

統計計算與模擬期末報告

模擬氣候因子影響蔬果價格之探討 以蘋果為例

111354022 鄭雲昊

摘要

蔬果價格在各種不同天氣氣候影響下 (如: 當地雨量、溫度等...) 時常會有所變化，在此次報告中，我使用蘋果作為母體，將現有的蔬果價格作為對照組，透過課堂上所學到的模擬隨機亂數，根據台灣的台北地區條件自行模擬出當地現有的氣候因素類別變數，再透過這些因素配適出模擬的蔬菜水果價格作為實驗組，使用不同的假設檢定來探討兩組間是否有明顯的差異，並觀察兩組蔬菜水果價格的波動差異，做出結論。

壹、引言

蘋果是屬於異花授粉植物類，大部分品種自花不能結成果實 (維基百科)，蘋果的生產是透過種植種子於土壤上，種子利用土壤的養分及水分來生長，開花並最終成樹，結成果實。

蘋果開花期是基於各地氣候而定，但一般集中在 4-5 月份。一般蘋果栽種後，於 2-3 年才開始結果。果實成長期之長短，一般早熟品種為 65-87 天，中熟品種為 90-133 天，晚熟品種則為 137-168 天。(維基百科)。

台灣蘋果的產季與產地，由於蘋果屬於溫帶水果，需要有長時間的低溫及特殊日照條件，因此台灣蘋果主要都種植在海拔 1500-2500 公尺以上的高山，例如大梨山地區的梨山、武陵農場、福壽山農場、清境農場、大禹嶺等地都是本土蘋果的故鄉。至於台灣蘋果的產季大部分會落在 8-10 月的夏秋之際才會開始，而且因為氣溫變化大，可能一兩週內就會全部成熟。(Luise_Lin , 2021)

在台灣，平常蘋果生長條件主要可分為以下四點：

1、溫度：

前段有提到，由於蘋果是溫帶水果，所以栽培蘋果的地區溫度，必須要低於常溫下，全年平均的氣溫約於 8-14 度，成長期 4-10 月約為 12-18 度，冬季最冷的 1 月份在 1-10 度之間，才可滿足蘋果冬眠的最低要求；如果在成長期 6-8 月份間平均氣溫為 18-24 度，則有利於果實成長、著色及糖分的累積，可影響價格上幅波動。

2、雨量和空氣濕度：

蘋果的成長周期間，所需要的水量相當於 540 毫米 (mm) 的降水量。而我國大部分蘋果產區降水量在 500-800 毫米 (mm) 之間，其中降水有 70 80% 間集中在 6-8 月。

3、光照：

蘋果的成長周期年日照不低於 1500 小時，適當的年日照在 2200 小時以上。

4、土壤：

適合蘋果生長的土壤 pH 值最好在 5.6-6.7 之間最佳，但在實際上，吻合需求的土壤較少，需要透過一些改良技術去優化。(濃情農意, 2017)

從蘋果生長條件中可以了解，適當的雨量及濕度對於蘋果生長佔有重要的一環，但如果持續下雨量過多，可能會影響到有關樹幹、枝或果實很多層方面的問題，像是當日光照量下降、果實出現些許裂痕、著色不鮮豔、掉落，長期下過量的雨會造成果樹爛根、死根甚至死樹，造成蘋果出口量下降，影響價格。

貳、研究特色

蘋果資料詳情：我主要使用來預測與模擬的蔬果是蘋果；蔬果的農產批發市場來源主要為 109_ 台北市場一，地點於”臺北市萬華區萬大路 533 號”，近水源快速交流道、華中河濱公園及新店溪；期間為民國 101 年 1 月 民國 112 年 4 月；蘋果品種主要以在台灣內部出產，如果考慮國外進口，則還要另外考慮在運輸工具上時的內部溫度、濕度，使用什麼儀器進行蘋果的保藏等等皆會影響到蘋果的新鮮程度，所以本篇暫不考慮，只考慮台灣本土出口蘋果；蘋果單位使用及價格為上中下盤的平均價 (台幣元/公斤)。

蘋果來源地的部分，原本的蘋果出產地在台北市場一裡大多是來自各個不同山上的水果商提供，各個山的生長環境大有不同，但是在這份報告中會假設出產地在台灣地區台中的梨山地區出產，以控制溫度、空氣濕度、降雨量、光照及土壤等詳情條件一致。使用的相關資料主要以梨山觀測站 (C0F861) 收集的資料為主。

