



# 統計應用分析報告

臺北市道路路面損壞修補數量

相關因素探討

土木建築科曾俊傑、統計室范汝欣

臺北市政府工務局

110年1月

## 摘要

道路品質的測量是十分專業的領域，且全面性的測量普查歷時較長，較難即時提供決策應用。考量現有資料中，道路查報修補數量與道路維護成本有關，亦一定程度的反映道路品質，因此本研究利用臺北市政府新建工程處「道路管理系統」資料，預測道路損壞修補面積，並進行其影響因素探討，供道路維護決策參考。

經利用梯度提升機(GBM, Gradient Boosting Machine)計算影響權重發現，「去年修補件數」為預測道路修補面積的最重要變數，其與「道路名稱」、「去年修補面積」合計占有73%的重要性。其中「道路名稱」推測可能隱含車流量資訊，在無法取得車流量資料的情況下，可能是個替代變數。另在「修補面積」較大的道路中，預測排名與實際排名非常一致，故建議 GBM 預測結果用於道路修補成本排名，而非道路修補成本數值。

## 目 次

壹、前言 .....	1
貳、研究方法 .....	1
一、以路寬 8 公尺以上 AC 路面查報修補面積為研究對象 .....	1
二、以道路銑鋪單元(簡稱道路單元)為分析單位 .....	2
三、以 11 個變數建立模型 .....	3
四、以 GBM 研究因素關係 .....	4
參、研究結果 .....	5
一、「去年修補件數」為最重要變數 .....	5
二、前 3 大變數占 73%的重要性.....	5
三、「道路名稱」可能隱含車流量資訊 .....	6
四、變數重要性存在局部差異 .....	6
五、「修補面積」較大的道路中，僅 5%至 13.33%預測排名與實際 排名差距較大 .....	8
肆、結論與建議 .....	10
伍、參考資料 .....	10
附錄、市府路道路單元及 11 個建模變數資料 .....	附 1

## 表 目 次

表 測試資料預測排名比較 .....	9
--------------------	---

## 圖 目 次

圖 1 AC 路面查報件數 .....	1
圖 2 路寬 8 公尺以上 AC 路面修補數量 .....	2
圖 3 道路單元 .....	3
圖 4 道路修補面積相關因素重要性 .....	5
圖 5 市府路預測案例 .....	7

# 臺北市道路路面損壞修補數量相關因素探討

## 壹、前言

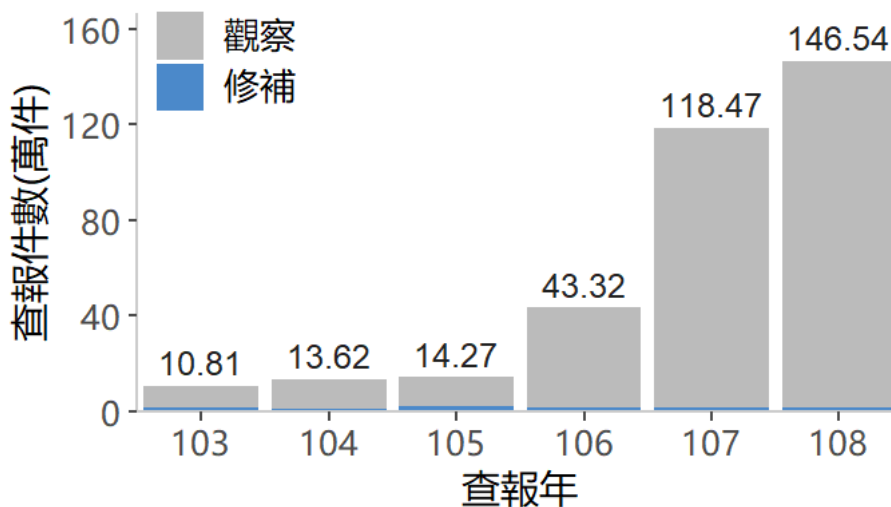
道路品質的測量是十分專業的領域，且全面性的測量普查歷時較長，較難即時提供決策應用。考量現有資料中，道路查報修補數量與道路維護成本有關，亦一定程度的反映道路品質，因此本研究對臺北市路寬 8 公尺以上 AC 路面損壞修補數量，進行相關因素探討，利用梯度提升機(GBM, Gradient Boosting Machine)計算因素的影響權重，供道路維護決策參考。

