Applying Machine Learning Methods to the Airframe Structural Design Cost Estimation – A Case Study of Wing-Box Project

S. Deng

Department of Power Vehicle and Systems Engineering, Chung-Cheng Institute of Technology, National Defense University

sgdeng@ccit.edu.tw

Tsung-Han Yeh

Graduate School of Defense Science, Chung Cheng Institute of Technology, National Defense University

oasis.yeh@msa.hinet.net

Abstract.

This research used two machine leaning methods, the Support Vector Regression (SVR) and Back-Propagation Neural Network (BPN), to create the cost prediction models for airplane wing-box structural design, and verified the feasibility and efficiency for both methods. In the case study, four different main structural part groups of the wing-box, Spars/Ribs/Skins/Stringers, were chosen. In the parts data base, the part dimensions were included and used for classifying the part groups. Each part group has 150 bill of parts, 100 bill of parts used for training samples, 50 bill of parts used for predicting samples, to test there accuracy. After verified through wing-box case study, the results showed either SVR or BPN can precisely predicting the design costs. But compare to the BPN, SVR can get the global optimal solution while using less decision parameters. This can save lots of time for searching the best parameters combination when creating the prediction model.

Keywords: cost estimation, airframe structure, support vector regression, back-propagation neural network

本研究運用機器學習中的支援向量迴歸與倒傳遞類神經網路建立飛機機翼結構件設計成本預測模型，並透過案例驗證兩種方法之可行性以及預測績效。在案例探討部份，本研究針對機翼主要結構件Spars、Ribs、Skins以及Stringers等結構件進行探討。每種零件的資料庫內均含有尺寸，並用以為分類的資料。本研究使用各主要結構件共150筆資料，100筆資料作為訓練樣本；50筆作為測試樣本來測試其準確性。經由機翼結構的案例測試結果顯示，SVR與BPN兩種方法皆能準確預測結構件設計成本，而SVR之優勢在於求解過程為尋求全域最佳解，加上所需決定之參數較BPN少許多，因此在建立預測模型時可節省參數組合試誤之時間成本。

關鍵字：成本估測，飛機結構， 支援向量回歸，倒傳遞類神經網路

Introduction

成本預測的準確性對於專案規劃以及產品研發等領域都是極為重要的 (Bode, 1997)，航太

產業亦是如此。成本數據可分為直接成本與間接成本，直接成本係指直接用於產品製造的成本，包括原料、生產設備以及工資等；間接成本係指非直接關係產品製造之成本，如工廠照明費用、各部門營運成本等(Carr, 1989)。飛機製造業成本預測模型大多是由馬力、重量及尺寸等相關特徵參數所建立，也就是所謂的成本預測關係式(Cost Estimate Relation, CER)。然而傳統的CER成本預測模型係基於統計回歸模型所建立，在面對快速變動的社會因素無法即時更新預測模型(Smith & Mason, 1997)，藉此本研究將運用機器學習中的支援向量迴歸(SVR)以及倒傳遞類神經網路(BPN)兩種方法進行飛機結構設計成本預測，最終目的係提供飛機產業有效的成本預測方法。

Design to Cost. 航太產業在早期是以按產品性能設計(Design to Performance)為主軸，但長久下來發現花費龐大的費用，因此如何有效控管成本成為重要課題，專案成功與否除了績效時程以外成本也是重要的考量因素。現今專案在進行規劃預算的方法可分為從上而下(Top-Down)與從下而上(Bottom-Up)兩種方法，前者係由高階主管訂定預算再由上至下劃分至各階層，而後者係由各階層主管預估概略預算，再由下至上統整，但此兩種模式都係以經驗人為進行成本預測，往往因估測不夠準確而使專案最終難以執行成功。

欲在專案執行前準確規劃以及控制成本必須分析各單位運作實際所需之成本，彙整後訂定一預算門檻，專案必須在預算門檻內達到目標，因此有按成本設計(Design to Cost, DTC)的概念提出。根據美國國防部文件DoD5000.28對按成本設計(DTC)定義為一種管理概念，即是在發展過程中要確定嚴格的成本目標，並且通過在使用能力、性能、成本以及進度之間進行權衡，對系統的採購、使用及支援成本予以控制，已達到既定的目標。過去產業界是以產品規格去決定成本多寡，此生產概念因為沒有成本限制考量所以經常造成浪費。而近年來因應全球產業趨勢，都以成本為優先考量下進行生產作業，所以必須以按成本設計(Design to Cost)的管理概念執行。若以專案管理角度解釋DTC即是在專案編列的設計成本限制內完成所應達到績效之設計工作，必須按成本設計方能同時兼顧績效與成本。執行按成本設計成功案例可以美國為例，美國太空總署NASA的 Mars Pathfinder-Rover計畫執行時提出新的政策，即在有限控管的成本下完成所規劃的任務，推行此政策之效益使得NASA在1976年的Viking-Mars Lander計畫所耗費30億美元降為新計劃的1.75億美元，總共節省了94%成本。此案例顯示目前全球產業趨勢已經由按產品性能設計(Design to Performance) 轉變為以按成本設計 (Design to Cost) 為考量下，完成既定規劃任務(Mantel et al., 2005)。

Airframe LCC. 飛機全壽期成本大致可分為以下部分，包括概念設計、初步設計、細部設計、製造與裝配、運作與支援、除役等階段。飛機全壽期成本各階段成本比例指出，飛機全壽期成本中，概念設計階段以及初步設計階段至細部設計階段結束對總壽期成本之影響高達95%，雖然實際成本產生從生產製造階段才大幅提升，但欲有效控管整體飛機全壽期成本之效益必須從概念設計階段開始規劃，才能達到節省最大成本效果。在規劃飛機全壽期成本時，必須依據成本分解結構（Costs Breakdown Structure, CBS）來進行，以時間階段劃分飛機全壽期成本可分為研發費用、生產費用、使用維修以及保障費用，以飛機各次系統劃分可分為飛機機體、動力系統、武器系統、電子系統等費用，其中研發費用又可分為設計費用、材料費用、測試費用、設備費用等(Roskam, 1990)，本研究將對飛機機體研發階段中的飛機結構件設計成本進行探討。至今國外已陸續發展出適應於各種產業的成本預測軟體，較具代表性的包括：PRICE H、ACEIT、SEER H以及COCOMO II等套裝軟體，而上述各成本預測分析軟體之理論基礎由統計參數迴歸分析法為主軸(Parametric Estimation

Handbook, 2008)。

Cost Estimation Methods. 成本預測方法可分為定性法與定量法(Layer, 2002)，本研究在此僅針對定量法中的統計參數法進行介紹。統計參數預測法中最廣泛被使用的是統計學中的迴歸分析法，其理論是藉由實驗數據中影響反應值的獨立變數（$x_1,x_2,\ldots,x_i$）以及反應變數的誤差（ε）對所欲得到之反應值（Y）進行預測，透過最小平方法可配適一階迴歸預測模型(Parametric Estimation Handbook, 2008)。針對一般簡單的預測問題運用一階迴歸預測模型即可達到良好成效，但隨著預測問題複雜度增加，若只使用變數本身作為參數進行預測反應值可能得到較差的預測成效，意即預測模型參數之解釋能力較差，因此必須將變數平方項以及變數間交互作用納入考量以建立二階預測模式。統計回歸預測模型在面對參數固定且變動不大的情況下可達到良好的預測效果，但以飛機製造產業為例，其預測模型通常會受到通貨膨脹、技術成長、產量產能變化等社會環境影響而失去原有的準確性，因此必須經常修改迴歸分析參數才能確保預測精確度，而修改回歸預測模型造成時間以及成本的浪費，因此傳統統計回歸分析已無法因應社會上的變動因素。

Machine Learning Methods. 近年來，至許多學者運用機器學習 (Machine Learning) 方法來取代傳統的回歸分析法。機器學習主要概念係仿造人腦神經單元通過神經系統進行學習與記憶，亦可稱為類神經網路(Artificial Neural Network, ANN)，其中最具代表性的為倒傳遞類神經網路(Back-propagation Neural Network, BPN)。此網路概念是藉由訓練過程不斷的進行誤差遞回，調整權重重新運行，以達到誤差最小化，最終獲得一準確的預測模式(Haykin, 1999)。倒傳遞類神經網路已被廣泛運用在各領域的成本預測方面，如Greese與Li(1993)運用類神經網針對木造橋樑專案成本預測並與傳統參數預測法進行比較、Zhang與Fuh(1998)運用倒傳遞類神經網路於產品包裝離型系統開發所需之成本預估、Murat與Zeynep(2004)運用類神經網路於強化混凝土架構建築結構系統早期設計階段成本預估。

雖然類神經網路已被廣泛運用在成本預測問題並且有良好的成效，但類神經網路仍存在許多待改善之處，包括本身在進行網路建立時需要設定很多的網路參數，如傳遞函數、隱藏層層數以及神經元數等，欲得到準確的預測模式必須花費許多時間決定網路參數；類神經網路在進行運算時容易陷入區域最小值的問題，所以許多研究運用啟發式演算法與類神經網路結合以解決跳脫區域最小值的問題；類神經網路缺乏一套完整的數學理論為基礎，且預測模型較複雜。

在1995年，Vapnik與研究團隊於AT&T實驗室發展出一套新穎的機器學習方法，稱之為支援向量機（Support Vector Machine, SVM）。支援向量機的理論架構是以統計學習理論（Statistical Learning Theory, SLT）為基礎之VC維理論（Vapnik – Chervonenkis dimension）以及依循結構風險最小化原則（Structural Risk Minimization Principle, SRM）。此法則是以最小化預測誤差的上界為目標（結構化風險），而不同於類神經網路以最小化訓練誤差（經驗風險）。因此，支援向量機可運用經驗誤差以及VC維之信賴區間進行調整獲得最佳化的模型。支援向量機在訓練過程係解決線性，並且具有限制式的二次規劃問題(Quadratic Programming, QP)。而解決二次規劃問題所代表的涵義即是支援向量機求得的解為最佳唯一解，因此避免了陷入區域最小值的可能性。支援向量機最初設計是對分類問題進行探討，由統計學習理論發展出學習演算法，從簡易向量分類器(Simple Vector Classifiers)逐漸發展成為超平面分類器(Hyperplane Classifiers)，此分類器已被運用在臉型識別(Guo et al.,

