國立臺灣大學工學院土木工程學系

博士論文

Department of Civil Engineering
College of Engineering

National Taiwan University

Doctoral Dissertation

應用支援向量機於颱風雨量及洪水預報

Typhoon Rainfall and Flood Forecasting Using Support

Vector Machine

周揚敬

Chou, Yang-Ching

指導教授：林國峰 教授

Advisor: Lin, Gwo-Fong

中華民國 102 年 7 月

July, 2013

國立臺灣大學博士學位論文
口試委員會審定書

應用支援向量機於颱風雨量及洪水預報

Typhoon Rainfall and Flood Forecasting Using Support
Vector Machine

　　本論文係周揚敬（D98521013）在國立臺灣大學土木工程
學系博士班完成之博士學位論文，於民國 102 年 7 月 12 日承下
列考試委員審查通過及口試及格，特此證明

口試委員：

林　國　峰	
（指導教授）	
陳　主　惠	
游　保　杉	
陳　明　杰	
賴　進　松	

系主任　　　呂　良　正

誌謝

首先誠摯的感謝指導教授林國峰老師悉心的指導，使我得以順利的完成博士論文，老師對創新研究知識的追求以及嚴謹的做事態度，使我在這些年中獲益匪淺。而論文初稿也承蒙老師與四位口試委員游保杉教授、陳主惠教授、陳明杰教授、賴進松教授的指導，使內容更加完善，在此也至上十二萬分的謝意。

非常感謝谷榕學長帶我進入學術領域，陪著我克服許多困難，也讓我第一次感受到文章被接受的喜悅。同樣非常感謝明璋學長幫研究室承擔了許多壓力，使我可以更專注在學術研究上。珮瑜學姐的細心，亦讓我在論文撰寫上不會犯太多錯。而在這漫長而苦悶的研究生涯中，感謝柏凱學長帶我們去光華散心，並帶領我們進入快打領域。在準備畢業口試的過程中，更感謝學弟軒宇的大力幫忙。學弟秉宸、明瑞、致瑋、信華、爵廷和學姊宜欣及學妹靚芸在研究上的協助，在此亦致上由衷的感謝。

感謝所有家人的支持，以及恩威並施的爸媽。最後，要感謝一路陪伴，當我最低潮時總是會鼓勵及支持我的老婆昭欣，因為有妳和寶貝兒女抖寶和菈菈，我才可以堅持完成博士學位，非常感謝。

中文摘要

本論文的主要目標為將支援向量機應用於洪災消減及災害預警上，主要可分為以下兩個部分：

當颱風來襲時，雨量預報在大部分災害預警系統中皆扮演了非常關鍵的角色。為了能更快速的得到準確的降雨預報，各個防災單位總是積極研發各種新式的預報模式。本研究提出一種稱為支援向量機(support vector machine, SVM)的類神經網路，並以此為基礎架構有效的颱風時雨量預報模式。相較於傳統上較常被使用的倒傳遞類神經網路，基於統計學習理論的支援向量機具有三項優勢。第一、支援向量機具備了更佳的學習能力(generalization ability)，第二、支援向量機在架構和權重的決定上保證有唯一解並且為全域最佳解，最後、支援向量機大量減少架構模式所需的訓練時間。本研究以實際案例來說明支援向量機所具備的優勢。研究結果顯示支援向量機相較於倒傳遞類神經網路不但能得到更加準確的預報結果，並且有更佳的強健性，其中最大的優勢是能大幅的縮短架構模式所需的時間。除了模式間的比較，為了能進一步提升長期預報的準確度，本研究更是新增了颱風因子做為降雨預報模式的輸入項，並與沒加入颱風因子做為輸入項的模式進行比較，以探討颱風因子對於雨量預報的影響。研究結果也證明了颱風因子可以有效提升中長期預報的準確度。總結來說，本研究提出以支援向量機為基礎納入颱風因子做為輸入項的預報模式確實能提升颱風時期雨量預報的準確度。而本模式亦

預期能為洪水預報、土石流警戒等災害預警系提供幫助。

　　對於洪水預警來說準確的流量預報是非常重要的關鍵。因此在第二階段的研究中提出一個以支援向量機為基礎，整合型的洪水預報模式來提升洪水預報的準確度。整合型的洪水預報模式可以分為兩個部分，雨量預報單元及流量預報單元。在第一階段，將以雨量及颱風因子作為輸入項發展雨量預報單元。接著將預報雨量及觀測流量作為輸入項發展流量預報單元。為了驗證整合型洪水預報模式的能力，本研究另外架構了直接納入觀測流量、雨量及颱風因子的洪水預報模式進行比較。並以實際發生的颱風事件作為研究案例並預報未來 1 至 6 小時的流量。研究結果顯示第一階段的雨量預報單元可以得到合理的預報結果。而將此預報雨量納入輸入項的整合型洪水預報模式，相較於直接納入各因子的洪水預報模式能得到更為準確的預報結果，甚至連尖峰流量亦有顯著的改善。值得注意的是，本研究提出的模式更是顯著的提升了中長延時的預報準確度。歸納結果，本研究提出的模式有效的減少了輸入項和輸出項間，隨著預報時間延長所帶來的負面影響，因此才能在中長延時仍能維持一定的準確度。而此一優勢將對於提升颱風時期洪水預警的反應時間有所幫助。
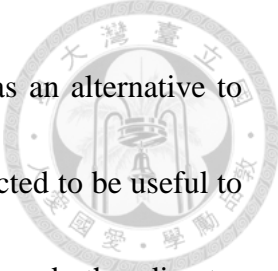
　　關鍵字：雨量預報，洪水預報，支援向量機，颱風因子，災害預警系統

Abstract

The objective of this dissertation is to apply support vector machine for flood mitigation and disaster warning. There are two major parts in this paper, which are summarized in the following manner.

Typhoon rainfall forecasting plays a critical role in almost all kinds of disaster warning systems during typhoons. To obtain more effective forecasts of hourly typhoon rainfall, novel models with better ability are desired. Based on support vector machines (SVMs), which is a kind of neural networks (NNs), effective hourly typhoon rainfall forecasting models are constructed. As compared with back-propagation networks (BPNs) which are the most frequently used conventional NNs, SVMs have three advantages: (1) SVMs have better generalization ability; (2) the architectures and the weights of the SVMs are guaranteed to be unique and globally optimal; (3) SVM is trained much more rapidly. An application is conducted to clearly demonstrate these three advantages. The results indicate that the proposed SVM-based models are more well-performed, robust and efficient than the existing BPN-based models. To further improve the long lead-time forecasting, typhoon characteristics are added as key input to the proposed models. The comparison between SVM-based models with and without typhoon characteristics confirms the significant improvement in forecasting performance due to the addition of typhoon characteristics for long lead-time

forecasting. The proposed SVM-based models are recommended as an alternative to the existing models. The proposed modeling technique is also expected to be useful to support reservoir operation systems and flood, landslide, debris flow, and other disaster warning systems.
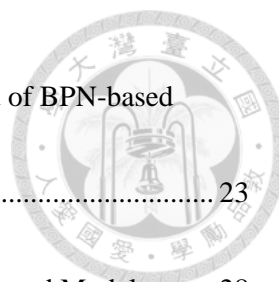
Accurate runoff forecasts are required to provide early warning of impending floods. In this part, an integrated flood forecasting model based on the support vector machine (SVM) is proposed to improve the flood forecasting performance. In the first stage, the observed typhoon characteristics and rainfall are used to produce rainfall forecasts. Then the forecasted rainfall and observed runoff are used to yield runoff forecasts. An actual application is performed to yield 1- to 6-h lead time runoff forecasts. The results show that the rainfall forecasting in the first stage can generate reliable rainfall forecasts, and the proposed model can provide accurate runoff forecasts, especially for the peak values. It is worth noting that the proposed model can significantly improve the 4- to 6-h lead time flood forecasting performance. In conclusion, the proposed model effectively mitigates the negative impact of increasing forecast lead time and is useful to improve the long lead time flood forecasting during periods of typhoon.

Keywords: rainfall forecasting, flood forecasting, support vector machines, typhoon characteristics, disaster warning systems.

Contents

List of tables

List of figures

Chapter 1 Introduction

1.1 Motivations

The island of Taiwan is situated in one of the main paths of the north-western Pacific typhoons. Each year, three to four typhoons attack the island on average. The torrential rain brought by typhoons (tropical cyclones occurring in the western Pacific Ocean) frequently lead to serious disasters, such as flooding, landslide or debris flow. To mitigate disasters due to typhoons, the development of flood warning systems are always needed. However, the highly non-linear and complex processes of typhoon rainfall and runoff make it difficult to construct a reliable physically-based model. To obtain more effective forecasts of typhoon rainfall and runoff, the development of better models has always been regarded as an important task.

An attractive alternative to the physically-based models is neural networks (NNs), which is a kind of information processing system with great flexibility in modeling nonlinear processes. The conception of NNs was inspired by a desire to understand human brain. Comprehensive reviews of the applications of NNs in hydrology have been presented by ASCE Task Committee (2000a, 2000b) and Maier and Dandy (2000). More recently, hydrologists, water resources engineers and managers have inspected more various applications of NNs for hydrologic forecasting, such as streamflow

forecasting (e.g., Tingsanchali and Gautam, 2000; Lin and Chen, 2004; Toth and Brath, 2007), tidal level forecasting (e.g., Supharatid, 2003), and groundwater level forecasting (e.g., Lin and Chen, 2005a).   In addition to the great flexibility in modeling nonlinear systems, NNs are very suitable for being integrated with decision- support systems due to their high computational efficiency.   For instance, NN-based models have been integrated with reservoir operation systems (e.g., Khalil et al., 2005; Chaves and Kojiri, 2007; Tu et al., 2008), city flood control systems (Chang et al., 2008), as well as debris flow warning systems (Chang et al., 2007).

Because of their flexibility in modeling nonlinear systems and their computational efficiency, NNs have gained a considerable attention.   More recently, a powerful kind of NNs named Support Vector Machines (SVMs) have attracted the attention of some hydrologists but only limited application are examined, such as hydrologic time series analysis (Liong and Sivapragasam, 2002; Yu et al., 2004; Asefa et al., 2005; Sivapragasam and Liong, 2005; Yu and Liong, 2007), reservoir inflow forecasting (Lin and Chen, 2009a), and streamflow forecasting (Yu et al., 2006; Kalra and Ahmad, 2009). Based on statistical learning theory, SVMs have advantages over back-propagation networks (BPNs) which are the most frequently used conventional NNs.   Firstly, SVMs have better generalization ability to relate the relatively irrelative input to the desired output.   This advantage is very helpful to decrease the negative impact when

2

increasing forecast lead-time. In other words, SVMs are capable of producing acceptably accurate forecasts for longer lead-time. Secondly, the optimization algorithm for SVMs is more robust, and the architectures and the weights of the SVMs are guaranteed to be unique and globally optimal. The performance of SVM-based models is more reliable because of the robust optimization algorithm. Finally, SVMs are trained much more rapidly. Thus, SVM-based models are more suitable to be integrated with disaster warning systems and decision support systems.

Due to the aforementioned attractive advantages, SVMs have emerged as an alternative data-driven tool in many conventional NN dominated fields. In this dissertation, SVMs were introduced and applied to typhoon rainfall and flood forecasting.

1.2 Backgrounds and Inspiration

1.2.1 Effective forecasting of hourly typhoon rainfall

Rainfall forecasting plays a critical role in almost all kinds of disaster warning systems during typhoons. In the Taiwan area, typhoon rainfall often causes casualties and has major economic impacts; however, it is an important water resource. As a typhoon approaches the island, the major goal of reservoir operation is to control floods. But when the typhoon leaves, the goal switches to restore sufficient water for future

usage. To achieve these two goals, reservoir operation should be appropriately conducted. More effective (or more accurate and reliable) forecasts of hourly rainfall are required as a vital reference for hourly reservoir inflow forecasting and for making important reservoir operation decisions. In addition, an improved hourly rainfall forecasting is expected to be useful to support flood, landslide, debris flow and other disaster warning systems.

As to rainfall forecasting, applications of NNs have also been presented (e.g., Luk et al., 2001; Chiang et al., 2007), but studies on NN-based models for hourly typhoon rainfall forecasting are still limited. To provide effective forecasts of hourly typhoon rainfall for being integrated with decision support systems, Lin and Chen (2005b) have assessed the potential of BPNs. The results indicated that BPN-based models yield acceptable forecasts for a lead time of one to two hours only. To provide effective warnings, longer lead-time forecasting is needed. However, as the forecast lead-time increases, the correlation between desired output and available input decreases. The data used for long lead-time forecasting include more relatively irrelative information which seriously undermines the performance of BPN-based models. Based on previous studies (Lin and Chen, 2005b, 2008), the generalization ability of BPNs was not good enough. To obtain effective forecasts of hourly rainfall for longer lead-time forecasting, it is justified to propose novel models with better generalization ability.

4

In this study, SVMs were used to construct typhoon rainfall forecasting models. Hong and Pai (2007) have constructed a SVM-based typhoon rainfall forecasting model with only antecedent rainfall as input. Their forecasts were acceptably accurate but only for one-hour ahead forecasting. With only antecedent rainfall as input, the performance of models usually decreases rapidly with increasing forecast lead-time. To further enhance the long lead-time forecasting, typhoon characteristics were regarded as key input and added to the proposed SVM-based models. Lin and Chen (2005b) have confirmed that the trend of rainfall could be demonstrated by typhoon characteristics when a typhoon was nearby. It is reasonable to speculate that typhoon characteristics are capable of providing valuable information for longer lead-time forecasting. Such a speculation has prompted an investigation into the influence of typhoon characteristics on rainfall forecasting, in particular, for long lead-time forecasting.

The objective of this study was to provide effective forecasts of hourly rainfall for supporting reservoir operation systems during typhoons. For this purpose, SVM-based, instead of BPN-based models with typhoon characteristics were proposed to yield 1-to 6-h lead time forecasts. In order to compare SVMs and BPNs, BPN-based models with same input are also constructed. Moreover, to assess the improvement in forecasting performances due to the addition of typhoon characteristics, two types of

model input (with and without typhoon characteristics) are designed for SVM-based models. Finally, an application was conducted and 11 typhoon events were used in this study. To reach just conclusions, cross validations were applied to evaluate the overall performance of the models and the statistical significance of the improvement in forecasting performance was identified by paired comparison t-tests. The results demonstrated the superiority of the proposed models more clearly.

1.2.2 Typhoon flood forecasting using integrated SVM

To mitigate disasters caused by typhoons, accurate and reliable flood forecasts are essential to provide early warning of impending floods and their improvement has been verified as a crucial task. In recent years, NNs have been successfully employed in various hydrologic modeling applications (e.g., de Vos and Rientjes, 2005; Hu et al., 2007; Lin and Chen, 2004; Wu and Chau, 2011) and specifically for rainfall and flood forecasting (e.g., Chang et al., 2004; Chiang et al., 2007; Lin and Chen, 2005b; Lin et al., 2010; Luk et al., 2001; Pramanik et al., 2011; Rathinasamy and Khosa, 2012; Toth and Brath, 2007). The major advantage of NNs is their capability to simulate complex relationship between desired output and available input given the existence of sufficient training datasets.

However, flood forecasting performance of most NNs decreases rapidly with

increasing of the forecast lead time. Operational agencies which are responsible for

flood mitigation and warnings could well benefit from improved forecast accuracy of

the longer lead times. Multi-stage NN-based models were developed in attempt to

achieve a longer as well as accurate forecast lead time (Chang et al., 2007; Lin and Wu,

2011). The concept of multi-stage NN-based models is that two or more NN-based

models are connected. Using the two-stage as an example, the connection between the

two stages is that forecasted values from the first-stage module are used as input to the

second-stage module. It is widely known that rainfall is one of the most important

inputs to flood forecasting model and the accuracy of long lead time flood forecasting

can be advanced with more accurate rainfall forecasts. Lin et al. (2009c) improved

longer lead time streamflow forecast by adopting BPNs to predict rainfall as input to a

Radial Basis Function (RBF)-based reservoir inflow forecasting model. Chiang and

Chang (2009) used Quantitative Precipitation Forecasting (QPF) information as input to

Recurrent Neural Network (RNN)-based flood forecasting model and reported a similar

finding, that is, the forecasted rainfall was capable of providing useful information for

flood forecasting, especially for long lead time.

In previous studies, multi-stage NN-based flood forecasting models were based on

conventional NNs. The architecture and the weights of these conventional NNs were

determined by a trial and error procedure which consisted of iterative time-consuming

process. Although the selection of NN-based models generally disregard the efficiency of the model training, it is essential to develop a well-performing model that can be quickly trained.

In this study, an integrated support vector machine (SVM)-based model was developed to yield 1- to 6-h lead time runoff forecasts. SVMs have been used for hydrologic time series forecasting (Liong and Sivapragasam, 2002; Sivapragasam and Liong, 2005; Wu et al., 2009; Yu and Liong, 2007; Yu et al., 2004). More recently, Rasouli et al. (2012) employed SVM and other machine learning methods with weather and climate inputs to forecast daily streamflow. According to statistical learning theory, SVM has better generalization ability and requires less training time than conventional NNs (e.g., BPN). For both rainfall and inflow forecasting, Lin et al. (2009a, 2009b) demonstrated that SVM-based models outperform BPN-based models. Moreover, the development of SVM-based models is efficient and thus expected to be suitable for development of the integrated model presented herein.

The objective of this study is to demonstrate an integrated SVM-based model for typhoon flood forecasting. A rainfall forecasting module was established in the first stage to pre-process the typhoon information (namely, typhoon characteristics and rainfall) as well as to produce rainfall forecasts. Afterwards, the rainfall forecasts

along with the observed runoff were used as input to the flood forecasting module in the

second stage. This procedure was expected to reduce the input dimensionality and

improve the performance of the longer forecast lead times, especially for the prediction

of peak runoff.
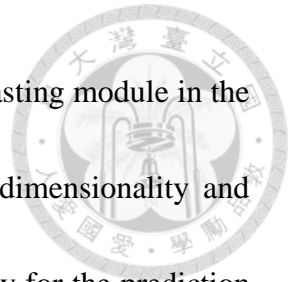
Chapter 2 Support vector machine

In the early 1990s, Vapnik developed SVMs for classification and then extended for regression (Vapnik, 1995). There are two major differences between the SVMs and the BPNs. Firstly, instead of empirical risk minimization (ERM), the structural risk minimization (SRM) induction principle is used to construct SVMs. For training BPNs, the only one objective is to minimize the total error (or empirical risk). As to SVMs, according to the SRM induction principle, both the empirical risk and the model complexity should be minimized simultaneously. The use of SRM induction principle results in the better generalization ability of SVMs. Another major difference is the determination of the model architecture and the weights. For BPNs, the architecture and the weights are respectively determined by a trial-and-error procedure and an iterative process (the error back-propagation algorithm), which both are very time-consuming. Vapnik (1995) dispensed with the time-consuming training process and expressed the determination of the architecture and weights of SVMs in terms of a quadratic optimization problem which can be rapidly solved by a standard programming algorithm. In this section, the methodology of the support vector regression (SVR) used in this paper is briefly described and more mathematical details about SVR can be found in several text books (Vapnik, 1995; Vapnik, 1998; Cristianini and Shaw-Taylor, 2000).