## 貳、研究方法

### 一、以路寬 8 公尺以上 AC 路面查報修補面積為研究對象

臺北市道路巡查對象主要分為 AC 路面、人行道及相關設施 2 大類，其中 AC 路面查報件數自 107 年起突破 100 萬件，108 年計有 146.54 萬件。查報案件經評估無須立即修補者即列為觀察案件，在後續的巡查中重複查報記錄，故包含觀察案件的查報數量不適合用以衡量道路損壞缺失的嚴重程度。(圖 1)

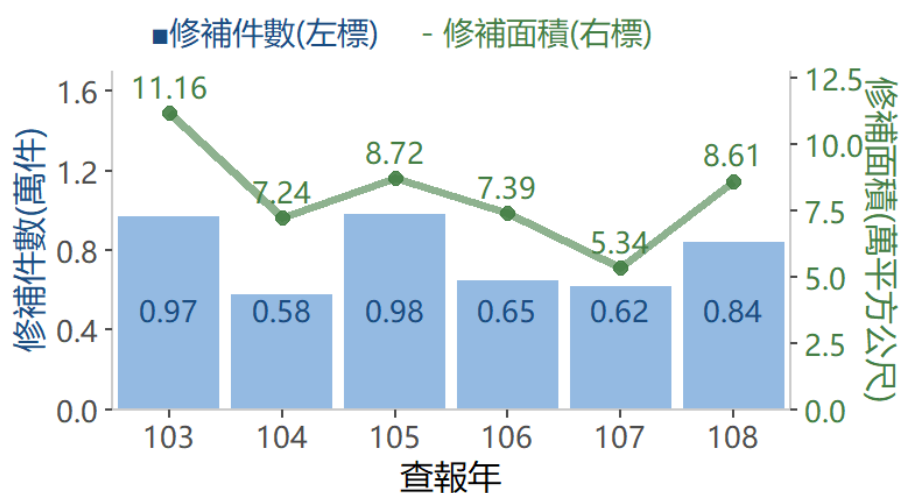
圖 1 AC 路面查報件數



資料來源：臺北市政府新建工程處「道路管理系統」。

相對的，有修補的數量與道路維護成本有關，亦一定程度的反映道路品質，因此本研究以有修補者為分析對象。另 105 年起寬度未滿 8 公尺道路(鄰里道路)維護工作由區公所移撥至工務局新建工程處，考量比較基礎的一致性，僅就路寬 8 公尺以上者進行探討，故研究預測標的為「路寬 8 公尺以上 AC 路面查報修補面積」(簡稱修補面積)。(圖 2)