2001)、文字分類(Zhang et al., 2008)、生物科技（蛋白質結構分類）(Cai et al., 2002)等分類問題，且皆有優異的成效。

在1997年，Vapnik等人導入ε-不敏感損失函數，選擇此損失函數係因能保持稀疏性(sparseness)，意即能用少量的支援向量來表示決策函數，之後支援向量機開始擴展運用於非線性預測問題，此預測模式稱為為支援向量迴歸（Support Vector Regression, SVR），許多學者已將SVR廣泛運用於各領域的預測問題，如Tay與Cao(2001)運用支援向量迴歸於時間序列財務預測問題，Hua等人(2007)運用支援向量迴歸整合邏輯迴歸進行企業財務困境預測、Chen 與 Wang(2007)運用支援向量迴歸結合基因演算法於旅遊產業需求預測以及Xi等人(2007)運用支援向量迴歸於空調廠房控制等研究。

本研究目的係運用機器學習方法構建準確的飛機結構件設計成本預測模式。本研究運用支援向量回歸(SVR)，以及倒傳遞類神經網路(BPN)進行預測模型績效比較。最終驗證使用機器學習中的SVR與BPN兩種方法，在飛機產業成本預測方面皆是可行的。本文架構在前言介紹成本的重要性、按成本設計(DTC)之概念、飛機全壽期成本架構、以及成本預測方法的演進；第二部份將介紹機器學習方法中的支援向量迴歸，與倒傳遞類神經網路之理論；第三部份為案例探討及結果分析；最後一部分為結論未來研究方向。

Methodology

Support Vector Regression. 支援向量迴歸 (Support Vector Regression, SVR) 主要是利用已知的訊息對未來未知的變數進行預測，主要是利用訓練資料建立預測模型，使測試資料與實際值達到誤差最小化的預測績效。(Vapnik, 1995)(Smola, 2002)假設給定一資料集S = (x_i,y_i) i=1…n，其中x_i為輸入向量、y_i為目標值、n為樣本數量，目標函數如(1)所示。基本概念係將資料x透過Φ非線性映射函數投影至高維度特徵空間進行線性迴歸。

$$f(x) = \Phi(x)\cdot w + b \quad (1)$$

式(1)中為w權重向量，b為門檻值，Φ為高維度特徵空間，意即輸入空間x的非線性映射，因此在高維度特徵空間進行線性迴歸等同於在低維度輸入空間進行非線性迴歸，可省略在高維度直接計算$\Phi(x)$以及w的內積值。權重向量w代表目標函數在高維度空間中的平坦程度，$\|w\|^2$為歐基里德長度，可用來衡量決策函數的平坦度亦可代表模型複雜程度，而在選定高維度特徵空間Φ之後，為使$f(x)$較為平坦，因此必須最小化$\|w\|^2$。而在建立迴歸方程式過程同時遵循結構風險最小化之下，並非每一筆訓練資料對建立迴歸方程式皆有正面效益，訓練資料中也存在著雜訊資料，而這些雜訊資料會影響支援向量迴歸最終預測結果，此時可結合損失函數以及導入懲罰係數C的觀念來處理。ε-不敏感損失函數主要功能是檢驗迴歸方程式與訓練資料之間的距離，若預測值與實際值間距離小於或等於ε，則損失函數為0；若大於ε，則損失函數不為0。

$$y - f(x) = \begin{cases} |y - f(x)| \le \varepsilon, 0 \\ |y - f(x)| > \varepsilon, otherwise \end{cases} \quad (2)$$

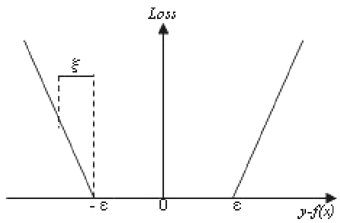

Figure 1 ε-insensitive loss function

支援向量回歸懲罰風險最小化目標式加入正的寬鬆變數(slack variable)，以ξ表示，同時在將鬆弛變數納入考量時必須給予一懲罰係數作為將雜訊資料納入迴歸函數的條件後可如式(3)所示。

$$\text{Minimize：} \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{n}\left(\xi_i + \xi_i^*\right) \tag{3}$$

subjected to: $y_i - \Phi(x_i)\cdot w - b \le \varepsilon + \xi_i$

$$\Phi(x_i)\cdot w + b - y_i \le \varepsilon + \xi_i^* \quad i = 1,\ldots,n \quad \xi_i, \xi_i^* \ge 0$$

藉由 Lagrange Multiplier 最佳化方法來處理，同時將目標式與限制式納入考量(Bertsekas, 1995)，式(3)可改寫成式(4)，Lagrange Function 如下：

$$\begin{aligned} L &= \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{n}\left(\xi_i + \xi_i^*\right) - \sum_{i=1}^{n}\left(\eta_i\xi_i + \eta_i^*\xi_i^*\right) \\ &- \sum_{i=1}^{n}\alpha_i\left(\varepsilon + \xi_i - y_i + \Phi(x_i)\cdot w + b\right) \\ &- \sum_{i=1}^{n}\alpha_i^*\left(\varepsilon + \xi_i^* - y_i - \Phi(x_i)\cdot w - b\right) \end{aligned} \tag{4}$$

式(4)中的 $\eta_i, \eta_i^*, \alpha_i, \alpha_i^*$ ≧0，這些稱為 Lagrange 係數，在鞍點(saddle point)上會產生極值，w, b, ξ_i, ξ_i^*作偏微分皆為 0，因此可得式(5)

$$\begin{aligned} &\partial_w L = w - \sum_{i=1}^{n}\left(\alpha_i - \alpha_i^*\right)x_i = 0 \\ &\partial_b L = \sum_{i=1}^{n}\left(\alpha_i - \alpha_i^*\right) = 0 \\ &\partial_{\xi_i} L = C - \alpha_i - \eta_i = 0 \\ &\partial_{\xi_i^*} L = C - \alpha_i^* - \eta_i^* = 0 \end{aligned} \tag{5}$$

將式(5)重整後代入式(4)中，並使用 Lagrange 對偶最佳化問題來解決，以 L_D 表示，而在原始問題 L 求最小值轉換至對偶問題 L_D 則轉換成求解最大值，如式(6)所示。

$$\text{Maximize：} L_D = -\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\left(\alpha_i - \alpha_i^*\right)\left(\alpha_j - \alpha_j^*\right)\Phi(x_i)\Phi(x_j) - \varepsilon\sum_{i=1}^{n}\left(\alpha_i + \alpha_i^*\right) + \sum_{i=1}^{n}y_i\left(\alpha_i - \alpha_i^*\right)$$

$$\text{subjected to: } \sum_{i=1}^{n}\left(\alpha_i - \alpha_i^*\right) = 0 \text{ and } \alpha_i, \alpha_i^* \in [0, C] \tag{6}$$

在此式(6)中 $\Phi(x_i)\Phi(x_j)$可由核函數 $K(x_i,x_j)$代替根據 Karush Kuhn-Tucher(KKT)理論並進行求解二次規劃問題(Bertsekas, 1995)，由式(5)可得最佳化 w，如式(7)所示，將式(7)代入式

(1)可得式(8)。

$$w^* = \sum_{i=1}^{n} \left(\alpha_i - \alpha_i^*\right) \Phi(x_i) \qquad (7)$$

$$f\left(x, \alpha_i, \alpha_j\right) = \sum_{i=1}^{n} \left(\alpha_i - \alpha_i^*\right) K\left(x \cdot x_i\right) + b \qquad (8)$$

最後可依據互補鬆弛條件(complementary slackness conditions)求解最佳 b 值，如式(9)所示。

$$\begin{aligned} &\alpha_i\left(\varepsilon + \xi_i - y_i + \Phi(x_i) w + b\right) = 0 \\ &\alpha_i^*\left(\varepsilon + \xi_i^* - y_i - \Phi(x_i) w - b\right) = 0 \\ &\left(C - \alpha_i\right) \xi_i = 0 \\ &\left(C - \alpha_i^*\right) \xi_i^* = 0 \end{aligned} \qquad (9)$$

Kernel Functions. 常用的核函數包括線性核函數(Linear Kernel Function) 式(10) 、多項式核函數（Polynomial Kernel Function）式(11)、以及徑向基核函數 (Radial Basis Kernel Function, RBF) 式(12)等三種(Steve Gunn, 1998)，其公式如下列所示。

$$k\left(x_i, x_j\right) = x_i \cdot x_j^T \qquad (10)$$

$$k\left(x_i, x_j\right) = \left(1 + x_i \cdot x_j\right)^d \qquad (11)$$

$$k\left(x_i, x_j\right) = \exp\left(-\gamma \left\|x_i - x_j\right\|^2\right) \qquad (12)$$

每個核函數皆有屬於本身的核參數，在建立預測模型時必須設定。核參數調整關係到最終預測模式的準確度，因此在運用支援向量迴歸時，核參數選擇將是重要課題。本研究採用徑向基核函數（RBF），因目前支援向量迴歸相關研究中若探討的問題本身為非線性，則都選擇使用徑向基核函數。因為核函數本身屬於非線性，它可在原始空間轉換到高維度處理非線性問題。當原始資料屬性為非線性，其使用效果較佳。另外若對資料性質無任何認知時，選用徑向基核函數可得到較好的結果(Smola, 1998)。