此報告會利用梨山上歷年來的平均資料，這部分去建立後面的模型以進行模擬氣候因子，但其中由於土壤，數值上比較難以去量化且也難以去收集每月的土壤的 pH 值，所以會假設以真實地區的條件去處理。會主要針對前三個生長條件：每月平均溫度、每月降雨量及每月日照時數 (小時) 去模擬兩個月後的數值去配適價格之迴歸模型。(要考慮到從蘋果包裝到批發的過程時，此篇報告預計二個月後蘋果到達市場)

根據梨山觀測站歷年收集的資料分成冬、春秋和夏季。

平均氣溫：冬季 12 月-3 月為 7 度-13 度；春秋季 4、5 月及 10、11 月為 11-16.3 度；夏季 6-9 月為 14.6-17.6 度。

平均降雨量：冬季 12 月-3 月為 5-552 毫米；春秋季 4、5 月及 10、11 月為 3-676 毫米；夏季 6-9 月為 37-1027.5 毫米。

平均日照時數：冬季 12 月-3 月為 100.6-225.1 小時；春秋季 4、5 月及 10、11 月為 143.8-256.1 小時；夏季 6 月-9 月為 112-222.1 小時。

順帶一提，由於此報告會著重在於模擬氣候因素去控制價格影響，所以有關市場供需的部分，例如蘋果每日交易量或其他會影響蘋果價格漲幅的條件，一樣是假設完全比照實際市場的情況，不考慮進去此次報告模擬價格計算的部分。

在 2015-2019 年時原本台灣氣象局預計四季應當平衡，但卻出現了夏季持續天數大於冬季持續天數，在未來這種極端氣候會越來越多，本次報告為了瞭解在未來極端氣候對於蘋果的影響為何，所以選擇 2017 年 1 月做為切割資料點，2012-2016 年 12 月為 training data；2017 年 1 月-2023 年 5 月 1 日為 testing data。

參、文獻探討與研究方法

在與此次報告相關模擬蔬菜價格的部分，有一篇是研究四季溫度對於大眾蔬菜，例如：甘藍菜、結球白菜及花椰菜，對產量和價格變化模式，使用不同的多元迴歸方式去配適其相關性並檢驗，其生長溫度是使

用前人的文獻所得去建立模型關係，發現使用 multiple regression 配適之模型其複判定係數 (coefficient of multiple determination) R^2 value，甘藍菜、結球白菜與花椰菜分別為 0.029、0.129 及 0.163。與本篇報告最大差別於，本篇生長溫度使用台灣實際觀測站所得資料進行多元迴歸模型配適，其模擬的蔬果價格準確性較高。

在此次報告中，首先將整筆資料透過某個極端氣候作為分割點分成兩部分，第一部分實際價格及其資料去配適多元迴歸模型，其中 y = 每月蘋果價格、 x_1 = 每月降雨量、 x_2 = 每月平均溫度、 x_3 = 每月日照時數 (小時)。

第二部分再細分成兩部分，一個部分的實際價格資料作為對照組、另一個部分就利用前言所給予的蘋果生成條件去模擬不同的數值，其中模擬的數值皆假設天氣穩定、無發生任何一個極端氣候的情況。再將其套入第一部份的多元迴歸模型，去模擬出 y 值蘋果價格，作為實驗組。

最後再去比較其實際價格與模擬價格有無差異，而後利用時間序列方法，例如：arima，套用兩組價格分別去預測未來一年內可能價格。

一、模擬蔬果價格

將當月蘋果價格的每日資料整理成月份資料，則資料的筆數如下表所示：

季節	筆數
春、秋季	21
夏季	26
冬季	23

總共 70 筆觀測值，由於蘋果量產季大多在夏季 8-10 月之間，所以夏季的月資料會較多，在夏季前幾個月是蘋果出產的冷淡期，則資料較少；然而為了後續方便比較，此次模擬組各做 9 組即可。