圖 2 路寬 8 公尺以上 AC 路面修補數量

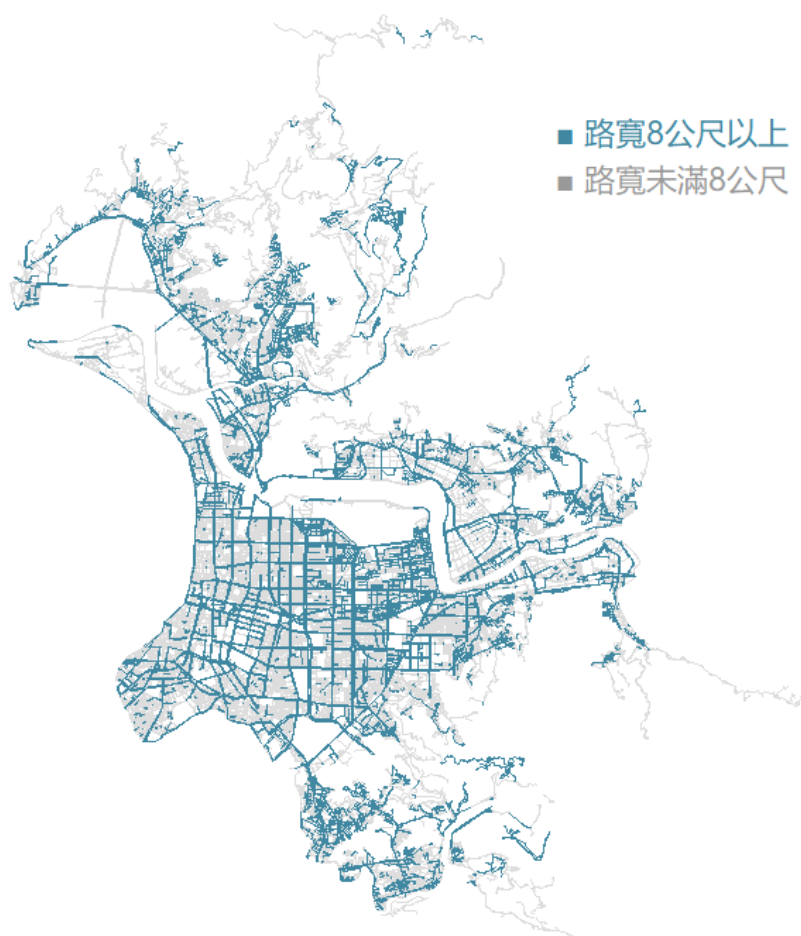


資料來源：臺北市政府新建工程處「道路管理系統」。

## 二、以道路銑鋪單元(簡稱道路單元)為分析單位

臺北市所有道路劃分為 1 萬 9,081 個道路單元，其中路寬 8 公尺以上者計 5,303 個(占 27.8%)，平均每個道路單元面積為 1,915 平方公尺。本研究以其 106 年至 108 年共計 1 萬 5,909 筆資料，隨機抽取 70%訓練資料(Training Data)建立預測模型，其餘 30%為測試資料(Testing Data)，用於驗證結果。(圖 3、附錄)

圖 3 道路單元



資料來源：臺北市政府新建工程處「道路管理系統」。

### 三、以 11 個變數建立模型

本研究取得的道路單元資料為臺北市政府新建工程處「道路管理系統」的「道路名稱」、「道路單元長度」、「道路單元面積」、「行政區」、「最近一次銑鋪日」、「IRI 值」、「修補件數」、「修補面積」。其中「最近一次銑鋪日」及「IRI 值」並非完整資料，僅 1,359 個道路單元(占 25.63%)有最近一次銑鋪紀錄；僅 1,714 個道路單元(占 32.32%)有 IRI 測量值(International Roughness Index，為一道路平坦度指標)，且無測量時間資料，故不一定能反映修補當時的道路平坦度。

本研究將道路單元的「最近一次銑鋪日」轉換為修補年的「距銑鋪日數」，以「道路名稱」、「道路單元長度」、「道路單元面積」、「行政區」、「距銑鋪日數」、「IRI 值」、「去年修補件數」、「去年修補面積」、「前年修補件數」、「前年修補面積」、「修補年」等 11 個變數建模，預測道路單元的「修補面積」。(附錄)

#### 四、以 GBM 研究因素關係

梯度提升(Gradient Boosting)是一種機器學習(Machine Learning)方法，通常以決策樹(Decision Tree)作為基礎，遞迴優化損失函數(Loss Function)，最後集成預測模型。

由於尋找最佳化損失函數參數通常不存在簡單方法，因此以遞迴方式逐步逼近最佳參數，而 GBM 是最常用的遞迴演算法。此外，尚有下列原因選擇 GBM：

- (一)能同時處理數值變數與類別變數：用於建模的 11 個變數中，除了數值變數外，尚包含「行政區」、「道路名稱」等類別變數，GBM 能同時納入模型。
- (二)能處理類別眾多的變數：「道路名稱」可能隱含道路使用量資訊，在未取得車流量資料下，可能是個替代變數。傳統模型較難處理數量眾多的路名，故利用能處理上千種分類的 GBM 建模。
- (三)能計算變數重要性：變數重要性為本研究探討重點，GBM 能輕易計算。
- (四)演算法適合預測排名：決策樹是一種分類演算法，GBM 以其為基礎，即使對「修補面積」數值預測結果不理想，對其排名的預測結果應不差。