SVR Procedure. 支援向量迴歸運作流程如(Figure 2)所示，每階段說明如下：

<u>Step1 資料蒐集與尺度化</u>：

此階段將欲輸入進行運作之原始資料進行初步整理以及尺度化(scaling)，尺度化主要係為了增加支援向量迴歸預測模型的準確性，同時將資料控制並轉換至特定範圍內，多數研究將所有資料轉換至(-1~1)或(0~1)之間。之後將資料分為訓練資料及測試資料，訓練資料用意在於建立支援向量迴歸預測模型，而測試資料用意在於對建立之預測模型進行驗證並可得到預測結果和績效。

<u>Step2 格子點算法搜尋最佳參數組合</u>：

支援向量迴歸建立預測模型時必須決定參數組合(C, γ, ε)，其中 C 為懲罰係數；γ 為 RBF 核參數；ε 為損失函數界限，選擇參數組合直接影響到支援向量迴歸預測模型的精確度，因此本研究使用格子點算法(Grid Search)搜尋最佳參數組合以提升預測模型精確度。將上階段劃分出的訓練資料進行格子點搜尋法，將可能之參數組合進行之支援向量迴歸運算並判定是否達到終止條件(本研究設定為 MSE 最小化)，若未達到則產生另一組參數重複運作，最終可得到一組最佳參數組合。

<u>Step3 建立支援向量迴歸模型</u>：

使用最佳參數組合投入支援向量迴歸進行建構預測模型，再將測試資料進行預測模型績效驗證，由績效評量準則可得知預測模型之準確度以及未來推廣能力。

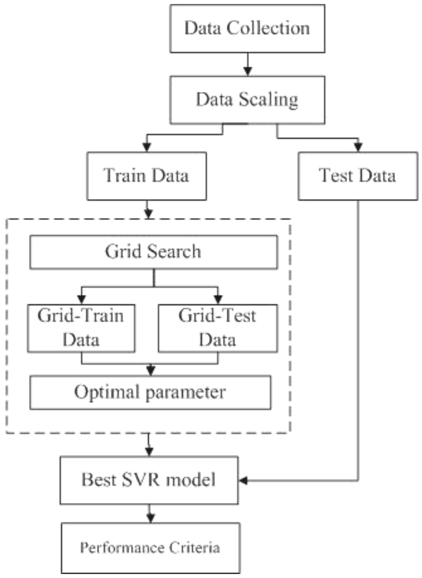

Figure 2 SVR procedure

Back-propagation Neural Network. 倒傳遞類神經網路(BPN)係屬於監督式學習多層前向式網路，主要架構包括輸入層(Input Layer)、隱藏層(Hidden Layer)以及輸出層(Output Layer)，架構圖如(Figure 3)所示。倒傳遞類神經網路運算過程可分為前向傳遞階段、誤差計算階段以及誤差遞迴階段，各階段內容如下：

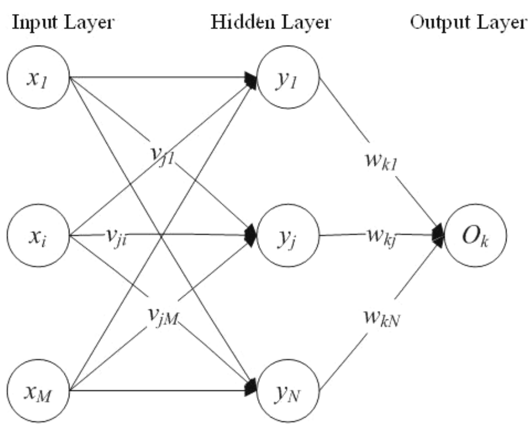

Figure 3 BPN framework

Step1 前向傳遞階段(forward pass)：式(13)將輸入層每個單元(x_i)透過權重(v_{ji})以及傳遞函數(f)計算至隱藏層單元(y_i)；式(14)將隱藏層每個單元(y_j)透過權重(w_{kj}) 以及活化函數(f)計算至輸出層單元(O_k)。

$$y_i = f(net_j = \sum_{i=1}^{M} v_{ji} x_i) \qquad (13)$$

$$O_k = f(net_k = \sum_{j=1}^{N} w_{kj} y_j) \qquad (14)$$

Step2 誤差計算階段(error computation)：定義誤差函數並計算網路輸出值與期望值之間的誤差，式(15)中d_k為期望輸出值；O_k為網路輸出預測值。

$$E = \frac{1}{2}\sum_{k=1}^{L}(d_k - O_k)^2 \qquad (15)$$

Step3 誤差遞迴階段(error back-propagation)：至此階段若誤差未滿足終止條件則利用梯度下降法(gradient descent)更新單元傳遞權重，η為梯度下降係數，式(16)係更新從隱藏層至輸出層之權重；式(17)係更新連結輸入層至隱藏層間的權重，重覆此過程達到中止條件，即網路輸出值與期望輸出間誤差最小化。

$$w_{kj} = w_{kj} + \Delta w_{kj} = w_{kj} - \eta \frac{\partial E}{\partial w_{kj}} \qquad (16)$$

$$v_{ji} = v_{ji} + \Delta v_{ji} = v_{ji} - \eta \frac{\partial E}{\partial v_{ji}} \qquad (17)$$

倒傳遞類神經網路(BPN)運用於預測模型建立時有許多參數必須給定，且不同於支援向量迴歸只需決定(C, γ, ε)參數組合，BPN必須決定輸入層(Input Layer)神經元數；隱藏層(Hidden Layer)層數以及神經元數；輸入層(Output Layer)神經元數；神經元傳遞函數；訓練函數以及學習函數等，由於所需決定參數過多，目前多數研究皆透過經驗試誤法決定可行之參數組合，也因此BPN的預測模型無法達到最佳化。

Performance Criteria. 評量準則係針對支援向量迴歸(SVR)以及倒傳遞類神經網路(BPN)演算過後得出之成本預測模式進行績效評量，主要概念係將預測值與實際值進行比對，以證實預測模式之可行性。本研究使用最小平方法(Mean Squared Error, MSE)(18)以及平均絕對誤差(Mean Absolute Percentage Error, MAPE)(19)統計評量值進行評估(Lewis, 1982)，MSE和MAPE即用來衡量真實質與預測值之間的差異程度，當MSE與MAPE值愈小時代表預測值愈接近真實值。以下列出此兩種統計評量值之計算公式其中為n樣本數；為y_i實際值；為y_i'預測值。

$$MSE = \frac{1}{n}\sum_{i-1}^{n}\left(y_i - y_i'\right)^2 \qquad (18)$$

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i - y_i'}{y_i}\right| \times 100\% \qquad (19)$$

Case Study

案例探討部份，假設欲規劃執行Wing-Box專案所需之結構設計成本，因此預測目標為結構件設計成本(Design Cost，單位：美元)，而影響預測目標的特徵參數為結構件尺寸(單位：

英吋)，包括結構件長度(Length)、寬度(Width)以及厚度(Thickness)。若以函數關係式來表示結構件設計成本與結構件尺寸之間的關係可如(20)所示。然而因過去結構件成本資料皆由經驗人為紀錄並無一預測模型，本研究藉由過去經驗資料透過機器學習方法建立一結構件設計成本預測模型。

Design Cost = f (Length, Width, Thickness)   (20)

飛機結構機體部分可分為，機身(Fuselage)、機翼(Wing)以及尾翼(Empennage)等部份，其中機翼中主要的組件為Wing-Box，其所需之結構件為Spar、Ribs、Stringers以及Skins等，透過蒐集過去經驗累積之各種Wing-Box結構件設計數據範圍(單位：英吋)如Table 1所示。

Table 1 過去經驗累積之結構件資料

Parts	Size	Range	Parts	Size	Range	Parts	Size	Range	Parts	Size	Range
Spar	L	15~50	Ribs	L	10~25	Skins	L	10~30	Stringers	L	30~50
	W	2~10		W	5~10		W	2~10		W	30~50
	T	2~10		T	2~5		T	2~10		T	0.1~1

匯整各種Wing-Box結構件設計數據範圍後，本研究使用機器學習方法中的支援向量迴歸(SVR)以及倒傳遞類神經網路(BPN)針對各種結構件進行設計時間預測模型構建，建模過程本研究使用過去累積之結構件資料150筆，並將總資料100筆作為訓練資料，50筆作為測試資料。

建模過程中必需決定SVR與BPN兩方法的參數。在SVR參數決定方面，核函數選用徑向基核函數(RBF)，最佳參數組合使用格子點搜尋算法運作，接著依據SVR流程進行運算，可得到各結構件設計成本預測模型訓練與測試階段績效(MSE, MAPE)以及最佳參數組合(C, γ, ε)。結構件Spar在SVR預測模型參數組合(C =1024, γ =0.25, ε =0.0078125)時，訓練階段績效MSE為0.025；測試階段積效MSE為0.11；MAPE為2.436%，結構件Ribs在SVR預測模型參數組合為(C =1024, γ =0.125, ε =0.0078125)時，訓練階段績效MSE為0.0002；測試階段績效MSE為0.064；MAPE為2.1064%，結構件Stringers在SVR預測模型參數組合為(C =1024, γ =0.125, ε =0.0078125)時，訓練階段績效MSE為0.0015；測試階段績效MSE為0.155；MAPE為2.078%，結構件Skins在SVR預測模型參數組合為(C =1024, γ =0.125, ε =0.0078125)時，訓練階段績效MSE為0.0007；測試階段績效MSE為0.05；MAPE為3.0146%，各結構件SVR運算結果綜整如Table 2所示。

Table 2 SVR 結果綜整表

Parts	Train Samples	kernel	C	γ	ε	MSE	Test Samples	MSE	MAPE(%)
Spar	100 samples per each part	RBF	1024	0.25	0.0078125	0.025	50 samples per each part	0.11	2.436
Ribs			1024	0.125	0.0078125	0.0002		0.064	2.1064
Stringers			1024	0.125	0.0078125	0.0015		0.155	2.078
Skins			1024	0.125	0.0078125	0.0007		0.05	3.0146

