

應用維度縮減方法於企業用電需求之預測

康益豪、王姿文、劉冠麟、李大維

指導教授 吳漢銘教授

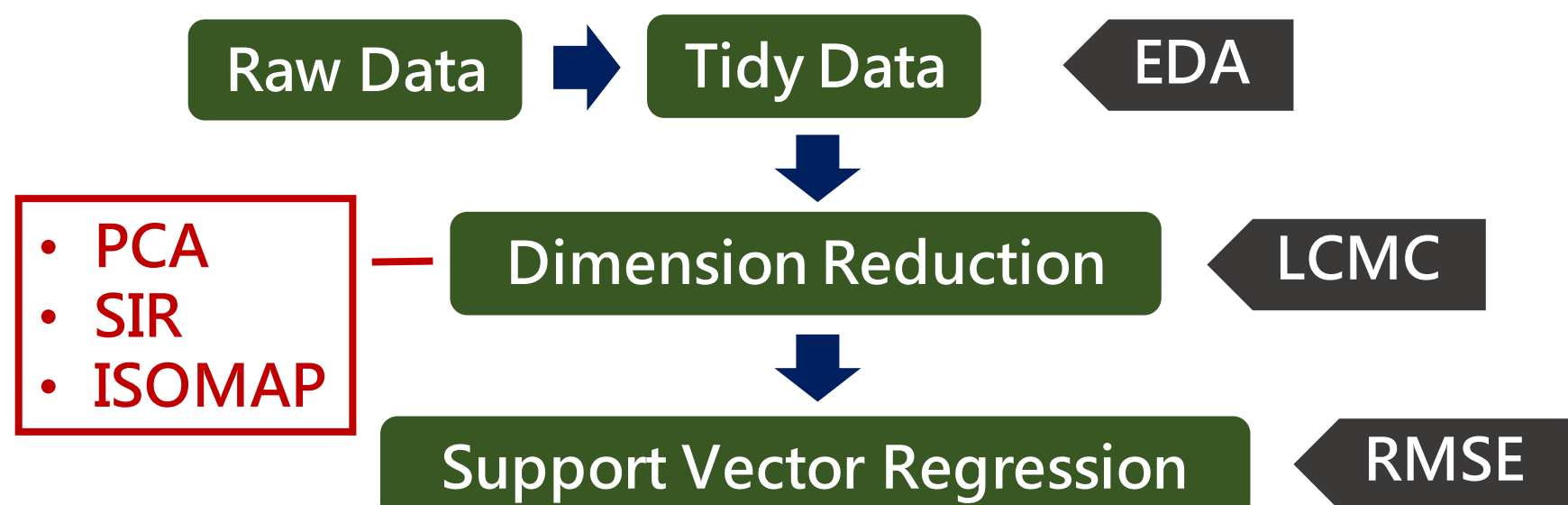
國立臺北大學經濟學系



研究目的

電力對於民生安全及經濟產業十分重要，因此我們決定藉由預測企業電力需求，探討主成分分析法、切片逆迴歸法和等軸距特徵映射法在支持向量迴歸模型的表現。

研究流程



研究資料

企業電力需求與產業發展及景氣循環有關，因此以進口總額、出口總額、領先景氣指標、工業生產指數、國際原油價格、各類消費者物價基本分類指數、就業人口數、經常性薪資、勞動生產力指數等多個經濟指標作為預測變數。

研究方法

- 動態時間扭曲法**
為一距離衡量方法。利用遞迴計算兩向量的最短累積距離，且適合用於序列資料。(Hui Ding等人，2008)
- 階層式分群法**
將距離相近的資料組成群組，並向上聚合而成一樹狀結構。本研究以動態時間扭曲法定義距離，並以此法及相關係數篩選預測變數的類別。(Johnson，1967)
- LCMC**
為一維度縮減法的衡量指標，藉由比較維度縮減前後鄰近資料點間相對距離的差異，進而判斷維度縮減的優劣。(Chen等人，2006)
- 主成分分析法**
為一線性維度縮減法。由特徵值分解將資料投影至變異最大之方向。本研究選擇累積解釋變異93%的四個主成分，意即將資料縮減至四維(圖 1)。(Pearson，1901)
- 切片逆迴歸法**
為一考量反應變數的線性維度縮減法。將資料依反應變數排序、切片，計算各切片平均，再由特徵值分解找出有效維度縮減方向。本研究以LCMC決定切片數為18，以卡方統計量決定將資料縮減至七維(表 1)。(Li，1991)
- 等軸距特徵映射法**
為一非線性維度縮減法。計算資料點間最近鄰居的距離，透過鄰居間距離衡量非鄰近點的距離，最後以多元尺度分析法找出維度縮減的方向。本研究以LCMC決定鄰居數為44，將資料縮減至四維(圖 2)。(J. C.，2000)

