

基於LSTM神經網路的線上訂房網站訂價策略研究

壹、研究目的

本研究希望開發一個在線上訂房網站上找出最佳訂房時機的工具，協助消費者用划算的價錢訂到心目中理想的飯店。為了達到此目的，本研究藉由線上訂房網站的公開房價資料，透過資料分析找出影響房價的關鍵因素，再訓練LSTM模型預測未來房價。

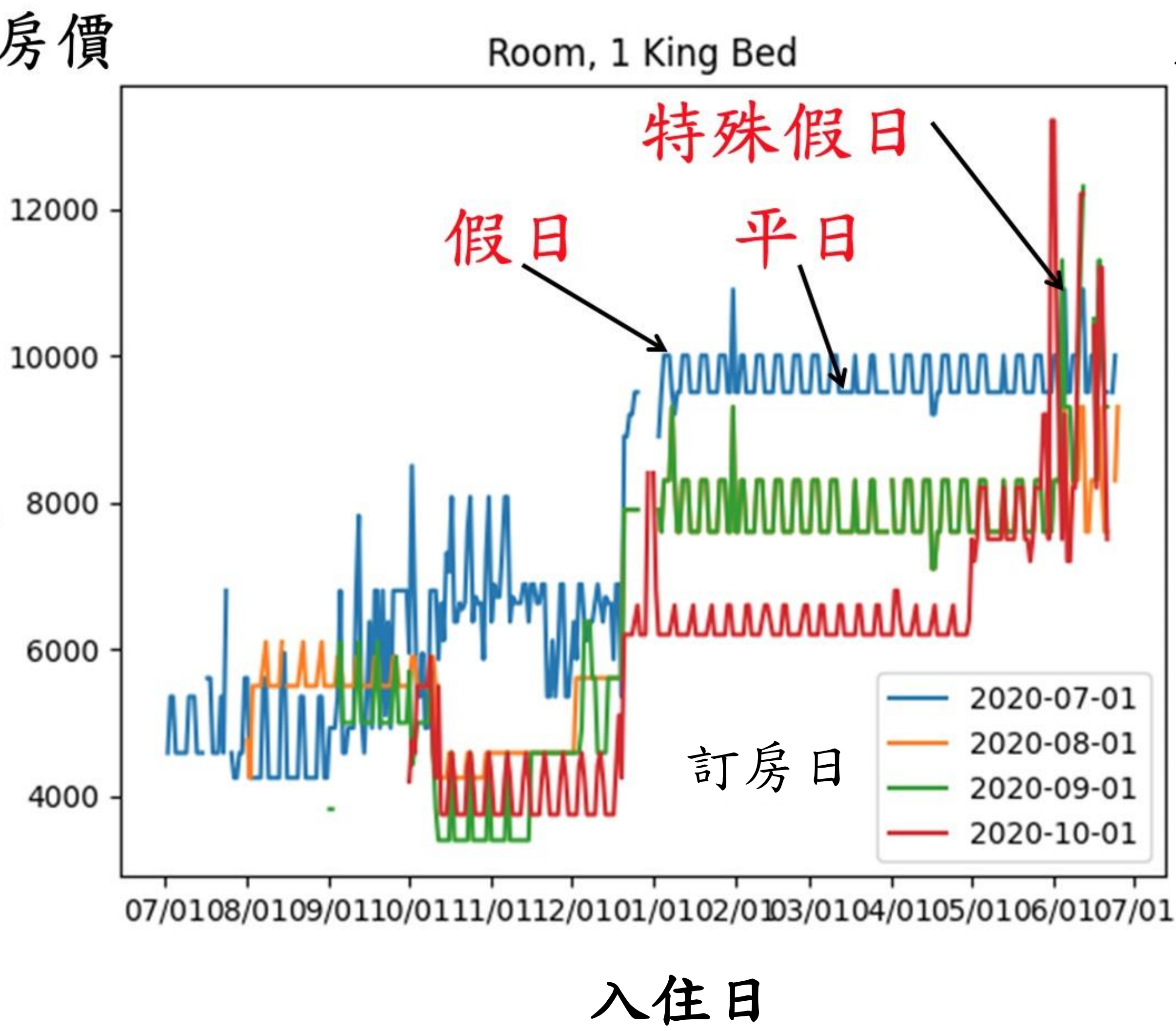
貳、研究流程

- (一) 開發爬蟲程式建立房價資料庫。
- (二) 視覺化分析，找出影響房價關鍵因素。
- (三) 訓練LSTM模型、預測未來房價走勢。

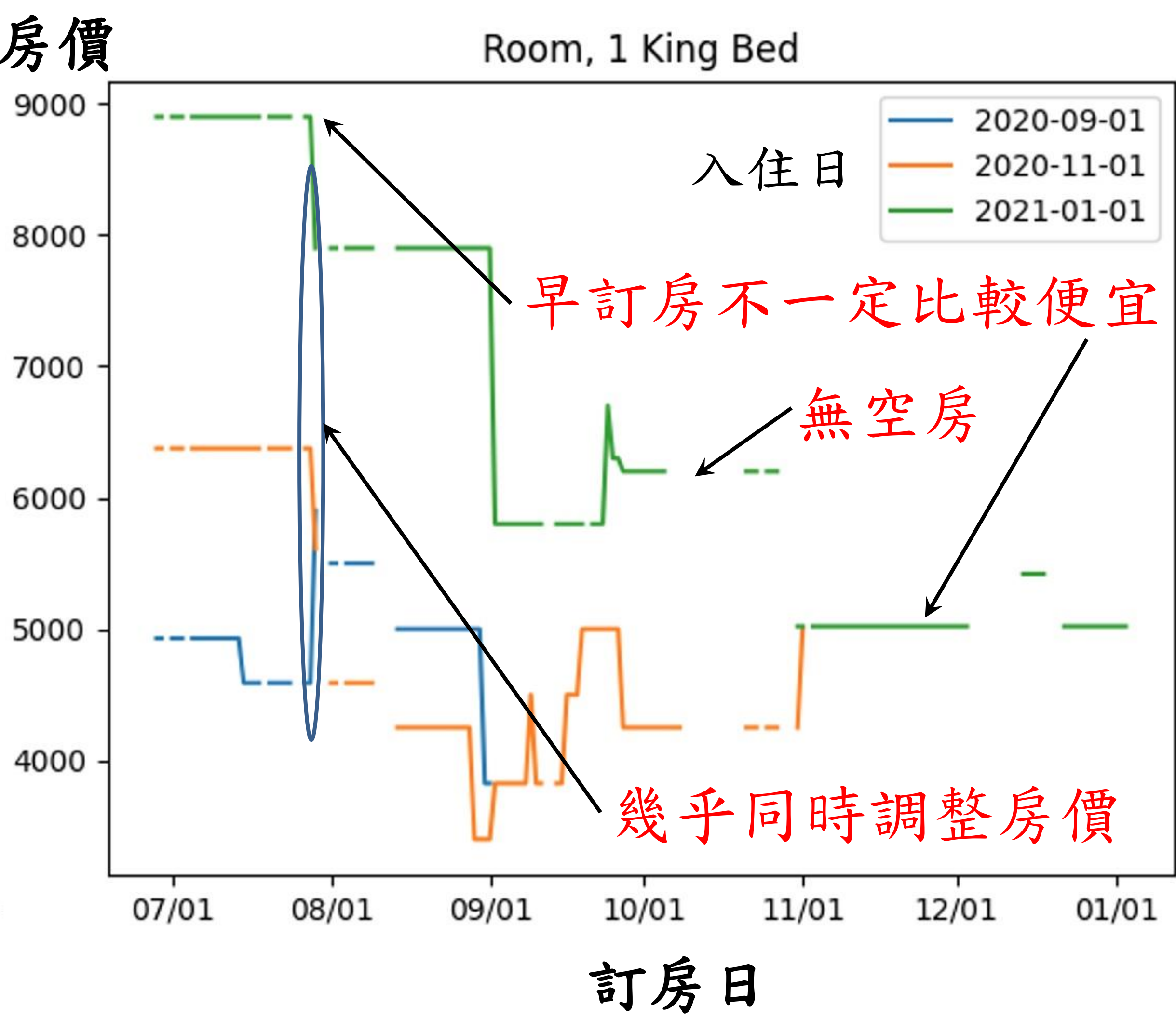
參、研究成果

一、資料分析