Based on N_d training data $[(\mathbf{x}_1, y_1), (\mathbf{x}_2, y_2), ..., (\mathbf{x}_{N_d}, y_{N_d})]$, the objective of the support vector regression is to find a non-linear regression function to yield the output $\hat{y}$, which is the best approximate of the desired output y with an error tolerance of ε. Firstly, the input vector $\mathbf{x}$ is mapped onto a higher dimensional feature space by a non-linear function $\phi(\mathbf{x})$. Then the regression function that relates the input vector $\mathbf{x}$ to the output $\hat{y}$ can be written as

$$\hat{y} = f(\mathbf{x}) = \mathbf{w}^{\mathrm{T}}\phi(\mathbf{x}) + b \tag{2.1}$$

where $\mathbf{w}$ and b are weights and bias of the regression function, respectively. Based on the SRM induction principle, $\mathbf{w}$ and b are estimated by minimizing the following structural risk function:

$$R = \frac{1}{2}\mathbf{w}^T\mathbf{w} + C\sum_{i=1}^{N_d}L_\varepsilon(\hat{y}_i) \tag{2.2}$$

where the Vapnik's ε-insensitive loss function L_ε is defined as

$$L_\varepsilon(\hat{y}) = |\ y - \hat{y}\ |_\varepsilon = \begin{cases} 0 & \text{for } |y - \hat{y}| < \varepsilon \\ |y - \hat{y}| - \varepsilon & \text{for } |y - \hat{y}| \geq \varepsilon \end{cases} \tag{2.3}$$

The first and second terms in Eq. (2) represent the model complexity and the empirical error, respectively. The trade-off between the model complexity and the empirical error is specified by a user-defined parameter C and $C = 1$ is set herein.

Vapnik (1995) expressed the SVR problem in terms of the following optimization

problem:

Minimize

$$R(\mathbf{w},b,\xi,\xi') = \frac{1}{2}\mathbf{w}^T\mathbf{w} + C\sum_{i=1}^{N_d}(\xi_i + \xi_i')$$

subject to

$$y_i - \hat{y}_i = y_i - (\mathbf{w}^T\phi(\mathbf{x}_i)+b) \le \varepsilon + \xi_i \qquad\qquad (2.4)$$
$$\hat{y}_i - y_i = (\mathbf{w}^T\phi(\mathbf{x}_i)+b) - y_i \le \varepsilon + \xi_i'$$
$$\xi_i \ge 0$$
$$\zeta_i' \ge 0$$
$$i = 1,2,...,l$$

where ξ and ξ', which are slack variables, represent the upper and the lower training

errors, respectively.   The above optimization problem is usually solved in its dual form

using Lagrange multipliers.   Rewriting Eq. (4) in its dual form and differentiating with

respect to the primal variables ($\mathbf{w},b,\xi,\xi'$) gives

Maximize

$$\sum_{i=1}^{N_d} y_i(\alpha_i - \alpha_i') - \varepsilon\sum_{i=1}^{N_d}(\alpha_i + \alpha_i') - \frac{1}{2}\sum_{i=1}^{N_d}\sum_{j=1}^{N_d}(\alpha_i - \alpha_i')(\alpha_j - \alpha_j')\phi(\mathbf{x}_i)^T\phi(\mathbf{x}_j)$$

subject to

$$\sum_{i=1}^{N_d}(\alpha_i - \alpha_i') = 0 \qquad\qquad (2.5)$$
$$0 \le \alpha_i \le C$$
$$0 \le \alpha_i' \le C$$
$$i = 1,2,...,N_d$$

where α and α' are the dual Lagrange multipliers.   Note that the solution to the

optimal problem (Eq. (5)) is guaranteed to be unique and globally optimal because the

objective function is a convex function.

The optimal Lagrange multipliers α^* are solved by the standard quadratic programming algorithm and then the regression function can be rewritten as

$$f(x) = \sum_{i=1}^{N_d} \alpha_i^* K(\mathbf{x}_i, \mathbf{x}) + b \tag{2.6}$$

where the kernel function $K(\mathbf{x}_i, \mathbf{x})$ is defined as

$$K(\mathbf{x}_i, \mathbf{x}) = \phi(\mathbf{x}_i)^{\mathrm{T}} \phi(\mathbf{x}) \tag{2.7}$$

The kernel function used in this paper is the radial basis function:

$$K(\mathbf{x}_i, \mathbf{x}) = \exp(-\frac{1}{n_x} |\mathbf{x}_i - \mathbf{x}|^2) \tag{2.8}$$

where n_x is the number of components in input vector $\mathbf{x}$.

Some of solved Lagrange multipliers $(\alpha - \alpha')$ are zero and should be eliminated form the regression function. Finally, the regression function involves the nonzero Lagrange multipliers and the corresponding input vectors of the training data, which are called the support vectors. The final regression function can be rewritten as

$$f(\mathbf{x}) = \sum_{k=1}^{N_{sv}} \alpha_k K(\mathbf{x}_k, \mathbf{x}) + b \tag{2.9}$$

where $\mathbf{x}_k$ denotes the k th support vector and N_{sv} is the number of support vectors.

The architecture of a SVM is presented in Fig. 2.1.
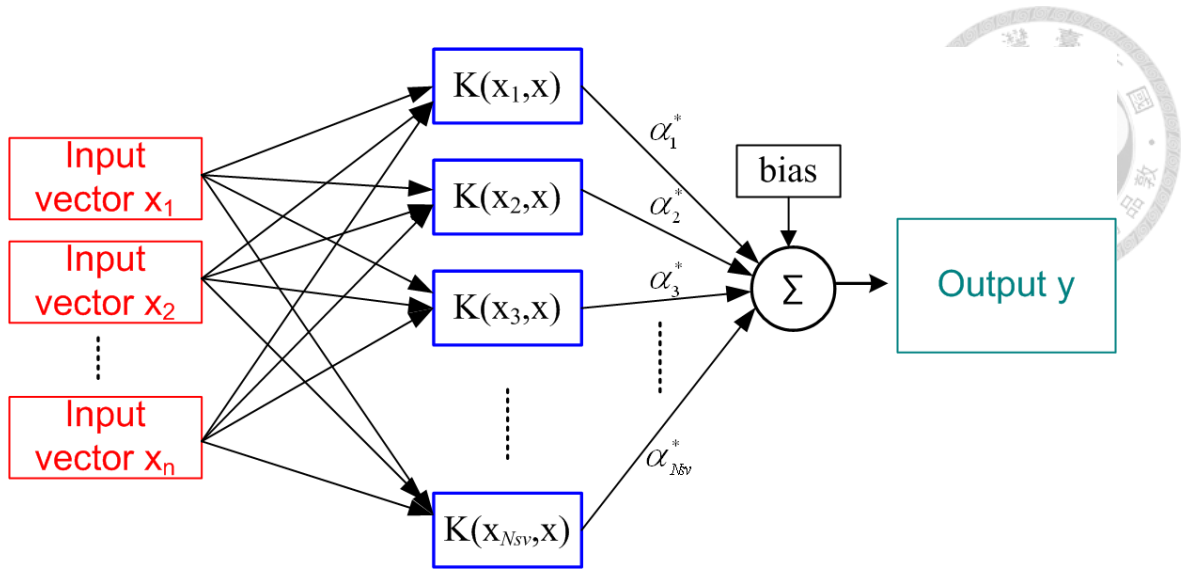
Figure 2.1　　　Architectural graph of SVM

Chapter 3 Effective forecasting of hourly typhoon rainfall

3.1 Application

3.1.1 The Study Area and Data

The study area is the Fei-Tsui Reservoir Watershed in northern Taiwan. The Fei-Tsui Reservoir is located downstream of three major tributaries (the Kingkwa Creek, the Diyu Creek and the Peishih Creek). The reservoir has a surface area of 10 km^2, a mean depth of 40 m, a maximum depth of 120 m, a full capacity of 406 million m^3, and a total watershed area of 303 km^2. From 1988 to 2007, the maximum and average yearly rainfall is 5736.6 mm and 3808.6 mm, respectively. Fig. 3.1 shows the study area and the locations of six rain gauges (Fei-Tsui, Pin-Lin, Shi-San-Ku, Chiu-Chiung-Ken, Bi-Hu and Tai-Pin). The rainfall data are obtained from the Water Resources Agency and the typhoon characteristics are collected from the Central Weather Bureau. The time periods of the data of rainfall and typhoon characteristics are hourly. A total of 11 typhoon events with typhoon characteristics and rainfall data available simultaneously are used herein. Table 3.1 summarizes the date of occurrence and duration of these 11 typhoon events. The typhoon characteristics include the position of the typhoon center, the distance between the center and the reservoir, the maximum wind speed near the center, the atmospheric pressure of the center, the radius of winds over 15 m/s, and the speed of the typhoon movement.

Table 3.1 Descriptions of typhoon events used in the modeling

Name of typhoon	Date	Duration (hour)	Maximum wind speed (m/s)	Maximum hourly rainfall (mm)
Ted	1992/09/20	69	108	19.6
Tim	1994/07/09	49	191	16.3
Gladys	1994/08/31	32	126	31.0
Seth	1994/10/08	75	184	18.7
Herb	1996/07/30	70	191	46.7
Haiyan	2001/10/15	28	130	14.1
Rammasun	2002/07/03	32	165	22.5
Nock-Ten	2004/10/24	44	155	35.4
Haitang	2005/07/17	66	198	42.5
Matsa	2005/08/04	48	144	31.2
Talim	2005/08/31	43	191	42.6
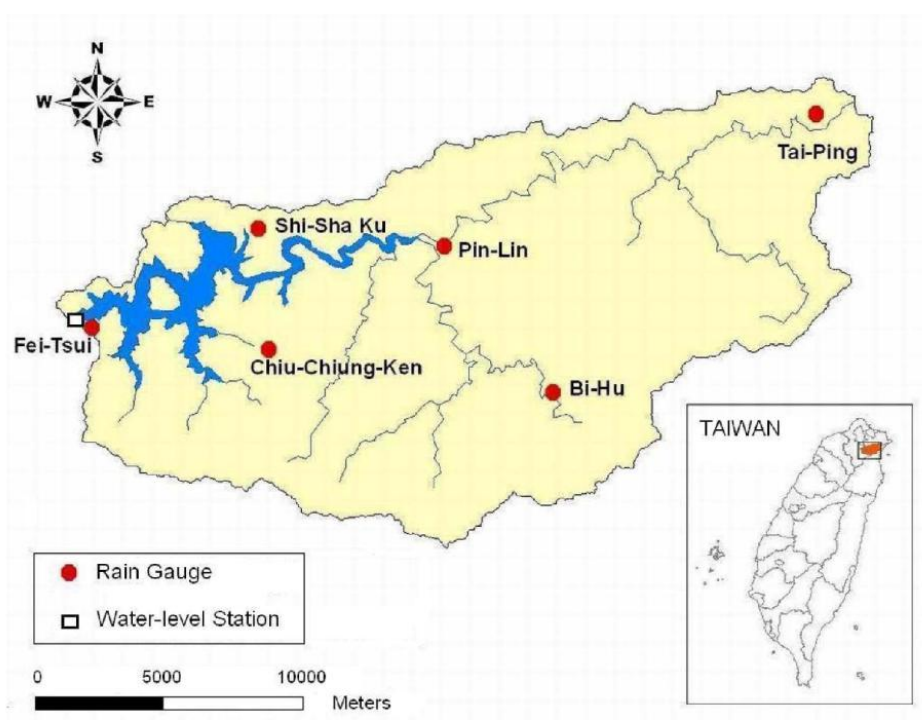
Figure 3.1　　The study area
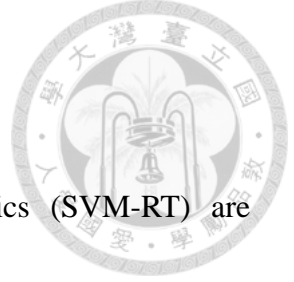
3.1.2 Development of Models

In this study, SVM-based models with typhoon characteristics (SVM-RT) are proposed to yield 1- to 6-h lead time rainfall forecasts. To make comparisons between SVMs and BPNs, BPN-based models with typhoon characteristics (BPN-RT) are constructed. Otherwise, SVM-based models without typhoon characteristics (SVM-R) are also constructed to evaluate the improvement in forecasting performance due to the addition of typhoon characteristics. Hence, a total of 18 NN-based forecasting models (SVM-RT, BPN-RT and SVM-R for 1- to 6-h lead time forecasts) are constructed.

The SVM-R has a general form as

$$R_{t+\Delta t} = f\left(R_t, R_{t-1}, \cdots, R_{t-(L_R-1)} \right) \tag{3.1}$$

where t is the current time, Δt is the lead-time period (from one to six hours), R_t is rainfall at time t, and L_R denotes the lag length of rainfall. The model development is schematized in Fig. 3.2(a). Firstly, only rainfall is used as input and the lag length of rainfall (L_R) is determined by the process shown in Fig. 3.2(b). The criterion for selecting the lag lengths is the relative percentage error (RPE):

$$\text{RPE} = \frac{E(L) - E(L+1)}{E(L)} \times 100 \tag{3.2}$$

where $E(L)$ and $E(L+1)$ are the RMSEs for models with L and $L+1$ lag
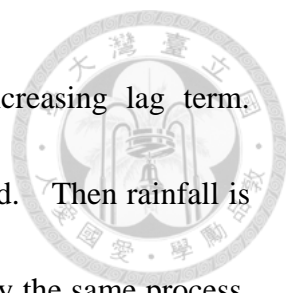
lengths, respectively. In general, the RMSE decreases with increasing lag term.

When the RPE is less than 5%, the increase of lag lengths is stopped. Then rainfall is

added to the input and the lag length of rainfall (L_R) is determined by the same process.

Once L_R is determined, the inputs of SVM-R are completely specified.
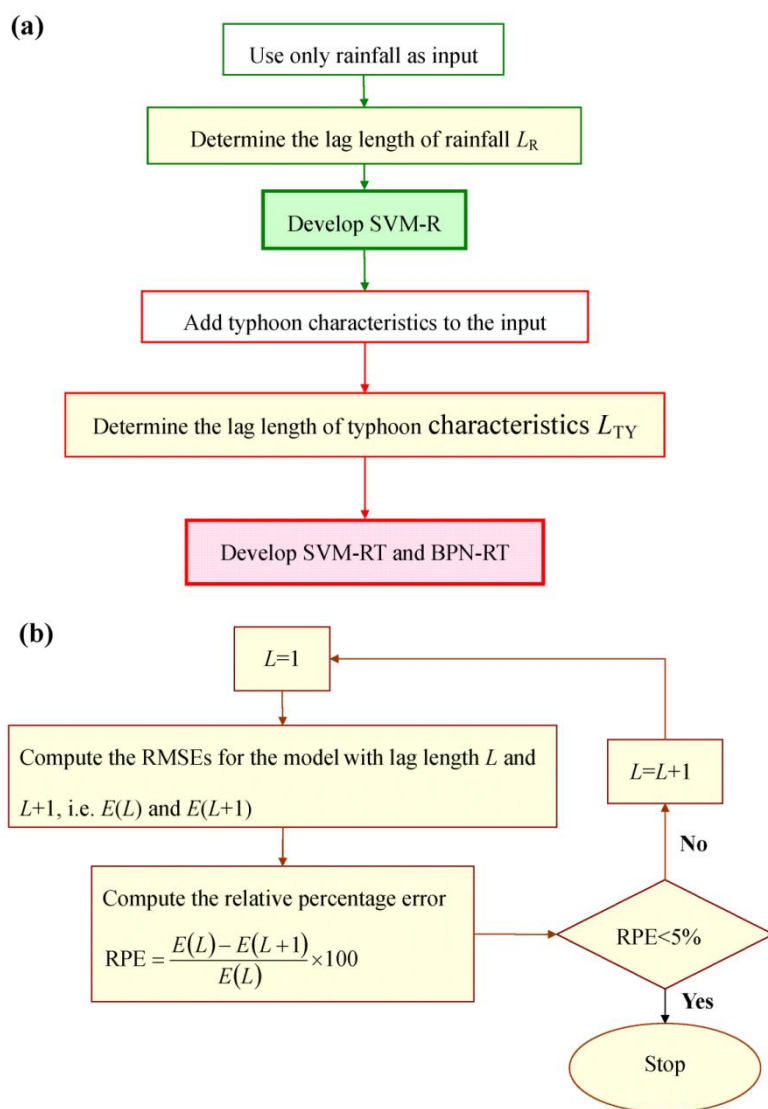
Figure 3.2   Flowcharts of (a) the model development and (b) the lag length

determination
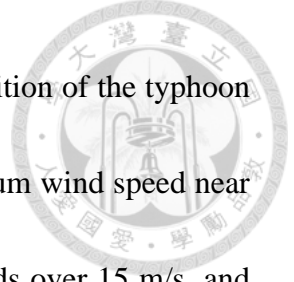
Based on the SVM-R, all typhoon characteristics (include the position of the typhoon center, the distance between the center and the reservoir, the maximum wind speed near the center, the atmospheric pressure of the center, the radius of winds over 15 m/s, and the speed of the typhoon movement) are then added to develop the SVM-RT and BPN-RT. The form of the SVM-RT and BPN-RT is

$$R_{t+\Delta t} = f\left(R_t, R_{t-1}, \cdots, R_{t-(L_R-1)}, TY_t, TY_{t-1}, \cdots, TY_{t-(L_{TY}-1)} \right) \qquad (3.3)$$

where TY_t is typhoon characteristics at time t, and L_{TY} denotes the lag length of typhoon characteristics which is determined by the same process shown in Figure 3.2(b). To further investigate the influence of each typhoon characteristic on rainfall forecasting, a total of six models with single typhoon characteristic are constructed by individually adding each typhoon characteristic in turn to the model without typhoon characteristics (SVM-R).

3.1.3 Cross Validation and Performance Measures

For event-based data, the collected events are usually separated into two sets of data: training and testing. Some events are chosen as training data and used to construct NN-based models. Then the performance of the NN-based models is tested by the remaining events which are not used in the training process. Different selections of training data and testing data produced different results and sometimes lead to different
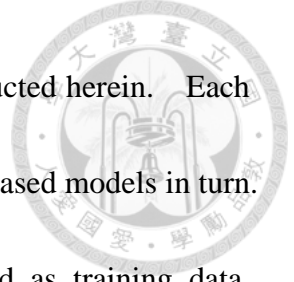
conclusions.   To reach just conclusions, cross validations are conducted herein.   Each

single typhoon event (except Typhoon Herb) is used to test the NN-based models in turn.

