

# 大數據就是大洞察力—以風電場營運為例

Graham Dudgeon 博士，能源產業行銷技術經理，MathWorks 公司

## 簡介

進行風電場操作性能下降之相關風險管理，是實現生產與收入目標的關鍵所在。愈能從風電場營運獲得更多洞察與理解，風險管理則愈有效。洞察和理解的根本立基，需要數據資料的採集和存儲。現代風力發電場的大量數據，以及風電場含有許多相似規格的資產之獨特特性，意味著風電場運營商能夠從他們的大數據獲得大洞察力。

本文介紹如何進行資料解析(analytic)任務，以洞察瞭解最大可用的電力是否可達成，或者是否/哪裡可能發生短缺。

透過回答“是什麼”以及“在哪裡，”風電場工程師可以更了解實現生產目標所伴隨的風險，並從上述的了解建議補救措施以解決短缺所帶來的風險。

我們考量自然工作流程進行三項類別的資料分析，據此建立洞察並提供決策支援：描述性分析、事後診斷和預測。雖然每個類別的定義有很多，細節也各不相同，不過下列的較高層級定義可適用於本文：

**描述性分析(Descriptive analysis)**，對數據資料集的主要特性建立統計計算並視覺化呈現出來。描述性分析的例子包括如均值和方差的計算，以及視覺化資料的技術，如散點圖、直方圖，以及箱線圖等。

**事後診斷分析(Post-event diagnostic analysis)**，力求找出系統變化的因果。在這種前後關係下，系統的變化通常是突發變化，會超過事前設定的閾值(threshold)。

**預測分析(Prognostic analysis)**，試圖預測系統的變化。該系統的狀態若被觀察出相對於基準行為，則此微妙變化則被認為是趨向於某些非期望的情況。

這些定義考量了包含類型三(Type III)風力渦輪機，詳細呈現了模擬電場電力輸出及影響輸出關鍵之輸入值間的關係的探討。下列用於演示數據分析工作流程的進展所呈現的例子僅是說明性的。

## 描述性分析

從系統操作收集來之數據的主要特性，是透過標準的統計計算並以視覺化的方法呈現出來。描述性統計的主要價值是告知並形成後續分析的核心。描述性分析更豐富、更準確的大數據資料集，意味著風電場的營運可以更長，更能獲得你想要的資訊。

我們首先考慮風速和電力輸出的關係，通常類似於圖 1。請注意，輸入和輸出之間並無一對一的關係，而是在於給定的風速，從可能的電力輸出值，從該平均值和標準差可以計算一個區值。如圖 1 所示的鼓出像船殼般的形狀，這是一種方便和快速的計算來定義表面的邊界。