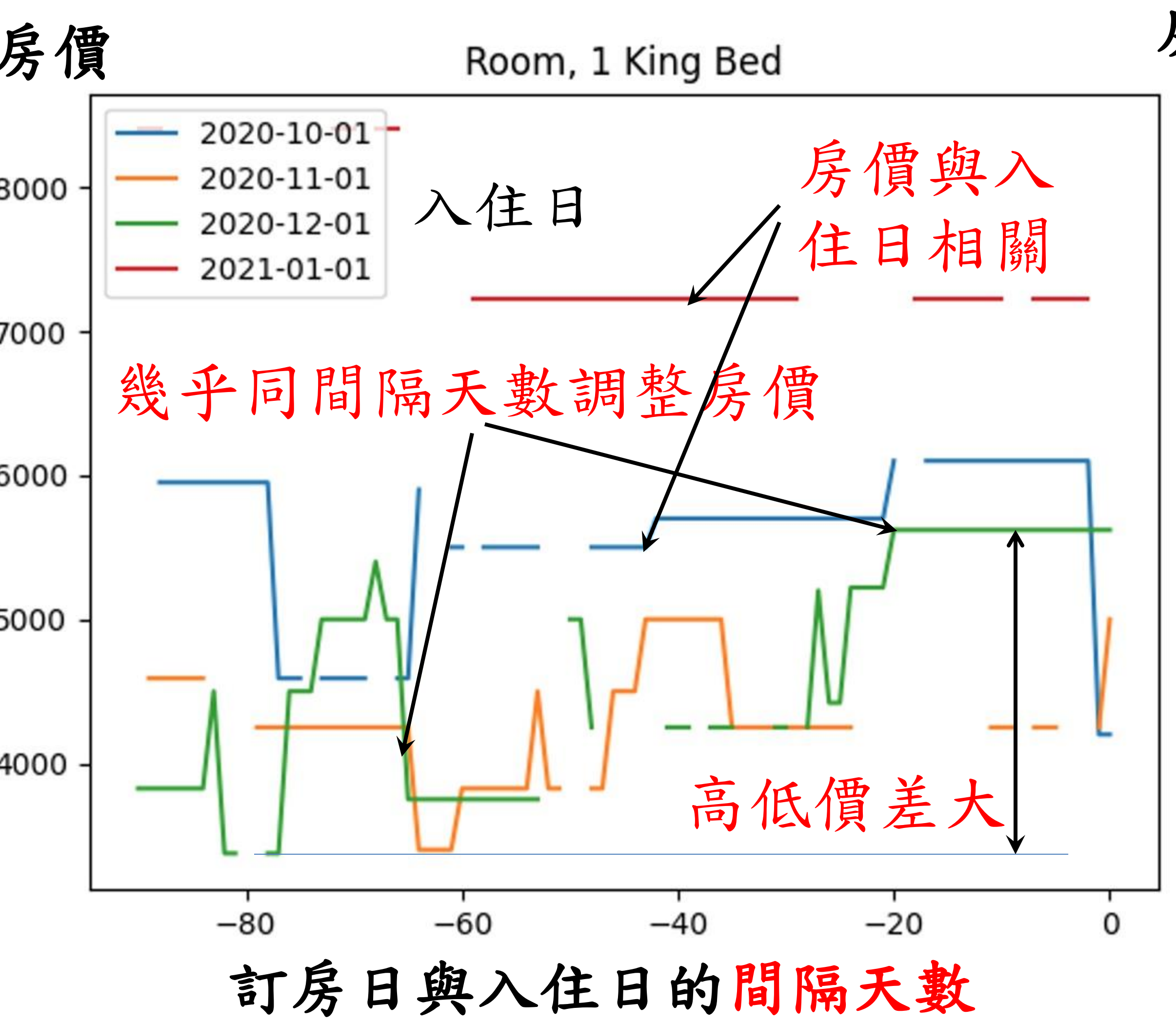
(一) 不同入住日對房價的影響



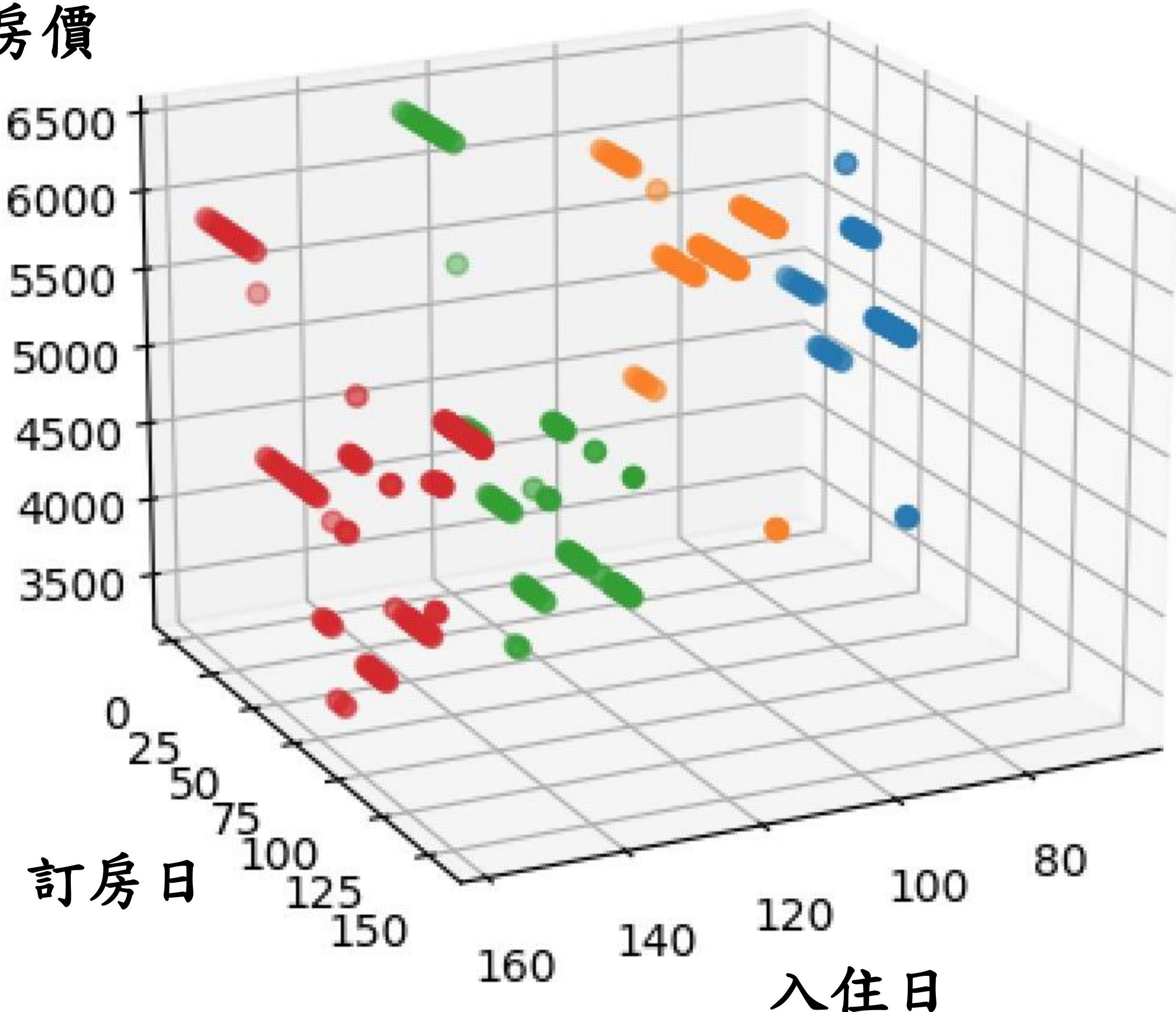
(二) 不同訂房日對房價的影響



(三) 間隔天數對房價的影響



(四) 入住/訂房日對房價的影響



小結：