Typhoon Herb yielded the maximum rainfall and should be used as training data.

Then conclusions are drawn based on the overall performance for the 10 testing events.

Three performance measures are used to evaluate the model performance herein.

1. Coefficient of efficiency (CE):

$$CE = 1 - \frac{\sum_{t=1}^{n}(R_t - \hat{R}_t)^2}{\sum_{t=1}^{n}(R_t - \overline{R})^2} \tag{3.4}$$

where R_t and $\hat{R}_t$ denote the observed and forecasted rainfall at time t, respectively,

$\overline{R}$ is the average of the observed rainfall, and n is the number of forecasts.   If the

CE value is equal to one, the forecasts are perfect.

2. Mean absolute error (MAE):

$$MAE = \frac{1}{n}\sum_{t=1}^{n}\left|R_t - \hat{R}_t\right| \tag{3.5}$$

3. Root mean square error (RMSE):

$$RMSE = \sqrt{\frac{1}{n}\sum_{t=1}^{n}\left(R_t - \hat{R}_t\right)^2} \tag{3.6}$$

3.2 Results and Discussion

Table 3.2 presents the three performance measures (CE, MAE and RMSE) of three NN-based models (SVM-RT, BPN-RT and SVM-R) for 1- to 6-h lead time rainfall forecasts. According to the three performance measures, the proposed SVM-based model with typhoon characteristics (SVM-RT) produces the best performance and BPN-RT performs the worst among all models. To further identify whether SVM-RT performs significantly better than BPN-RT and SVM-R for the same testing event, paired comparison t-tests are conducted at the 1% significance level. The equation of t-tests is defined as

$$t = \frac{\left|\overline{X} - \mu_0\right|}{S_{\overline{X}}} = \frac{\left|\overline{X} - \mu_0\right|}{s/\sqrt{n}} \tag{3.7}$$

The results listed in Table 3.3 show that SVM-RT significantly yields higher CE, lower RMSE, and lower MAE values than both BPN-RT and SVM-R. To clearly demonstrate the superiority of the proposed models (SVM-RT), more performance comparisons are discussed in depth in the rest of this section.

Table 3.2 Coefficient of efficiency (CE), mean absolute error (MAE) and root mean square error (RMSE) for various models

Lead time (hour)	CE		
	SVM-RT	BPN-RT	SVM-R
1	0.44	0.35	0.43
2	0.32	0.17	0.26
3	0.24	0.04	0.17
4	0.24	0.02	0.10
5	0.21	-0.12	0.01
6	0.21	-0.19	-0.02
Lead time (hour)	MAE (mm)		
	SVM-RT	BPN-RT	SVM-R
1	3.21	3.85	3.29
2	3.59	4.34	3.77
3	3.91	4.67	4.12
4	3.93	4.82	4.31
5	4.04	5.08	4.58
6	4.20	5.32	4.75
Lead time (hour)	RMSE (mm)		
	SVM-RT	BPN-RT	SVM-R
1	5.14	5.54	5.21
2	5.69	6.27	5.95
3	6.04	6.81	6.30
4	6.06	6.88	6.62
5	6.23	7.40	6.96
6	6.28	7.69	7.11

Table 3.3 Paired comparison *t*-tests of three performance measures (CE, MAE and RMSE) resulting from SVM-RT and BPN-RT and from SVM-RT and SVM-R

Model	Alternate hypothesis	*t*-statistic	Statistically significant at the 1% level
SVM-RT and BPN-RT	$CE_{SVM-RT} > CE_{BPN-RT}$	3.18	Yes
	$MAE_{SVM-RT} < MAE_{BPN-RT}$	4.35	Yes
	$RMSE_{SVM-RT} < RMSE_{BPN-RT}$	3.37	Yes
SVM-RT and SVM-R	$CE_{SVM-RT} > CE_{SVM-R}$	6.47	Yes
	$MAE_{SVM-RT} < MAE_{SVM-R}$	7.20	Yes
	$RMSE_{SVM-RT} < RMSE_{SVM-R}$	6.85	Yes

Note: The critical *t*-value is 2.39.

3.2.1 The Improvement Due to the Use of SVM-based Models Instead of BPN-based Models

To highlight the improvement in forecasting performance due to the use of SVM-based models instead of BPN-based models, we first focus on the comparison between the proposed SVM-based models (SVM-RT) and BPN-based models with the same input (BPN-RT). As shown in Fig. 3.3(a), the CE values of both SVM-RT and BPN-RT decrease with increasing forecast lead-time. However, it is clear that SVM-RT yields significantly higher CE values than BPN-RT for 1- to 6-h lead-time forecasting. Thus, it is concluded that SVM-based models perform better than BPN-based models. Fig. 3.3(a) also shows that the CE values of SVM-RT decrease more slowly than those of BPN-RT. For 1- to 2-h lead time forecasts, the CE values of SVM-RT only decrease from 0.44 to 0.32, but those of BPN-RT rapidly decrease from 0.35 to 0.17. Then, for 3- to 6-h lead time forecasts, the performance of BPN-RT gets worse and the CE values are almost equal or even lower than zero. It is clear that BPN-RT cannot yield effective forecasts when the forecast lead-time is greater than two hours. As to SVM-RT, the performance is still acceptable for long lead-time forecasting. For 3- to 6-h lead time forecasts, the CE values only decrease from 0.24 to 0.21. The use of SVM-based models instead of BPN-based models effectively decreases the negative impact of increasing forecast lead-time.

The improvement in CE due to the use of SVM-based models instead of BPN-based models presented in Fig. 3.3(b) more clearly shows that the proposed SVM-based models (SVM-RT) effectively improve the forecasting performance.   For 1- to 6-h lead time forecasts, the improvement in CE increases from 9% to 40%.   It is clear that SVM-based models are more appropriate for long lead-time forecasting than BPN-based models.   There are reasonable explanations for the results presented in Fig. 3.3.   As the forecast lead time increases, the correlation between desired output and available input decreases.   Hence the data used for long lead-time forecasting include more relatively irrelative information and the models require better generalization ability to relate the input to the desired output.   Based on the statistical learning theory, SVMs have better generalization ability than BPNs.   Thus, a greater improvement in performance could be obtained by using SVM-based models instead of BPN-based models, especially for long lead-time forecasting.
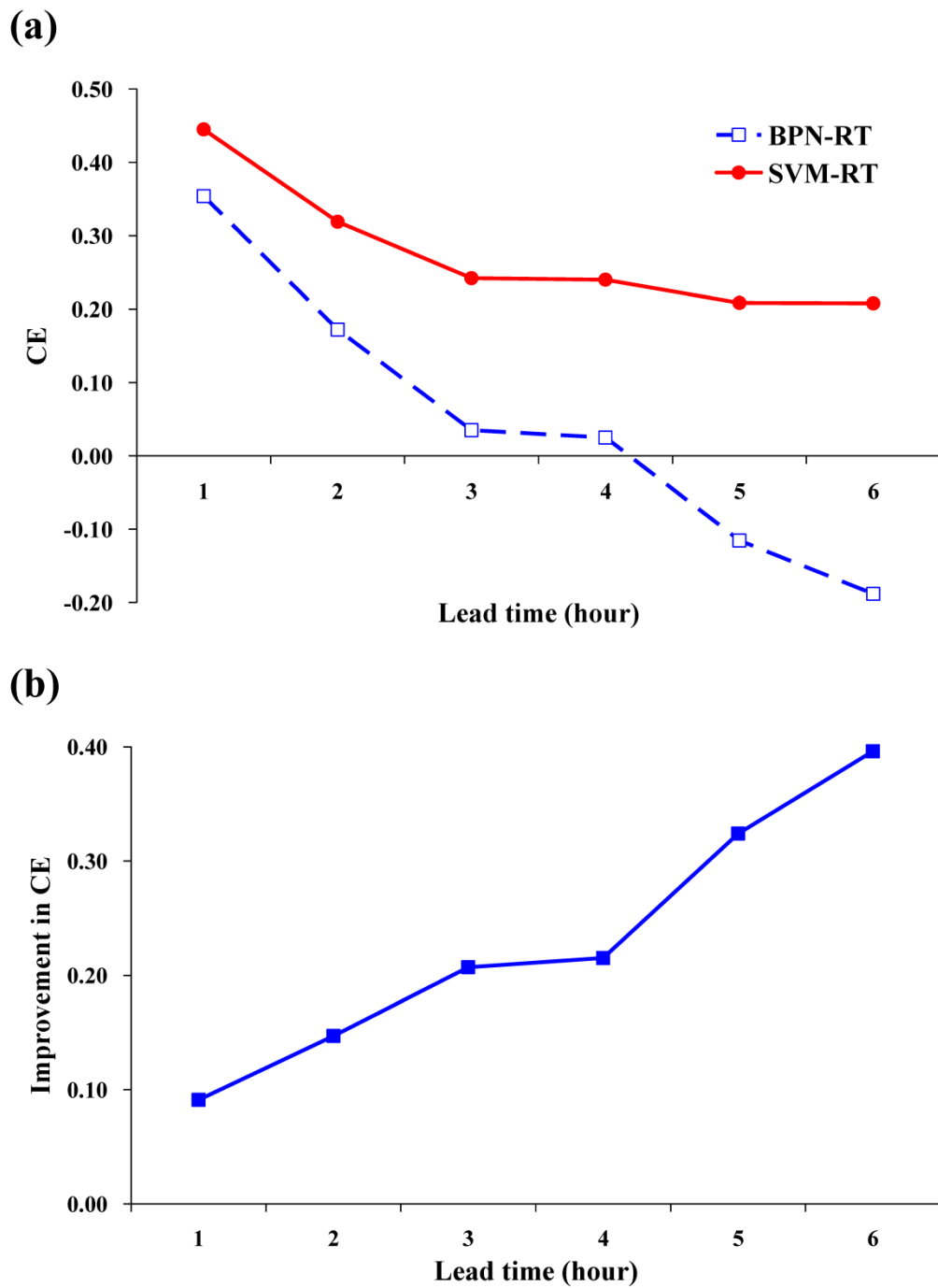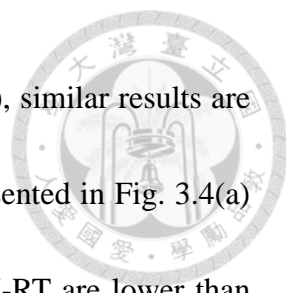
Figure 3.3   (a) CE values of SVM-RT and BPN-RT and (b) the improvement in CE

due to the use of SVM-based models instead of BPN-based models

According to other two performance measures (MAE and RMSE), similar results are also obtained. For 1- to 6-h lead time forecasts, the bar charts presented in Fig. 3.4(a) and 3.4(b) clearly show that both MAE and RMSE values of SVM-RT are lower than those of BPN-RT. In addition, the MAE and RMSE values of SVM-RT increase more slowly than those of BPN-RT with increasing forecast lead-time. The negative impact of increasing forecast lead-time has been effectively decreased by using SVM-based models instead of BPN-based models. The percentages of decrease in MAE and RMSE due to the use of SVM-based models instead of BPN-based models are presented in Fig. 3.4(c). For 1- to 6-h lead time forecasts, SVM-RT respectively decreases MAE and RMSE values from 17% to 21% and from 7% to 18% as compared to BPN-RT. Again, the results confirm that the proposed SVM-based models (SVM-RT) effectively improve the forecasting performance.

(a)
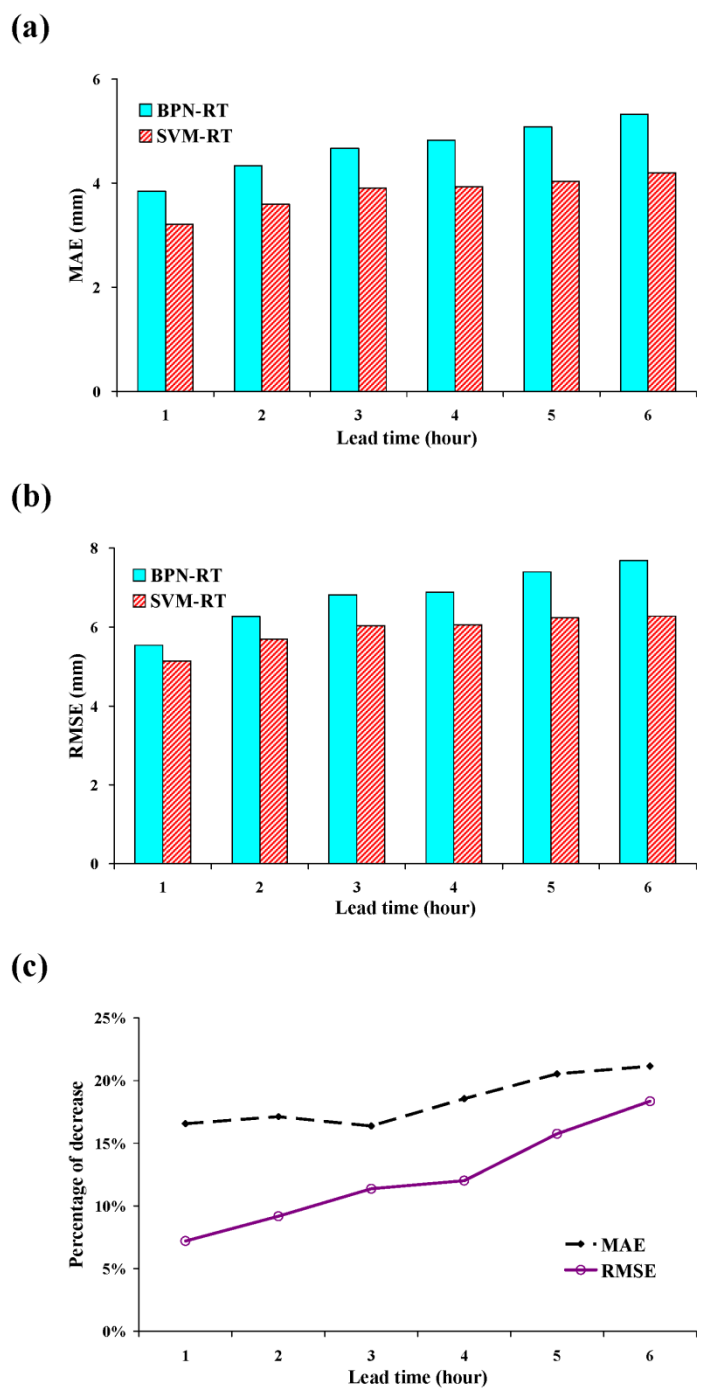

(b)

(c)

Figure 3.4   (a) MAE, (b) RMSE values of SVM-RT and BPN-RT, and (c) the

percentages of decrease in MAE and RMSE due to the use of SVM-based models

instead of BPN-based models

3.2.2 The Comparison of Robustness between SVM-based and BPN-based Models
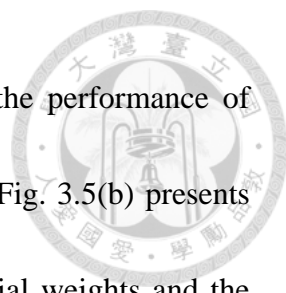
A robust parameter optimization algorithm is required to construct a robust model which can yield reliable forecasts. But for hydrological models, there are very limited studies on the related topics. In addition to better generalization ability, the robustness of optimization algorithm for SVMs is one of the major advantages over BPNs. For a robust optimization algorithm, the obtained optimal weights are slightly influenced by the initial conditions, such as initial weights. On the contrary, the optimization algorithm is less robust, if obtained optimal weights highly depend on the initial weights. For BPNs, the weights are determined by an iterative process and the optimal weights depend on the initial weights which are usually set randomly. Even when a BPN is trained with the same training data, but different initial weights may lead to different optimal weights and, of course, different forecasting performance. Hence, enough sets of initial weights should be tried to get the globally optimal weights. Based on our experiences, at least 30 sets of initial weights are required to obtain a reliable forecasting performance. As to SVMs, the determination of the architecture and the weights is expressed in terms of a quadratic optimization problem with a convex objective function, and the weights are solved by the standard programming algorithm. The solved weights are guaranteed to be unique and globally optimal. Thus, the same

optimal weights could be always obtained when a SVM is trained with the same

training data. Theoretically, the performance of SVM-based models is more reliable

than that of BPN-based models

To demonstrate the robustness of SVM-based models more clearly, SVM-RT and

BPN-RT are taken for example herein. Being trained with 10 typhoon events and

tested by the remaining one event, SVM-RT yields a constant RMSE. As to BPN-RT,

different initial weights lead to different RMSE values even when the same training and

testing data are used. In this case, BPN-RT is trained with 30 different sets of initial

weights and hence yields 30 different RMSE values for each testing typhoon event.

The lack of robustness can be observed by the variation in RMSE, which is evaluated

by the coefficient of variation (CV). A higher CV value of RMSE represents the

higher variation in RMSE and also indicates that the performance of BPN-RT is less

reliable. For each typhoon event, the CV is calculated from a data set of 30 RMSE

values resulting from BPN-RT trained with 30 different sets of initial weights. Fig.

3.5(a) presents the CV values of 10 typhoon events. Among 10 events, 6 events have

CV values greater than 20% and 2 events exceeding 40%. For Typhoon Gladys,

BPN-RT yields the highest CV value of 50%. The result clearly shows that BPN-RT

lacks robustness.

