

利用 BMA 進行颱風系集降雨預報

黃煜鈞¹ 張惠玲² 洪景山³ 蕭朱杏¹

台灣大學流行病學與預防醫學研究所¹ 中央氣象局科技中心²

中央氣象局資訊中心³

摘要

為攫取預報過程中的各種不確定性，利用系集方式進行颱風降雨量的預報，是現今世界各國的預報趨勢。其中，貝氏模型平均(Bayesian Model Averaging, BMA)能針對不同系集成員給予不同權重，不僅較為彈性，也能產生機率型的預報。然而，隨著科技的進步，產生的系集成員數量越來越多，可能導致貝氏模型的建構因為系集成員數的增加，而影響預報最終的表現。本研究對於台灣不同地區的平地與山區降雨量進行研究，利用 BMA 進行颱風系集降雨預報。首先，在訓練資料中以平均格點均方根誤差(Region-based root mean squared error, R-RMSE)來評估各個系集成員的預報表現，再根據 R-RMSE 數值之大小排序，逐步增加系集成員數於 BMA 分析中，並同步應用於相同侵台路徑的颱風降雨之預報。最後，本研究將利用驗證資料，評估 BMA 的預報表現，並與其他系集預報方法進行比較。

關鍵字：貝氏模型平均、颱風系集降雨預報、平均格點均方根誤差

一、前言

對於颱風降雨量的預報，系集預報可以提供我們對於預報的不確定性，彙整系集成員的預報值來做預報則是現今常使用之方法之一。

BMA 方法則是給予各個系集成員不同之權重，並且以混合機率模型的形式呈現，可以提供機率型的預報，也是現今常使用之方法。隨著科技的進步，如今可以產生的系集成員數量越來越多，以提供我們更多對於預報的不確定性，然而，這些系集成員的預報模式可能會有些共通之處，例如對於平地地區，某些成員的預報可能較為接近，或是某些成員的預報較為準確；反之，有某些成員的預報表現可能較差，甚至帶有雜訊，因此影響最終的預報結果。如果可以將這些系集成員在過去的預報表現納入之後的預報當中，可以使得往後進行預報時，能擁有更多的資訊，對於系集成員預報值的使用也可以更有彈性。本研究將利用訓練

集之資料對於台灣不同區域的平地與山區的降雨量進行分析，研究不同地區與不同地形格點的系集成員的預報表現，並以此結果進行 BMA 分析，及應用於相同侵台路徑之颱風降雨預報。

舉例來說，2013 年蘇利颱風與 2015 年杜鵑颱風同屬第二類侵台颱風路徑，因此本研究先將蘇利颱風過境台灣的累積降雨量資料當作訓練集，杜鵑颱風過境台灣的累積降雨量資料當作驗證集；在訓練集中，以平均格點均方根誤差來評估各個系集成員的預報表現，並將此結果應用於杜鵑颱風，以一次加入一個系集成員的預報方式，對於驗證集進行系集降雨預報。最後，本研究將利用驗證資料，評估 BMA 的預報表現，並與系集平均之預報結果比較。

我們的研究目標分為兩個部分，第一個部分利用蘇利颱風當作訓練集，使用 BMA 對於杜鵑颱風的累積降雨量進行預報，比較 BMA 與系集平均的預報表現；第二個部分利用訓練集

資料所求出之 R-RMSE 數值，以一次加入一個系集成員的方式逐步增加系集成員數量，研究順序性(sequential) BMA 與系集預報的預報表現，並討論是否可以不需使用到所有系集成員的預報值，或是將預報表現較差的成員給予較低的權重，即可得到更好的預報結果。

二、資料來源

在以下的分析中，我們採用兩個颱風的降雨量預報資料檔，分別為蘇利颱風，過境時間為 2013 年 7 月 11 日至 2013 年 7 月 14 日，與杜鵑颱風，過境時間為 2015 年 9 月 26 日至 2015 年 9 月 29 日。先選擇這兩個颱風進行分析的原因為，根據氣象局颱風資料庫所提供的資訊，這兩個颱風的侵台路徑均屬於 2 類；臺灣地區的颱風降雨深受地形鎖定效應(terrain-locking effect)所影響，所以這兩個颱風的降雨特性可能有相似之處，因此，我們希望可以藉由蘇利颱風的資料，從中獲得各個系集成員的預報表現，以應用於對於杜鵑颱風的降雨量預報。