## 參、研究結果

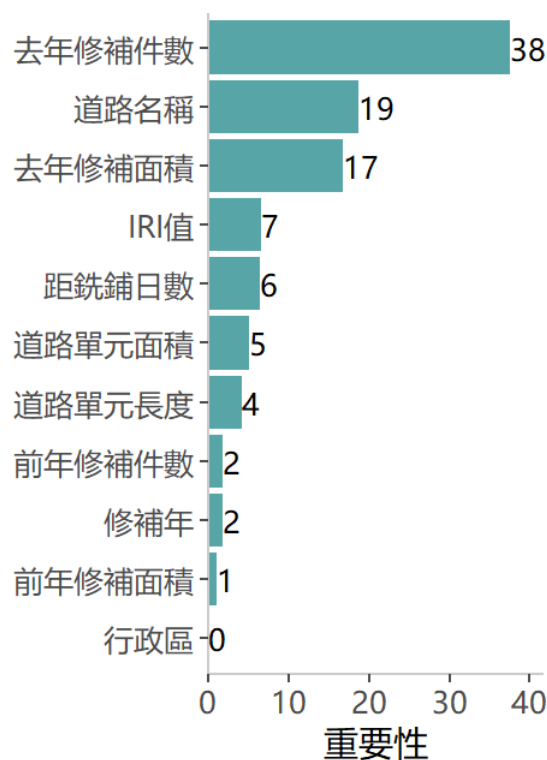
### 一、「去年修補件數」為最重要變數

變數重要性有多種衡量方法，本研究以變數在各個樹模型的平均預測誤差改善值代表其重要性。最重要變數為具有最大改善值的變數，其他變數的重要性則是相對最重要變數計算而得。由 106 年至 108 年資料發現，影響「修補面積」的最重要因素為「去年修補件數」。(圖 4)

### 二、前 3 大變數占 73%的重要性

若重要性總計以 100 計算，「去年修補件數」為 38，其次依序為「道路名稱」19、「去年修補面積」17，前 3 個變數即占了 73%的重要性。其餘 8 個變數重要性都在 10 以下，分別為「IRI 值」、「距銑鋪日數」、「道路單元面積」、「道路單元長度」、「前年修補件數」、「修補年」、「前年修補面積」、「行政區」。(圖 4)

圖 4 道路修補面積相關因素重要性



資料來源：本研究。

### 三、「道路名稱」可能隱含車流量資訊

車流量能反映道路的使用量，應為道路品質高度相關的變數，惟本研究並未取得資料。然而分析結果發現，「道路名稱」為第 2 大重要變數，推測其背後隱含有車流量資訊。

另影響「修補面積」的最重要因素為「去年修補件數」而非「去年修補面積」，推測原因為修補件數是相對穩定的變數，少數大範圍(面積)的修補案件可能降低「去年修補面積」的預測能力。此外，「距銑鋪日數」、「IRI 值」應為道路品質高度相關的變數，惟本研究使用的資料較不完整，可能造成其重要性排名後退。(圖 4)