To discuss the robustness of SVM-RT and BPN-RT in depth, the performance of SVM-RT and BPN-RT tested by Typhoon Gladys is highlighted.   Fig. 3.5(b) presents the RMSE values of BPN-RT trained with 30 different sets of initial weights and the constant RMSE value of SVM-RT.   As shown in Fig. 3.5(b), the constant RMSE yielded by SVM-RT indicates that SVM-RT is a robust model and the performance is reliable.   On the contrary, it is found that the variation in RMSE for BPN-RT is very significant.   The great variation in RMSE confirms that the initial weights have great influence on the forecasting performance of BPN-based models.   Thus, enough sets of initial weights should be tried to ensure the satisfactory and reliable results.   For the case of Typhoon Gladys, 16 sets (more that 50% of all sets) have significantly worse performances (with RMSE>10 mm).   It is reasonable to speculate that the iterative processes have been trapped in the local optimal solutions and such a situation can be completely avoided by the use of SVMs.   Based on the above results, it is concluded that SVM-based models are more robust than BPN-based models.
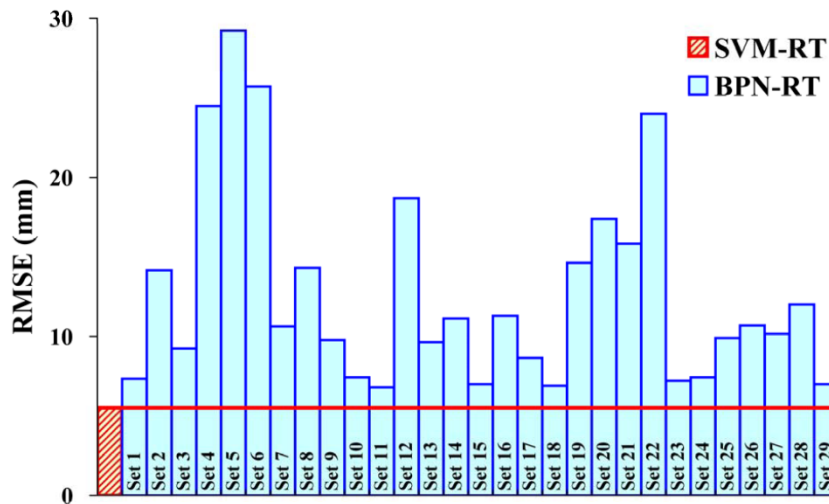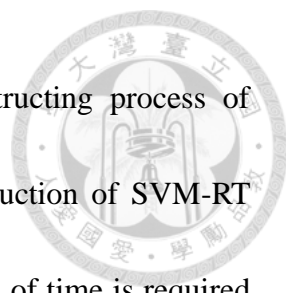
(a)
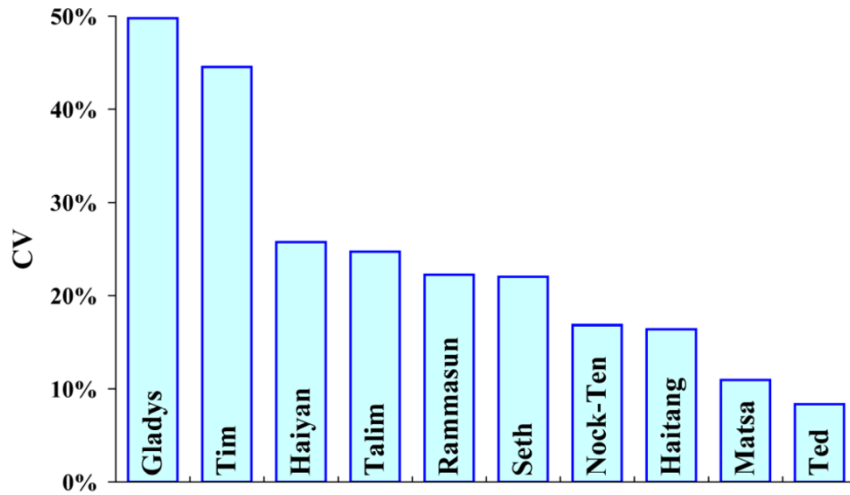
(b)

Figure 3.5   (a) CV values for each typhoon event resulting from BPN-RT trained

with 30 different sets of initial weights.   (b) RMSE values of BPN-RT trained with 30

different sets of initial weights and the constant CE value of SVM-RT (taking Typhoon

Gladys as an example)

3.2.3 The Comparison of Efficiency between SVM-based and BPN-based Models

Efficiency is an important issue for models, but the efficiency of hydrological models is only assessed in limited studies.   Lin and Chen (2005c, 2006) have examined the efficiency of BPN-based models and the results clearly show the time-consuming training process of BPNs.   In many conventional NN dominated fields, SVMs defeat BPNs because of much higher efficiency.   However, the high efficiency of SVMs has also received little attention in the hydrologic domain.   According to the previous study (Lin and Chen, 2009), SVMs are trained much more rapidly than BPNs.   In fact, the optimal architecture and weights of SVMs are rapidly "solved", not "searched". On the contrary, BPNs are trained by the error back-propagation algorithm which is a very time-consuming iterative process.   Based on the methodologies of SVMs and BPNs, it is obvious that the development of SVM-based models could be more efficient than that of BPN-based models.

To demonstrate the high efficiency of SVM-based models more clearly, SVM-RT and BPN-RT are taken for example.   For BPN-RT, the architectures with one to ten hidden neurons are tested to determine the optimal architecture (the most appropriate number of hidden neurons).   For each number of hidden neurons, 30 different sets of initial

weights are tried during the training process. The whole constructing process of BPN-RT requires 20000 seconds (about 5.5 hours), but the construction of SVM-RT needs only 2 seconds. As compared to BPN-RT, only about 0.01% of time is required for SVM-RT. Obviously, the development of SVM-based models is much more efficient than that of BPN-based models. In addition, as compared to BPN-based models, the SVM-based models could be more rapidly retrained with real-time data and are more suitable to be integrated with the decision support system for real-time rainfall forecasting or real-time reservoir operation.

3.2.4 The Improvement Due to the Addition of Typhoon Characteristics

To highlight the influence of typhoon characteristics on rainfall forecasting, we focus on the performance comparison between SVM-R and SVM-RT. As shown in Fig. 3.6(a), the CE values of both SVM-R and SVM-RT decrease with increasing forecast lead-time. However, the CE values of SVM-R decrease more rapidly than those of SVM-RT. The results show that SVM-RT forecasts rainfall depth more accurately than SVM-R, especially for long lead time forecasting. For 1- to 3-h lead time forecasts, the CE values of SVM-R and SVM-RT respectively decrease from 0.43 to 0.17 and from 0.44 to 0.24. Then, the performance of SVM-R gets worse and the CE values

decrease rapidly from 0.10 to -0.02 for 4- to 6-h lead time forecasts. It is clear that SVM-R cannot yield effective forecasts (i.e. CE<0) when the forecast lead-time is greater than four hours. As to SVM-RT, the performance does not get worse for long lead-time forecasting. For 4- to 6-h lead time forecasts, the CE values only decrease from 0.24 to 0.21. The addition of typhoon characteristics effectively decreases the negative impact of increasing forecast lead-time.

The improvement in CE due to the addition of typhoon characteristics is presented in Fig. 3.6(b). As shown in Fig. 3.6(b), the improvement increases with increasing forecast lead-time. For 1- to 3-h lead time forecasts, the improvement in CE only increases from 0% to 7%. The improvement due to the addition of typhoon characteristics is not very significant for one- to three-hour ahead forecasts. A much greater improvement in CE is obtained for the long lead-time forecasting. The improvement in CE increases from 15% to 23% for 4- to 6-h lead time forecasts. Obviously, the addition of typhoon characteristics significantly improves the long lead-time forecasting.

(a)

(b)
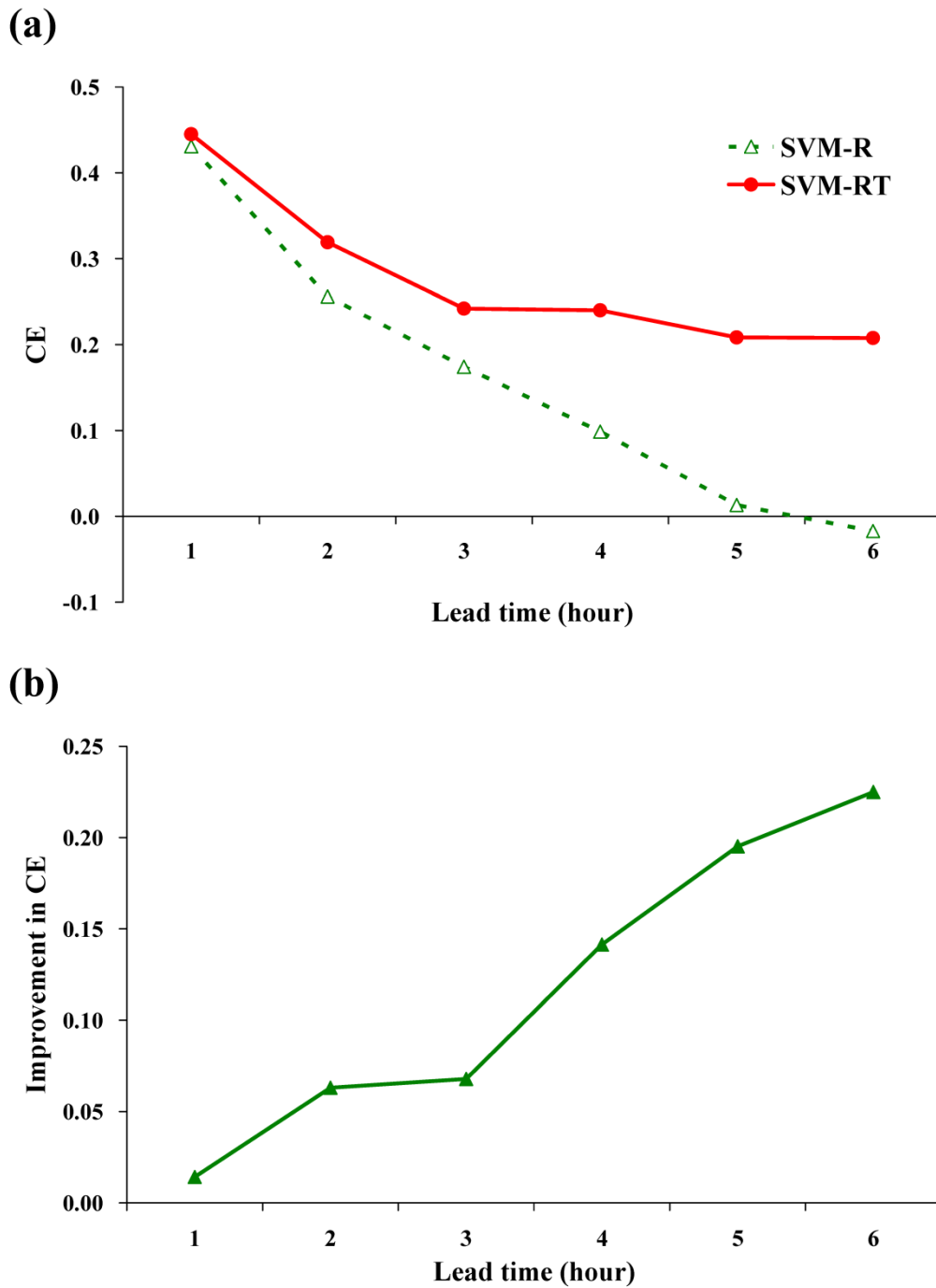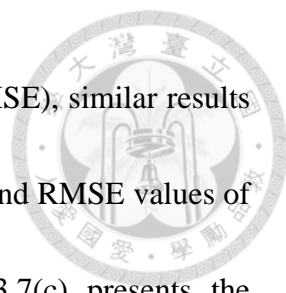
Figure 3.6　　(a) CE values of SVM-RT and SVM-R and (b) the improvement in CE

due to the addition of typhoon characteristics

According to the other two performance measures (MAE and RMSE), similar results are also obtained. Fig. 3.7(a) and 3.7(b) show that both the MAE and RMSE values of SVM-R increase more rapidly than those of SVM-RT. Fig. 3.7(c) presents the percentages of decrease in MAE and RMSE due to the addition of typhoon characteristics. For 1- to 3-h lead time forecasts, SVM-RT respectively decreases MAE and RMSE values from 2% to 5% and from 1% to 4% as compared to SVM-R. The percentages of decrease in MAE and RMSE values are not very significant. However, the addition of typhoon characteristics brings a greater decrease in MAE and RMSE values for the long lead time forecasting. As compared to SVM-R, SVM-RT respectively decreases MAE and RMSE values from 9% to 12% and from 8% to 10% for 4- and 6-h lead time forecasts. Again, the results confirm that typhoon characteristics effectively improve the long lead-time forecasting and decrease the negative impact of increasing forecast lead-time.
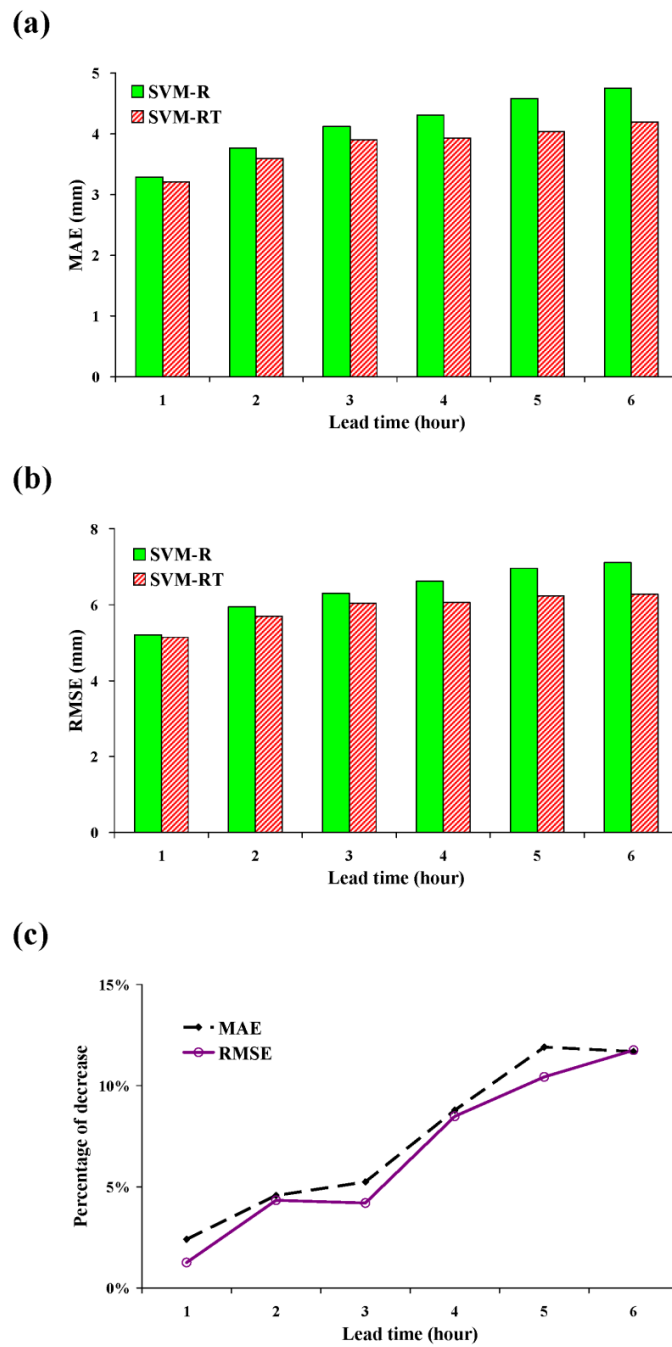
(a)

(b)

(c)

Figure 3.7 (a) MAE, (b) RMSE values of SVM-RT and SVM-R, and (c) the

percentages of decrease in MAE and RMSE due to the addition of typhoon

characteristics

To investigate the influence of each typhoon characteristic on rainfall forecasting, performance comparisons between models with single typhoon characteristic and models without typhoon characteristics (SVM-R) for 10 testing events are made. The number of events for which the model with single typhoon characteristic yields a higher CE value than SVM-R is presented in Fig. 3.8. Though only single typhoon characteristic is added to the model, improvement in performance could still be obtained for at least 40% of total events. The results show that each collected typhoon characteristics are effective for typhoon rainfall forecasting. Among all collected typhoon characteristics, the position of typhoon center is the most effective typhoon characteristic for improving rainfall forecasting. The addition of the position of typhoon center improves forecasting performance for 90% of total events. In addition, the distance between the center and the reservoir and the speed of the typhoon movement improves forecasting performance for 87% and 73% of total events, respectively. It is noted that the aforementioned three typhoon characteristics (the position of typhoon center, the distance between the center and the reservoir, and the speed of the typhoon movement) are highly related to the typhoon path. Thus, more accurate forecasts of typhoon path may be required to further improve the accuracy level of typhoon rainfall forecasting.
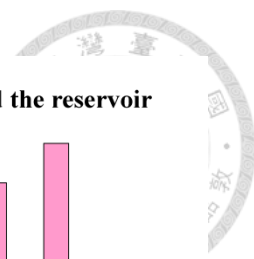
Position of the typhoon center
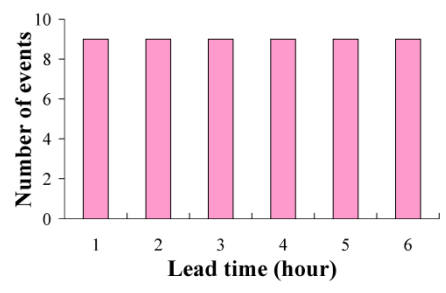
Distance between the center and the reservoir
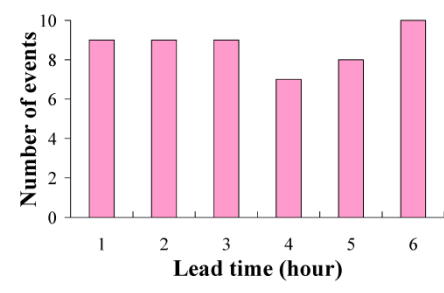
Maximum wind speed near the center
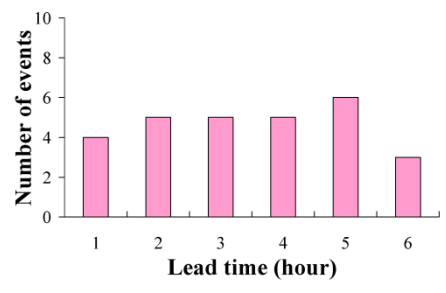
Atmospheric pressure of the center
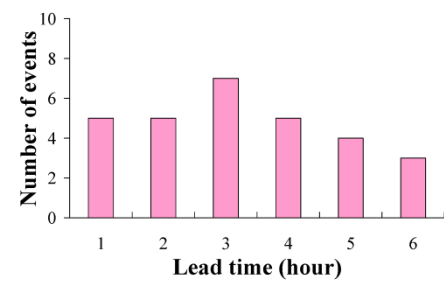
Radius of winds over 15 m/s
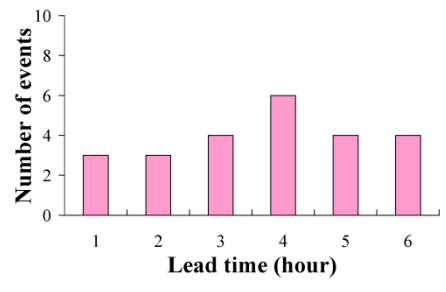
Speed of the typhoon movement
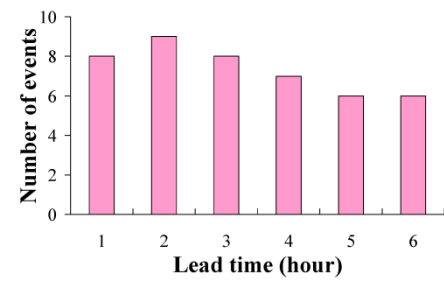
Figure 3.8   The number of events for which the model with each single typhoon

characteristic yields a higher CE value than SVM-R

3.3 Summary

To provide effective forecasts of hourly rainfall for supporting reservoir operation systems during typhoons, more accurate, robust and efficient models are proposed in this chapter. For this purpose, SVMs, instead of BPNs, are adopted to construct forecasting models. In addition to using SVMs instead of BPNs, typhoon characteristics are added to the proposed model to further improve the long lead-time forecasting. An application is conducted to demonstrate the superiority of the proposed models. Firstly, the forecasting performance of SVM-based models is compared with that of BPN-based models. The results clearly show that SVM-based models yield acceptably accurate forecasts for 1- to 6-h lead time forecasts, but BPN-based models produce effective 1- to 2-h lead time forecasts only. Hence SVM-based models perform much better than BPN-based models. In addition, the other two major advantages of SVMs over BPNs are the robustness and the efficiency, which are very important but have received little attention in the hydrologic domain. Representative examples given in this paper indicate that SVM-based models are more robust and the development of SVM-based models is much more efficient.