在BPN參數決定部分，本案例係屬於3個輸入值(結構件長度、寬度以及厚度)對應1個輸出值(設計時間)之預測問題，隱藏層層數設定為1層；隱藏層神經元依據經驗公式設定為7個神經元，傳遞函數使用雙曲正切與線性傳遞函數，網路訓練函數使用收斂速度較快的

Levenberg-Marquardt訓練函數，訓練迭代次數設定為1000次；訓練終止條件設定為0.1，當訓練誤差收斂低於0.1即停止訊練。

接著進行 BPN 運算，可得到各結構件設計成本預測模型訓練與測試階段績效(MSE, MAPE)。結構件 Spar 在 BPN 預測模型訓練階段績效 MSE 為 0.095；測試階段積效 MSE 為 0.1156；MAPE 為 2.4803%，結構件 Ribs 在 BPN 預測模型訓練階段績效 MSE 為 0.085；測試階段積效 MSE 為 0.077；MAPE 為 2.3917%，結構件 Stringers 在 BPN 預測模型訓練階段績效 MSE 為 0.098；測試階段積效 MSE 為 0.1696；MAPE 為 3.1631%，結構件 Skins 在 BPN 預測模型訓練階段績效 MSE 為 0.099；測試階段積效 MSE 為 0.073；MAPE 為 2.08%，各結構件 BPN 運算結果綜整如 Table 3 所示。

Table 3 BPN 結果綜整表

Parts	Train Samples	MSE	Test Samples	MSE	MAPE(%)
Spar	100 samples per each part	0.095	50 samples per each part	0.1156	2.4803
Ribs		0.085		0.077	2.3917
Stringers		0.098		0.1696	3.1631
Skins		0.099		0.073	2.08

藉由比較 SVR 與 BPN 兩種方法對於各結構件進行訓練後測試結果可知，SVR 在每個結構件測試結果績效皆優於 BPN，結構件 Spar 測試結果經過 SVR 運算之測試績效 MSE 為 0.11，MAPE 為 2.436%；經過 BPN 運算之測試績效 MSE 為 0.1156，MAPE 為 2.4803%。結構件 Ribs 測試結果經過 SVR 運算之測試績效 MSE 為 0.064，MAPE 為 2.1064%；經過 BPN 運算之測試績效 MSE 為 0.077，MAPE 為 2.3917%。結構件 Stringers 測試結果經過 SVR 運算之測試績效 MSE 為 0.155，MAPE 為 2.078%；經過 BPN 運算之測試績效 MSE 為 0.1696，MAPE 為 2.08%。結構件 Skins 測試結果經過 SVR 運算之測試績效 MSE 為 0.05，MAPE 為 3.0146%；經過 BPN 運算之測試績效 MSE 為 0.073，MAPE 為 3.1631%，SVR 與 BPN 訓練後測試結果比較如 Table 4 所示。

Table 4 SVR 與 BPN 訓練後測試結果比較表

Parts	Performance	SVR	BPN
Spar	MSE	0.11	0.1156
	MAPE(%)	2.436	2.4803
Ribs	MSE	0.064	0.077
	MAPE(%)	2.1064	2.3917
Stringers	MSE	0.155	0.1696
	MAPE(%)	2.078	2.08
Skins	MSE	0.05	0.073
	MAPE(%)	3.0146	3.1631

至此已完成各結構件設計成本預測模型，接著假設進行一個 Wing-Box 專案所需的結構件尺寸與數量，例如編號 1 的 Spar 結構件尺寸(L43,W5,T4)，數量 5 條；編號 2 的 Spar 結構件尺寸(L38,W5,T5)，數量 3 條，詳細結構尺寸數量如 Table 5 所示。

Table 5 Wing-Box 所需結構件尺寸數量表

Parts	Number	L	W	T	Quantity
Spar	1	43	5	4	5
	2	38	5	5	3
Ribs	3	25	7	4	3

	4	23	6	4	3
	5	21	6	3	3
Stringers	6	12	7	6	15
Skins	7	50	34	0.5	10
	8	36	31	0.4	5

將結構件尺寸資訊投入設計工時預測模型可得到預測結果，如編號 1 的 Spar 結構件由 SVR 預測之設計成本需要 14.682 美元，由 BPN 預測之設計成本需要 14.825 美元，詳細預測結果如 Table 6 所示。

Table 6 各結構件設計成本預測結果比較(單位：美元)

Parts	Num	SVR results	BPN results
Spar	1	14.682	14.825
	2	16.377	16.449
Ribs	3	9.596	9.681
	4	7.461	7.614
	5	6.407	6.534
Stringers	6	7.039	7.837
Skins	7	7.675	7.847
	8	5.558	5.629

將各結構件設計成本進行數量加總可得到設計成本總和，由預測結果(Table 7)顯示使用SVR預測模型計算出Wing-Box專案所需之設計成本為403.058美元，使用BPN預測模型所需之設計成本為418.728美元。若以按成本設計(DTC)的概念執行專案，欲將專案成本控制在一個門檻條件下完成既定工作，SVR建立的預測模型較BPN能控制較低的成本。

Table 7 設計成本總和預測結果比較(單位：美元)

Parts	SVR results	BPN results
Spar	122.541	123.452
Ribs	70.392	71.106
Stringers	105.585	117.555
Skins	104.54	106.615
Design Cost	403.058	418.728

Conclusion

在基於按成本設計(DTC)的概念下執行專案，不論過去採用從上而下(Top-Down)或從下而上(Bottom-Up)的人為成本預測模式，都因預測不夠準確而最終使專案難以成功，所以準確的專案成本預測將成為未來執行DTC的趨勢。過去統計參數成本預測方法在面對全球環境快速變遷的情況下，無法快速反應並且更新成本預測模型。因此，本研究使用機器學習方法中的支援向量迴歸(SVR)與倒傳遞類神經網路(BPN)，透過訓練測試以及調整參數的機制建立成本預測模型，並且驗證機器學習方法運用飛機設計成本預測，係可行且達到良好效果。

本研究以假設規劃一飛機機翼結構件Wing-Box專案並預測所需之設計成本。藉由預測結果顯示，無論是使用SVR或BPN進行預測的準確度皆能控制在約3%以內。雖然SVR與BPN預測結果並未有相當顯著差異，但在建立預測模型時 BPN需要花費更多時間成本進

行參數調整以及試誤，而SVR只需決定核函數以及參數組合(C, γ, ε)，即能得到最佳預測模型。因此，以機器學習方法取代傳統統計參數預測法將成為未來的趨勢。本研究未來發展將進行在「少量訓練樣本」的情況下SVR與BPN的預測準確度比較，另外進一步探討其他飛機機體結構之設計成本預測，結構件製造成本與組裝成本預測，最終目標完成整體飛機全壽期成本預測。

References

Bode, J. 1997. Decision support with neural networks in the management of research and development: Concepts and application to cost estimation, *Information & Management*, 34: 33-40.

Carr, R. I. 1989. Cost Estimating Principles, *Journal of Construction Engineering and Management*, 115 (4).

Smith, A.E. and Mason, A.K. 1996. Cost Estimation Predictive Modeling: Regression versus Neural Network, The Engineering Economist.

Mantel, S. J., Meredith, J. R., Shafer, S. M. and Sutton, M. M. 2005. *Project Management in Practice 2nd Edition*, Wiley.

Roskam, J. 1990. *Airplane Design Part VIII：Airplane cost estimation：Design, Development, Manufacturing and Operation*.

Layer, A., Brinke, E. T., Houten, F. V., Kals, H. and Haasis, S. 2002. Recent and Future trends in cost estimation, *Computer Integrated Manufacturing*, 15 (6): 499-510.

Parametric Estimating Handbook 4th Edition, International Society of Parametric Analysis (ISPA), 2008.

Greese, R. C. and Li, L. 1993. Cost Estimation of Timber Bridges Using Neural Network, *Cost Engineering*, 37 (5): 17-22.

Zhang, Y. F. and Fuh, J. Y. H. 1998. A Neural Network Approach for Early Cost Estimation of Packaging Products, *Comput Ind Eng*, 34 (3): 433-450.

Murat, H. and Zeynep, S. 2004. A neural network approach for early cost estimation of structural systems of buildings, *International Journal of Project Management*, 22: 295-602.

Vapnik, V. 1995. *The nature of statistical learning theory*, Springer.

Guo, G., Li, S. Z. and Chan, K. L. 2001. Support vector machines for face recognition, *Image and Vision Computing*, 19: 631-638.

Zhang, W., Yoshida, T. and Tang, X. 2008. Text classification based on multi-word with support vector machine, *Knowledge-Based Systems*.

Cai, Y. D., Liu, X. J., Xu, X. B. and Chou, K. C. 2002. Prediction of protein structural classes by support vector machines, *Computers and Chemistry*, 26: 293-296.

Vapnik, V., Golowich, S. and Smola, A. 1997. Support Vector Method for Function Approximation, Regression Estimation, and Signal Processing, *Advance in Neural Information Processing System*, 9: 281-287.

Tay, E. H. and Cao, L. 2001 Application of support vector machines in financial time series forecasting, *Omega*, 29: 309-317.

Hua, Z., Wang, Y., Xu, X., Zhang, B. and Liang, L. 2007. Predicting corporate financial distress based on integration of support vector machine and logistic regression, *Expert Systems with Applications*, 33: 434-440.

Chen, K. Y. and Wang, C. H. 2007. Support vector regression with genetic algorithms in forecasting tourism demand, *Tourism Management*, 28: 215-226.

Xi, X. C., Poo, A. N. and Chou, S. K. 2007. Support vector regression model predictive control on a HVAC plant, *Control Engineering Practice*, 15: 897-908.

Smola, A. J. and Schölkopf, B. 2002. *Learning with Kernels – Support Vector Machines, Regularization, Optimization, and Beyond*, MIT Press, London.