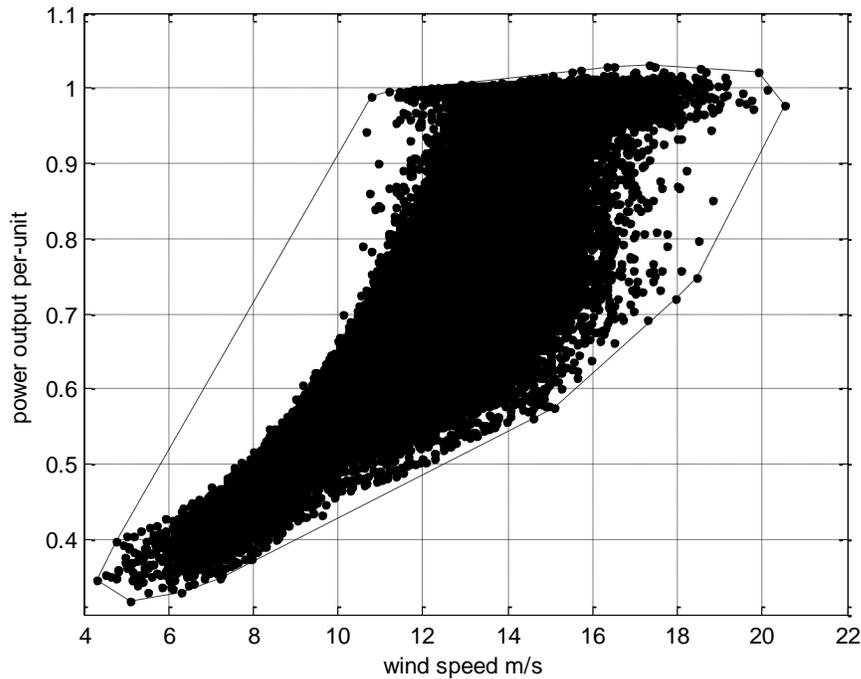


圖 1. 風力渦輪機風速與電力輸出之離散點圖範例，鼓出的船殼形以虛線標示。

許多風電場具有多個同類規格的風力渦輪機，這事實意味著描述性分析的輸出將構成事後診斷和預測分析一個堅實的基礎。這樣做的原因是，以電力輸出與風速之間關係的比較基準做為之後的黃金參考是較為容易的。透過對多個相似規格的渦輪機進行比較，那麼最有效邊界條件的渦輪機則可被識別出來。圖 2 顯示出了具有類似規格的兩個渦輪機邊界的例子。邊界條件若關掉，在給定的風速下，渦輪機 1 和渦輪機 2 相比則呈現出電力輸出降低的情形，渦輪機 1 比渦輪機 2 效率較低。在預測部分我們將返回這個範例再討論。

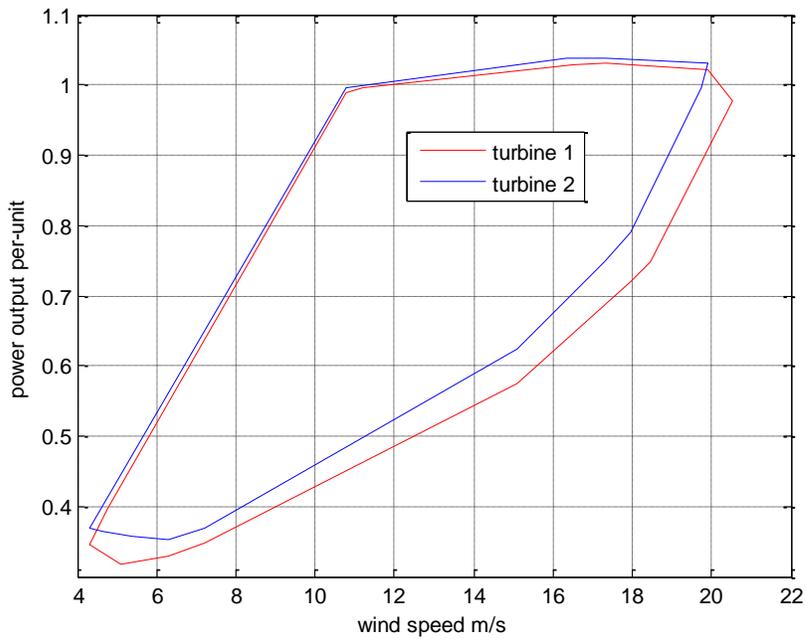


圖2，兩個具有類似規格之風力渦輪機的邊界條件圖。

### 事後診斷分析

圖3顯示出風速與電力輸出的表面圖，其中可以看出，電力輸出已經下降到所預期的邊界條件以外。這表明一些事件已發生，而且該事件是不利於預期效能的。圖3並沒有洞察到影響電力減少的原因，因此更進一步的輸入條件必須加以考慮進去。在這種情況下，標準診斷將該被應用和輸入到該系統中進行事件前和事件後的檢查。