Finally, the comparison between the SVM-based models with and without typhoon characteristics is presented to confirm that the addition of typhoon characteristics

effectively decrease the negative impact of increasing forecast lead-time and significantly improves the forecasting performance, especially for long lead-time forecasting. In conclusion, the proposed SVM-based models are more accurate, robust and efficient than existing BPN-based models, and the typhoon characteristics should be used as input to the typhoon rainfall forecasting models for long lead-time forecasting. The proposed SVM-based models with typhoon characteristics are recommended as an alternative to the existing models. The proposed modeling technique is useful to improve the hourly typhoon rainfall forecasting and is suitable to be integrated with reservoir operation systems. The proposed modeling technique is also expected to be helpful to support flood, landslide, debris flow and other disaster warning systems.

Chapter 4 Typhoon flood forecasting using integrated SVM

4.1 Model development

4.1.1 Model construction

The architecture of the proposed two-stage SVM-based model (named SVM-QR$_f$ herein) is illustrated in Fig. 4.1a. In the first stage, the rainfall forecasting module, which is developed based on SVMs, is used to pre-process the typhoon information (namely, typhoon characteristics and rainfall) and to produce rainfall forecasts. Then in the second stage, the forecasted rainfall and the observed runoff are used as input to the flood forecasting module which is also developed based on SVMs. For comparison with the proposed model, another SVM-based flood forecasting model (named SVM-QRT) with observed runoff, rainfall and typhoon characteristics is also constructed. It should be noted that rainfall and typhoon characteristics are directly used as input to SVM-QRT without any processing (Lin et al., 2009a). The architecture of SVM-QRT is illustrated in Fig. 4.1b.

(a)

(b)
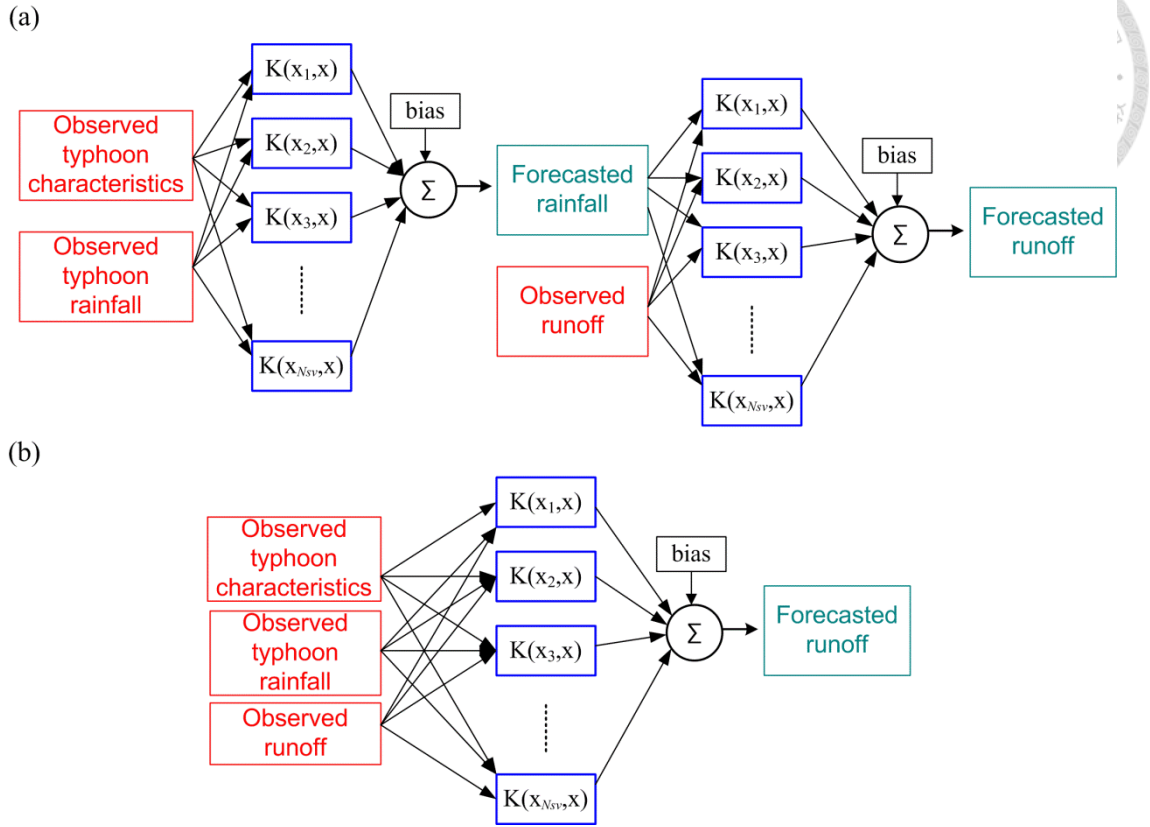
Figure 4.1   Architectural graphs of (a) the proposed model and (b) the existing

model

The construction of the proposed model, SVM-QR$_f$, is summarized below.   First, the

rainfall and typhoon characteristics are used as input to the rainfall forecasting module.

The general form of the rainfall forecasting module is

$$R_{t+\Delta t} = f(R_t, R_{t-1}, \cdots, R_{t-(L_R-1)}, TY_t, TY_{t-1}, \cdots, TY_{t-(L_{TY}-1)}) \tag{4.1}$$

where t is the current time, Δt is the lead-time period (from 1 to 6 h), R_t is

observed rainfall at time t, and L_R denotes the lag length of rainfall, TY_t is typhoon

characteristics at time t, L_{TY} denotes the lag length of typhoon characteristics, and

$R_{t+\Delta t}$ is the forecasted rainfall at time $t + \Delta t$.

Then, the forecasted rainfall ($R_{t+\Delta t}$) and observed runoff data are used as input to the

flood forecasting module in the second stage.   The general form of the proposed model

is

$$Q_{t+\Delta t} = f(Q_t, Q_{t-1}, \cdots, Q_{t-(L_Q-1)}, R_{t+\Delta t}) \tag{4.2}$$

where Q_t is observed runoff at time t, and L_Q denotes the lag length of runoff,

$Q_{t+\Delta t}$ is the forecasted runoff at time $t + \Delta t$.

As to the model without using forecasted rainfall (SVM-QRT), it has a general form

as

$$Q_{t+\Delta t} = f(Q_t, Q_{t-1}, \cdots, Q_{t-(L_Q-1)}, R_t, R_{t-1}, \cdots, R_{t-(L_R-1)}, TY_t, TY_{t-1}, \cdots, TY_{t-(L_{TY}-1)}) \tag{4.3}$$

The flowchart of the development of SVM-QR$_f$ and SVM-QRT is shown in Fig. 4.2.

In model construction, determination of the appropriate lag lengths of input is an

important step.   A trial-and-error procedure is applied to determine the lag lengths of

input.   The lag lengths, $L_{TY,}$ L_R and L_Q, are determined by the same process.   The

criterion for selecting the lag lengths is the relative percentage error (RPE):

$$RPE = \frac{E(L) - E(L+1)}{E(L)} \times 100 \tag{4.4}$$
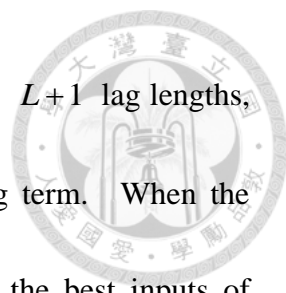
where $E(L)$ and $E(L+1)$ are the RMSEs for models with L and $L+1$ lag lengths, respectively. In general, the RMSE decreases with increasing lag term. When the RPE is less than 5%, the increase of lag lengths is stopped and the best inputs of forecasting models are selected.

```
┌─────────────────────────────┐
│ Use only typhoon characteristics │
│          as input           │
└─────────────────────────────┘
              │
              ▼
┌─────────────────────────────┐
│ Determine the lag length of typhoon │
│   characteristics L_TY      │
└─────────────────────────────┘
              │
              ▼
┌─────────────────────────────┐
│  Add rainfall to the input  │
└─────────────────────────────┘
              │
              ▼
┌─────────────────────────────┐
│ Determine the lag length of rainfall L_R │
└─────────────────────────────┘
```
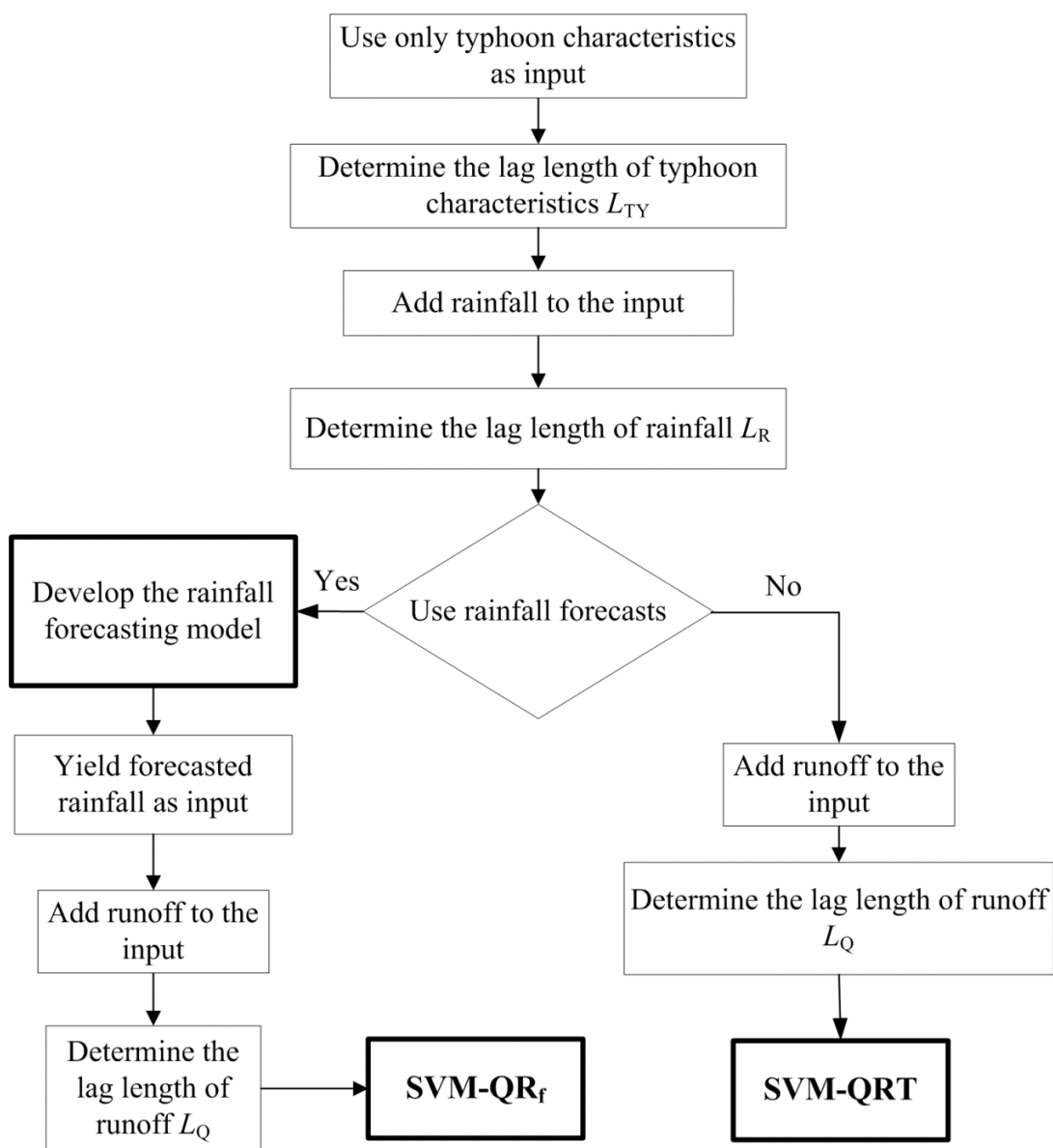
Figure 4.2   Flowchart of the model development

Decision: Use rainfall forecasts — Yes → Develop the rainfall forecasting model → Yield forecasted rainfall as input → Add runoff to the input → Determine the lag length of runoff L_Q → **SVM-QR$_f$**

No → Add runoff to the input → Determine the lag length of runoff L_Q → **SVM-QRT**
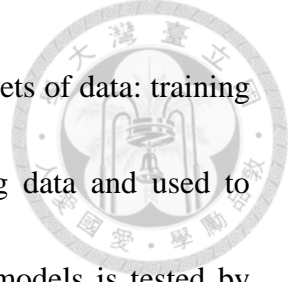
For event-based data, the collected events are separated into two sets of data: training and testing. Some of the collected events are chosen as training data and used to construct NN-based models. The performance of the NN-based models is tested by the remaining events. Different selections of training data and testing data yield different results and sometimes lead to different conclusions. In this study, we used cross validation and each single typhoon event (except the event with the maximum-runoff) is used to test the NN-based models in turn. Hence, for N typhoon events, a total of N-1 testing results are obtained. The conclusions are drawn on the basis of the overall performance for these testing results.

4.1.2 Performance measures

To evaluate the forecasting performance of models, three criteria are used. They are listed below.

1. Mean coefficient of efficiency (MCE)

For a single testing event, the coefficient of efficiency (CE) is written as

$$CE = 1 - \frac{\sum_{t=1}^{n}(Q_t - \hat{Q}_t)^2}{\sum_{t=1}^{n}(Q_t - \overline{Q})^2} \qquad (4.5)$$

where Q_t and $\hat{Q}_t$ denote the observed and forecasted runoff at time t, respectively,

$\overline{Q}$ is the average of the observed runoff, and n is the number of time steps. If the CE value is equal to one, the forecasts are perfect. Because the cross validations are used herein, the mean CE of N testing events is written as

$$\text{MCE} = \frac{1}{N} \sum_{j=1}^{N} \text{CE}_j \tag{4.6}$$

where CE_j is the CE for the jth testing event.

2. Root mean square error (RMSE)

The RMSE is a measure which represents the errors between two sets of data. The smaller the RMSE value, the better the forecasts. The RMSE is written as

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{t=1}^{n} (Q_t - \hat{Q}_t)^2} \tag{4.7}$$

3. Mean error of peak runoff (MEPR)

For a single testing event, the error of peak runoff (EPR) is defined as

$$\text{EPR} = \left| \frac{\hat{Q}_p - Q_p}{Q_p} \right| \tag{4.8}$$

where $\hat{Q}_p$ and Q_p is the forecasted peak runoff and the observed peak runoff respectively. The mean error of peak runoff of N testing events is written as

$$\text{MEPR} = \frac{1}{N} \sum_{j=1}^{N} \text{EPR}_j \tag{4.9}$$

where EPR_j is the error of peak runoff for the jth testing event.

4.2 Application, results and discussion

4.2.1 Application

Taiwan is located in one of the main paths of the north-western Pacific typhoons. During the past 100 years, on average, approximately four typhoons have hit Taiwan each year. To mitigate disasters due to typhoons, accurate and reliable runoff forecasts, especially for long lead time, are required to provide early warning of impending floods. The study area is the Wu River basin in central Taiwan. The elevation of this basin varies from 10 m to. 3500 m. The basin with an area of 2026 km^2 ranks fourth in Taiwan. The length of the main river is 119 km and the average slope is 1/92. In addition, the study area is abundant in rainfall during the rainy season (May to October). Heavy rainfalls brought by typhoons frequently caused flood disasters in the Wu River basin. In 2008, two typhoons (Kalmaegi and Fung-Wong) successively hit the central Taiwan. The direct economic loss caused by these two typhoons is estimated to be about 3 billion USD. Another important matter is that Taichung with a population of about 3 million people is located downstream of the Nan-Pei Bridge on the Wu River. Therefore, a well-performing and efficient flood forecasting model is needed.