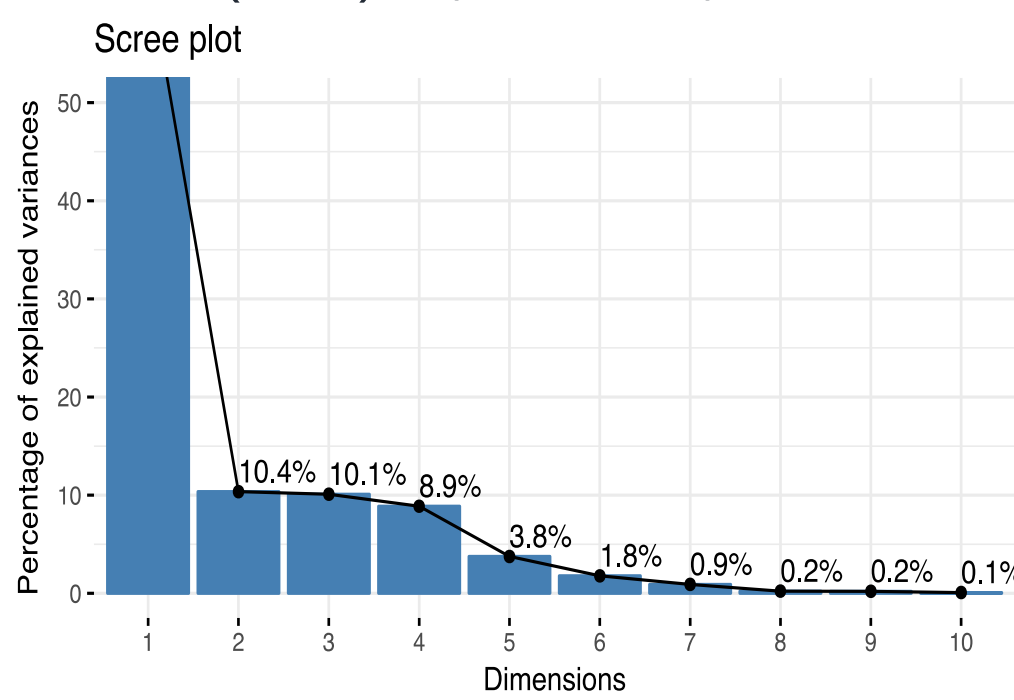


圖 1

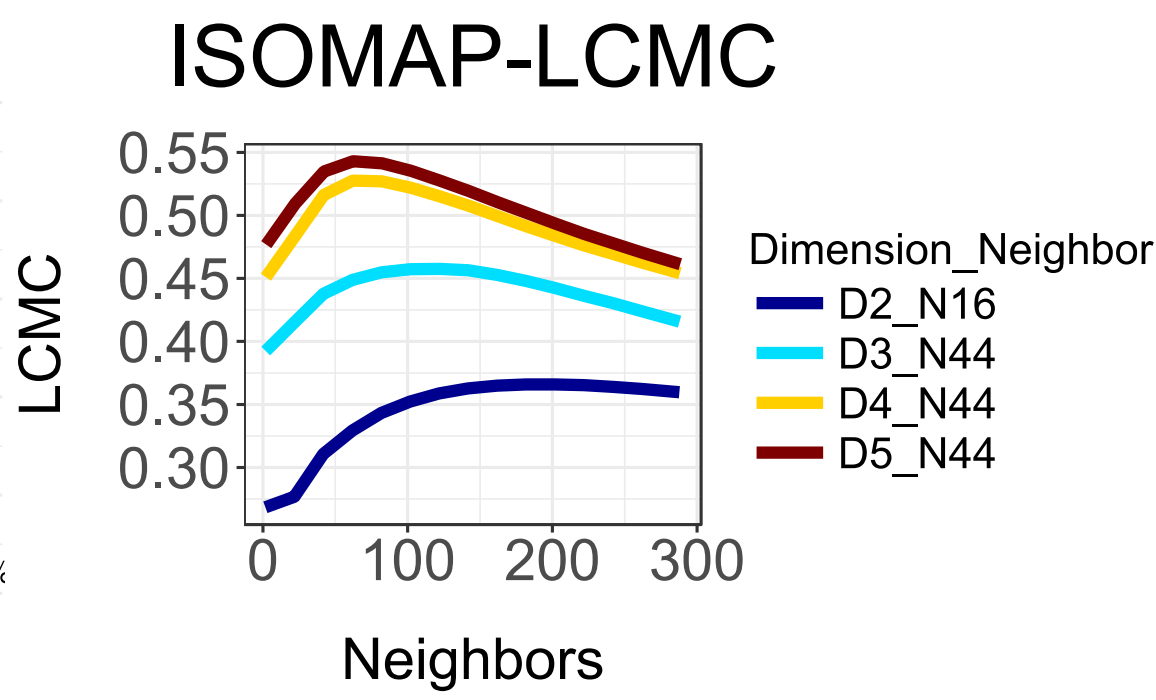


圖 2

表 1

	Stat	df	p-value
0D vs >= 1D	8361	170	0.000e+00
1D vs >= 2D	3689	144	0.000e+00
2D vs >= 3D	2472	120	0.000e+00
3D vs >= 4D	1499	98	0.000e+00
4D vs >= 5D	824.5	78	0.000e+00
5D vs >= 6D	292.6	60	0.000e+00
6D vs >= 7D	113.4	44	4.924e-08
7D vs >= 8D	31.67	30	0.3829

模型建立

本研究建立的未來一期 至三期SVR 模型，如下：

$$Y_{t+1} = \beta_1 Y_t + DR(X_t^1, X_t^2, \dots, X_t^{10})$$

$$Y_{t+2} = \beta_1 Y_{t+1} + \beta_2 Y_t + DR(X_t^1, X_t^2, \dots, X_t^{10})$$