首先，根據梨山觀測站統計的平均氣溫模擬出每個季 (冬、春秋與夏) 的氣溫，配適出直方圖：

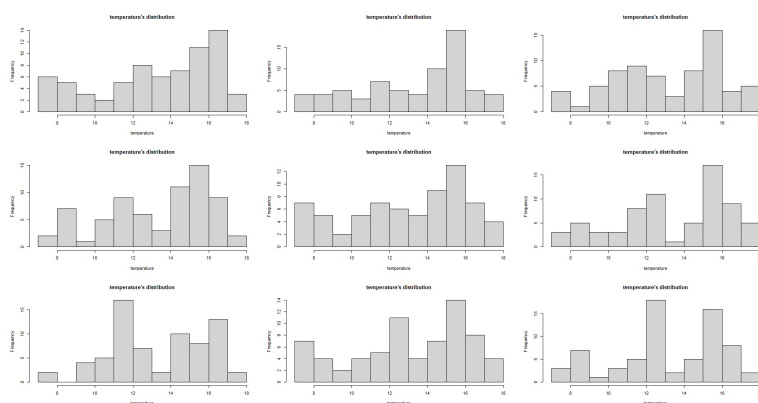


Figure 1: 模擬溫度之直方圖

圖從左上至右下分別為第 1-9 組，從這些直方圖可以了解大多數資料分布呈左尾分布，為了了解模擬的溫度資料比較符合哪種分配，所以進行分布檢定，而由於溫度的分布情況較特殊，大多都會像是隨機過程裡面的隨機漫步模型 Random Walk Model，所以我對這 9 組觀測值進行 ADF 檢定：

H0: 此組資料具有單位根 (Unit Root) 或資料非平穩性

H1: 此組資料不具單位根 (Unit Root) 或資料具平穩性

其 9 組資料 p-value 值如下所示：

組別	第 1 組	第 2 組	第 3 組	第 4 組	第 5 組	第 6 組	第 7 組	第 8 組	第 9 組
p-value	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
結果	拒絕 H0：資料具平穩性								

代表這 9 組模擬溫度都可能存在隨機漫步模型的情況，溫度資料 9 組皆穩定，所以任意取一組當作實驗組。

再來，根據梨山觀測站統計的平均降雨量模擬出每個季 (冬、春秋與夏) 的降雨量，配適直方圖：

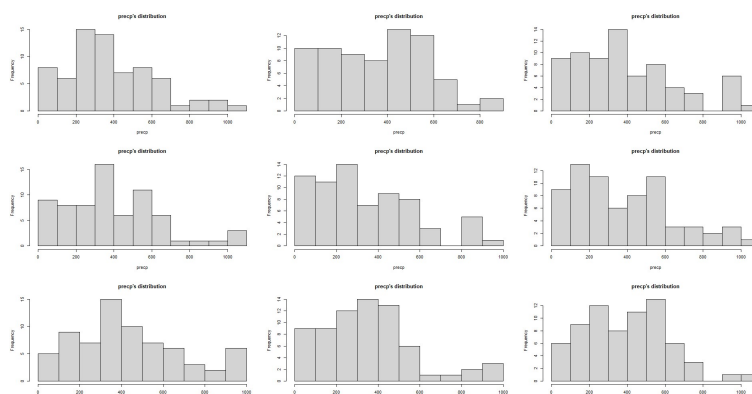


Figure 2: 模擬降雨量之直方圖

圖從左上至右下為第 1-9 組，從這些圖則可以得知其資料分布

則像右尾分布，而降雨量可以考慮使用常態分佈或是 gamma 分布檢定是否服從哪一種，所以採用 ks 檢定檢驗，其中 gamma 分布透過最大概似估計法估計出來為 $\alpha=2, \beta=\frac{1}{20}$ ，並用 ks.test 檢驗發現 p-value 皆為 <0.05 ；檢驗常態分佈發現 p-value 亦皆為 <0.05 ；最後再檢驗 uniform(0,1000) 發現 p-value 也皆為 <0.05 ，意即無一個檢測為拒絕 H_0 ，代表降雨量可能沒有服從的機率分配。