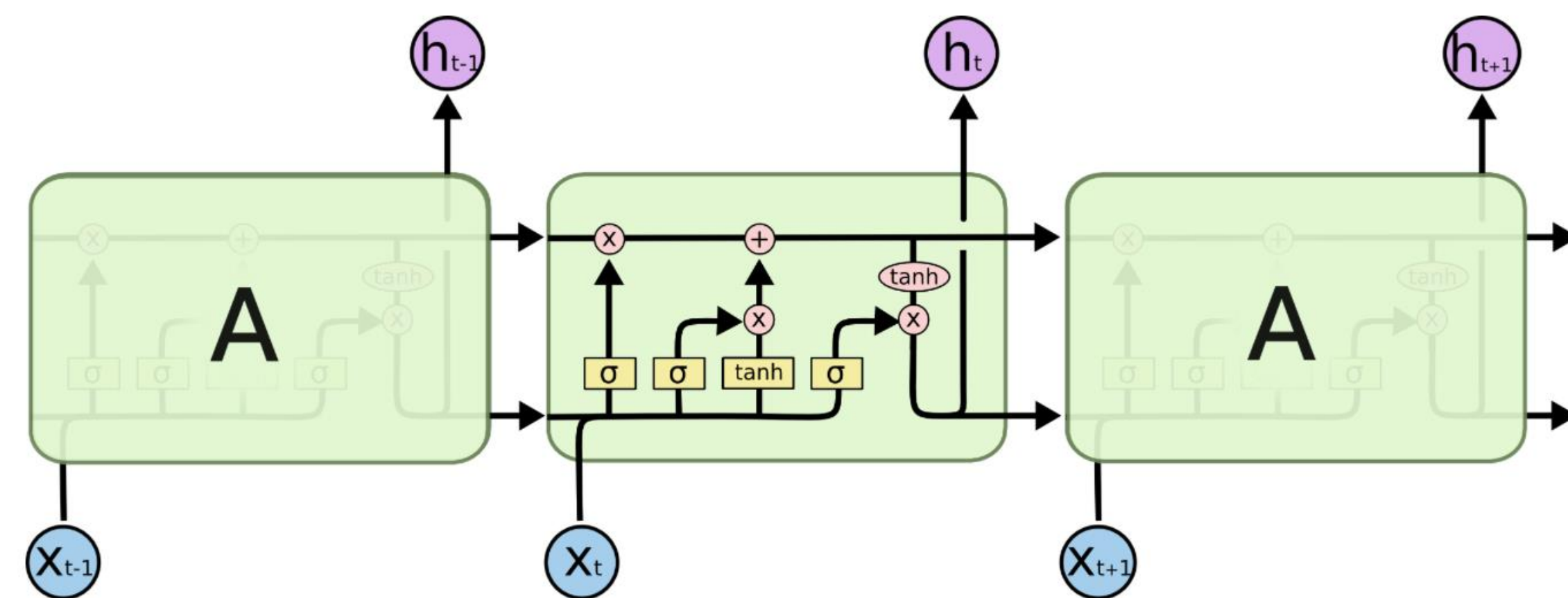
入住日(星期幾、是否為特殊假日)與訂房日是兩項影響房價的重要因素，但根據資料分析的結果，看不出房價調整的明顯規律性。此外，影響房價因素還有飯店星級、房型等其他因素，使結果更加複雜，並無法用目前所學的數學、統計方法去近似，所以我們決定借助機器學習來解決此問題。