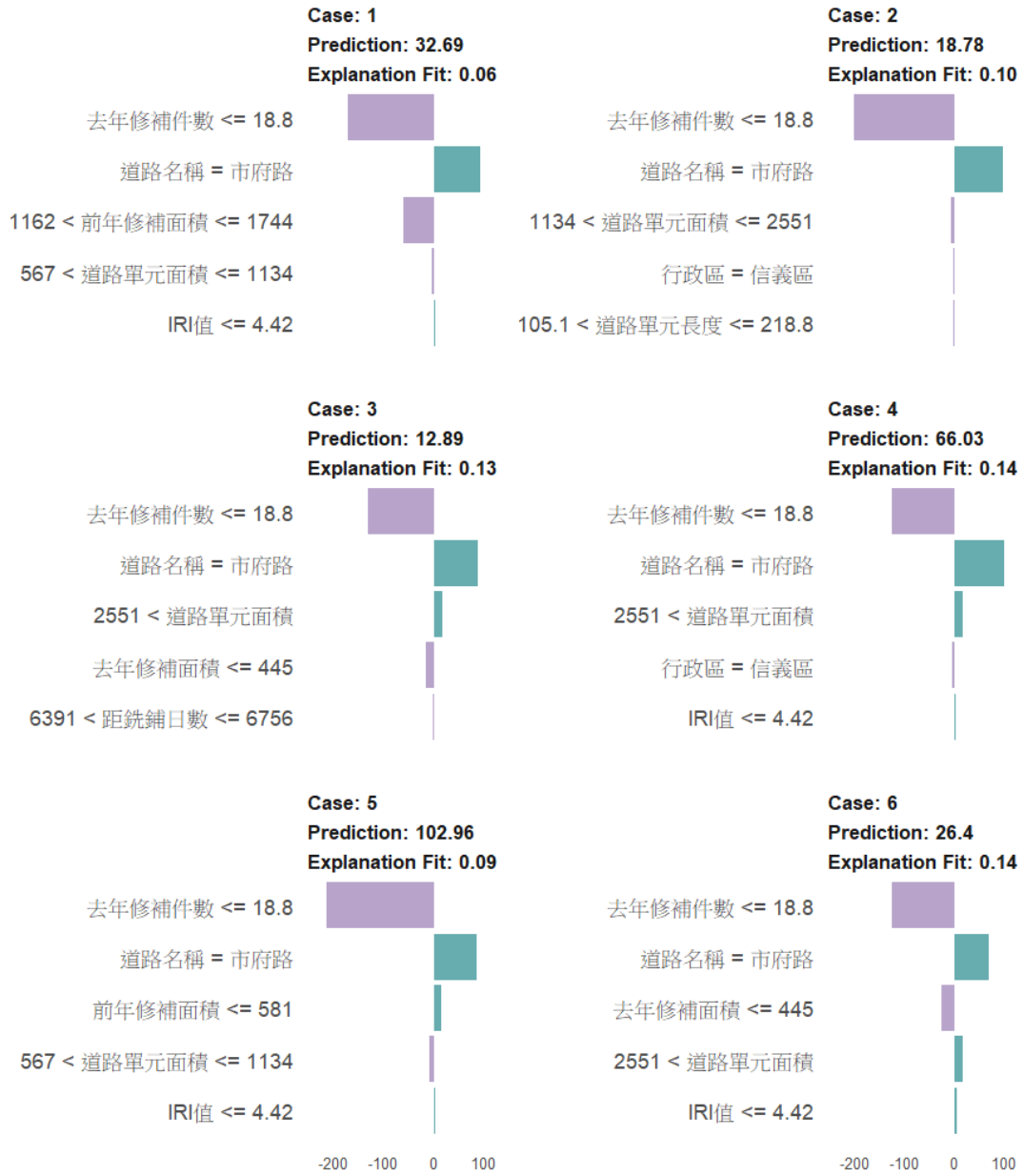
### 四、變數重要性存在局部差異

圖 5 為市府路的 6 個預測案例，圖中依序列出重要性前 5 大的變數。全部 6 個案例的最重要變數及次要變數皆與整體平均結果一樣為「去年修補件數」、「道路名稱」，且兩者占了大部分重要性。此外，6 個案例皆屬於「去年修補件數」的最小級距( $\leq 18.8$ )，與「修補面積」是負向關係，即屬於此級距所預測的「修補面積」數值會較小。路名為「市府路」則與「修補面積」是正向關係，即市府路的「修補面積」預測值會較大。

對整體平均而言，「去年修補面積」的重要性與「道路名稱」相差不多，但市府路 6 個案例僅 Case 3，6 排進前 4 名，且重要性遠遠較低，顯示變數重要性存在局部差異。