模擬降雨量盒鬚圖如下：

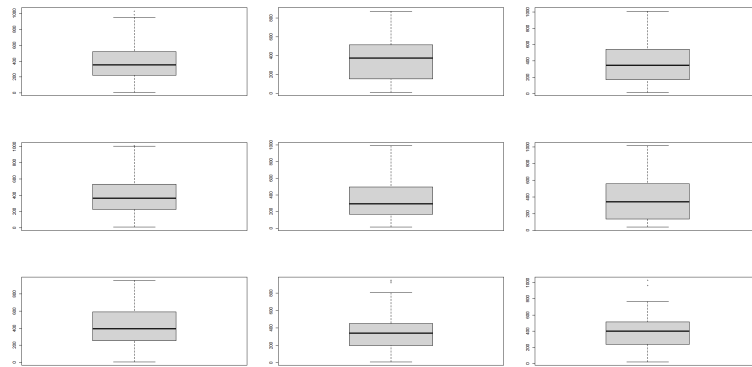


Figure 3: 模擬降雨量之盒鬚圖

由於模擬降雨量沒有服從的機率分配，所以會根據離群值多寡及蘋果對於水的需求量去選擇，於是選擇第 2 組當作實驗組。

最後，根據梨山觀測站統計的平均照光時數模擬出每個季 (冬、春秋與夏) 的照光量，配適直方圖：

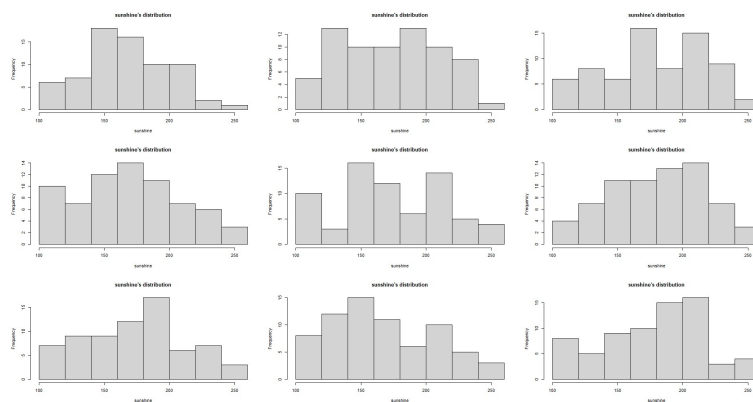


Figure 4: 模擬照光時數之直方圖

圖從左上至右下為第 1-9 組，從這些圖則可以得知其資料分布比較前兩個模擬組群來說平均，像等尾分配，所以針對這些模擬組去檢查是否服從常態分配，使用 `ks.test` 結果發現 p-value 皆為 <0.05 ；然後再使用 `ks.test` 去檢驗 `uniform(100,250)` 發現其 p-value 如下表所示：

組別	第 1 組	第 2 組	第 3 組	第 4 組	第 5 組	第 6 組	第 7 組	第 8 組	第 9 組
p-value	0.0217	0.4871	0.2874	0.3020	0.4236	0.1908	0.3020	0.1193	0.1915
結果	拒絕 H_0	不拒絕 H_0							

由上表可知，9 組裡面有 8 組服從 $U(100,250)$ ，於是擴大到抽取 100、1000 及 10000 次樣本再去測試得下表：

試驗次數	100 次	1000 次	10000 次
>0.05 次數	92	853	8652
<0.05 次數	8	147	1348

上表可以得知模擬這個日照光條件去跑試驗的次數越多，其服從 $U(100,250)$ 的情況會越來越小，只是此模擬結果檢定力算強大，所以可以假設模擬的日照光條件服從 $Uniform(100,250)$ 。

因為其模擬光照時數服從 $U(100,250)$ ，代表資料都會穩定在此區間產生，所以任意取一組作為實驗組。

二、複迴歸分析

複迴歸模型如下：

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \varepsilon$$

其中 $y =$ 價格, $x_1 =$ 溫度, $x_2 =$ 降雨量, $x_3 =$ 光照時數

在考慮交互作用下去配適複迴歸模型三次，其交互作用 p-value 表如下：

交互作用	溫度 * 降雨量	降雨量 * 光照時數	光照時數 * 溫度
p-value	0.5216	0.8402	0.0979
結果	不拒絕 H_0		

由上表可以得知其 p-value 皆 >0.05 達不顯著，所以此次複迴歸模型採用不考慮交互作用。

所以在不考慮交互作用的情況下，透過程式運算得出模型如下：

$$\hat{y} = 97.9369 - 3.4086x_1 + 0.0013x_2 + 0.0253x_3$$

此複迴歸結果表示，其複迴歸係數為 0.4161，如果單與前次文獻探討所討論之甘藍菜、結球白菜與花椰菜的複迴歸係數相比 (0.029、0.129 及 0.163)，確為較高。

三、模擬價格及比較

將前面得到的模擬氣候因子帶進去不考慮交互作用之複迴歸模型得到以下的基本資料 (實驗組)：

function	mean	median	variance	range(min,max)
value	58.07	56.47	98.91	42.34~77.76

而將批發市場提供之真實價格基本資料 (對照組) 則為：

function	mean	median	variance	range(min,max)
value	46.82	44.224	153.06	25.30~88.05