參、研究成果

二、房價預測

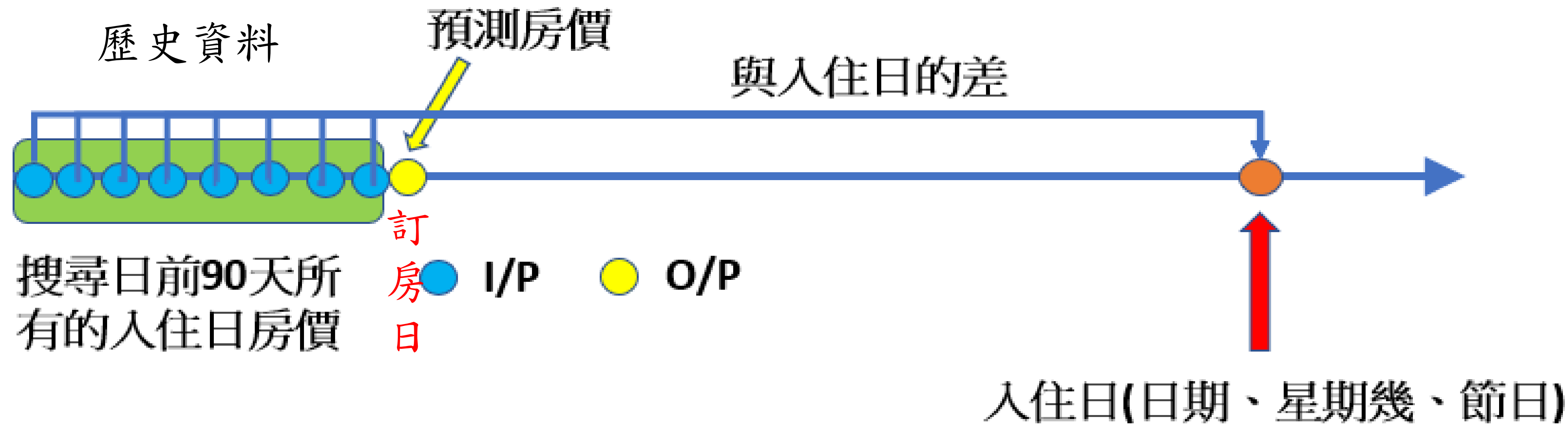
(一)長短期記憶(Long Short-Term Memory, LSTM)模型