本研究所採用的預報資料來自中央氣象局 WRF 系集預報系統(WRF ensemble prediction system, WEPS)所提供的 0-24 小時雨量預報，共有 20 組系集成員預報。蘇利颱風過境期間共有 7 報，而杜鵑颱風過境期間有 13 報。因為考慮到報與報之間資料的獨立性問題，在本研究中，暫時先使用初始時間相同的資料集進行分析，對於蘇利颱風使用了 4 報的資料，均是初始時間為 00UTC 對於未來 24 小時的累積降雨量預報檔，對於杜鵑颱風亦是使用了 4 報的資料，初始時間為 18UTC 對於未來 24 小時的累積降雨量預報檔。

為了避免資料分析之結果受到地區與地形的影響，以下分析將對不同地區與不同地形分別進行研究。將台灣分為 6 個地區，分別為北部、中部、南部、東北部、東部與東南部；格點之地形高度以 500 公尺做為劃分，高於 500

公尺定義為山區格點，低於 500 公尺定義為平地格點。因為本研究的目標聚焦在台灣陸地區域降雨量的預報與預報評估，因此離島地區之資料將先不納入以下的分析當中。

三、研究方法

本研究使用的方法為 BMA 與傳統系集預報(即系集平均, Ensemble mean)。BMA 是一種後處理方法，對於每一個系集成員建構預測分配，且加上對每個系集成員所估計出來的權重，得到最終的預測混合模型。本研究使用 BMA 方法的原因有兩個，第一個是因為經由建構預測分配，可以取得更多對於預報的不確定性，且亦可以產生機率型的預報；第二個原因為，BMA 的建構可以經由權重來調整各個系集成員在預報時所提供的資訊貢獻度，預報表現較好的成員會給予較高之權重，反之則給予較低的權重；這不只納入了更多系集成員預報的資訊，在預報上也較為彈性。

以下將經由三個部分來介紹研究方法：(一)先利用 BMA 進行颱風降雨量預報 (Slougher et al. 2007)，BMA 的預測模型如下方方程式，其中 y 代表觀測值， f_k 代表系集成員預報值，本分析中有 20 個系集成員。式子中的， w_k 代表權重； $I[\]$ 為指標函數(Indicator function)，中括弧內條件成立即為 1，不成立即為 0。

$$p(y|f_1, \dots, f_{20}) = \sum_{k=1}^{20} w_k \{p(y = 0|f_k) I[y = 0] + p(y > 0|f_k) g_k(y|f_k) I[y > 0]\}$$

由於累積降雨量資料中會有數值為 0 的情況，即沒有下雨，只使用常見之伽碼分配(Gamma distribution)來建模並不恰當，因此所使用的 BMA 模型中會以羅吉斯迴歸(logistic regression)來估計累積降雨量為 0 之機率，對於觀測值大於 0 的部分，則維持使用伽碼分配來建模。羅吉斯迴歸模型如下， δ_k 為指標函

數，若系集預報值為 0，則 $\delta_k=1$ ，其餘為 0， a_{0k} ， a_{1k} 與 a_{2k} 為必須估計之參數。

$$\text{logit } p(y = 0|f_k) = a_{0k} + a_{1k}f_k^{1/3} + a_{2k}\delta_k, \\ k = 1, 2, \dots, 20$$

累積降雨觀測值大於 0 的部分，會利用伽碼分配來建模，伽碼分配如下，其中， α_k 與 β_k 為必須估計之參數。

$$g_k(y|f_k) = \frac{1}{\beta_k^{\alpha_k} \Gamma(\alpha_k)} y^{\alpha_k-1} \exp\left(-\frac{y}{\beta_k}\right), \\ k = 1, 2, \dots, 20$$