Bertsekas, D. P. 1995. *Nonlinear Programming*, Belmont, MA:Athenas.

Gunn, S. 1998. *Support Vector Machines for Classification and Regression*, MIT press.

Haykin, S. 1999. *Neural Networks – A Comprehensive Foundation 2nd Edition*, Prentice Hall.

Lewis, C. D. 1982. *Industrial and Business Forecasting Methods : a practical guide to exponential smoothing and curve fitting*, Butterworth-Heinemann.

Niu, M. C. Y. 1988. *Airframe Structural Design*, Hong Kong: Conmilit Press.

BIOGRAPHY

Professor S.Deng received his Master and Ph. D degrees of Mechanical Engineering from North Carolina State University, Raleigh, North Carolina, in 1994. Since then, he has been working for Aeronautical Research Lab. (ARL) of Chung-Shan Institute of Science & Technology (CSIST), and Center for Aero & Space Technology (CAST) of Industrial of Technology Research Institute (ITRI). After working for these two main research centers in Taiwan, he then turned himself into the college education. He was teaching in the Industrial Engineering Department of Da-Yeh University, Taiwan as Associate Professor. He is currently Professor of the Power Vehicle and Systems Engineering Department, Chung-Cheng Institute of Technology, National Defense University of Taiwan. His major current research interests lie in several areas of Systems Engineering and Project Management.

Mr. Tsung-Han Yeh received his Master degree in Department of Industrial Engineering and Technology Management, from Da-Yeh University, Taiwan, in 2006, and presently pursuing his Ph.D at Chung Cheng Institute of Technology, National Defense University, Taiwan. His research areas include Machine Learning, Project Management and System Engineering, etc.

Applying Machine Learning Methods to the Airframe Structural Design Cost Estimation – A Case Study of Wing-Box Project

S. Deng

Department of Power Vehicle and Systems Engineering, Chung-Cheng Institute of Technology, National Defense University

sgdeng@ccit.edu.tw

Tsung-Han Yeh

Graduate School of Defense Science, Chung Cheng Institute of Technology, National Defense University

oasis.yeh@msa.hinet.net

Abstract. This research used two machine learning methods, the Support Vector Regression (SVR) and Back-Propagation Neural Network (BPN), to create the cost prediction models for airplane wing-box structural design, and verified the feasibility and efficiency for both methods. In the case study, four different main structural part groups of the wing-box, Spars/Ribs/Skins/Stringers, were chosen. In the parts data base, the part dimensions were included and used for classifying the part groups. Each part group has 150 bill of parts, 100 bill of parts used for training samples, 50 bill of parts used for predicting samples, to test their accuracy. After verified through wing-box case study, the results showed either SVR or BPN can precisely predicting the design costs. But comparing with the BPN, SVR can get the global optimal solution while using less decision parameters. This can save lots of time for searching the best parameters combination when creating the prediction model.

Keywords: cost estimation, airframe structure, support vector regression, back-propagation neural network

Introduction

The precision of cost estimation is essential for areas including project planning, product development (Bode, 1997), as well as aeronautical industries. The data of cost can be categorized into direct cost and indirect cost; direct cost refers to the cost directly related to the manufacturing of the product, such as raw materials, equipments, and wages; whereas indirect cost refers to the cost not directly related to the production of the product, such as factory illumination cost, operation cost of departments, and so on (Carr, 1989). Most of the cost prediction models for airframe industries are developed using horsepower, weight, dimension and some other feature parameters, that is so called Cost Estimation Relation (CER). However, the traditional CER models are based on statistical regression model, which are not able to update the estimation models instantaneously facing the fast-changing social factors (Smith & Mason, 1997). Therefore, this research uses two machine learning methods, the Support Vector Regression (SVR) and Back-Propagation Neural Network (BPN), to create the cost prediction models for airplane wing-box structural design, for the purpose of providing efficient cost estimation methods for airframe industries.

Design to cost. In the early stage, aeronautical industries centred on Design to Performance, which turned out to be costly. Therefore, how to efficiently control the cost has become an important issue, and cost has become an important index besides the schedule performance in terms of the success of product. Nowadays the methodology of project cost estimation can be categorized as Top-Down and Bottom-Up; in the former one the top management determines the budget and then allocates it downwards to each level; whereas in the latter one the managers in each level estimate the cost and then integrate it upwards. However, both methods estimate the cost arbitrarily relying on the past experience, which may cause the failure of project due to imprecise cost estimation.

In order to precisely plan and control cost before the execution of projects, a budget threshold should be decided from the analysis and integration of the actual operating cost for each segment, within which the project should be fulfilled. Therefore, the concept of Design to Cost (DTC) is proposed. According to the document of U.S. Department of Defense DoD5000.28, DTC is defined as a management concept to fulfill the project with a pre-specified cost constraint through controlling the cost of purchasing and logistics to achieve the tradeoff between the ability, performance, cost and the progress. The industries used to determine the cost based on the product specification, which neglected the consideration of cost constraint and usually caused waste. In the recent years, following the international industrial trend, cost is the most priority for manufacturing and operating, therefore DTC concept is employed. From the perspective of project management, DTC is to design the product to fulfill the requirement of performance with the constraint of the defined cost threshold; both performance and cost are taken into account in the mean time. Taking U.S. as a successful example employing DTC, NASA proposed the new idea of fulfilling the planned target with cost limitations when executing the project of Mars Pathfinder-Rover, which benefited NASA in cutting down the cost from 3 billion US dollars for Viking-Mars Lander project in 1976 to 175 million US dollars for the new project, total 94% cost savings. This case showed the international industrial trend has changed from Design to Performance to Design to Cost to meet the appointed target (Mantel et al., 2005).

Airframe LCC. Airframe Life Cycle Cost can be briefly divided into the following parts: conceptual design, preliminary design, detailed design, manufacturing and resembling, operating and logistics, and elimination. It is pointed out that, in Airframe LCC, conceptual design, preliminary design and detailed design affect up to 95% of the total cost. Although the actual cost rises dramatically from the stage of manufacturing, in order to efficiently control the overall Airframe LCC, we need to plan from the stage of conceptual design to minimize the total cost. When planning Airframe LCC, according to Cost Breakdown Structure (CBS), airframe LCC should be broken down into R&D cost, manufacturing cost, maintenance and warranty cost according to the time frame; it can also be broken down into cost of main body, cost of power system, cost of weapon system, cost electronic system and so on, according the structure of airframe. Among these, R&D cost can be further broken down into cost of design, cost of materials, cost of test, and cost of equipment (Roskam, 1990). This research will focus on the airframe structural design cost in the stage of R&D. Currently there are cost estimation software available for different industries, the representative ones including PRICE H, ACEIT, SEER H and COCOMO II. The theoretical fundamental of the above cost estimation software are all Statistical Parametric Regression (Parametric Estimation Handbook, 2008).

Cost Estimation Methods. Cost estimation methods can be divided into qualitative methods and quantitative methods (Layer, 2002), and this research will only introduce quantitative statistical parametric method. Most widely adopted method in statistical parametric estimation is statistical regression analysis, which is the technique to model and predict response variable (Y) with independent variable $(x_1, x_2, ..., x_i)$ and the error term (ε) using Least Square Method to determine the parameters in the linear regression model (Parametric Estimation Handbook, 2008). For simple estimation problems, linear regression model is adequate. However, as the estimation problem becomes more complicated, it may cause poor performance only using independent variables as input to forecast the response variable, which may lead to poor ability of explanation. Therefore, the square and the interaction of independent variables should be taken into account to establish second order estimation model. Statistical regression estimation model can achieve good results when the parameters are relatively stable; however, taking airframe manufacturing industry as example, the estimation model is usually subject to the impact of inflation, growth of technology, change of process capability and so on thus becomes imprecise. Thus the regression parameters need to be revised and updated frequently to ensure the precision, which could be quite time-consuming and costly. Therefore, traditional statistical regression estimation model is not adequate to suit the change in the society.

Machine Learning Methods. In recent years, many scholars adopt Machine Learning Methods to replace traditional Regression Analysis. The main concept of Machine Learning is to mimic the process of human brain neurons learning and memorizing through nervous system, also known as Artificial Neural Network (ANN), among which the most representative one is Back-propagation Neural Network (BPN). It is an iterative procedure to perform error back-propagation and adjust the weights to minimize the error and eventually to obtain an optimal estimation model (Haykin, 1999). BPN has been widely applied in cost estimation problems in many aspects. To name a few, Greese and Li (1993) apply Neural Network to cost estimation of Timber Bridges and compare the results with traditional parametric estimation. Zhang and Fuh (1998) apply BPN to develop early cost estimation model of packaging products. Murat and Zeynep (2004) also apply Neural Network to early design cost estimation of reinforced concrete structural systems of buildings.

Although Artificial Neural Network has been widely used in cost estimation problem and performed efficiently, there are still several limitations. Many network parameters are required when building the network, such as transfer function, the number of hidden layer, and the number of neurons; it is time-consuming to determine the network parameters in order to obtain a precise estimation model. ANN tends to be faced the local minimum problem when performing computation, therefore many papers use the combination of heuristic algorithm and ANN in order to avoid the local minimum problem. Besides, ANN lacks of a comprehensive mathematical theoretical foundation, and the estimation model is quite complex.

In 1995, Vapnik and his research team in AT&T Lab developed a new machine learning method, called Support Vector Machine (SVM). The theoretical foundation of SVM is Vapnik-Chervonenkis Dimension Theory (VCD) based on Statistical Learning Theory (SLT) and Structural Risk Minimization Principle (SRM). This principle targets at minimizing the upper bound of estimation error (i.e. structural risk), whereas ANN minimizes training error (i.e. empirical risk). Therefore, SVM can obtain the optimal model by adjusting empirical error and VC dimension confidence interval. SVM solves Quadratic Programming (QP) problem with linear constraints in the training process. The meaning of solving QP is that the solution obtained by SVM is the unique optimal solution, which can avoid the possibility of facing local optimal problem. SVM was first applied to the classifying problem, developed from statistical learning theory to learning method, from Simple Vector Classifiers to Hyperplane Classifiers. This Classifier has been applied to many aspects, such as Face recognition (Guo et al., 2001), text classification (Zhang et al., 2008), Biology technology (Protein structural classification) (Cai et al., 2002) and so on, and performed efficiently.