兩組盒鬚圖為：

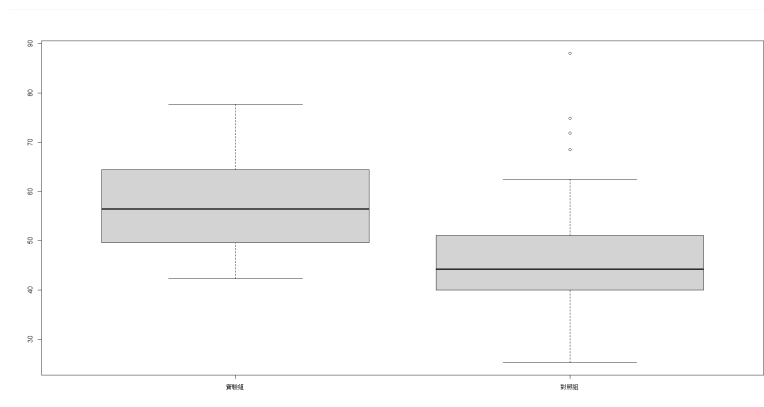


Figure 5: 兩組價格之盒鬚圖

其中左邊為實驗組 (模擬價格)；右邊為對照組 (實際價格)。

由上述可以得知，實驗組在假設天氣穩定的情況下，透過比較變異數及其最大、最小值範圍可知，價格其離散程度較對照組穩定，從盒鬚圖也可以看出，實驗組不

存在離群值，吻合天氣穩定之設定。

兩組之價格走勢圖為：

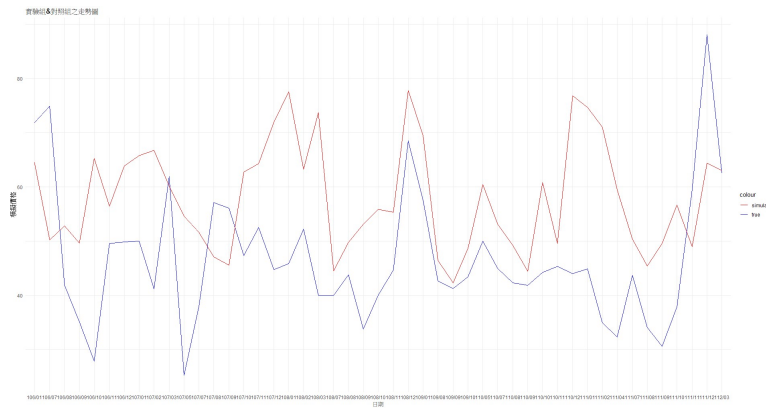


Figure 6: 兩組價格之走勢圖

其中紅線為實驗組 (模擬價格)；藍線為對照組 (實際價格)。

從這個走勢圖可以看出，實驗組模擬價格的部分價格波動幅度比較穩定，然而對照組實際價格則價格波動幅度較大些，可能受到一些氣候因素影響而造成，像是颱風、短期乾旱...

再來，為了印證前面的結果是否符合自己判斷的部分，於是檢定兩組平均數是否有差異，即 $H_0: \mu_s = \mu_t$ ，其中 μ_s = 實驗組之平均數, μ_t = 對照組之平均數。使用獨立樣本 t 檢定檢驗，其結果 $p\text{-value} < 0.05$ ，代表拒絕 H_0 ：兩組平均數有差異，再搭配走勢圖可給予適當之結論，即天氣穩不穩定對於價格有一定的影響。

而後再檢定對照組的變異數是否大於實驗組， $H_0 : \sigma_s^2 < \sigma_t^2$,

其中 $\sigma_s^2 =$ 實驗組之變異數, $\sigma_t^2 =$ 對照組之變異數。

使用 F 檢定得出其 p-value=0.07123 > 0.05，則不拒絕 $H_0 : \sigma_s^2 < \sigma_t^2$ ，則代表對照組離散程度不一定大於實驗組。

四、預測未來價格

價格預測的部分，主要使用 arima 配合 forecast 做預測，由於實驗組模擬價格資料由前面走勢圖可看出具規律平穩性，沒有季節性存在，所以其預測出來的未來價格不具參考性，所以此篇報告不做實驗組之預測價格。