Figure 4.3 shows the study area and the locations of four hourly rainfall stations (Pei-Shan, Chin-Liu, Hui-Suen and Tsui-Luan) and one hourly water-level station

(Nan-Pei Bridge). The rainfall and runoff data were obtained from the Water Resources Agency and the data of typhoon characteristics were collected from the Central Weather Bureau. In this paper, the effect of areal rainfall was considered. The areal rainfall of the watershed was calculated by the Thiessen method. The typhoon characteristics include the latitude and longitude (degree) of the typhoon center, the distance (km) between the typhoon center and the water-level station, the near-center maximum wind speed (m/s), the central pressure (hPa), the storm radius (km) and the speed (km/hr) of the typhoon movement. For this study, typhoon events that include the typhoon characteristics, rainfall and runoff data were listed in Table 4.1.
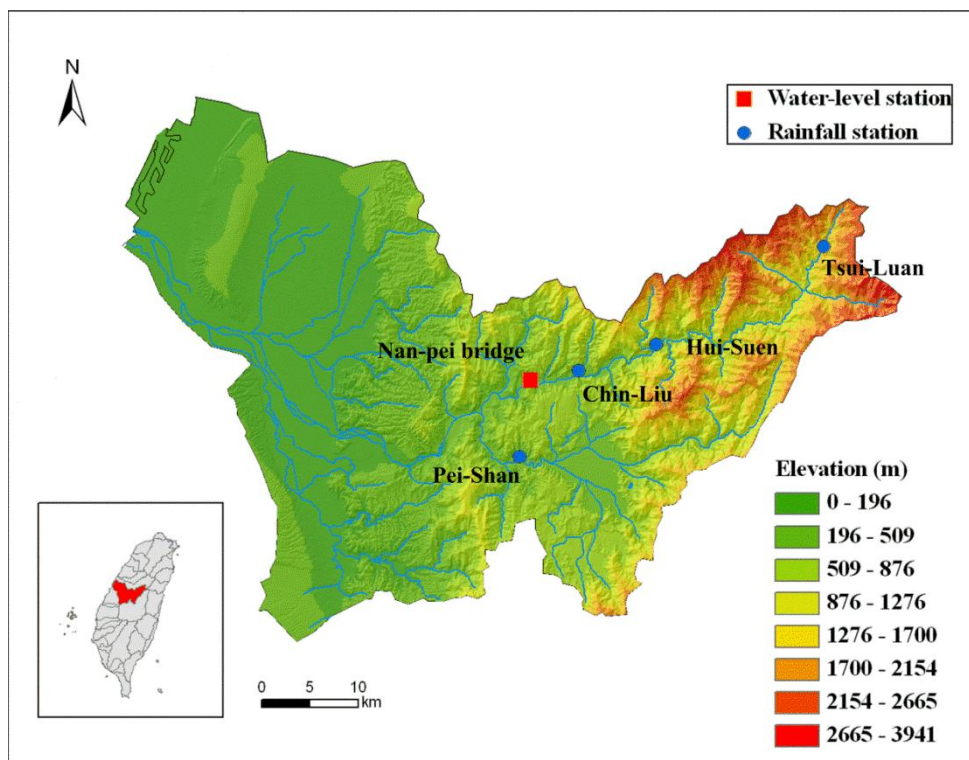

Figure 4.3　　The study area and locations of rainfall and water-level stations

Table 4.1 Description of typhoon events used in the modeling

Name	Date	Duration (h)	Scale	Maximum hourly rainfall (mm)	Peak runoff (m^3/s)
Tim	10 Jul 1994	50	Intense typhoon	11.48	87.8
Doug	8 Aug 1994	30	Intense typhoon	38.03	749
Toraji	29 Jul 2001	62	Moderate typhoon	85.07	1530
Nari	16 Sep 2001	111	Moderate typhoon	19.00	84.5
Nakri	9 Jul 2002	96	Minor typhoon	21.09	244
Nanmadol	3 Dec 2004	37	Moderate typhoon	13.21	61.6
Haitang	17 Jul 2005	81	Intense typhoon	38.94	916.6
Talim	31 Aug 2005	55	Intense typhoon	14.36	262.4
Longwang	1 Oct 2005	40	Intense typhoon	16.73	208.4
Sepat	17 Aug 2007	48	Intense typhoon	21.14	228.2
Wipha	17 Sep 2007	52	Moderate typhoon	20.02	183
Krosa	4 Oct 2007	79	Intense typhoon	27.01	267.5
Kalmaegi	16 Jul 2008	58	Moderate typhoon	67.56	370.2
Fung-wong	26 Jul 2008	73	Moderate typhoon	30.25	388.2
Sinlaku	11 Sep 2008	127	Intense typhoon	47.86	662
Jangmi	27 Sep 2008	52	Intense typhoon	29.44	696

Note: According to the classification system of the Taiwan Central Weather Bureau, the intensities of minor, moderate and intense typhoons are 34-63, 64-99, and $\geq$100 knot, respectively.
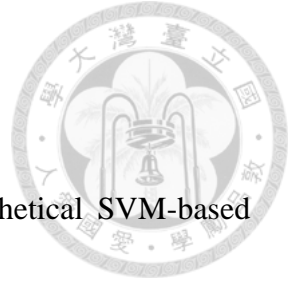
4.2.2 Results of rainfall forecasts

Before the rainfall forecasting module is constructed, a hypothetical SVM-based model (named SVM-QR$_i$) is first tested to show the influence of rainfall forecasts on flood forecasting.   It should be noted that the rainfall inputs to the SVM-QR$_i$ are the ideal values.   That is, the gauge measurements (ideal values) are used as perfect rainfall forecasts.   Table 4.2 provides the list of inputs that are used to construct SVM-based models.   The MCE values of SVM-QR$_i$ and the conventional model (SVM-QRT) are presented in Fig. 4.4.   The comparison between SVM-QRT and SVM-QR$_i$ shows the influence of ideal rainfall on flood forecasting.   Additionally, the result of a BPN-based model (named BPN-QRT), which uses the same inputs as SVM-QRT, is also presented in Fig. 4.4.   As shown in Fig. 4.4, SVM-QRi performs the best among all models.   Furthermore, SVM-QRT clearly yields higher MCE than BPN-QRT, which is consistent with the conclusion of Lin et al. (2009a, 2009b) that SVM mostly outperforms BPN.   However, both SVM-QRT and BPN-QRT cannot yield effective forecasts when the forecast lead time is greater than 3 h, whereas SVM-QR$_i$ give accurate flood forecasts up to 6 h.   The results confirm that if the perfect rainfall forecasts are available, the SVM-based model can effectively mitigate the negative impact of increasing forecast lead time.

In reality, perfect rainfall forecasts do not exist. So we need the rainfall forecasting module to obtain the future rainfall information. Lin et al. (2009b) confirmed that the addition of typhoon characteristics significantly improves the rainfall forecasting performance, especially for long lead time forecasting. Hence, data of rainfall and typhoon characteristics are used to develop a SVM-based rainfall forecasting module. The RMSE values resulting from the rainfall forecasting module are presented in Fig. 4.5. As shown in Fig. 4.5, the RMSE values increase from 4.3 mm to 6.3 mm for 1- to 3-h lead time forecasts, while they only increase slightly from 6.6 mm to 7 mm for 4- to 6-h lead time forecasts. We note that in this region, the maximum yearly rainfall is higher than 2000 mm and the maximum hourly rainfall is higher than 80 mm. The RMSE values of 1- to 6-h lead time forecasts are all lower than 8 mm, which indicates the SVM-based rainfall forecasting module can yield quite accurate rainfall forecasts.

Table 4.2 Input variables to the NN models

Lead time (h)	Input		
	SVM-QRT	SVM-QR$_f$	SVM-QR$_i$
1	$Q(t)$, $Q(t\text{-}1)$, $R(t)$, $Ty(t)$	$Q(t)$, $Q(t\text{-}1)$, $\hat{R}(t+1)$	$Q(t)$, $Q(t\text{-}1)$, $R(t+1)$
2	$Q(t)$, $Q(t\text{-}1)$, $R(t)$, $Ty(t)$	$Q(t)$, $Q(t\text{-}1)$, $\hat{R}(t+2)$	$Q(t)$, $Q(t\text{-}1)$, $R(t+2)$
3	$Q(t)$, $Q(t\text{-}1)$, $R(t)$, $Ty(t)$	$Q(t)$, $Q(t\text{-}1)$, $\hat{R}(t+3)$	$Q(t)$, $Q(t\text{-}1)$, $R(t+3)$
4	$Q(t)$, $Q(t\text{-}1)$, $R(t)$, $Ty(t)$	$Q(t)$, $Q(t\text{-}1)$, $\hat{R}(t+4)$	$Q(t)$, $Q(t\text{-}1)$, $R(t+4)$
5	$Q(t)$, $Q(t\text{-}1)$, $R(t)$, $Ty(t)$	$Q(t)$, $Q(t\text{-}1)$, $\hat{R}(t+5)$	$Q(t)$, $Q(t\text{-}1)$, $R(t+5)$
6	$Q(t)$, $Q(t\text{-}1)$, $R(t)$, $Ty(t)$	$Q(t)$, $Q(t\text{-}1)$, $\hat{R}(t+6)$	$Q(t)$, $Q(t\text{-}1)$, $R(t+6)$

Note: Q: observed runoff; R: observed rainfall; $\hat{R}$: forecasted rainfall; Ty: observed typhoon characteristics.
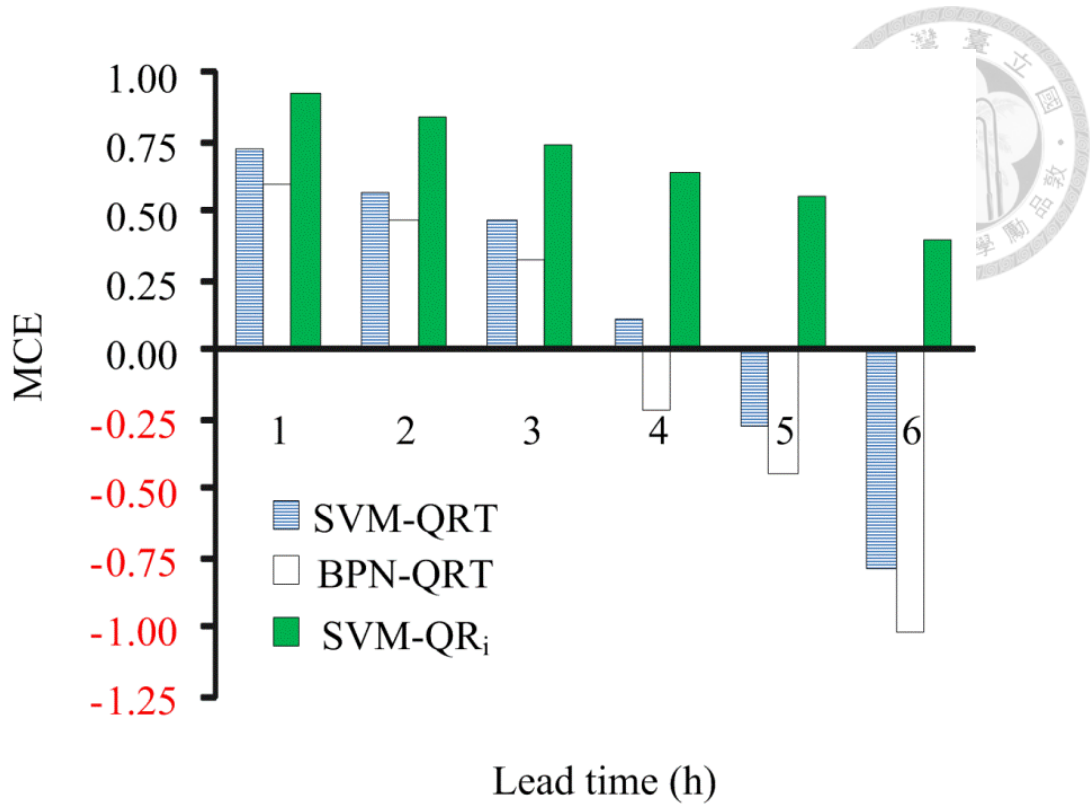
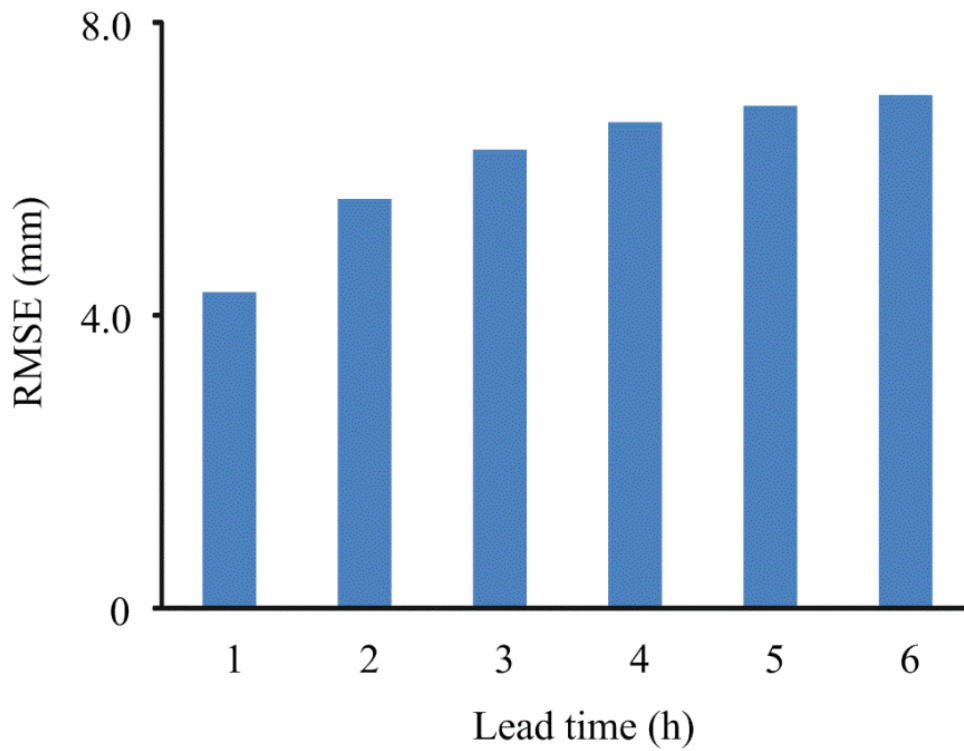Figure 4.4   MCE values of SVM-QRT, BPN-QRT and SVM-QR$_i$

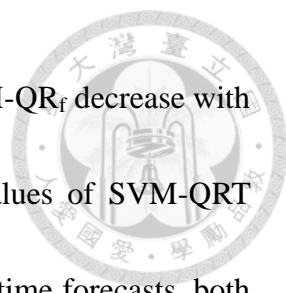Figure 4.5   RMSE values of the rainfall forecasts

4.2.3 Influence of forecasted rainfall on flood forecasting

The MCE and MEPR of three SVM-based models (SVM-QRT, SVM-QR$_f$ and SVM-QR$_i$) for 1- to 6-h lead time forecasts are summarized in Table 4.3. The input data of SVM-QR$_i$ are the perfect rainfall forecasts and antecedent runoff. However, SVM-QR$_i$ is a hypothetical model and cannot be used in practice. As for the input data of both SVM-QRT and SVM-QR$_f$, the antecedent runoff, rainfall and typhoon characteristics are used. However, in the SVM-QR$_f$ model, the rainfall forecasting module is used to pre-process typhoon information (namely, typhoon characteristics and rainfall) and to provide the forecasted rainfall. For SVM-QRT, the rainfall and typhoon characteristics are directly used as inputs without any processing. In this subsection we focus on the comparison between SVM-QR$_f$ and SVM-QRT to highlight the advantage of the proposed model.
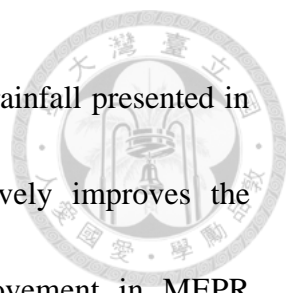
Table 4.3 MCE and MEPR for various models

Lead time (h)	SVM-QRT	SVM-QR$_f$	SVM-QR$_i$
MCE			
1	0.73	0.93	0.93
2	0.56	0.85	0.84
3	0.47	0.76	0.74
4	0.12	0.65	0.63
5	-0.28	0.58	0.55
6	-0.80	0.46	0.39
MEPR (%)			
1	7.28	4.42	3.74
2	13.23	8.16	4.06
3	16.41	11.49	5.21
4	22.28	13.58	6.73
5	28.12	14.28	7.92
6	32.51	15.10	11.78

The MCE values for runoff forecasts of both SVM-QRT and SVM-QR$_f$ decrease with increasing forecast lead time (Fig. 4.6a). However, the MCE values of SVM-QRT decrease more rapidly than those of SVM-QR$_f$. For 1- to 3-h lead time forecasts, both models provide reasonable runoff forecasts. For 4- to 6-h lead time forecasts, the performance of SVM-QRT gets worse and the MCE values are almost equal or even lower than zero. It is clear that SVM-QRT cannot yield effective forecasts when the forecast lead time is greater than 3 h. As to SVM-QR$_f$, the performance is still acceptable for long lead time forecasting. Regardless of the forecast lead time, the proposed model can provide more accurate runoff forecasts than the model without using forecasted rainfall. Furthermore, the improvement in MCE due to the use of SVM-QR$_f$ instead of SVM-QRT is presented in Fig. 4.6b. It is also concluded that SVM-QR$_f$ outperforms SVM-QRT.

As shown in Fig. 4.7a, a similar trend is observed that the MEPR values of both SVM-QRT and SVM-QR$_f$ increase with increasing forecast lead time. However, it is clear that SVM-QR$_f$ yields significantly lower MEPR values than SVM-QRT for 1- to 6-h lead time forecasts. Fig. 4.7a also shows that the MEPR values of SVM-QR$_f$ increase more slowly than those of SVM-QRT. For 1- to 6-h lead time forecasts, the MEPR values of SVM-QR$_f$ only increase from 4.4% to 15.1%, but those of SVM-QRT rapidly increase from 7.3% to 32.5%. The improvement in MEPR due to the use of

the proposed model instead of the model without using forecasted rainfall presented in Fig. 4.7b more clearly shows that the proposed model effectively improves the forecasting performance. For 1- to 6-h lead times, the improvement in MEPR increases from 2.9% to 17.4%. This indicates that SVM-QR$_f$ is more appropriate for forecasting peak runoff than SVM-QRT, especially for long lead time forecasting.

In this study, it is founded that the runoff forecasts cannot be improved by using raw typhoon characteristics as input to an SVM-based model. Because of the short concentration time in the study basin, the direct use of observed data (runoff, rainfall and typhoon characteristics) in model development cannot provide useful information for long lead time forecasting. When the forecast lead time increases, the data used for long lead time forecasting include more complex noise and the correlation between desired output and available input decreases rapidly. Because the rainfall forecasting module successfully reduces the complication of typhoon characteristics, the proposed model effectively improves the long lead-time forecasting.
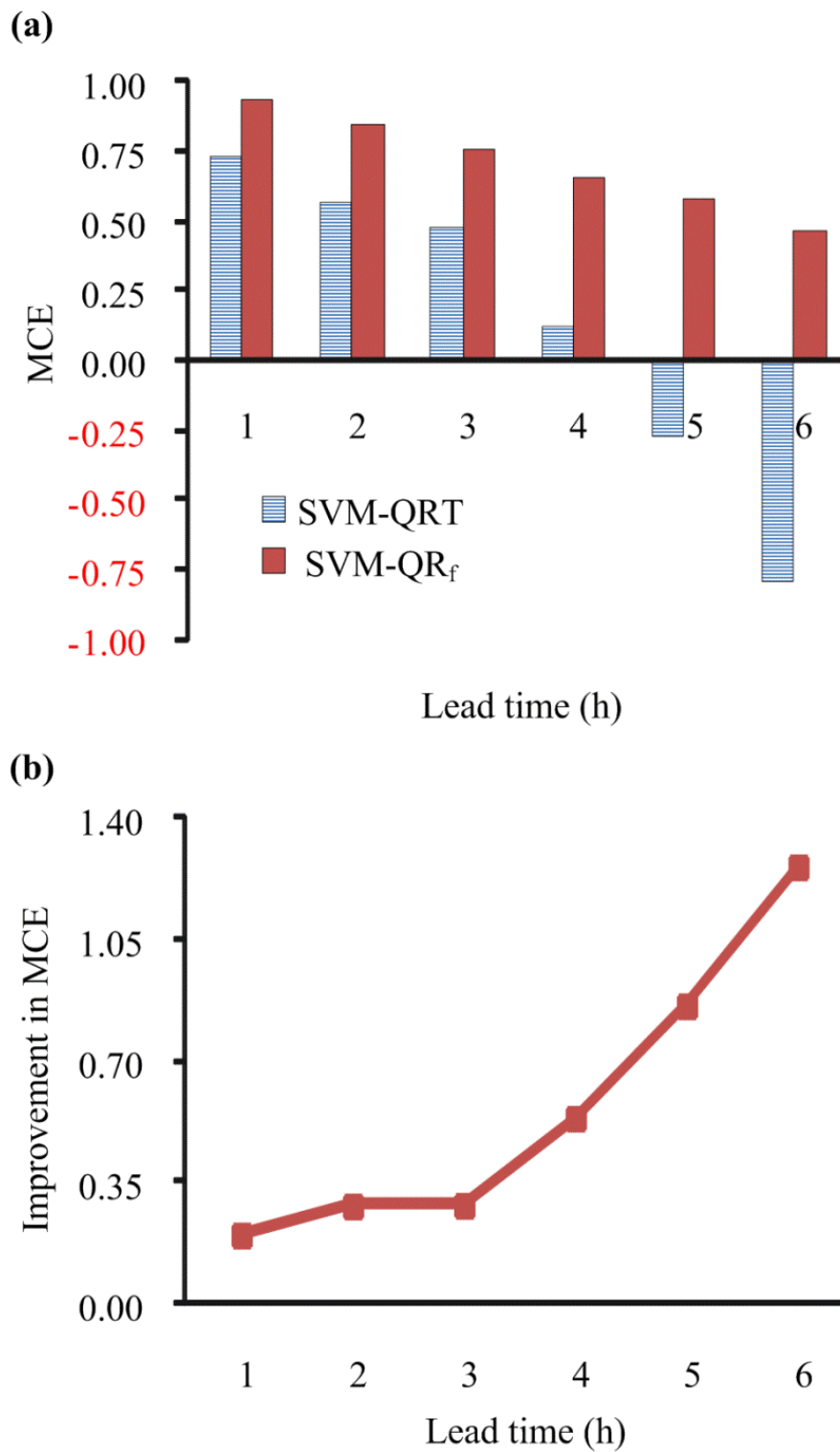
Figure 4.6   (a) MCE values of SVM-QRT and SVM-QR$_f$ and (b) the improvement in