In 1997, Vapnik et al. introduce ε-insensitive loss function to maintain the sparseness, in other words, to use less support vectors to represent the decision function. After that, SVM starts to be widely used in non-linear estimation problems, which is called Support Vector Regression (SVR). Many scholars have widely applied SVR to estimation problems in many aspects. For example, Tay and Cao (2001) apply SVR to financial time series analysis; Hua et al. (2007) integrate SVR and logistic regression in predicting corporate financial distress; Chen and Wang (2007) integrate SVR and genetic algorithms in forecasting tourism demand; Xi et al. (2007) apply SVR to provide predictive control on a HVAC plant.

The purpose of this research is to apply machine learning methods to establish airframe structural design cost estimation model. Both SVR and BPN are applied to develop estimation models and the efficiency of them are compared. It is verified that both SVR and BPN are feasible for cost estimation problems in the area of airframe industry. The outline of the paper is as follows: in the introduction part, the importance of cost estimation, the concept of DTC, airframe LCC structure, and the evolution of cost estimation methods are discussed; the second part will introduce the theory of SVR and BPN in machine learning methods; the third part is case study and result analysis; and the last part concludes our research and proposes the future research direction.

Methodology

Support Vector Regression. Support Vector Regression (SVR) is a technique to forecast unknown change using available information, to establish estimation model using training data, and to minimize the errors of training data and actual value (Vapnik, 1995) (Smola, 2002). Given a data set $S=(x_i, y_i)\ i=1...n$, among which is

x_i the input vector, y_i is the target value, n is the sample size, then the objective function is shown in (1). The basic concept is to map data x through Φ non-linear mapping function to high dimensional feature space to perform linear regression.

$$f(x) = \Phi(x) \cdot w + b \quad (1)$$

In (1), w is the weight vector, b is the threshold value, Φ is the high dimensional feature space, that is, the non-linear mapping of input space x. Performing linear regression in high dimensional feature space equals to performing non-linear regression in low dimensional input space, except that we do not need to calculate directly Φ(x) and the integral mean value of w. Weight vector w represents the flatness of the target function in high dimensional space; $\| w \|^2$ is Euclidean Length, which is the measurement of the flatness of the decision function or the complexity of the model. After determining high dimensional feature space Φ, in order to make $f(x)$ flatter, $\| w \|^2$ should be minimized. In the process of establishing regression function following SRM, not every bill of training data contributes positively. There is noise in the training data. For the noise which may affect the efficiency of SVR estimation model, it should be dealt with using loss function integrated with penalty parameter C. The main purpose of ε-insensitive loss function is to examine the distance between the regression function and the training data. If the distance between the forecast value and the actual value is smaller than or equals to ε, then the loss function is 0; if it is greater than ε, then the loss function is not 0.

$$y - f(x) = \begin{cases} |y - f(x)| \le \varepsilon, 0 \\ |y - f(x)| > \varepsilon, otherwise \end{cases} \quad (2)$$

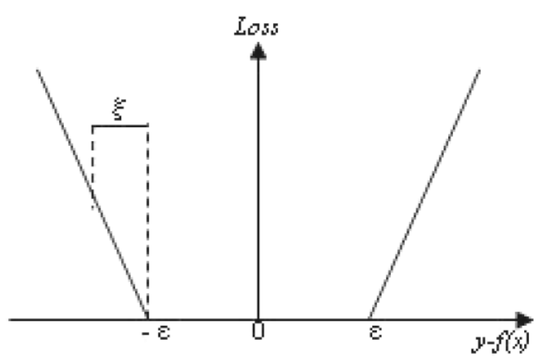

Figure 1. ε-insensitive loss function

SVR adds positive slack variable to regularized risk minimization function, denoted as ξ; in the mean time, when taking slack variable into account a penalty parameter should be defined as the constraints to include the noise in the regression function, as shown in (3).

$$\text{Minimize: } \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{n}\left(\xi_i + \xi_i^*\right) \tag{3}$$

$$\text{subjected to: } y_i - \Phi(x_i)\cdot w - b \le \varepsilon + \xi_i$$

$$\Phi(x_i)\cdot w + b - y_i \le \varepsilon + \xi_i^* \quad i = 1,\ldots n \quad \xi_i, \xi_i^* \ge 0$$

By Lagrange Multiplier optimization and taking decision function and constraints into consideration (Bertsekas, 1995), (3) can be rewritten as Lagrange Function as shown in (4):

$$\begin{aligned} L = {} & \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{n}\left(\xi_i + \xi_i^*\right) - \sum_{i=1}^{n}\left(\eta_i\xi_i + \eta_i^*\xi_i^*\right) \\ & - \sum_{i=1}^{n}\alpha_i\left(\varepsilon + \xi_i - y_i + \Phi(x_i)\cdot w + b\right) \\ & - \sum_{i=1}^{n}\alpha_i^*\left(\varepsilon + \xi_i^* - y_i - \Phi(x_i)\cdot w - b\right) \end{aligned} \tag{4}$$

In (4), $\eta_i, \eta_i^*, \alpha_i, \alpha_i^* \ge 0$, all called Lagrange parameters. Optimal value exists in the saddle point, where the partial differential of w, b, ξ_i, ξ_i^* all equal to 0. Thus we have (5):

$$\begin{aligned} & \partial_w L = w - \sum_{i=1}^{n}\left(\alpha_i - \alpha_i^*\right)x_i = 0 \\ & \partial_b L = \sum_{i=1}^{n}\left(\alpha_i - \alpha_i^*\right) = 0 \\ & \partial_{\xi_i} L = C - \alpha_i - \eta_i = 0 \\ & \partial_{\xi_i^*} L = C - \alpha_i^* - \eta_i^* = 0 \end{aligned} \tag{5}$$

(5) can be reformulated and integrated into (4), and solved by Lagrange dual optimization, denoted as L_D. Therefore the original problem to minimize L becomes its dual problem to maximize L_D, as shown in (6).

$$\text{Maximize: } L_D = -\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\left(\alpha_i - \alpha_i^*\right)\left(\alpha_j - \alpha_j^*\right)\Phi(x_i)\Phi(x_j) - \varepsilon\sum_{i=1}^{n}\left(\alpha_i + \alpha_i^*\right) + \sum_{i=1}^{n}y_i\left(\alpha_i - \alpha_i^*\right)$$

$$\text{subjected to: } \sum_{i=1}^{n}\left(\alpha_i - \alpha_i^*\right) = 0 \text{ and } \alpha_i, \alpha_i^* \in [0, C] \tag{6}$$

In (6), $\Phi(x_i)\Phi(x_j)$ can be replaced by kernel function $K(x_i, x_j)$ according to Karush Kuhn-Tucher (KKT) conditions and solved as quadratic programming problem (Bertsekas, 1995). From (5) we can obtain optimal value of w, as shown in (7). We then can obtain (8) by integrating (7) into (1).

$$w^* = \sum_{i=1}^{n} (\alpha_i - \alpha_i^*) \Phi(x_i) \qquad (7)$$

$$f(x, \alpha_i, \alpha_j) = \sum_{i=1}^{n} (\alpha_i - \alpha_i^*) K(x \cdot x_i) + b \qquad (8)$$

In the end, we can obtain optimal *b* value according to complementary slackness conditions, as shown in (9).

$$\begin{aligned} &\alpha_i(\varepsilon + \xi_i - y_i + \Phi(x_i)w + b) = 0 \\ &\alpha_i^*(\varepsilon + \xi_i^* - y_i - \Phi(x_i)w - b) = 0 \\ &(C - \alpha_i)\xi_i = 0 \\ &(C - \alpha_i^*)\xi_i^* = 0 \end{aligned} \qquad (9)$$

Kernel Functions. Commonly used kernel functions include three types, Linear Kernel Function in (10), Polynomial Kernel Function in (11), and Radial Basis Kernel Function (RBF) in (12) (Steve Gunn, 1998), shown as follows.

$$k(x_i, x_j) = x_i \cdot x_j^T \qquad (10)$$

$$k(x_i, x_j) = (1 + x_i \cdot x_j)^d \qquad (11)$$

$$k(x_i, x_j) = \exp\left(-\gamma \|x_i - x_j\|^2\right) \qquad (12)$$

Each kernel function has its own kernel parameter, which should be specified when building the estimation model. Adjusting Kernel parameters could affect the precision of the final estimation model, therefore when using SVR, how to select kernel parameters is an important issue. RBF is adopted in our research. In the recent SVR related research, RBF are chosen if the problem is non-linear. The reason is that kernel function itself is non-linear, which can transform from the original space to high dimensional space to solve non-linear problem. When the raw data is non-linear, it can perform efficiently. In addition, if the characteristic of raw data is not identified, using SVR can generally guarantee a good result (Smola, 1998).

SVR Procedure. SVR procedure is shown in Figure 2, with the statement of each step as follows:

<u>Step 1 Data collection and scaling:</u>

In this stage, raw data is preliminarily organized and scaled. The purpose of scaling is to increase the accuracy of SVR estimation model, in the mean time, to restrict and convert data into a specified interval. In many papers data is converted into (-1~1) or (0~1). Then the data is split into training data and test data. Training data is used to develop SVR model, whereas test data is used to verify the model and obtain the estimation result and examine the performance.