(二) 我們利用平均格點均方根誤差(R-RMSE)來評估各個系集成員在訓練集的預報表現。R-RMSE 公式如下，對於每一位系集成員，計算研究地區的所有格點在訓練集四天中預報值之均方根誤差，並且考慮每個地區的平地格點與山區格點數量不同，因此除上格點數以標準化 SSE 數值。計算出各個系集成員在訓練集的 R-RMSE 數值之後，再以 R-RMSE 數值大小來評估各個系集成員的預報表現，數值越小，代表預報表現越好。

$$R - RMSE_{member\ i} \\ = \sqrt{\frac{\sum_{k \in \{selected\ grid\}} \sum_{j=1}^4 (f_{ijk} - y_{ijk})^2}{number\ of\ selected\ grid * 4}}, \\ i = 1, 2, \dots, 20$$

(三) 藉由第二步驟所計算出各個系集成員的 R-RMSE 數值，可以將各個系集成員在訓練集的預報表現進行排序，以一次一位系集成員的方式逐步增加系集成員數量，對於驗證集進行順序性的 BMA 預測與順序性系集平均預報。舉例來說，第一次放入在訓練集預報表現最好的系集成員預報值，建構 BMA 與系集平均預報以對驗證集進行預報，第二次放入在訓練集預報表現前兩名的系集成員預報值，以此方式至放入所有系集成員預報值，研究順序性 BMA 與系集平均預報的預報表現。預報表現以 CRPS 值當作評估指標，CRPS 值定義如下， F 為累積機率函數， y 為觀測雨量， x 為系集預報值。

$$CRPS(F, y) = \int_{-\infty}^{\infty} (F(x) - I\{x \geq y\})^2 dx$$

四、分析結果

本研究的分析使用 R 軟體中的“ensembleBMA”套件(Fraley et al. 2018)，以下解說之圖表亦是由 R 軟體所取得之結果。這裡對於台灣陸地六個地區之平地與山區格點進行分析與研究，並利用 BMA 與系集平均兩種方法對於杜鵑颱風 2015 年 9 月 28 日至 2015 年 9 月 29 日的累積雨量進行預報。在 BMA 的建構上，將蘇利颱風視為訓練集以對 BMA 模型參數進行估計。

以下研究結果顯示，大致上 BMA 的預報表現較系集平均來的好，只有南部山區、東部平地與東南部地區之系集平均的預報表現較 BMA 來的好。順序性 BMA 與系集平均的結果顯示，有時並不需要納入所有系集成員預報值，即可達到較好的預報表現。

圖 1 為六個地區 BMA 與系集平均在平地與山區格點的 CRPS 數值比較，平地格點以實線表示，山區格點以虛線表示，紅線代表 BMA，藍線代表系集平均。以圖 1 的 (a) 北部地區為例，北部地區一共選取了 243 個格點，145 個平地格點，98 個山區格點，平地格點的部分，紅色實線均在藍色實線之下方，放入所有系集成員時，由 CRPS 值可以看出，BMA 的預報表現較系集平均來的好，山區格點也有相同的結果。逐步放入系集成員的研究結果顯示：(1)不管是 BMA 或是系集平均，並不需要採用所有的系集成員預報，即可以得到較小的 CRPS 數值；(2)採用 5 至 10 個系集成員後，再繼續增加成員數，CRPS 值並沒有明顯的變動。

中部地區、南部平地、東北部地區與東部山區採用所有系集成員時，BMA 的預報表現均比系集平均來的好；但在南部山區與東部平地採用所有系集成員時，系集平均比 BMA 來的

好。順序性 BMA 與系集平均亦顯示，並不需要採用所有系集成員，即可得到較小 CRPS 值的資訊；東南部地區的系集平均優於 BMA，但兩者之 CRPS 數值沒有明顯的差距。表 1 為各個系集成員在訓練集的 R-RMSE 排名與數值，順序性 BMA 與系集平均均是以此結果進行分析，以北部地區平地格點為例，在訓練集預報表現最好的為 6 號成員，R-RMSE 數值為 64.6，在訓練集預報表現最差的為 10 號成員，R-RMSE 數值為 136.7。