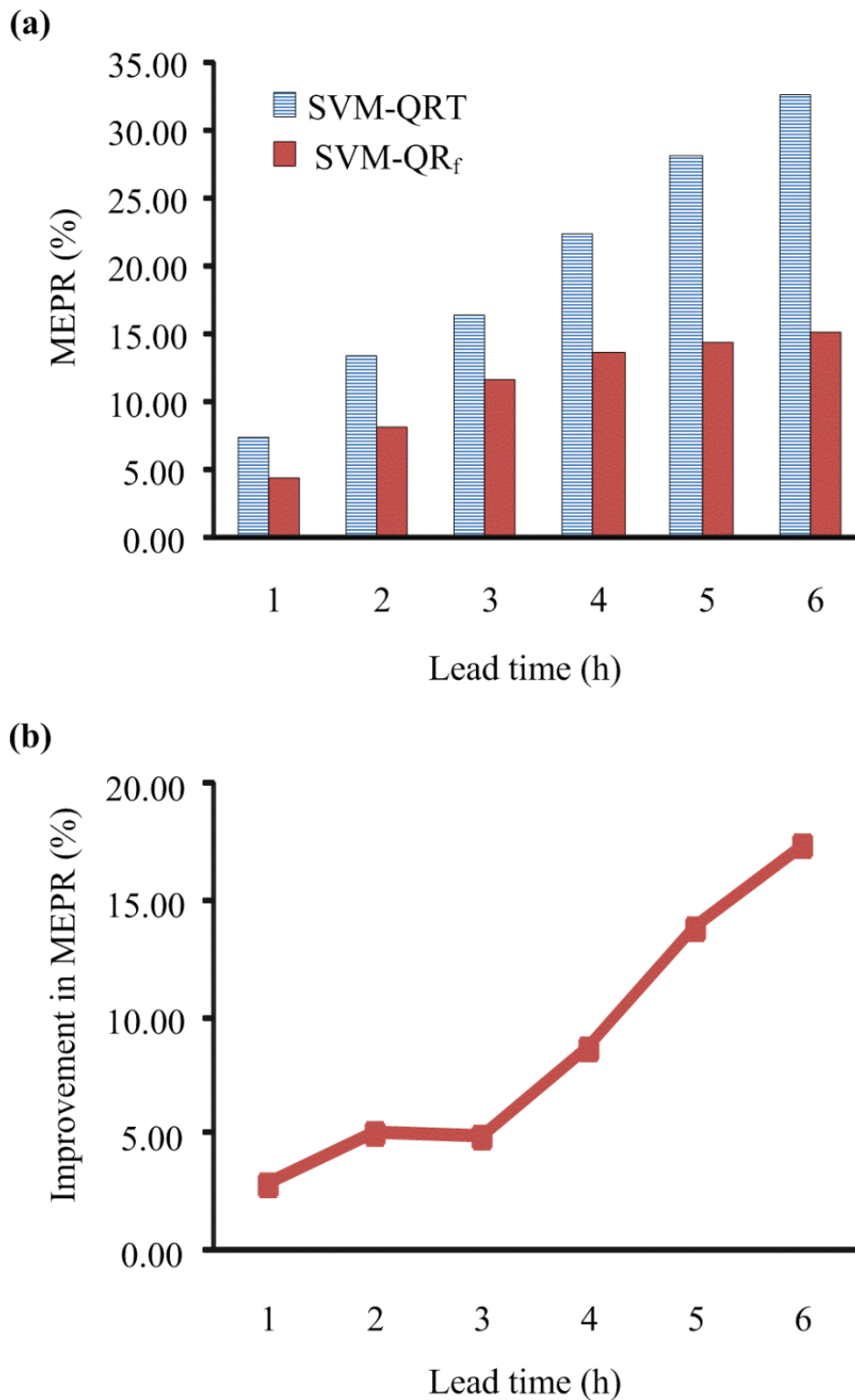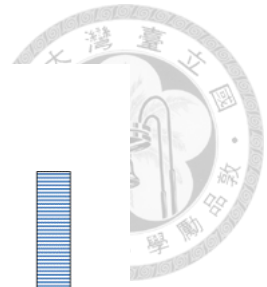MCE due to the use of SVM-QR$_f$ instead of SVM-QRT

Figure 4.7   (a) MEPR values of SVM-QRT and SVM-QR$_f$ and (b) the improvement

in MEPR due to the use of SVM-QR$_f$ instead of SVM-QRT

In addition to the overall performance, evaluation of individual events is described herein. The number of events for which SVM-QR$_f$ yields a higher CE than SVM-QRT is counted and presented in Fig. 4.8a. In a like manner, Fig. 4.8b presents the results for another performance measure, EPR. Fig. 4.8 shows that SVM-QR$_f$ performs better than SVM-QRT for most of the events. To further assess whether SVM-QR$_f$ performs better than SVM-QRT for the same testing events, paired comparison t-tests are conducted at the 1% significance level. Table 4.4 shows that SVM-QR$_f$ yields significantly higher CE and lower EPR than SVM-QRT. To highlight the comparison, Fig. 4.9 shows the hydrographs of 1-h lead time forecasts resulting from SVM-QR$_f$ and SVM-QRT for the most extreme runoff event (resulting from Typhoon Haitang). As shown in Fig. 4.9, both SVM-QR$_f$ and SVM-QRT slightly underestimate the peak runoff, but reproduce low runoff appropriately because low runoff is more frequent in data set than high runoff. However, SVM-QR$_f$ captures the peak runoff better than SVM-QRT. It is concluded that for peak runoff forecast, SVM-QR$_f$ is capable of providing more accurate forecasts as compared to SVM-QRT. For 1- to 6-h lead times, the comparison of the observed runoff with the forecasts resulting from SVM-QR$_f$ is presented in Fig. 4.10. The result shows that the proposed two-stage model (SVM-QR$_f$) is able to make good forecasts because the forecasted hydrograph accurately matches the observed hydrograph. This indicates that SVM-QR$_f$ exhibits excellent overall performances.

Table 4.4 Paired comparison *t*-tests of two performance measures (CE and EPR) resulting from SVM-QRT and SVM-QR$_f$.

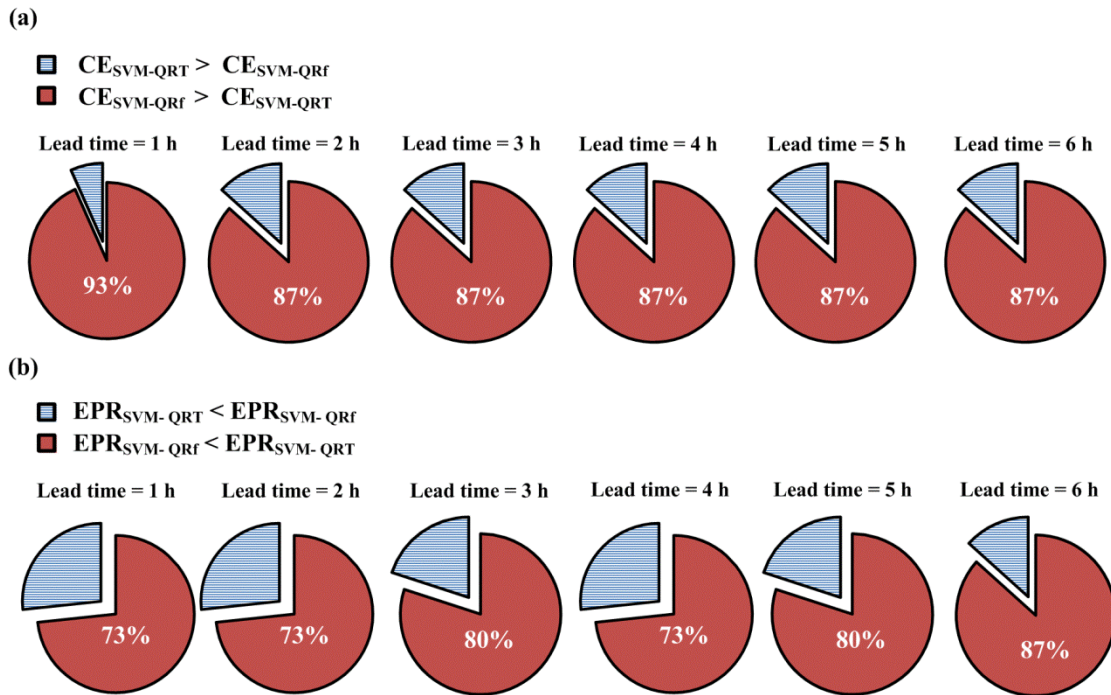| Alternative hypothesis	*t* Statistic	Critical *t* value	Statistically significant at the 1% level
CE$_{SVM-QRT}$ < CE$_{SVM-QRf}$	2.95	2.37	Yes
EPR$_{SVM-QRf}$ < EPR$_{SVM-QRT}$	4.19	2.37	Yes

Figure 4.8   Number of events for which (a) CE values of SVM-QR$_f$ are higher than those of SVM-QRT and (b) EPR values of SVM-QR$_f$ are lower than those of SVM-QRT
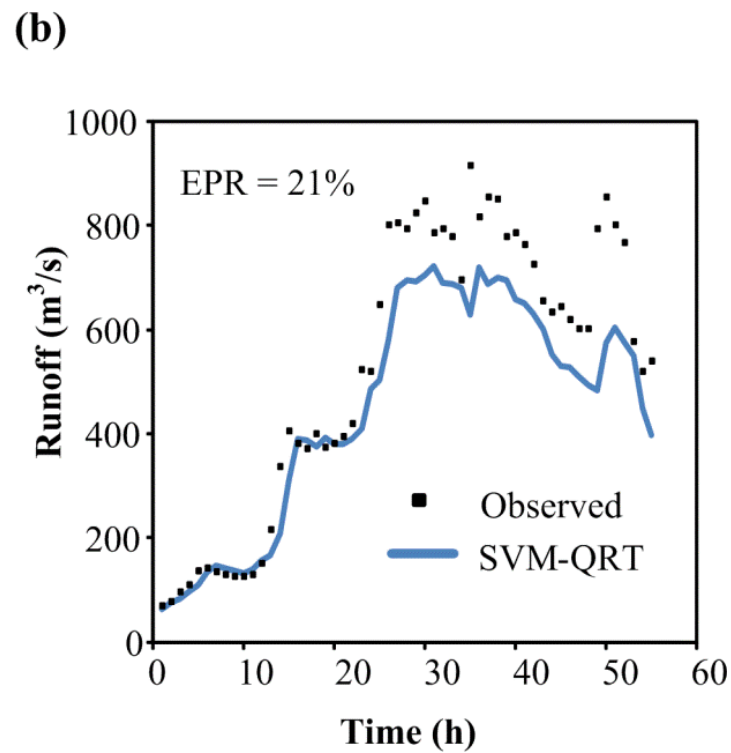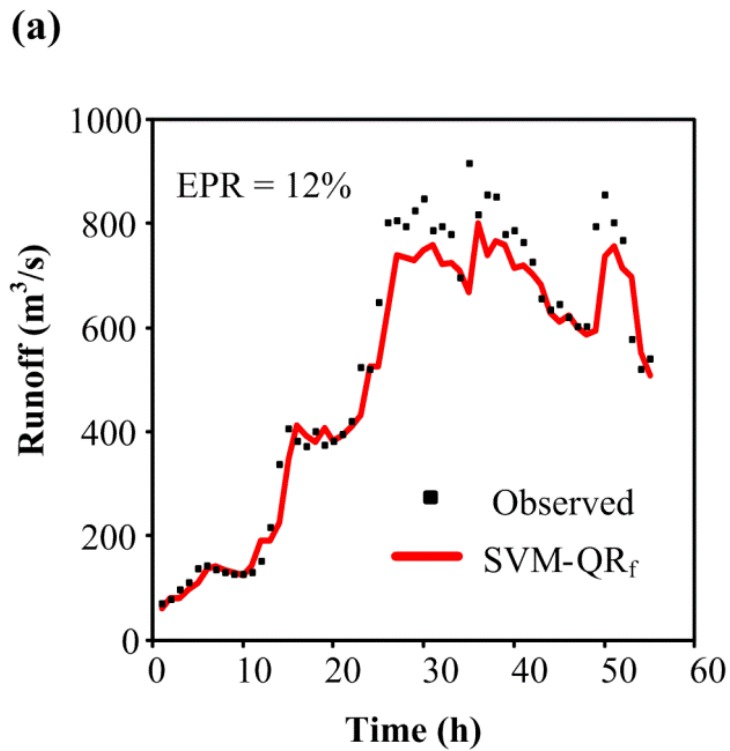
(a)

(b)

Figure 4.9   Comparison of the observed runoff with the 1-h lead time forecasts

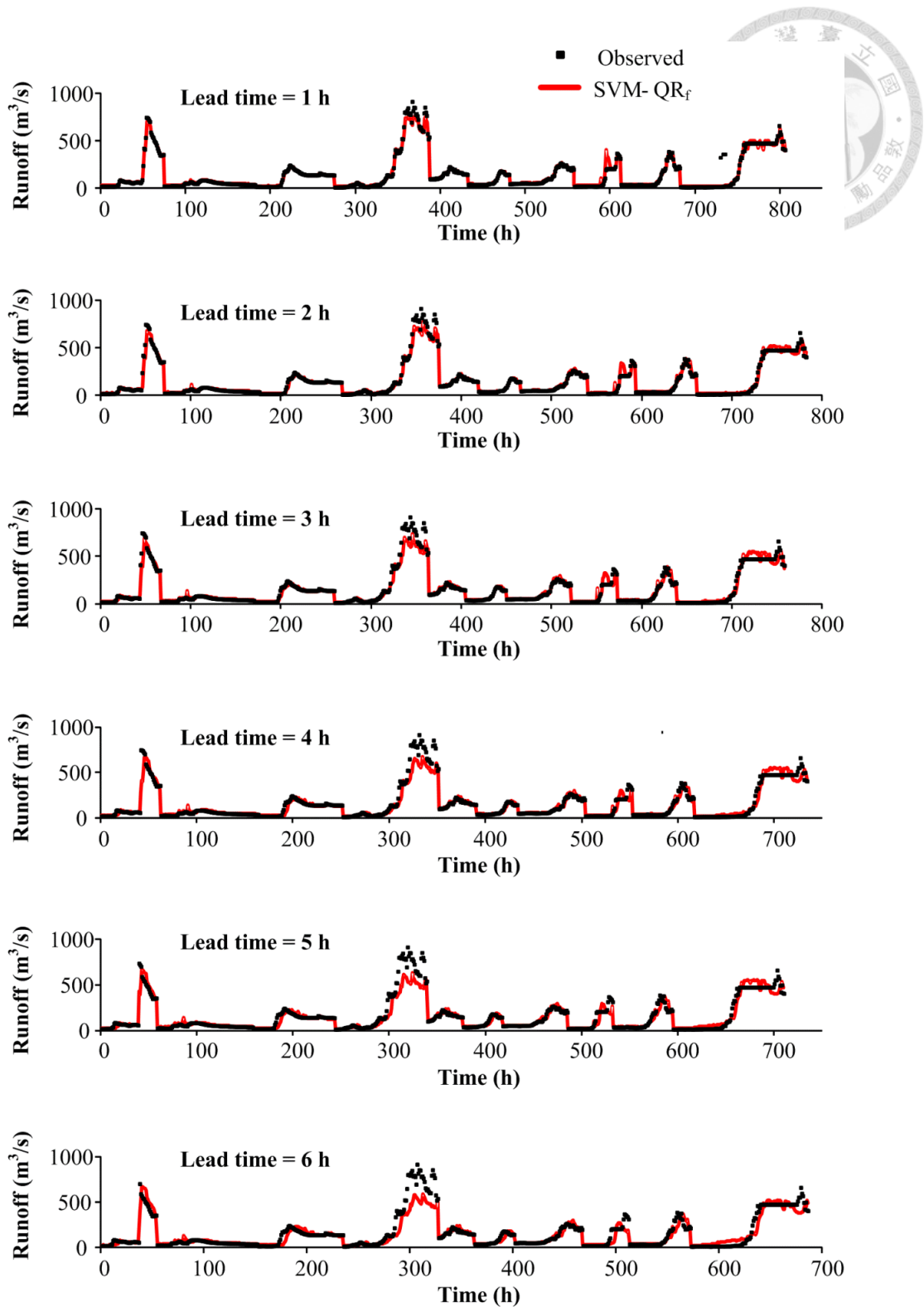resulting from (a) SVM-QR$_f$ and (b) SVM-QRT for Typhoon Haitang

Figure 4.10 Comparison of the observed runoff with the 1- to 6-h lead time forecasts

resulting from SVM-QR$_f$

4.3 Summary

In this chapter, a integrated SVM-based model is proposed for improving runoff forecast during typhoon events. In the first stage, the rainfall forecasting module is used to pre-process the typhoon information (namely, typhoon characteristics and rainfall) and to produce rainfall forecasts. Then, in the second stage, the forecasted rainfall and observed runoff are used as input to the flood forecasting module to yield runoff forecasts. A case study for the Wu River basin in central Taiwan is performed to assess the proposed model (i.e. SVM-QR$_f$). In addition, a single-stage SVM-based model (i.e. SVM-QRT), which directly uses observed runoff, rainfall and typhoon characteristics as input without any processing, is also constructed for comparison.