Step 2 Using Grid Search method to search for optimal parameter combination:

The parameter combination (C, γ, ε) should be determined when developing estimation model using SVR, in which C is the penalty parameter; γ is RBF kernel parameter; ε is the loss function threshold. Selection of parameter combination directly affects the precision of SVR estimation model; therefore in our research we use Grid Search method to search for the optimal parameter combination in order to improve the precision of estimation model. We apply Grid Search to the above mentioned training data, use possible parameter combination to perform SVR computation to determine whether it has reached the termination condition (in this research it is MSE minimization). If not, we generate another set of parameter combination to iteratively perform SVR computation until we achieve the optimal parameter combination.

Step 3 Developing SVR model:

We input optimal parameter combination to develop SVR estimation model, and use test data to verify the performance of the model. From the performance criteria we can know the precision and the future generalization ability of the model.

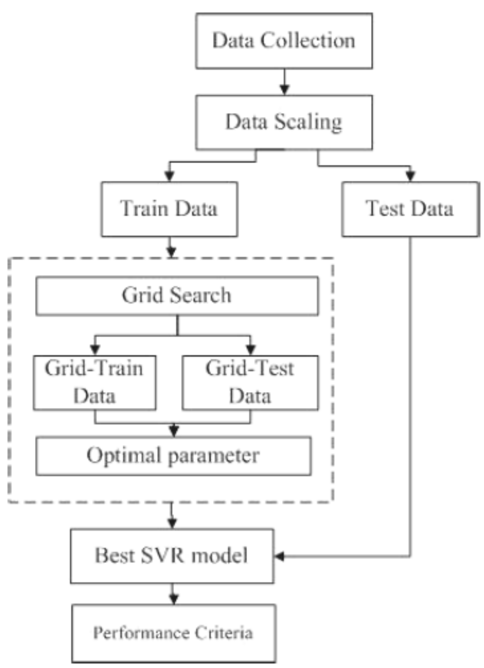

Figure 2. SVR procedure

Back-propagation Neural Network. BPN belongs to supervised learning for multilayer forward network, the main structure of which includes Input Layer, Hidden Layer, and Output Layer, as shown in Figure 3. The computation procedure of BPN can be divided into forward pass, error computation, and error back-propagation. Statement of each stage is as below:

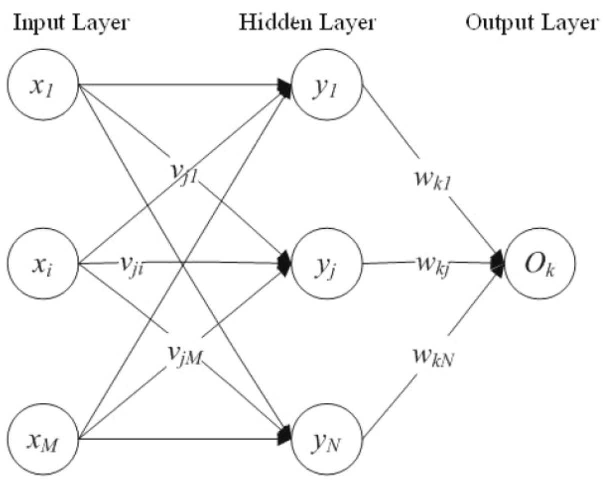

Figure 3. BPN framework

Step 1 Forward pass: In (13), each unit (x_i) in Input Layer is mapped to unit (y_i) in Hidden Layer with weight (v_{ij}) and propagation function (f); in (14), each unit (y_i) in Hidden Layer is mapped to unit (O_k) in Output Layer with weight (w_{kj}) and activation function (f).

$$y_i = f(net_j = \sum_{i=1}^{M} v_{ji} x_i) \qquad (13)$$

$$O_k = f(net_k = \sum_{j=1}^{N} w_{kj} y_j) \qquad (14)$$

Step 2 error computation: error function is defined and error of the output of the network and the expectation is calculated; in (15), d_k is the expected output; O_k is the forecasted network output.

$$E = \frac{1}{2} \sum_{k=1}^{L} (d_k - O_k)^2 \qquad (15)$$

Step 3 error back-propagation: in this stage, if the error does not satisfy the termination condition, then gradient descent method is used to update the weight of unit propagation. η is gradient descent parameter. In (16), the weight mapping from Hidden Layer to Output Layer is updated; in (17), the weight mapping from Input Layer to Hidden Layer is updated. This process is repeated until termination condition is satisfied, in other words, the error between the network output and the expected output is minimized.

$$w_{kj} = w_{kj} + \Delta w_{kj} = w_{kj} - \eta \frac{\partial E}{\partial w_{kj}} \qquad (16)$$

$$v_{ji} = v_{ji} + \Delta v_{ji} = v_{ji} - \eta \frac{\partial E}{\partial v_{ji}} \qquad (17)$$

There are many parameters needed to be specified when using BPN to develop estimation model. When using SVR, only (C, γ, ε) needs to be determined. Unlike

SVR, when using BPN, we need to decide the number of neurons in Input Layer, the number of Hidden Layer and the number of neurons in Hidden Layer, the number of neurons in Output Layer, the neuron propagation function, training function, learning function, and so on. Due to the large number of parameters, recent research all adopts trial-and-error method to decide the feasible parameter combination. Therefore BPN estimation model does not reach the optimal.

Performance Criteria. Performance criteria measure the efficiency of estimation model obtained from SVR and BPN. The main idea is to compare the estimated value with the actual output to verify the feasibility of the estimation model. This research employs Mean Square Error (MSE) (18) and Mean Absolute Percentage Error (MAPE) (19) to assess the models (Lewis, 1982). MSE and MAPE are used to measure the difference between the actual value and the forecast value; smaller MSE or MAPE means forecast value is closer to actual value. The function to calculate these two statistics is listed as follows, where n is the sample size, y_i is the actual value, and y_i' is the forecast value.

$$MSE = \frac{1}{n}\sum_{i=1}^{n}\left(y_i - y_i'\right)^2 \qquad (18)$$

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i - y_i'}{y_i}\right| \times 100\% \qquad (19)$$

Case Study

In this case study, assume that we want to plan and execute the structural design cost of Wing-Box project, then the forecasting target is the design cost (Unit: US dollar). The feature parameter affecting the target is the dimensions of the structural part, including length, width, and thickness. The relationship between the design cost and the dimensions can be represented by function as shown in (20). However, in past practice, the structural cost data is recorded arbitrarily and there is a lack of estimation model. This research attempts to develop an estimation model for structural design cost with historical data using machine learning method.

$$\text{Design Cost} = f\,(\text{Length, Width, Thickness}) \qquad (20)$$

The structure of airframe can be divided into several parts, including fuselage, wing, empennage, and the main component of the wing is Wing-Box, which can be further divided into spar, ribs, stringers, skins, and so on. The historical data range of structural part groups of wing-box is shown in Table 1 (Unit: Inch).

Table 1: Historical structural part data

Parts	Size	Range	Parts	Size	Range	Parts	Size	Range	Parts	Size	Range
Spar	L	15 ~ 50	Ribs	L	10 25	Skins	L	10 ~ 30	Stringers	L	30 ~ 50
	W	2 ~ 10		W	5 10		W	2 ~ 10		W	30 ~ 50
	T	2 ~ 10		T	2 5		T	2 ~ 10		T	0.1 ~ 1

After gathering the data range of the structural part groups of Wing-Box, this research applies machine learning methods SVR and BPN to develop estimation model for each individual structural part group. Historical data used in this research is 150 bill, among which 100 are used as training data, and the remaining 50 as test data.

The parameters of SVR and BPN should be specified during the process of model construction. As for SVR, RBF is employed, and Grid Search method is used to search the optimal parameter combination. Following the procedure of SVR, we can obtain the performance of each individual structural design cost estimation model (MSE, MAPE) and the corresponding optimal parameter combination (C, γ, ε). Under the parameter combination (C=1024, γ=0.25, ε=0.0078125) for SVR estimation model, structural part group Spar training stage MSE is 0.025; test stage MSE is 0.11; MAPE is 2.436%. Under the parameter combination (C=1024, γ=0.125, ε=0.0078125) for SVR estimation model, structural part group Ribs training stage MSE is 0.0002; test stage MSE is 0.064; MAPE is 2.1064%. Under the parameter combination (C=1024, γ=0.25, ε=0.0078125) for SVR estimation model, structural part Stringers training stage MSE is 0.0015; test stage MSE is 0.155; MAPE is 2.078%. Under the parameter combination (C=1024, γ=0.25, ε=0.0078125) for SVR estimation model, structural part group Spar training stage MSE is 0.025; test stage MSE is 0.11; MAPE is 2.436%. Under the parameter combination (C=1024, γ=0.125, ε=0.0078125) for SVR estimation model, structural part group Skins training stage MSE is 0.0007; test stage MSE is 0.05; MAPE is 3.0146%. The SVR computation result of different structural part is shown in Table 2.

Table 2 SVR computation result

Parts	Train Samples	kernel	C	γ	ε	MSE	Test Samples	MSE	MAPE(%)
Spar	100 samples per each part	RBF	1024	0.25	0.0078125	0.025	50 samples per each part	0.11	2.436
Ribs			1024	0.125	0.0078125	0.0002		0.064	2.1064
Stringers			1024	0.125	0.0078125	0.0015		0.155	2.078
Skins			1024	0.125	0.0078125	0.0007		0.05	3.0146

As for BPN parameter, the estimation problem in this case is to map three input (structural part length, width, and thickness) to one output (design duration). The number of Hidden Layer is set at 1; the number of neurons in Hidden Layer is set at 7 with past experience. Hyperbolic tangent linear propagation function is used; Lavenberg-Marquardt network training function is used due to high convergence speed. The number of iteration in the training stage is set at 1000; the termination condition is set at 0.1, in other words, the training process will stop when the training error converges to less than 0.1.