五、結論與討論

本研究討論 BMA 與系集平均的預報表現，且利用訓練集之資料，以 R-RMSE 來評估各個系集成員的預報表現，並以一次一個系集成員的方式逐步增加系集成員數量，分析順序性 BMA 與系集平均在 CRPS 數值上的變化。在 BMA 的建構上，我們利用蘇利颱風當作訓練集，預報杜鵑颱風 2015 年 9 月 28 日至 2015 年 9 月 29 日的累積降雨量，與系集平均的預報結果進行比較。在此研究中，對於驗證集的累積雨量預報，大致上 BMA 的預報表現較系集平均的表現來的好。且由順序性系集成員的研究結果可以得知，並不需要使用所有系集成員，即可以達到較好的預報表現，也就是往後進行系集預報時，可以考慮給予訓練集預報表現較差之系集成員較低的權重，甚至給予 0 之權重，以加權的概念，求得最終的系集預報。

本研究目前用來評估預報表現的準則為 CRPS 數值，接下來將會納入更多的評估準則，來加以驗證得到的分析結果。另外，本研究目前以氣象局颱風資料庫上所提供的颱風侵襲路徑資訊，先選取一個颱風的資料當作訓練集，接下來的研究會更進一步討論如何選取訓練集的資料，希望可以納入更多的颱風資料當作訓練集，增加統計上的代表性，以應用在未

來侵台颱風的降雨預報。此外，目前只針對侵台路徑為 2 類的颱風進行分析，後續也將持續針對其他不同侵台路徑的颱風進行分析，希望能得到與此研究一致的分析結果。

本研究顯示：利用訓練集各個系集成員預報表現，對於驗證集之降雨量預報有一定的預報能力；由此可知：有相同侵台路徑之颱風，在相同地區與相同的地形下，系集成員之預報能力會較為相似。如何將訓練集的系集成員預報表現指標，與驗證集的系集成員預報表現指標進行配對與確認，並且納入颱風資料時續性的特性，對於不同時間點進行預報，以及動態更新系集成員的預報表現，將會是未來研究的重要目標。

六、參考文獻

1. 張語軒、張庭槐、吳蕙如, 2012: “貝氏模型平均應用於臺灣地區溫度機率預報”, 氣象學報, 49 卷, 第 1 期, 19-37
2. Fraley, C., Raftery, A. E., Slughter, J. M., Gneiting, T., 2018: “Package ‘ensembleBMA”
3. Raftery, A. E., Gneiting, T., Balabdaoui, F., and Polakowski, M., 2005: “Using Bayesian Model Averaging to Calibrate Forecast Ensembles”, *Monthly Weather Review*, 133, 1155-1174
4. Slughter, J. M., Raftery, A. E., Gneiting, T., Frailey, C., 2007: “Probabilistic Quantitative Precipitation Forecasting Using Bayesian Model Averaging”, *Monthly Weather Review*, 135, 3209-3220
5. Slughter, J. M., Gneiting, T., Raftery, A. E., 2010: “Probabilistic wind speed forecasting using ensembles and Bayesian model averaging”, *Journal of the American Statistical Association*, 105(489), 25-35