另模型在市府路 6 個案例的解釋力( $R^2$ )皆低於 0.2，對修補面積「數值」配適結果不理想，惟後續分析發現，不影響預測「排名順序」的準確性。(圖 5)

圖 5 市府路預測案例



資料來源：本研究。

## 五、「修補面積」較大的道路中，僅 5%至 13.33%預測排名與實際排名差距較大

以所建模型預測未參加建模的 30%測試資料，將每條道路的平均預測值與平均實際值，以及所有資料(不分訓練、測試)的平均實際值進行排名比較<sup>1</sup>。所有資料實際排名(c)前 10 名的道路，其測試資料預測排名(a)有 6 條道路亦為前 10 名、1 條為第 12 名。另外 3 條為「潭美街」、「經園街」，及「南昌路一段」，其中「潭美街」因未抽選入測試資料集，無法進行比較。「經園街」依其他相近條件道路的預測結果(a 第 415 名)與實際排名(b 第 2 名、c 第 7 名)差距較大。「南昌路一段」預測排 49 名(a)，與測試資料實際排名(b 第 72 名)皆在 10 名之外。

所有資料實際排名(c)前 30 名的道路，有 22 條預測排名(a)亦為前 30 名，其餘 8 條中，有 2 條無測試資料，2 條測試資料實際排名(b)亦非前 30 名(南昌路一段、建國北路一段)，其餘 4 條(占 13.33%)與實際排名(b, c)差距較大者為「經園街」(a415, b2, c7)、「西藏路」(a43, b14, c22)、「松壽路」(a68, b8, c27)、「新生北路二段」(a72, b9, c28)。

綜上，前 30 名「修補面積」較大的道路中，有 4 條(占 13.33%)預測結果與實際排名差距較大。若僅觀察前 20 名道路，則僅 1 條(占 5%，經園街)差距較大。若僅觀察前 10 名道路，亦為 1 條(占 10%，經園街)差距較大。

雖然前述「市府路」案例之「數值」配適結果不理想，惟若僅預測排名，則結果與實際排名非常一致(a4, b6, c2)，顯示 GBM 的預測結果仍有一定的參考性。

---

<sup>1</sup> 相同道路名稱可能包含多個道路單元的多年(106 年至 108 年)資料，表中為其平均值。

表 測試資料預測排名比較

道路名稱	測試資料(未參加建模)				所有資料(不分訓練、測試)	
	預測值 (m <sup>2</sup> )	預測排名 a	實際值 (m <sup>2</sup> )	實際排名 b	實際值 (m <sup>2</sup> )	實際排名 c
思源街	36	6	2,076	1	401	1
市府路	42	4	65	6	255	2
潭美街	無測試 資料	無測試 資料	無測試 資料	無測試 資料	202	3
凱達格蘭大道	83	2	15	28	132	4
南京東路六段	53	3	55	10	132	5
館前路	131	1	276	3	118	6
經園街	3	415	293	2	98	7
忠孝東路五段	36	8	47	12	73	8
南昌路一段	7	49	3	72	54	9
成功路三段	30	12	89	5	52	10
重慶南路一段	23	15	105	4	49	11
長順街	17	18	64	7	49	12
金山南路一段	無測試 資料	無測試 資料	無測試 資料	無測試 資料	46	13
康定路	13	28	0	398	41	14
堤頂大道一段	14	26	0	398	40	15
環河南路一段	21	16	1	127	37	16
環河南路三段	14	24	0	398	35	17
桂林路	39	5	18	26	35	18
重慶南路三段	13	27	48	11	31	19
建國北路二段	15	23	38	15	30	20
長沙街二段	16	21	35	16	29	21
西藏路	8	43	40	14	26	22
研究院路一段	31	11	6	55	26	23
金山北路	12	32	2	81	25	24
建國北路一段	10	36	1	108	24	25
延壽街	24	14	0	398	23	26
松壽路	6	68	59	8	23	27
新生北路二段	5	72	56	9	21	28
安和路二段	11	34	24	21	21	29
西園路二段	10	35	6	53	20	30