$$Y_{t+3} = \beta_1 Y_{t+2} + \beta_2 Y_{t+1} + \beta_3 Y_t + DR(X_t^1, X_t^2, \dots, X_t^{10})$$

模型考量過去用電量的資訊，而DR為維度縮減法的函數。本研究進行13次Time Slice交叉驗證選擇最佳參數，並以RMSE作為模型預測誤差的衡量標準。

研究結果

不論是未來一期、二期或三期預測，預測值皆與實際值相差一個垂直截距項，本研究認為可能是因為我們運用的維度縮減方法皆無考慮序列關係，而運用SIR得到較好的預測結果可能是因為此法有考慮企業用電量的變異程度(圖 4~7)。從RMSE的結果可知，三種維度縮減法在未來二期的模型皆有嚴重過度配適的情形，且維度縮減法不僅無達到減緩維度災難問題，預測能力也沒有比無維度縮減的模型佳(表 2)。

Predicted Result (PCA)

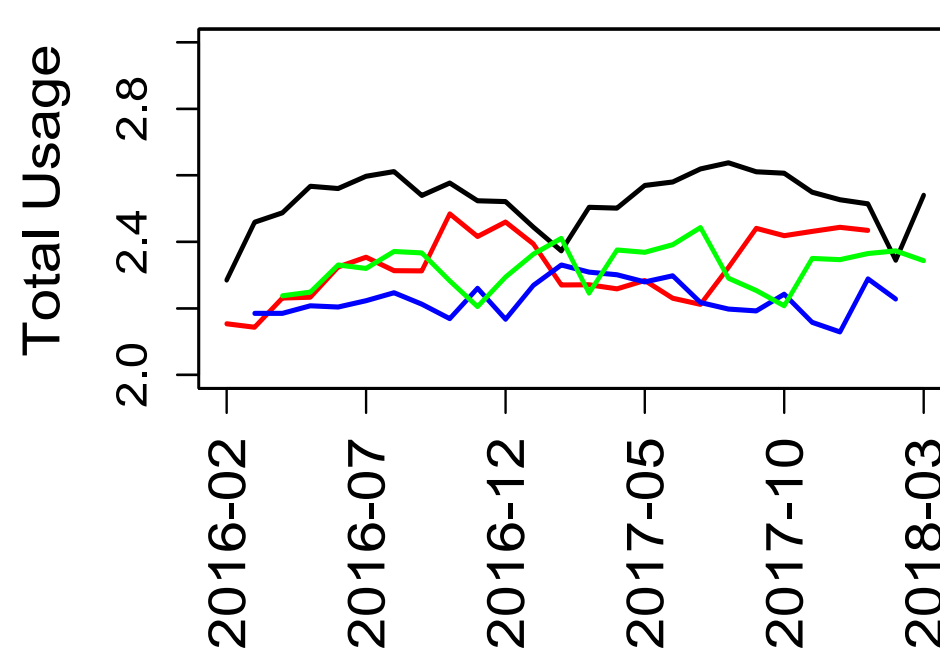


圖 4

Predicted Result (SIR)

