After that, the residual random and periodical errors are corrected through the exponential smoothing technique and Fourier series. By examining different indicators, the DEFGM (1,1) proves its superiority to other models, including the GM (1,1), DGM (1,1), and EFGM (1,1) for the successful predicting of rainfall brought about by typhoons. Conclusions are drawn below:

1. The prediction of the GM (1,1) has significant trends, periodicity and randomness. The predicted sequence is affected by neighboring data points, which then results in the time lag for predicted extreme values when compared to real data. This study improves the background value of the grey model by integration. It effectively catches the sequence's trend and solves the extreme value delays of the original grey model.
2. The predicted rainfall of the GM (1,1) is unable to predict the peak rainfall accurately. However, this problem can be solved when the GM (1,1) is combined with the exponential smoothing technique and the Fourier series, called the EFGM (1,1). The *MCC* and *MCE* of the DEFGM (1,1) are 0.994 and 0.991, respectively which are much closer to 1 than those of the EFGM (1,1) (0.987 and 0.973). Thus, the overall prediction performance for the DEFGM (1,1) is relatively good when compared with others. Moreover, the MER_P DEFGM (1,1) is closest to 0, which means that it can catch the timing of peak rainfalls. It is apparent that a DEFGM (1,1) can really increase prediction accuracy.
3. In order to comply with the exponential law of the grey system, the original sequence is converted into an exponential function sequence by the dynamic index technique, which can reduce the wavy pattern and change it to a sequence with a stable pattern. The dynamic index technique adopted in this study is a function of $t^{1/n}$ and $\beta^{(-1)^{\gamma}\xi}$. To decrease the effect of $t^{1/n}$, the transformation parameter n is set to a larger value. Through calibration of parameters, the wavy pattern of the original sequence is obviously decreased. Furthermore, in order to reduce the interference by ξ and to ensure sequence stability, ξ should be set to a small value. Therefore, future studies can focus on decreasing the effects of time, so as to increase the stability of the converted system.
4. The dynamic index adopted in this study is one of the methods used for smoothing the original sequence. The transformation parameters α, n, β, γ, and ξ can be investigated further and the focus of future studies.

REFERENCES

1. Bedient, P. B., Holder, A., Benavides, J. A., and Vieux, B. E., "Radar-based flood warning system applied to tropical storm Allison," *Journal of Hydrologic Engineering*, Vol. 8, No. 6, pp. 308-318 (2003).
2. Burlando, P., Rosso, R., Cadavid, L. G., and Salas, J. D., "Forecasting of short-term rainfall using ARMA models," *Journal of Hydrology*, Vol. 144, pp. 193-211 (1993).
3. Cheng, B. Q., Han, F. Q., and Gui, Z. H., "Application of wavelet transform-based grey model to fault forecasting of hydroelectric generating sets," *Power System Technology*, Vol. 29, pp. 40-44 (2005).
4. Deng, J. L., "Grey differential equation," *The Journal of Grey System*, Vol. 1, pp. 1-14 (1993).
5. Deng, J. L., "Introduction to grey system," *The Journal of Grey System*, Vol. 1, pp. 1-24 (1989).
6. Deng, J. L., "The control problems of grey systems," *Systems and Control Letters*, Vol. 5, pp. 288-294 (1982).
7. Guo, L. C., Wu, W., Guo, J. Q., Wang, P., and Zhou, B. S., "Applying grey swing model to predict the incidence trend of hemorrhagic fever with renal syndrome in Shenyang," *Journal of China Medical University*, Vol. 37, No. 6, pp. 839-842 (2008).
8. Kang, B. S., Kim, S. Y., and Ko, I. H., "Grid-based estimation of reservoir flood inflow using radar-observed precipitation," *Proceedings of Korea Water Resources Association*, Korea Water Resources Association, Seoul, South Korea, pp. 183-188 (2005).
9. Kang, M. G., Park, S. W., and Cai, X. M., "Integration of hydrologic grey model with global search method for real-time flood forecasting," *Journal of Hydrologic Engineering*, Vol. 14, No. 10, pp. 1126-1145 (2009).
10. Kayacan, E., Ulutas, B., and Kaynak, O., "Grey model-based models in time series prediction," *Expert Systems with Applications*, Vol. 37, pp. 1784-1789 (2010).
11. Labadie, J. W., Lazaro, R. C., and Morrow, D. M., "Worth of short-term rainfall forecasting for combined sewer overflow control," *Water Resources Research*, Vol. 17, No. 5, pp. 1489-1497 (1981).
12. Li, B., "Three-data modeling of grey model," *The Journal of Grey System*, Vol. 2, pp. 11-20 (1990).
13. Lin, C. B., Su, S. F., and Hsu, Y. T., "High-precision forecast using grey models," *International Journal of Systems Science*, Vol. 32, pp. 609-619 (2001).
14. Lin, Y. F., "A novel method of residual recognition to increase forecast precision of grey model," *National Conference of Grey Model and Applications*, pp. 65-71 (2001).
15. Lin, Y. H. and Lee, P. C., "Novel high-precision grey forecasting model," *Automation in Construction*, Vol. 16, pp. 771-777 (2007).
16. Lin, Y. H., Lee, P. C., and Chang, T. P., "Adaptive and high-precision grey forecasting model," *Expert Systems with Applications*, Vol. 36, pp. 9658-9662 (2009).
17. Obeysekera, J. T. B., Tabios III, G. Q., and Salas, J. D., "On parameter estimation of temporal rainfall models," *Water Resources Research*, Vol. 23, No. 10, pp. 1837-1850 (1987).
18. Shi, Y. F. and Ning, J. S., "Improved grey model by exponential smoothing and its applications in deformation data analysis," *Journal of China Coal Society*, Vol. 30, No. 2, pp. 206-209 (2005).
19. Su, S. F., Lin, C. B., and Hsu, Y. T., "A high precision global prediction approach based on local prediction approaches," *IEEE Transactions on Systems, Man, and Cybernetic-Part C: Applications and Review*, Vol. 32, pp. 416-425 (2002).
20. Xiang, Y. L., "A fit-modeling method for GM(1,1) grey swing sequence and its application," *Environment Protection of Chemical Industry*, Vol. 18, No. 5, pp. 299-302 (1998).
21. Yang, Y. C., "Future prospect of sea-air transport logistics across the Taiwan strait via the Kaohsiung offshore shipping center," *Journal of Marine Science and Technology*, Vol. 18, No. 4, pp. 546-552 (2010).
22. Yu, P. S., Chen, C. J., and Chen, S. J., "Application of grey and fuzzy methods for rainfall forecasting," *Journal of Hydrologic Engineering*, Vol. 5, No. 4, pp. 339-345 (2000).
23. Yu, P. S., Chen, C. J., Chen, S. J., and Lin, S. C., "Application of grey model toward runoff forecasting," *Journal of the American Water Resources Association*, Vol. 37, No. 1, pp. 151-166 (2001).