房價有時間序列的特性，使用RNN-based的演算法可以將時間的屬性考慮進去，而LSTM又改良了一些RNN的缺點，因此本研究選用LSTM模型進行訂房價格預測。



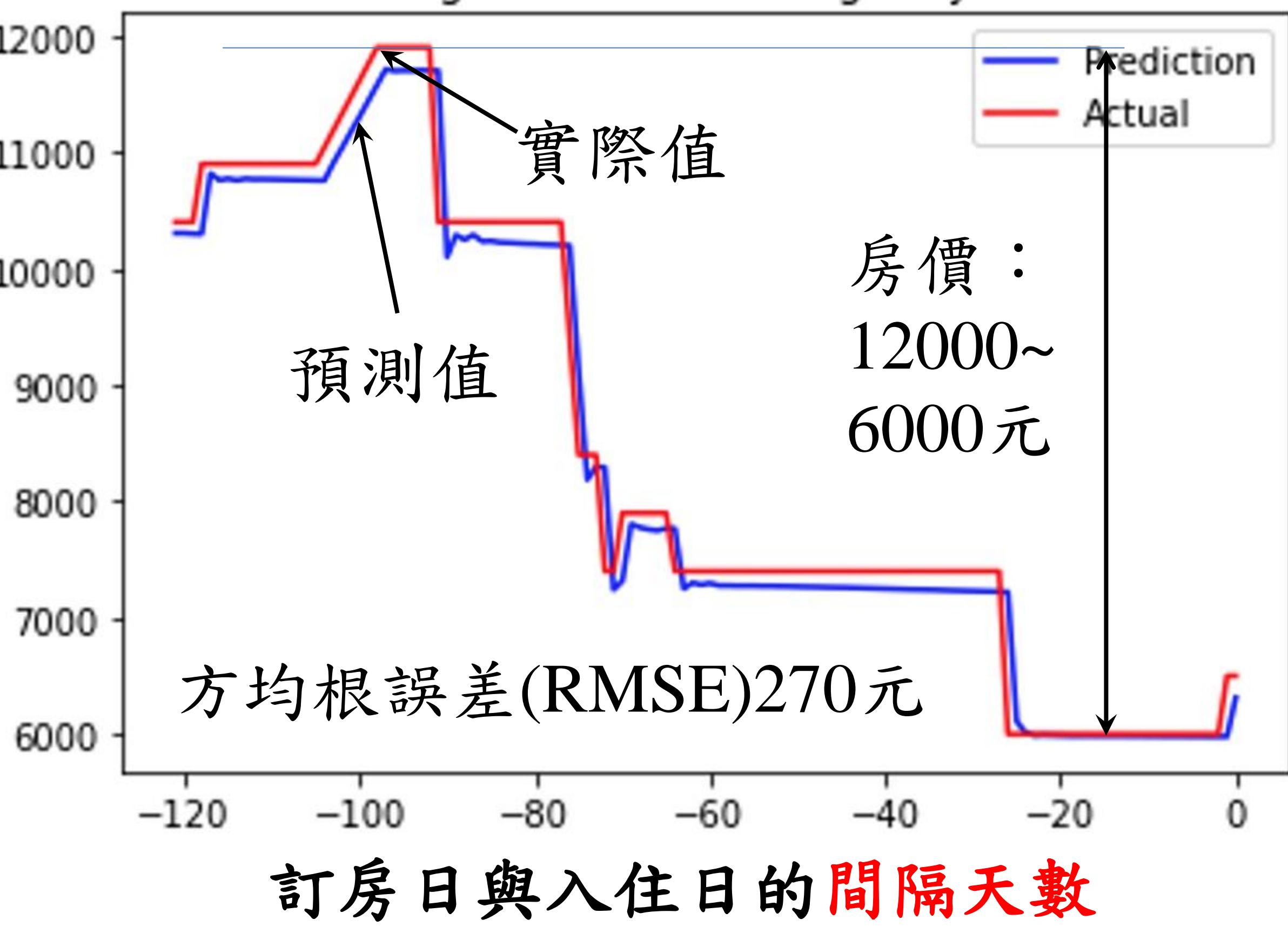
(二) 訓練模型

- 本研究使用雙層LSTM，其中：
- LSTM 輸入參數(I/P)為訂房日前90日的關鍵因素(ex:入住日、訂房日)與房價。
 - LSTM 輸出參數(O/P)為訂房日當天預測房價。
 - 模型利用預測值與實際值的方均根誤差進行調整、訓練。

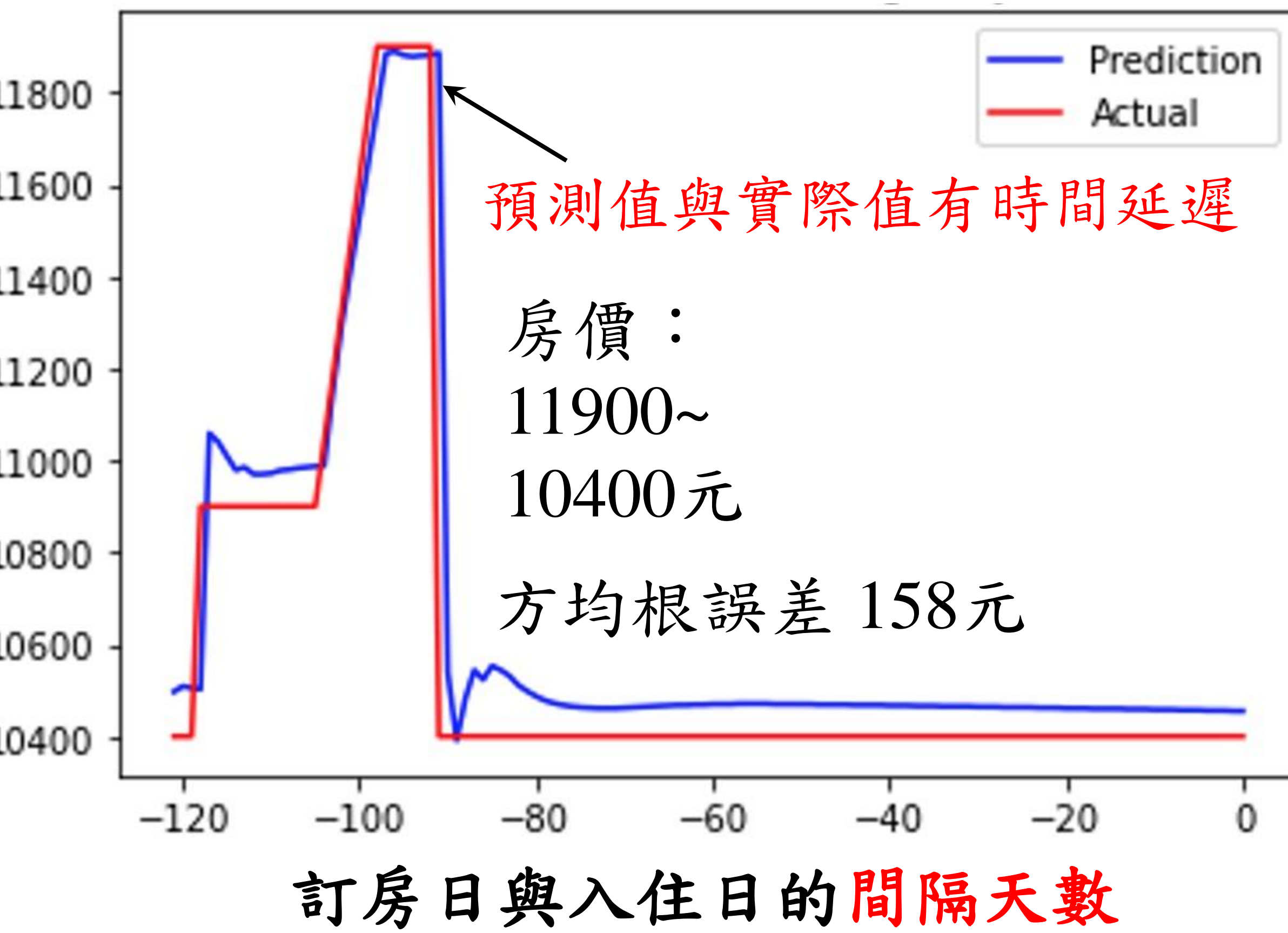


(三) 模型測試(單日預測)