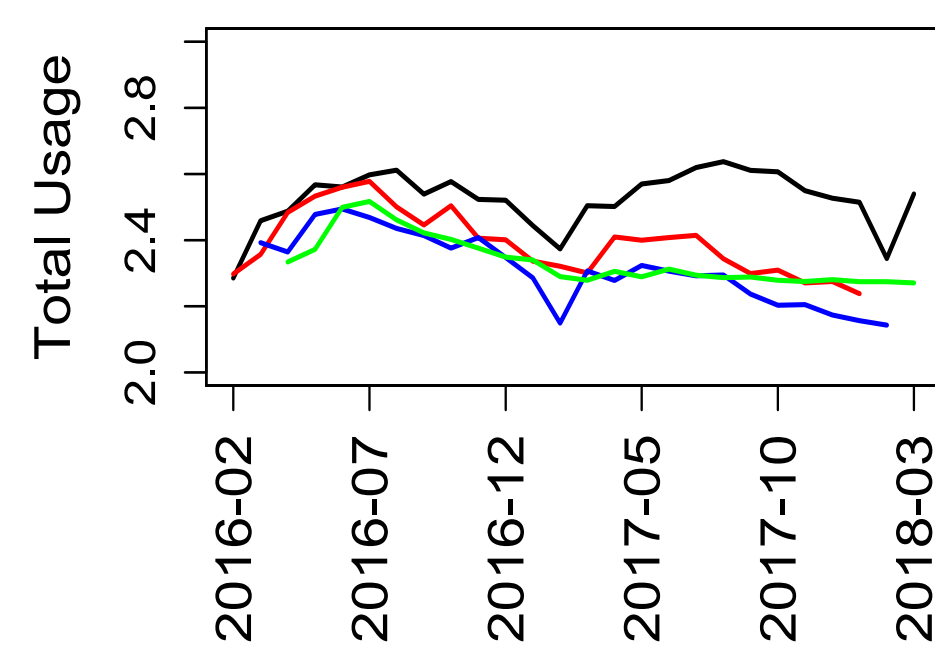


圖 5

Predicted Result (ISOMAP)

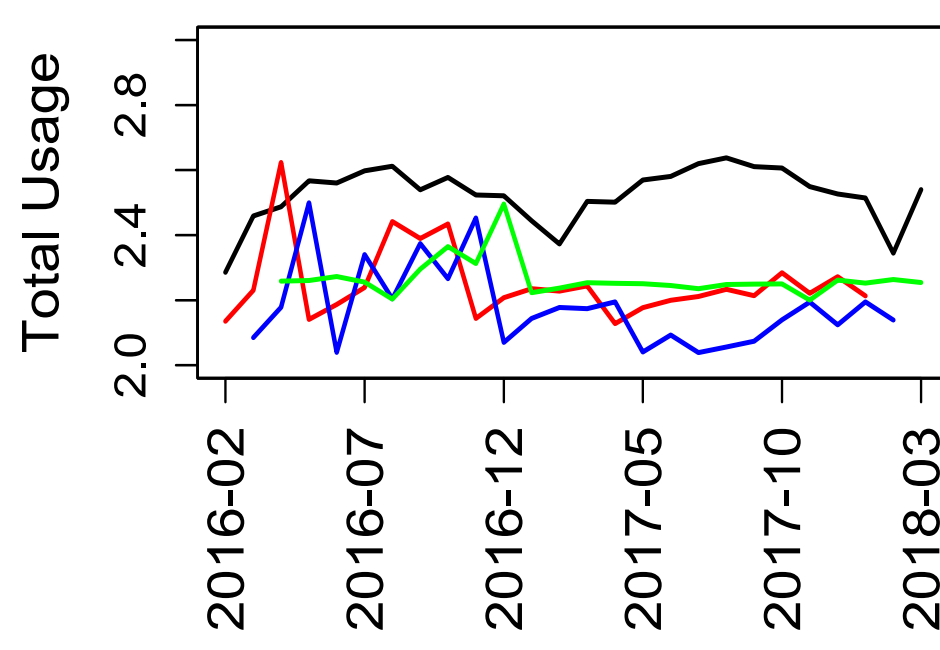


圖 6

Predicted Result (RAW)