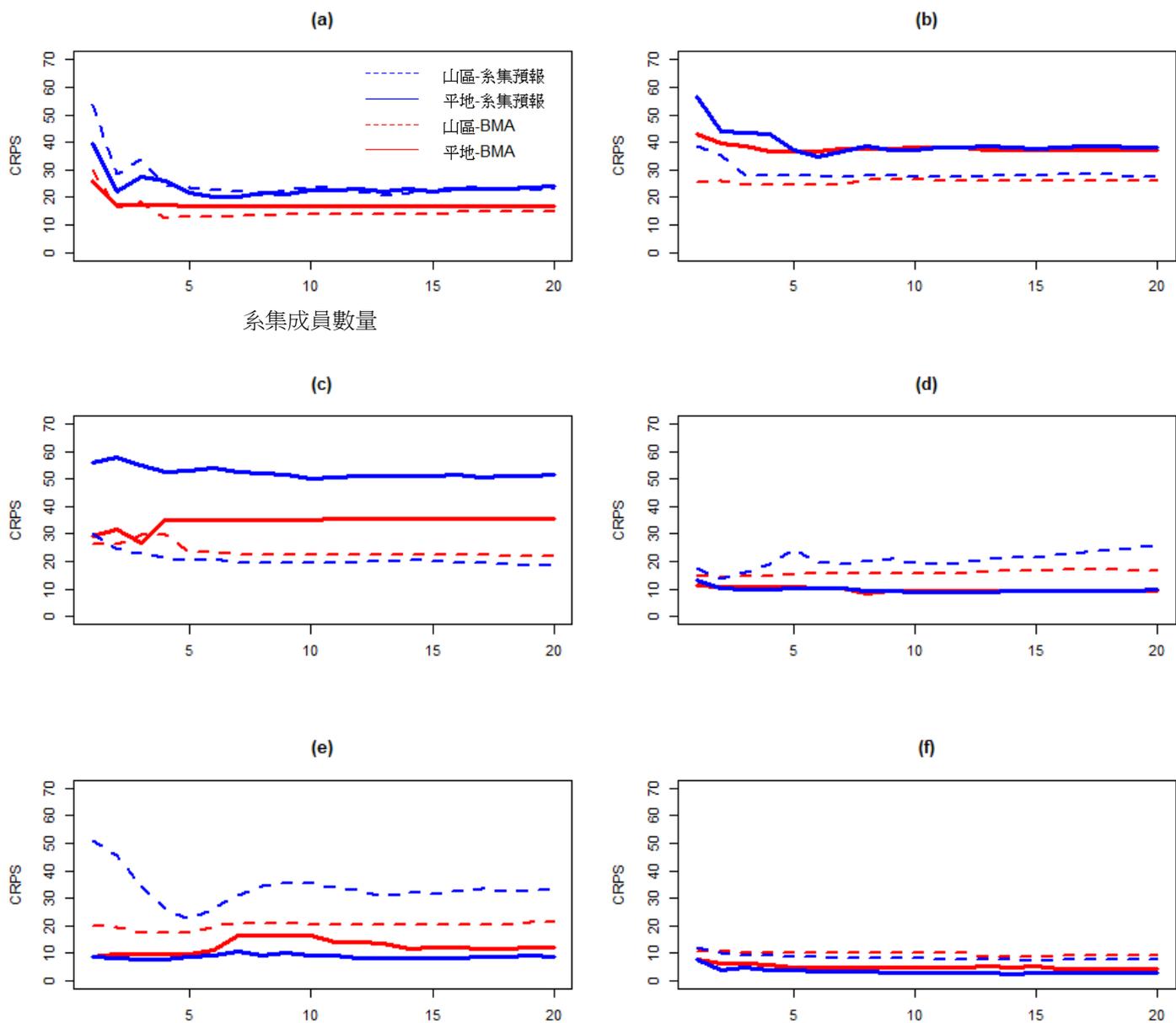


圖 1 BMA 與系集預報對於杜鵑颱風影響台灣期間(2015 年 9 月 28 日 18UTC ~ 2015 年 9 月 29 日 18UTC)平地與山區格點降雨量預報之 CRPS 數值，(a) 北部地區(145 個平地格點、98 個山區格點) (b) 中部地區(184 個平地格點、199 個山區格點) (c) 南部地區(157 個平地格點、90 個山區格點) (d) 東北部地區(27 個平地格點、55 個山區格點) (e) 東部地區(39 個平地格點、122 個山區格點) (f) 東南部地區(59 個平地格點、107 個山區格點)。