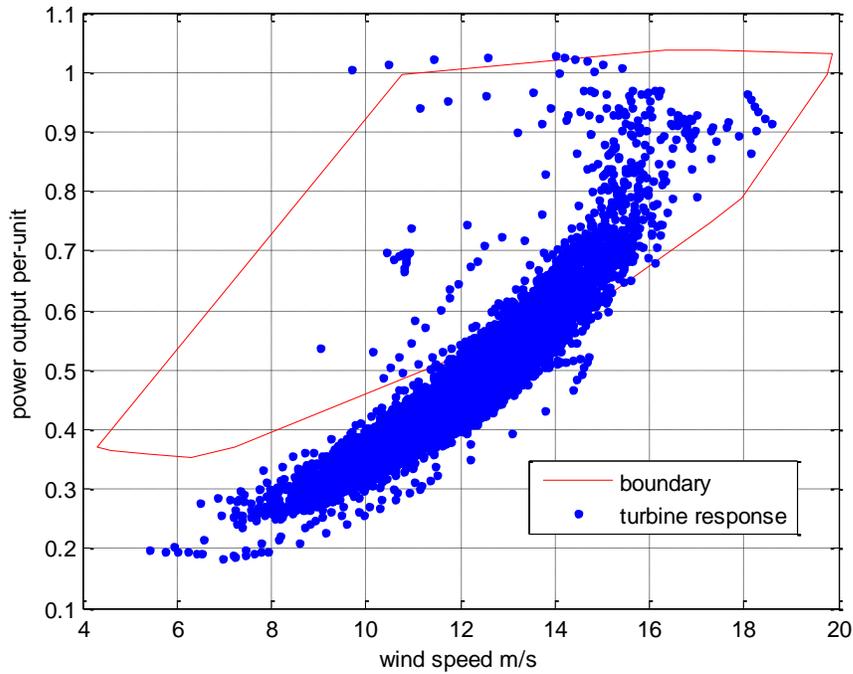


圖 3. 電力產能減少且超過預期運作的邊界條件之外。

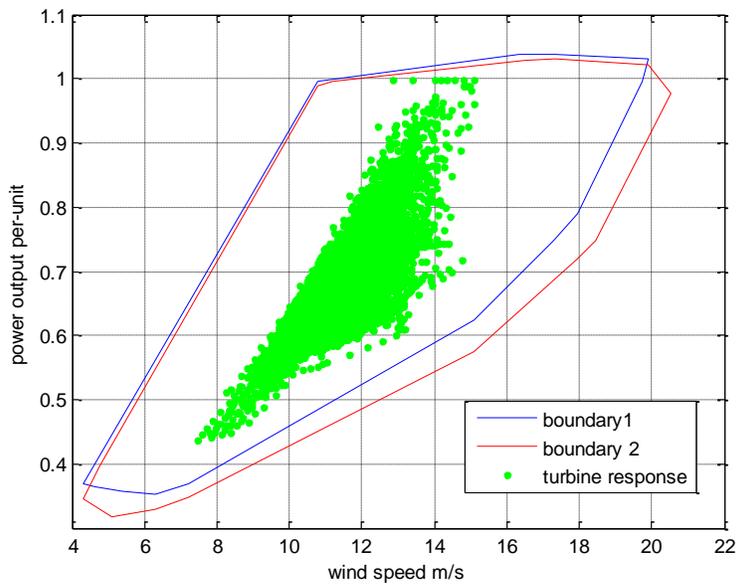


圖 4. 系統操作都在效能預測的邊界內。

圖 4 顯示出風力渦輪機的響應僅在邊界層內運作。從這個數據顯示，沒有任何跡象表明渦輪機是否的運行是否在退化的狀態，作為監視系統，只依賴於違反邊界條件與否將不足達成預警的功用。這只能做為發展預測模型時一種有價值的補充工具，以協助檢視風機性能退化的檢測。

## 預測分析

預測分析的核心是建立預測模型。儘管許多不同的建模典範存在，如 ARMAX、迴歸樹 (regression trees) 和類神經網路 (neural networks)，但它們仍都依賴於必須一數據資料集所呈現、操作時需要有足夠的輸入和輸出來建構，才能形成一個準確的模型。

資料數據的選擇須要具有所需操作條件的代表性，且基於邊界條件的計算如圖 2 所示。邊界可以由相似規格的各個風力渦輪所構成，所述具有最低表面積邊界的風力渦輪機，可以被選擇出來做為參考。