資料來源：本研究。

## 肆、結論與建議

影響道路品質的因素眾多，在資料取得成本高的情況下，本研究以現有資料，對道路路面損壞修補數量進行相關因素探討。結果發現「去年修補件數」為預測「修補面積」的最重要變數，其與「道路名稱」、「去年修補面積」合計占有73%的重要性。其中「道路名稱」推測可能隱含車流量資訊，在無法取得車流量資料的情況下，可能是個替代變數。

在「修補面積」較大的道路中，僅5%至13.33%預測排名與實際排名差距較大，雖然「市府路」案例的修補面積數值配適結果不理想，惟預測排名與實際排名非常一致，故建議GBM預測結果用於道路修補成本排名，而非道路修補成本數值。

## 伍、參考資料

1. Bradley Boehmke & Brandon Greenwell. 2020-02-01. “Hands-On Machine Learning with R.”
2. Alexey Natekin<sup>1</sup> and Alois Knoll. “Gradient boosting machines, a tutorial.”

## 附錄、市府路道路單元及 11 個建模變數資料



資料來源：臺北市政府新建工程處「道路管理系統」。

修補面積 (m <sup>2</sup> )	修補年	行政區	道路名稱	道路單元長度(m)	道路單元面積(m <sup>2</sup> )	距銑鋪日數(日)	IRI值	去年修補件數(件)	去年修補面積(m <sup>2</sup> )	前年修補件數(件)	前年修補面積(m <sup>2</sup> )
1.00	108	信義區	市府路	24.80	653.68	729	4.12	2	2.00	1	-
3,646.94	108	信義區	市府路	31.83	954.76	729	2.39	42	1,245.75	8	7.00
43.21	108	信義區	市府路	138.21	2,073.18	729	NA	5	47.38	0	-
5.03	108	信義區	市府路	138.58	2,078.72	729	NA	10	64.27	1	1.00
3.00	108	信義區	市府路	170.84	2,562.62	729	3.84	2	44.56	1	1.00
21.23	108	信義區	市府路	174.10	2,611.50	729	3.84	3	129.79	0	-
463.33	108	信義區	市府路	273.27	3,019.08	3,398	3.84	2	384.68	1	1.00
-	108	信義區	市府路	292.29	3,229.20	7,121	NA	0	-	0	-
2.00	107	信義區	市府路	24.80	653.68	364	4.12	1	-	1	1,165.70
1,245.75	107	信義區	市府路	31.83	954.76	364	2.39	8	7.00	4	4.00
47.38	107	信義區	市府路	138.21	2,073.18	364	NA	0	-	3	3.00
64.27	107	信義區	市府路	138.58	2,078.72	364	NA	1	1.00	6	6.00
44.56	107	信義區	市府路	170.84	2,562.62	364	3.84	1	1.00	0	-
129.79	107	信義區	市府路	174.10	2,611.50	364	3.84	0	-	1	1.00
384.68	107	信義區	市府路	273.27	3,019.08	3,033	3.84	1	1.00	2	10.00
-	107	信義區	市府路	292.29	3,229.20	6,756	NA	0	-	0	-
-	106	信義區	市府路	24.80	653.68	-1	4.12	1	1,165.70	2	16.90
7.00	106	信義區	市府路	31.83	954.76	-1	2.39	4	4.00	2	4.00
-	106	信義區	市府路	138.21	2,073.18	-1	NA	3	3.00	0	-
1.00	106	信義區	市府路	138.58	2,078.72	-1	NA	6	6.00	8	163.00
1.00	106	信義區	市府路	170.84	2,562.62	-1	3.84	0	-	2	4.00
-	106	信義區	市府路	174.10	2,611.50	-1	3.84	1	1.00	0	-
1.00	106	信義區	市府路	273.27	3,019.08	2,668	3.84	2	10.00	3	161.96
-	106	信義區	市府路	292.29	3,229.20	6,391	NA	0	-	0	-