表格 1 各個系集成員在訓練集之 R-RMSE 排名與數值

北部平地		北部山區		中部平地		中部山區		南部平地		南部山區	
Member	R-RMSE										
V6	64.6	V9	121.7	V11	30.7	V11	74	V15	21.2	V16	59.2
V13	66.9	V19	122.8	V18	36.2	V17	76.8	V13	21.9	V3	65.1
V5	70.5	V17	126.6	V19	36.5	V1	84.6	V16	22.7	V17	66.8
V16	75.2	V13	128.7	V2	39.8	V7	84.7	V2	24.2	V14	68.1
V1	75.3	V6	133.9	V9	41.4	V13	85.1	V11	24.6	V2	68.7
V11	79.1	V11	136.2	V17	47.8	V5	85.1	V4	26.7	V15	68.9
V14	80	V7	141.4	V13	49.7	V9	87.3	V5	26.8	V1	69.8
V9	80.8	V18	144.7	V7	49.8	V2	87.4	V18	27.1	V20	70.4
V18	82.1	V5	148.1	V15	51.8	V3	90.9	V14	28	V6	71.2
V15	84.3	V15	149.5	V16	52.1	V6	91.6	V9	28.1	V4	71.4
V7	84.5	V14	150.3	V1	53.1	V19	92.5	V1	28.4	V13	73.5
V2	86.2	V20	156.8	V5	53.5	V18	92.7	V3	28.7	V12	77.7
V20	87.6	V1	161.1	V4	54.2	V15	94.3	V12	29.9	V11	81.8
V3	89.9	V2	164	V8	54.7	V14	94.9	V20	30.6	V7	84.3
V19	92.1	V3	167.1	V10	55.2	V8	95.2	V19	30.9	V18	85.5
V8	92.9	V8	169.5	V6	55.9	V12	100.1	V7	31.2	V9	89.3
V12	95.7	V12	172.1	V20	58.2	V16	102.4	V17	32.4	V5	94.9
V4	98.9	V16	174.5	V3	60.3	V20	103.4	V6	39.8	V8	99.5
V17	111.2	V4	187.4	V14	60.8	V4	103.4	V8	55.8	V19	101.5
V10	136.7	V10	195.9	V12	66.4	V10	109.4	V10	57.8	V10	129.8

東北部平地		東北部山區		東部平地		東部山區		東南部平地		東南部山區	
Member	R-RMSE										
V14	38	V14	45.9	V1	6.7	V7	43.1	V1	21.2	V1	35.7
V8	40.3	V18	46.7	V4	6.9	V6	45.1	V7	25	V16	40.1
V20	40.4	V20	48.6	V13	6.9	V1	45.3	V3	26.1	V20	40.7
V5	41	V17	51.6	V6	7.6	V13	45.4	V16	28	V7	41.7
V19	42.2	V5	51.7	V17	7.9	V18	46.2	V19	28	V15	44.2
V9	43.9	V13	52.4	V15	7.9	V5	48.7	V15	29.1	V3	46.1
V13	44.5	V1	52.7	V3	8.7	V3	49.3	V17	29.5	V17	46.5
V11	44.7	V9	54	V20	9.5	V17	50.1	V11	29.5	V14	46.5
V18	46.2	V6	55.3	V16	10.7	V16	50.2	V13	29.5	V6	47.2
V1	47.2	V4	58.2	V18	12	V9	50.8	V6	29.8	V11	51.7
V16	48.2	V19	59.9	V19	12.3	V14	51.3	V14	30	V4	52.4
V6	50.1	V8	60.1	V11	13.3	V20	52.2	V20	30.3	V13	53.3
V7	51.1	V7	61	V10	14.7	V4	52.4	V4	31.6	V2	53.5
V17	51.3	V16	61	V2	15.5	V15	53.2	V18	32.3	V12	55.7
V10	53.5	V11	61.8	V8	15.9	V11	57.3	V9	34.7	V18	62.5
V12	59.1	V15	66.2	V7	17.6	V12	61.1	V2	35.5	V8	66
V2	59.3	V10	74.9	V9	19.2	V8	62.9	V12	38.2	V9	68.2
V4	59.3	V12	75.4	V12	24	V19	63.4	V5	47.1	V19	68.6
V15	63.4	V2	75.8	V5	24.6	V2	65.6	V8	52.1	V5	69
V3	68.9	V3	87.7	V14	27.7	V10	73.7	V10	68.2	V10	95.9