對照組：

針對對照組的真實批發價格進行 arima 模型配適，找出參數 p、d 和 q 值，由 adf.test 檢定求出其 p-value = 0.07037 > 0.05，代表此筆資料不具隨機性，即 d = 0。

再來由 acf 圖搭配 pacf 圖去判斷 p 值和 q 值，其圖如下：

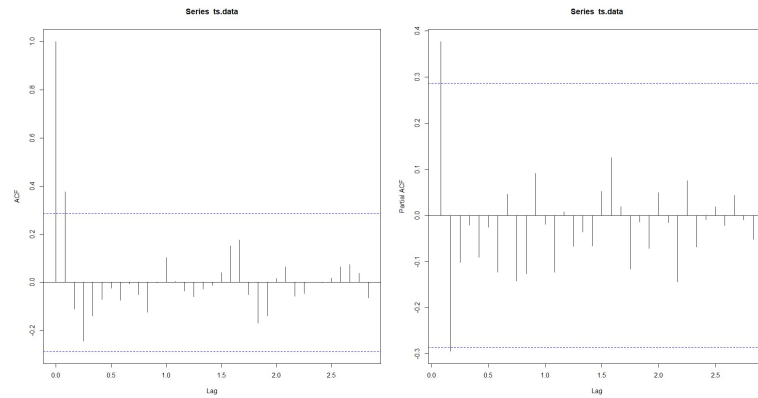


Figure 7: acf 及 pacf 圖

可由圖判斷出 $p=0$ 、 $q=1$ ，再考慮到季節性，所以配適 $ARIMA(0,0,1)(0,1,0)[12]$ 模型後再去預測其未來價格得下圖：

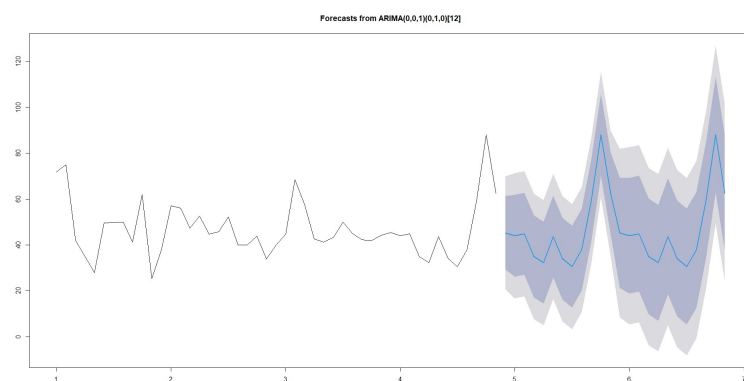


Figure 8: 預測未來價格走勢圖

其預測未來 12 個月之價格表為：

Month	1	2	3	4	5	6	7	8	9	10	11	12
Predict Price	44.0	44.9	35.0	32.3	43.7	34.1	30.6	37.8	59.5	88.0	62.5	45.3

肆、研究結果

1. 根據當地平時氣候模擬出的蘋果價格高於真實蘋果價格，可以得知如果有出現極端天氣時，其可能會影響到實際蘋果價格。
2. 實驗組的模擬價格，可以從前面的實驗過程得知，雖然可以成功透過收集當地高山資料，去模擬出蘋果價格，但其波動程度非常穩定，與實際價格之走勢圖相比，缺少季節性存在，則再使用時間序列模型做預測時，其結果會不準確。
3. 於最後預測未來價格表格中可以得知，在蘋果旺季 6-8 月時期價格最低，淡季時價格則有變高的趨勢。

伍、參考文獻

維基百科，取自 <https://zh.wikipedia.org/zh-tw/> + ”蘋果”

奧丁丁 (2021，0820)・高山蘋果盛典開賣啦 | 台灣蘋果品

種、口感必知圖鑑大公開・取自 <https://today.line.me/tw/v2/article/gm>

豆豆爺爺果園 (2019，0917)・持續降雨對蘋果樹及蘋果有

這麼多影響，你知道嗎？・取自 <https://kknews.cc/news/x6azj68.html>

中央氣象局 (2023)・COVIS 氣候資料服務系統・取自

<https://codis.cwb.gov.tw/StationData>

中央氣象局 (2023)・每月氣象・取自 <https://www.cwb.gov.tw/V8/C/C/S>

農產品批發市場交易行情站 (2023)・產品日交易行情・取

自 <https://amis.afa.gov.tw/fruit/FruitProdDayTransInfo.aspx>