接下來建立一個模型預測動力輸出，可以考慮多個輸入，如風速、葉片槳距、電磁轉矩參考，以及發電機軸的轉速等。使用不同組合的輸入值可以開發出許多不同的模型，並從中選擇最適當的模型結構。這些模型然後可以在附加的數據集進行測試，以確定其在識別性能下降模型時的穩健性。

圖 5 顯示量測數據的資料和兩個預測模型輸出之間的均方誤差 (MSE) — 一個迴歸樹和一個類神經網路。在這種情況下，渦輪機 20 - 30 與渦輪機 10 - 20 和渦輪機 30 - 40 相比之下性能退化，且是很大的下降量。在這種情況下，性能下降來自於軸摩擦力的增加；而模型和測得的輸出值之間的誤差，也指出性能退化的程度。因此工程團隊可以使用這些訊息來幫助做出何種適當的決定與行動，如進行定期維護計劃。

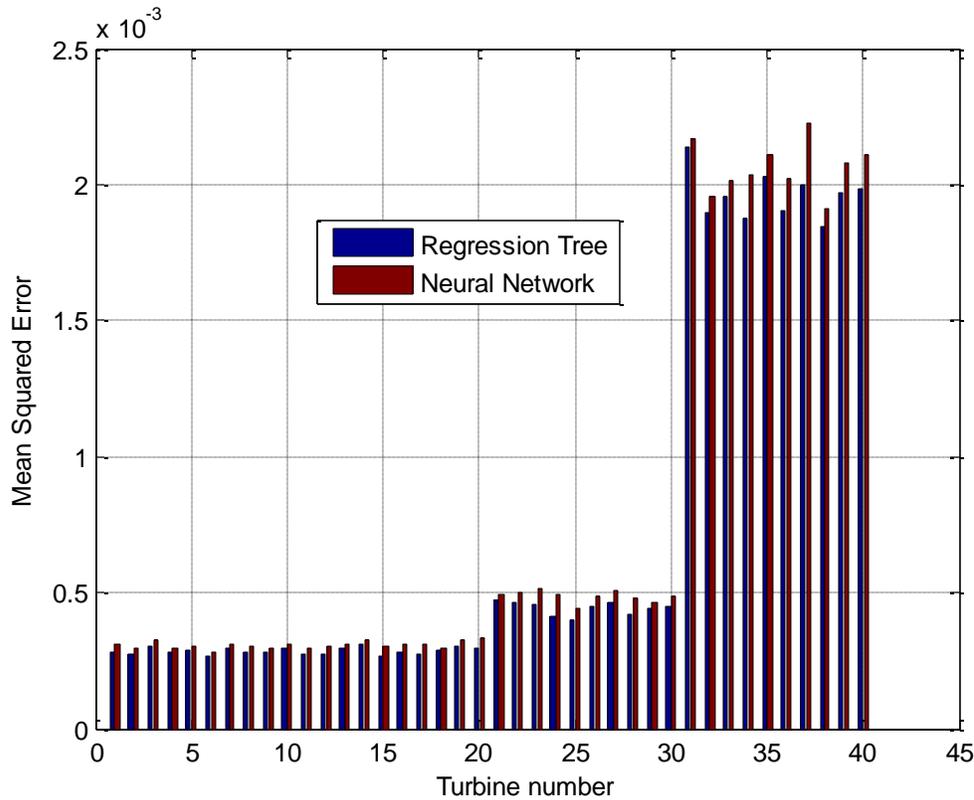


圖 5. 渦輪輸出與預測模型比較的均方誤差(MSE)

### 摘要

進行風電場操作性能下降之相關風險管理，是達成生產與收入目標的關鍵所在。愈能從風電場營運獲得更多洞察與理解，風險管理則愈有效。從現代風力發電場所獲得的浩大數據，再加上一個事實，即風電場擁有許多類似規格風機資產所產生的資料數據，構成了預測分析技術的強大基礎，其目的是能識別出風機的性能下降。性能降低的初期就可以被檢測到，這是工程團隊及時解決退化，採取必要行動和做出明智決定的最佳時機。