W飯店King房型2021年1月25日入住房價

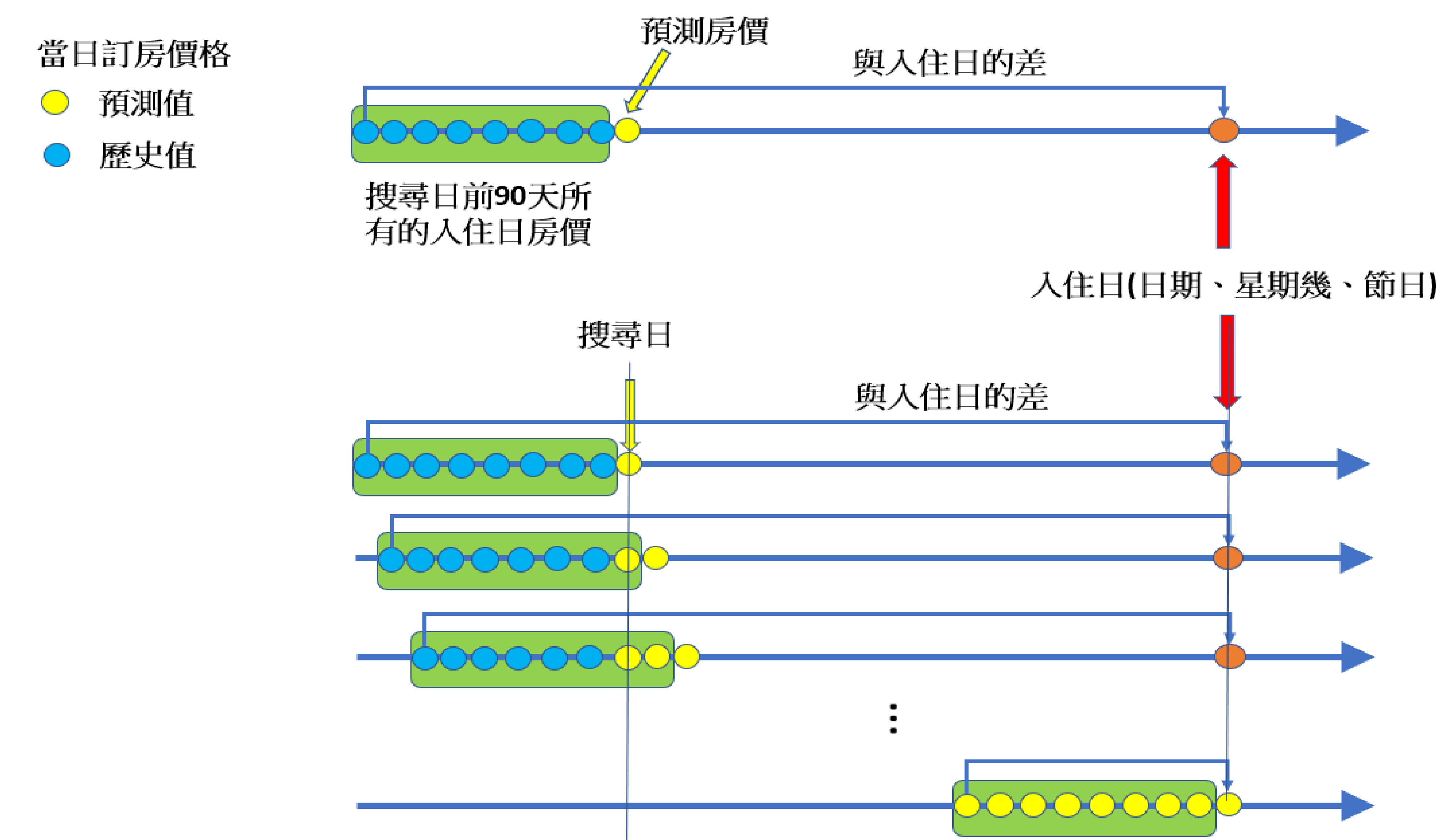


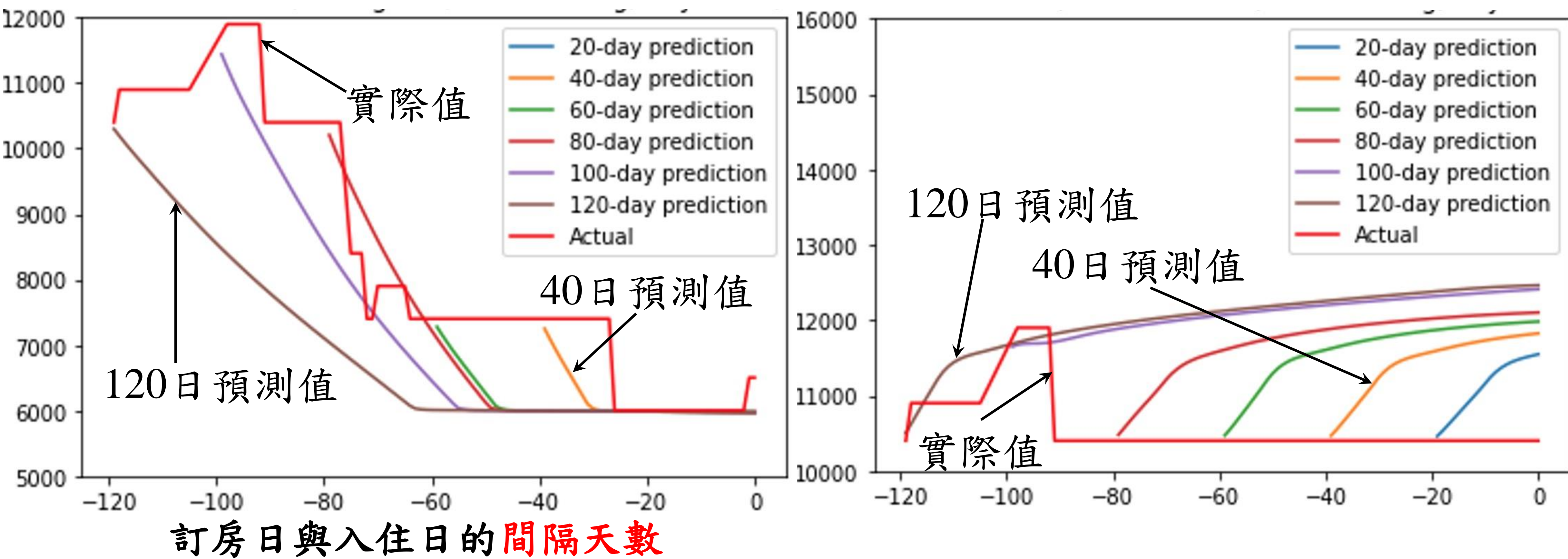
飯店Double房型2021年1月25日入住房價



(四) 實際應用

模型訓練與測試均為單日預測的結果，而本研究的目的是要提供「消費者從搜尋日到入住日之間進行訂房所得到的入住日房價的走勢預測」，因此實際應用方式修正如下：





	1day	20day	40day	60day	80day	100day	120day
RMSE(NTD)	158	803	1113	1287	1393	1676	1592

表一：預測結果與實際結果的方均根誤差

小結：

本研究基於單日預測的數據去訓練LSTM模型，以LSTM進行單日預測房價有顯著**的成果**，但可觀察到LSTM常見**預測結果與實際結果有一天延遲的問題**。儘管目前預測的誤差很大，但以現有模型應用於本研究所需要的長天期預測應是可行的。

肆、成果與展望

一、成果

- 1.以**自行開發的爬蟲程式**解決缺乏飯店房價資料的問題。
- 2.資料分析發現**入住日、訂房日、「星級」、「房型」**是影響房價的關鍵因素。
- 3.成功訓練LSTM模型，在**單日房價預測**上有不錯的效果。

二、展望

1. 優化爬蟲工具
 - (一)、放在伺服器每天自動蒐集資料