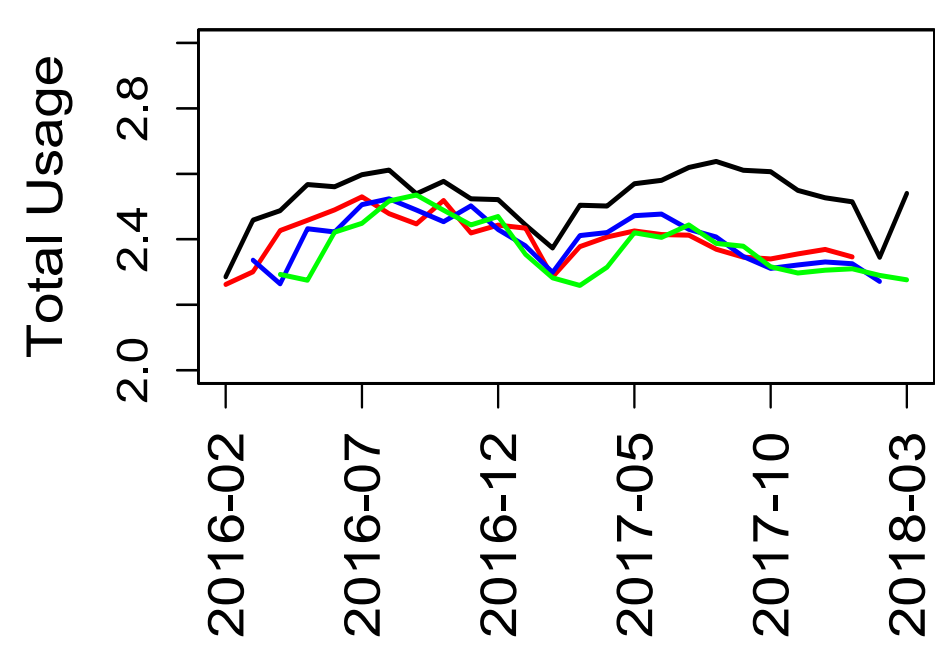


圖 7

● 一期預測 ● 二期預測 ● 三期預測 ● 實際值

表 2

		PCA	SIR	ISOMAP	RAW
One-Step Ahead	Train	0.1086	0.1712	0.1202	0.1773
	Test	0.2338	0.1802	0.2766	0.1506
Two-Step Ahead	Train	0.1263	0.1636	0.1270	0.1767
	Test	0.3287	0.3445	0.3485	0.1592
Three-Step Ahead	Train	0.1066	0.1493	0.1222	0.1667
	Test	0.2452	0.2723	0.2709	0.1889

註：RAW代表無使用維度縮減法

結論

三種維度縮減方法：PCA、SIR、ISOMAP，並無達到降低RMSE的效果，未來可以嘗試加入電價、氣溫等變數，及採用考慮序列關係的維度縮減方法，或許能夠獲得更好的預測結果。

參考資料

- Kavaklioglu, Kadir. (2010). Modeling and prediction of Turkey's electricity consumption using Support Vector Regression. *Applied Energy* Volume 88, Issue 1, January 2011, Pages 368-375.
- Lam, Joseph C., H.L. Tang, and Danny H.W. Li. (2007). Seasonal variations in residential and commercial sector electricity consumption in Hong Kong. *Energy* Volume 33, Issue 3, March 2008, Pages 513-523.
- Mohamed, Zaid, and Pat Bodger. (2003). Forecasting electricity consumption in New Zealand using economic and demographic variables. *Energy* Volume 30, Issue 10, July 2005, Pages 1833-1843.
- Xuchan Ju, Manjin Cheng, Yuhong Xia, Fuqiang Quo, Yingjie Tian. (2014). Support Vector Regression and Time Series Analysis for the Forecasting of Bayannur's Total Water Requirement. *Procedia Computer Science* Volume 31, 2014, Pages 523-531.