After this we perform BPN computation and obtain performance measure (MSE, MAPE) for cost estimation model in training and test stage. For BPN estimation model, structural part group Spar's training stage MSE is 0.095; test stage MSE is 0.1156; MAPE is 2.4803%. Structural part group Ribs' training stage MSE is 0.085; test stage MSE is 0.077; MAPE is 2.3917%. Structural part group Stringers' training stage MSE is 0.098; test stage MSE is 0.1696; MAPE is 3.1631 %. Structural part group Skins' training stage MSE is 0.099; test stage MSE is 0.073; MAPE is 2.08%. The computation result for BPN estimation model is shown in Table 3.

Table 3 BPN computation result

Parts	Train Samples	MSE	Test Samples	MSE	MAPE(%)
Spar	100 samples per each part	0.095	50 samples per each part	0.1156	2.4803
Ribs		0.085		0.077	2.3917
Stringers		0.098		0.1696	3.1631
Skins		0.099		0.073	2.08

We can see from the comparison of training and test results using SVR and BPN that, SVR outperforms BPN in each structural part group. Using SVR structural part group Spar's MSE is 0.11, MAPE is 2.436%; whereas using BPN its MSE is 0.1156, MAPE is 2.4803%. Using SVR structural part group Ribs' MSE is 0.064 MAPE is 2.1064%; whereas using BPN its MSE is 0.077, MAPE is 2.3917%. Using SVR structural part group Stringers' MSE is 0.155, MAPE is 2.078%; whereas using BPN its MSE is 0.1696, MAPE is 2.08%. Using SVR structural part group Skins' MSE is 0.05, MAPE is 3.1631%; whereas using BPN its MSE is 0.073, MAPE is 3.1631%. The comparison of the testing result of SVR and BPN is shown in Table 4.

Table 4 Comparison of testing result of SVR and BPN after training

Parts	Performance	SVR	BPN
Spar	MSE	0.11	0.1156
	MAPE(%)	2.436	2.4803
Ribs	MSE	0.064	0.077
	MAPE(%)	2.1064	2.3917
Stringers	MSE	0.155	0.1696
	MAPE(%)	2.078	2.08
Skins	MSE	0.05	0.073
	MAPE(%)	3.0146	3.1631

Until now we have completed development of cost estimation models for each structural part group, following which we assume the dimensions and quantity of a Wing-Box project. For example, we assume No. 1 Spar's dimension is (L43, W5, T4), quantity 5; No. 2 Spar's dimension is (L38, W5, T5), quantity is 3. The detailed dimension and quantity of structural part group are shown in Table 5.

Table 5 Dimension and quantity of Wing-Box structural part group

Parts	Number	L	W	T	Quantity
Spar	1	43	5	4	5
	2	38	5	5	3
Ribs	3	25	7	4	3
	4	23	6	4	3
	5	21	6	3	3
Stringers	6	12	7	6	15
Skins	7	50	34	0.5	10
	8	36	31	0.4	5

Applying the structural part data information to the estimation models we can obtain the estimation result. For example, No. 1 Spar requires 14.682 US dollars based on SVR cost estimation model, versus 14.825 US dollars based on BPN cost estimation model. The detailed result is shown in Table 6.

Table 6 Cost estimation result for individual structural part group (Unit: US dollar)

Parts	Num	SVR results	BPN results
Spar	1	14.682	14.825
	2	16.377	16.449
Ribs	3	9.596	9.681
	4	7.461	7.614
	5	6.407	6.534
Stringers	6	7.039	7.837
Skins	7	7.675	7.847
	8	5.558	5.629

To sum up the cost for each individual structural part group we can obtain the total design cost. As shown in Table 7, total structural design cost of the Wing-Box project is 403.058 US dollars by SVR estimation model, and 418.728 US dollars by BPN model. According to the concept of DTC, in order to achieve the target while restricting the project cost within a certain threshold, SVR model performs better than BPN model in terms of lower cost.

Table 7 Comparison of total design cost (Unit: US dollar)

Parts	SVR results	BPN results
Spar	122.541	123.452
Ribs	70.392	71.106
Stringers	105.585	117.555
Skins	104.54	106.615
Design Cost	403.058	418.728

Conclusion

In order to implement DTC concept for the project, in the past no matter whether we choose Top-Down or Bottom-Up method to arbitrarily estimate the cost, it is not sufficiently precise to guarantee the success of the project. Therefore, precise cost estimation becomes the future trend to implement DTC. The traditional statistical parametric cost estimation methods cannot response quickly or update immediately facing the fast-changing international environment. Therefore, this research employs machine learning methods SVR and BPN to develop cost estimation models through training, testing and adjusting parameters, and demonstrates that machine learning method is feasible and efficient developing airframe structural design cost estimation model.

This research assumes a project to estimate the structural design cost of Wing-Box. From the result it is shown that both SVR and BPN can guarantee the estimation precision within 3%. Although there is no significant difference between the result of SVR and the result of BPN, it takes more time to adjust parameters and perform trial-and-error in developing BPN model. Instead, only kernel function and parameter combination (C, γ, ε) are required for SVR to guarantee optimal cost estimation model. Therefore, it is inevitable trend that machine learning methods replace conventional statistical parametric method. Our future research will focus on comparison of estimation precision of SVR and BPN when there is only a few training sample, and further discuss design cost estimation models for other parts of airframe, manufacturing and resembling cost estimation, and eventually complete cost estimation for airframe LCC.

References

Bode, J. 1997. Decision support with neural networks in the management of research and development: Concepts and application to cost estimation, *Information & Management*, 34: 33-40.

Carr, R. I. 1989. Cost Estimating Principles, *Journal of Construction Engineering and Management*, 115 (4).

Smith, A.E. and Mason, A.K. 1996. Cost Estimation Predictive Modeling: Regression versus Neural Network, The Engineering Economist.

Mantel, S. J., Meredith, J. R., Shafer, S. M. and Sutton, M. M. 2005. *Project Management in Practice 2nd Edition*, Wiley.

Roskam, J. 1990. *Airplane Design Part VIII* : *Airplane cost estimation* : *Design, Development, Manufacturing and Operation*.

Layer, A., Brinke, E. T., Houten, F. V., Kals, H. and Haasis, S. 2002. Recent and Future trends in cost estimation, *Computer Integrated Manufacturing*, 15 (6): 499-510.

Parametric Estimating Handbook 4th Edition, International Society of Parametric Analysis (ISPA), 2008.

Greese, R. C. and Li, L. 1993. Cost Estimation of Timber Bridges Using Neural Network, *Cost Engineering*, 37 (5): 17-22.

Zhang, Y. F. and Fuh, J. Y. H. 1998. A Neural Network Approach for Early Cost Estimation of Packaging Products, *Comput Ind Eng*, 34 (3): 433-450.

Murat, H. and Zeynep, S. 2004. A neural network approach for early cost estimation of structural systems of buildings, *International Journal of Project Management*, 22: 295-602.

Vapnik, V. 1995. *The nature of statistical learning theory*, Springer.

Guo, G., Li, S. Z. and Chan, K. L. 2001. Support vector machines for face recognition, *Image and Vision Computing*, 19: 631-638.

Zhang, W., Yoshida, T. and Tang, X. 2008. Text classification based on multi-word with support vector machine, *Knowledge-Based Systems*.

Cai, Y. D., Liu, X. J., Xu, X. B. and Chou, K. C. 2002. Prediction of protein structural classes by support vector machines, *Computers and Chemistry*, 26: 293-296.

Vapnik, V., Golowich, S. and Smola, A. 1997. Support Vector Method for Function Approximation, Regression Estimation, and Signal Processing, *Advance in Neural Information Processing System*, 9: 281-287.

Tay, E. H. and Cao, L. 2001 Application of support vector machines in financial time series forecasting, *Omega*, 29: 309-317.

Hua, Z., Wang, Y., Xu, X., Zhang, B. and Liang, L. 2007. Predicting corporate financial distress based on integration of support vector machine and logistic regression, *Expert Systems with Applications*, 33: 434-440.

Chen, K. Y. and Wang, C. H. 2007. Support vector regression with genetic algorithms in forecasting tourism demand, *Tourism Management*, 28: 215-226.

Xi, X. C., Poo, A. N. and Chou, S. K. 2007. Support vector regression model predictive control on a HVAC plant, *Control Engineering Practice*, 15: 897-908.

Smola, A. J. and Schölkopf, B. 2002. *Learning with Kernels – Support Vector Machines, Regularization, Optimization, and Beyond*, MIT Press, London.
Bertsekas, D. P. 1995. *Nonlinear Programming*, Belmont, MA:Athenas.
Gunn, S. 1998. *Support Vector Machines for Classification and Regression*, MIT press.
Haykin, S. 1999. *Neural Networks – A Comprehensive Foundation 2nd Edition*, Prentice Hall.
Lewis, C. D. 1982. *Industrial and Business Forecasting Methods : a practical guide to exponential smoothing and curve fitting*, Butterworth-Heinemann.
Niu, M. C. Y. 1988. *Airframe Structural Design*, Hong Kong: Conmilit Press.

BIOGRAPHY

Professor S.Deng received his Master and Ph. D degrees of Mechanical Engineering from North Carolina State University, Raleigh, North Carolina, in 1994. Since then, he has been working for Aeronautical Research Lab. (ARL) of Chung-Shan Institute of Science & Technology (CSIST), and Center for Aero & Space Technology (CAST) of Industrial of Technology Research Institute (ITRI). After working for these two main research centers in Taiwan, he then turned himself into the college education. He was teaching in the Industrial Engineering Department of Da-Yeh University, Taiwan as Associate Professor. He is currently Professor of the Power Vehicle and Systems Engineering Department, Chung-Cheng Institute of Technology, National Defense University of Taiwan. His major current research interests lie in several areas of Systems Engineering and Project Management.

Mr. Tsung-Han Yeh received his Master degree in Department of Industrial Engineering and Technology Management, from Da-Yeh University, Taiwan, in 2006, and presently pursuing his Ph.D at Chung Cheng Institute of Technology, National Defense University, Taiwan. His research areas include Machine Learning, Project Management and System Engineering, etc.