 - (二)、解決被網站封鎖的問題
2. 解決長天期房價預測問題
未來會從兩方面著手：
 - (一). 深入了解LSTM背景知識，基於目前單日房價預測LSTM模型，從修正訓練特徵細部參數設定的規則等方式將模型進行優化，以提高預測準確度。
 - (二). 根據多個輸入時間序列預測多個輸出時間序列，重新設計可供長天期房價預測LSTM模型。

伍、參考資料

[1] Expedia幫旅宿業帶進60億營收的工具大揭密(<https://www.bnext.com.tw/article/55536/explore-%2719-rev+>)

[2] 盧慶龍, 郭曉怡, 陳善珮,“旅館產業住宿服務的訂價因素與特徵價格之研究,” 臺北城市大學學報 第36期, 2013 年 05 月

[3] 侯佩妤, 陳俊智, 包曉天,“電子口碑訊息和訂價：以臺灣連鎖經營酒店為例,” 商略學報, 9卷, 2期, 103-118頁, 2017年

[4] 如何解決LSTM預測延遲現象(<https://www.coder.work/article/3159785>)

[5] Understanding LSTM Networks(<http://colah.github.io/posts/2015-08-Understanding-LSTMs/>)

[6] 洪錦魁。Python網路爬蟲：大數據擷取、清洗、儲存與分析：王者歸來