Regarding the performance of rainfall forecasting, it is found that the first-stage of the proposed model yields quite accurate 1- to 6-h lead time rainfall forecasts. The use of typhoons characteristics can effectively reduce the negative impacts of increasing forecast lead time. As to the performance of flood forecasting, a comparison between the proposed two-stage model and the single-stage model shows that the proposed model significantly improves the runoff forecasts. In addition to the overall performance, the proposed model significantly improves the forecasts of peak runoff, especially for long lead time forecasting. The better performance of the proposed

model confirms that the processed typhoon information is more useful than the raw typhoon information. The use of forecasted rainfall and the proposed two-stage structures are justified. In conclusion, the proposed model is expected to improve hourly typhoon flood forecasting.
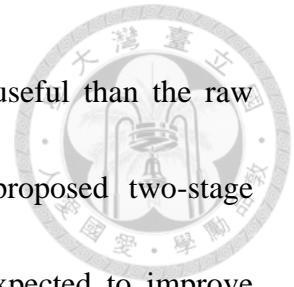
Chapter 5 Conclusions

In this dissertation, SVMs are investigated and applied on different hydrological problems, including typhoon rainfall forecasting and flood forecasting. These studies have indicated that SVMs are robust and efficient tools for flood warning system. Conclusions of these studies are summarized as follow:

5.1 Effective forecasting of hourly typhoon rainfall

1. In order to provide effective hourly rainfall forecasts, support vector machines (SVMs), instead of back propagation neural networks (BPNs), are presented to construct forecasting models.

2. An application in the Fei-Tsui Reservoir is conducted to demonstrate the three advantages of the proposed models.

3. As to forecasting performance, SVM-based models can yield accurately accurate forecasts for 1- to 6-h lead time forecasting, but BPN-based models produce effective 1- to 2-h lead time forecasting only. This result indicates that SVMs have better generalization ability than BPNs

4. The other two major differences between SVMs and BPNs are the robustness and the efficiency. The results of this study demonstrate that the SVMs are more robust and efficient than BPNs.

5. To further improve the long lead time forecasting, the comparison between the SVM-based models with and without typhoon characteristics is also presented in this study. The result confirms that the addition of typhoon characteristics effectively decrease the negative impact of increasing forecast lead time and significantly improves the forecasting performance, especially for long lead time forecasting.

6. It would be more interesting if the study compares the results of SVM with another model (e.g. Radial Basis Function Network and/or other Neural Networks) to assess the skill of SVM in a competing contest.

7. Research should continue on the selection of the effective typhoon characteristics to further improve the model performance.

5.2 Typhoon flood forecasting using integrated SVM

1. Based on the combination of the rainfall forecasting module and flood forecasting module, an integrated SVM-based model is proposed to improve long lead time flood forecasting.

2. An application in the Wu River basin is conducted to demonstrate the superiority of the proposed models.

3. According to the previous study, a SVM-based rainfall forecasting module with

rainfall and typhoon characteristics is proposed. In an area with abundant rainfall, the rainfall forecasting module still yields quite accurate 1- to 6-h lead time rainfall forecasts.

4. As to the performance of flood forecasting, it can be found that the proposed integrated model is capable of providing more accurate forecasts as compared to the single-stage model. The better performance of the proposed model confirms that the processed typhoon information is more useful than the raw typhoon information.

5. In addition to the overall performance, the hydrographs of 1-h lead time forecasts resulting from SVM-QR$_f$ and SVM-QRT for the most extreme runoff event (resulting from Typhoon Haitang) are also presented to show the performance of peak runoff forecasts. It is concluded that the proposed model significantly improves the forecasts of peak runoff, especially for long lead time forecasting.

6. Although the study uses the data of maximum-runoff Typhoon just in the training data set. It would be helpful to use that data in the testing data set either, to assess the ability of model in extrapolation as well.

7. The study shows that the two-stage SVM based approach improves the forecasting especially at long lead times. In the future work, it would be valuable to extend the

forecast to even longer lead times such as 9, 12 or even 24 hours as in most flood

forecasting applications.
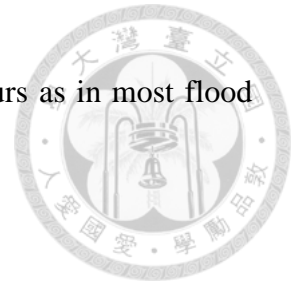
References
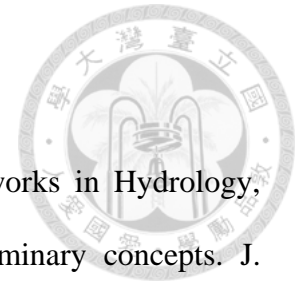
ASCE Task Committee on Application of Artificial Neural Networks in Hydrology, 2000a. Artificial neural networks in hydrology. I: preliminary concepts. J. Hydrol. Eng. 5 (2), 115-123.

ASCE Task Committee on Application of Artificial Neural Networks in Hydrology, 2000b. Artificial neural networks in hydrology. II: hydrologic applications. J. Hydrol. Eng. 5 (2), 124–137.

Asefa, T., Kemblowski, M., Lall, U., Urroz, G., 2005. Support vector machines for nonlinear state space reconstruction: Application to the Great Salt Lake time series. *Water Resour. Res.* 41, W12422, doi: 10.1029/2004WR003785.

Chang, F.J., Chang, K.Y., Chang, L.C., 2008. Counterpropagation fuzzy-neural network for city flood control system. *J. Hydrol.* 358 (1-2), 24-34.

Chang, F.J., Chiang, Y.M., Chang, L.C., 2007. Multi-step-ahead neural networks for flood forecasting. Hydrol. Sci. J.-J. Sci. Hydrol. 52 (1), 114–130.

Chang, F.J., Tseng, K.Y., Chaves, P., 2007. Shared near neighbours neural network model: a debris flow warning system. *Hydrol. Process.* 21 (14), 1968-1976.

Chang, L.C., Chang, F.J., Chiang, Y.M., 2004. A two-step-ahead recurrent neural network for stream-flow forecasting. Hydrol. Processes. 18 (1), 81–92.

Chaves, P., Kojiri, T., 2007. Stochastic fuzzy neural network: case study of optimal reservoir operation. *J. Water Resour. Plan. Manage.-ASCE*. 133, 509–518.

Chiang, Y.M., Chang F.C., 2009. Integrating hydrometeorological information for rainfall-runoff modelling by artificial neural networks. Hydrol. Processes. 23

(11), 1650–1659.

Chiang, Y.M., Chang, F.C., Jou, J.D.B., Lin, P.F., 2007. Dynamic ANN for precipitation estimation and forecasting from radar observations. J. Hydrol. 334 (1–2), 250–261.

Cristianini, N., Shaw-Taylor, J., 2000. An Introduction to Support Vector Machines and Other Kernel-Based Learning Methods. Cambridge University Press, New York.

de Vos, N.J., Rientjes, T.H.M., 2005. Constraints of artificial neural networks for rainfall–runoff modeling: trade-offs in hydrological state representation and model evaluation. Hydrol. Earth Syst. Sci. 9 (1-2), 111–126.

Hong, W.C., Pai, P.F., 2007. Potential assessment of the support vector regression technique in rainfall forecasting. *Water Resour. Manage.* 21, 495–513.

Hu, T.S., Wu, F.Y., Zhang, X., 2007. Rainfall–runoff modeling using principal component analysis and neural network. Hydrol. Res. 38 (3), 235–248.

Kalra, A., Ahmad, S., 2009. Using oceanic-atmospheric oscillations for long lead time forecasting. *Water Resour. Res. 45*, W03413, doi: 10.1029/2008WR006855.

Khalil, A., Mckee, M., Kemblowski, M., Asefa, T., 2005. Sparse Bayesian learning machine for real-time management of reservoir releases. *Water Resour. Res.*, *41*, W11401. doi: 10.1029/2004WR003891.

Lin, G.F., Chen, G.R., 2005a. Determination of aquifer parameters using radial basis function network approach. *J. Chin. Inst. Eng.* 28 (2), 241-249.

Lin, G.F., Chen, G.R., 2006. An improved neural network approach to the

determination of aquifer parameters. *J. Hydrol.* 316 (1-4), 281-289.

Lin, G.F., Chen, G.R., 2008. A systematic approach to the input determination for neural network rainfall-runoff models. *Hydrol. Process.* 22 (14), 2524–2530.

Lin, G.F., Chen, G.R., Huang, P.Y., Chou, Y.C., 2009a. Support vector machine-based models for hourly reservoir inflow forecasting during typhoon-warning periods. J. Hydrol. 372 (1–4), 17–29.

Lin, G.F., Chen, G.R., Wu, M.C., Chou, Y.C., 2009b. Effective forecasting of hourly typhoon rainfall using support vector machines. Water Resour. Res. 45, W08440. doi: 10.1029/2009WR007911.

Lin, G.F., Chen, L.H., 2004. A non-linear rainfall-runoff model using radial basis function network. J. Hydrol. 289 (1–4), 1–8.

Lin, G.F., Chen, L.H., 2005b. Application of artificial neural network to typhoon rainfall forecasting. Hydrol. Processes. 19 (9), 1825–1837.

Lin, G.F., Chen, L.H., 2005c. Time series forecasting by combining the radial basis function network and the self-organizing map. *Hydrol. Process.* 19 (10), 1925–1937.

Lin, G.F., Huang, P.Y., Chen, G.R., 2010. Using typhoon characteristics to improve the long lead-time flood forecasting of a small watershed. J. Hydrol. 380 (3–4), 450–459.

Lin, G.F., Wu, M.C., 2011. An RBF network with a two-step learning algorithm for developing a reservoir inflow forecasting model. J. Hydrol. 405 (3–4), 439–450.

Lin, G.F., Wu, M.C., Chen, G.R., Tsai, F.Y., 2009c. An RBF-based model with an
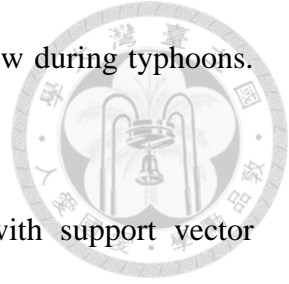
information processor for forecasting hourly reservoir inflow during typhoons. Hydrol. Processes. 23 (25), 3598–3609.

Liong, S.Y., Sivapragasam, C., 2002. Flood stage forecasting with support vector machines. J. Am. Water Resour. Assoc. 38 (1), 173–186.

Luk, K.C., Ball, J.E., Sharma, A., 2001. An application of artificial neural networks for rainfall forecasting. Math. Comput. Modell. 33, 683–693.

Maier, H.R., Dandy, G.C., 2000. Neural networks for the prediction and forecasting of water resources variables: a review of modeling issues and applications. Environ. Modell. Software. 15, 101-124.

Pramanik, N., Panda, R.K., Singh, A., 2011. Daily river flow forecasting using wavelet ANN hybrid models. J. Hydroinform. 13(1), 49–63.

Rathinasamy, M., Khosa, R., 2012. Multiscale nonlinear model for monthly streamflow forecasting: a wavelet-based approach. J. Hydroinform. 14(2), 424–442.

Rasouli, K., Hsieh, W.W., Cannon, A.J., 2012. Daily streamflow forecasting by machine learning methods with weather and climateinputs. J. Hydrol. 414–415, 284–293.

Sivapragasam, C., Liong, S.Y., 2005. Flow categorization model for improving forecasting. Nord. Hydrol. 36 (1), 37–48.

Supharatid, S., 2003. Tidal-level forecasting and filtering by neural network model. *Coast. Eng. J.* 45 (1), 119-137.

Tingsanchali, T., Gautam, M.R., 2000. Application of tank, NAM, ARMA and neural network models to flood forecasting. *Hydrol. Process.* 14 (14), 2473-2487.

Toth, E., Brath, A., 2007. Multistep ahead streamflow forecasting: Role of calibration data in conceptual and neural network modeling. Water Resour. Res., 43, W11405. doi: 10.1029/2006WR005383.
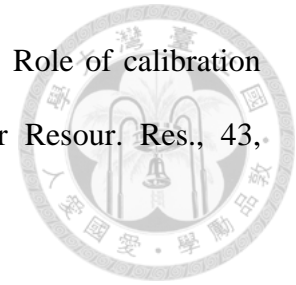
Tu, M.Y., Hsu, N.S., Tsai, F.T.C., Yeh, W.W.G., 2008. Optimization of hedging rules for reservoir operations. *J. Water Resour. Plan. Manage.-ASCE.* 134 (1), 3-13.

Vapnik, V., 1995. The Nature of Statistical Learning Theory. Springer, New York.

Vapnik, V., 1998. Statistical Learning Theory. John Wiley, New York.

Wu, C.L., Chau, K.W., 2011. Rainfall–runoff modeling using artificial neural network coupled with singular spectrum analysis. J. Hydrol. 399 (3–4), 394–409.

Wu, C.L., Chau, K.W., Li, Y.S., 2009. Predicting monthly streamflow using data-driven models coupled with data-preprocessing techniques. Water Resour. Res. 45, W08432. doi: 10.1029/2007WR006737.

Yu, P.S., Chen, S.T., Chang, I.F., 2006. Support vector regression for real-time flood stage forecasting. *J. Hydrol.* 328 (3-4), 704-716.

Yu, X.Y., Liong, S.Y., 2007. Forecasting of hydrologic time series with ridge regression in feature space. J. Hydrol. 332 (3-4), 290–302.

Yu, X.Y., Liong, S.Y., Babovic, V., 2004. EC-SVM approach for real-time hydrologic forecasting. J. Hydroinform. 6 (3), 209–223.
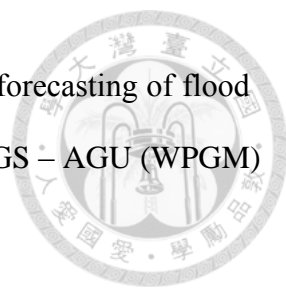
Publications

A. 期刊論文

1. G.F. Lin, G.R. Chen, P.Y. Huang, <u>Y.C. Chou</u>, 2009, "Support Vector Machine-based Models for Hourly Reservoir Inflow Forecasting during Typhoon-Warning Periods," *Journal of Hydrology*, Vol. 372, Issues 1-4, pp. 17-29. doi:10.1016/j.jhydrol.2009.03.032. (**SCI**)

2. G.F. Lin, G.R. Chen, M.C. Wu, <u>Y.C. Chou</u>, 2009, "Effective Forecasting of Hourly Typhoon Rainfall Using Support Vector Machines," *Water Resources Research*, Vol. 45, Article Number W08440. doi:10.1029/2009WR007911. (**SCI**)

3. G.F. Lin, <u>Y.C. Chou</u>, M.C. Wu, 2013, "Typhoon Flood Forecasting Using Integrated Two-stage Support Vector Machine Approach," *Journal of Hydrology*, Vol. 486, pp. 334-342. doi:10.1016/j.jhydrol.2013.02.012. (**SCI**)

4. G.F. Lin, H.Y. Lin, <u>Y.C. Chou</u>, 2013, "Development of a Real-time Regional Inundation Forecasting Model for the Inundation Warning System," *Journal of Hydroinformatics*. doi:10.2166/hydro.2013.202. (published online) (**SCI**)

B. 研討會論文

1. 林國峰, <u>周揚敬</u>, 陳谷榕, 黃珮瑜, 2008, "應用支持向量機於颱風洪水預報之研究", 第十七屆水利工程研討會論文集, pp. D8-1~D8-7.

2. 鍾秉宸, 林國峰, <u>周揚敬</u>, 陳谷榕, 2009, "支持向量機於颱風降雨預報之研究", 第十八屆水利工程研討會論文集, pp. L49-L57. (榮獲學生論文競賽第三名)

3. 林國峰, <u>周揚敬</u>, 鍾秉宸, 陳谷榕, 2009, "以支援向量機為基礎的颱風降雨
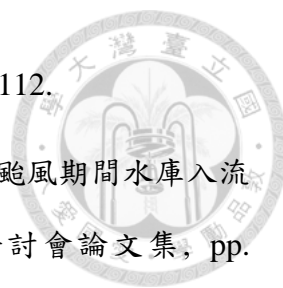
預報模式", 九十八年度農業工程研討會論文集, pp. 100-112.

4. 林國峰, <u>周揚敬</u>, 吳明璋, 陳谷榕, 2009, "支援向量機於颱風期間水庫入流量預報之研究", 第十三屆海峽兩岸水利科技交流研討會論文集, pp. B5-1~B5-9.

5. G.F. Lin, G.R. Chen, M.C. Wu, <u>Y.C. Chou</u>, B.C. Chung, 2010, "Application of Support Vector Machines for Effective Hourly Rainfall Forecasting during Typhoons," Proceedings of the 9th International Conference on Hydroinformatics, Tianjin, China, pp. 1778-1785.

6. 林國峰, <u>周揚敬</u>, 吳明璋, 黃珮瑜, 2010, "資訊處理器結合支援向量機模式於流量預報之研究", 九十九年度農業工程研討會論文集, , pp. 210-222.

7. 林國峰, 陳宜欣, <u>周揚敬</u>, 2010, "應用自組織映射線性輸出網路結合強化訓練法建置洪水預報模式", 第十九屆水利工程研討會論文集, pp. R10-R18. (榮獲學生論文競賽第二名)

8. 林國峰, 鄭家豪, 吳明璋, <u>周揚敬</u>, 2011, "應用新型人工智慧技術於增進雨量預估之研究", 第十五屆海峽兩岸水利科技交流研討會論文集, 第 A-22 論文.

9. G.F. Lin, J.M. Wang, M.C. Wu, <u>Y.C. Chou</u>, 2011, "A SOM-based clustering approach and its application to design hyetographs", Proceedings of the Second Civil Engineering Symposium of National Taiwan University and Tongji University, pp. 76-77.

10. 歐靚芸, 林國峰, 周揚敬, 2011, "以支援向量機為基礎的颱風降雨-淹水預報模式," 第二十屆水利工程研討會學生論文競賽手冊, pp 78-85. (榮獲學生論文競賽第一名)

11. G.F. Lin, Y.C. Chou, M.C. Wu, C.Y. Ou, 2012, "Effective forecasting of flood inundation depths using support vector machines", The AOGS – AGU (WPGM) Joint Assembly, Singapore.

12. Y.C. Chou, G.F. Lin, M.C. Wu, 2012, "Support Vector Machine with an Enforced Learning Algorithm for Flood Forecasting", Proceedings of the International Workshop on Typhoon and Flood, Taiwan.

13. G.F. Lin, Y.C. Chou, H.Y. Lin, M.J. Chang, 2013, "Typhoon Flood Forecasting Using a Data-driven Approach", AOGS 10th Annual Meeting, Australia.

簡歷

姓名：周揚敬 (Chou Yang-Ching)

生日：民國 73 年 4 月 26 日

學歷：

1. 國立武陵高級中學

2. 國立中興大學土木工程學系學士

3. 國立台灣大學土木工程學系研究所水利組碩士

4. 國立台灣大學土木工程學系研究所水利組博士