

藉由JMP機器學習達成 預測品質控制/預測維護及 智慧製造..以半導體為例

黃崇仁 博士

黃譯賢 博士

凱律科技股份有限公司

JEN@PHOENIC-TECH.COM.TW

886-963488517

DECEMBER 28,2021

JMP作為業界頂尖的資料分析品牌，提供多樣的預測建模和資料採擷手段，並持續被Forrester評為資料科學、機器學習和預測分析領域的領導者，其系列產品廣泛地被財富500強公司、世界50強半導體企業、主要政府機構所使用，並能夠幫助決策者、分析人員真正地從資料中挖掘出對企業經營決策有益的資訊

JMP – 全球先進企業的首選分析軟體



注：客戶包括但不限於此

凱律科技股份有限公司簡介

3

- ▶ 我們是大數據分析(Big Data)、模擬、建模、預測(機器學習 Machine Learning)軟體SAS JMP台灣代理商(A partner of SAS JMP in Taiwan) ;
- ▶ 我們是加速建設Industry 4.0 & AI之基礎：『建構數據分析模型(Modeling/Big Data)』及製程、品質、實驗設計(研發)、專利問題實務解決方案提供者；
- ▶ 我們的客戶包含半導體封裝測試產業、精密機械設備產業、電子電機產業、自動化產業、記憶體製造產業、化材產業及其他。

客戶實際改善

4

提高“效率”至關重要

未使用JMP

- 研發成本高
- 研發到量產週期長
- 資料分析手段有限
- 資料分析效率低
- 良率提升慢、良率不穩

目標 =

解決方案 =

實現結果 =

使用JMP後

- 持續縮短研發週期
- 降低研發成本
- 持續改善製程與品質流程

- 建立全球統一的資料分析平臺
- 持續提高工程師與管理層的資料分析與決策能力
- 讓研發、製程、生產、庫存等關鍵領域用資料溝通

- 自上而下驅動
- 形成量化驅動決策的企業文化，而JMP正是助力文化落地的標準工具之一

製造產業常見問題

- ▶ 品質缺陷
- ▶ 精度缺失
- ▶ 設備故障
- ▶ 整體運轉效率損失

製造業為什麼需要機器學習？

- ▶ 在資料無處不在、資料量爆炸式增長的今天，如何高效的利用資料、選擇適合的分析方法成為了關鍵，機器學習正是在這樣的背景下走入工程師的視野
- ▶ **機器學習**，是一種可以基於樣本資料自動化構建數學模型，以便能夠基於相同產品投入新批次生產時用來預測產出結果是否在規格內的分析方法

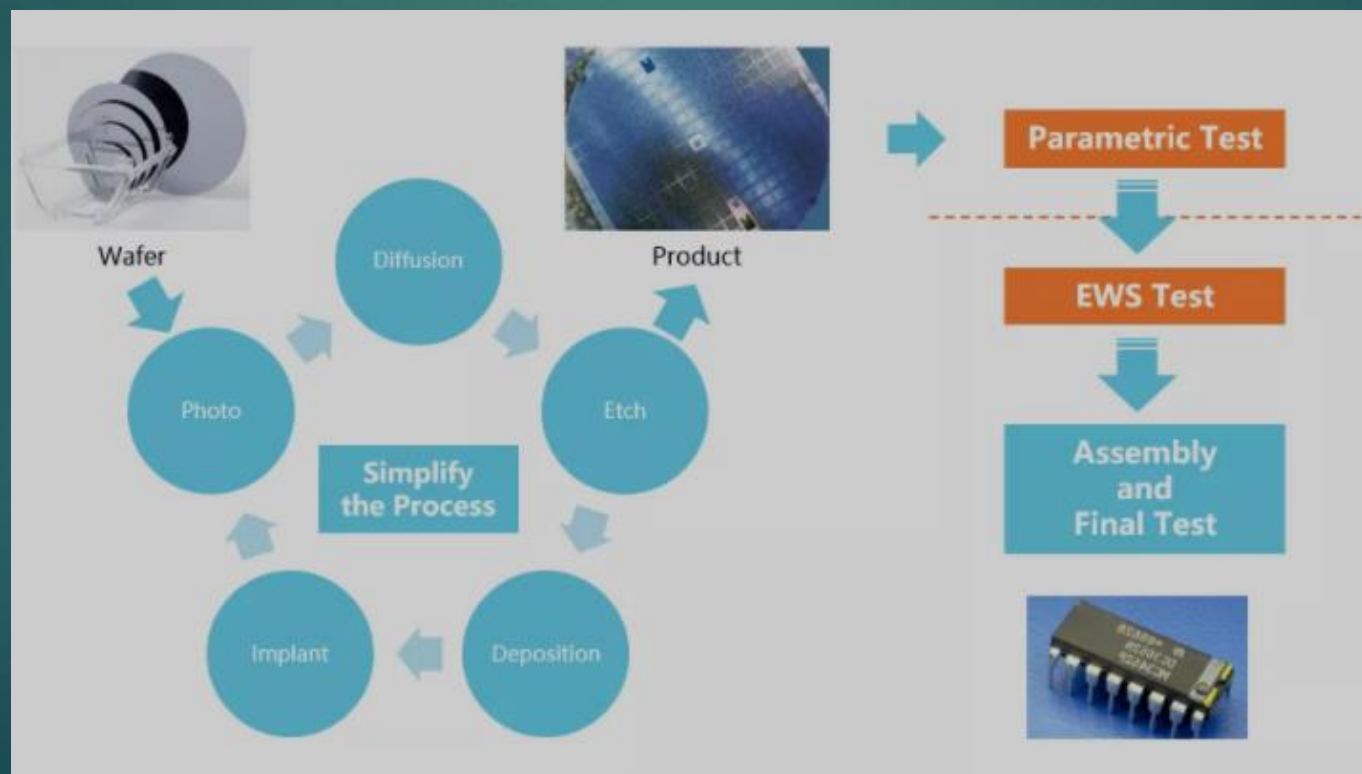
機器學習預測建模流程

7



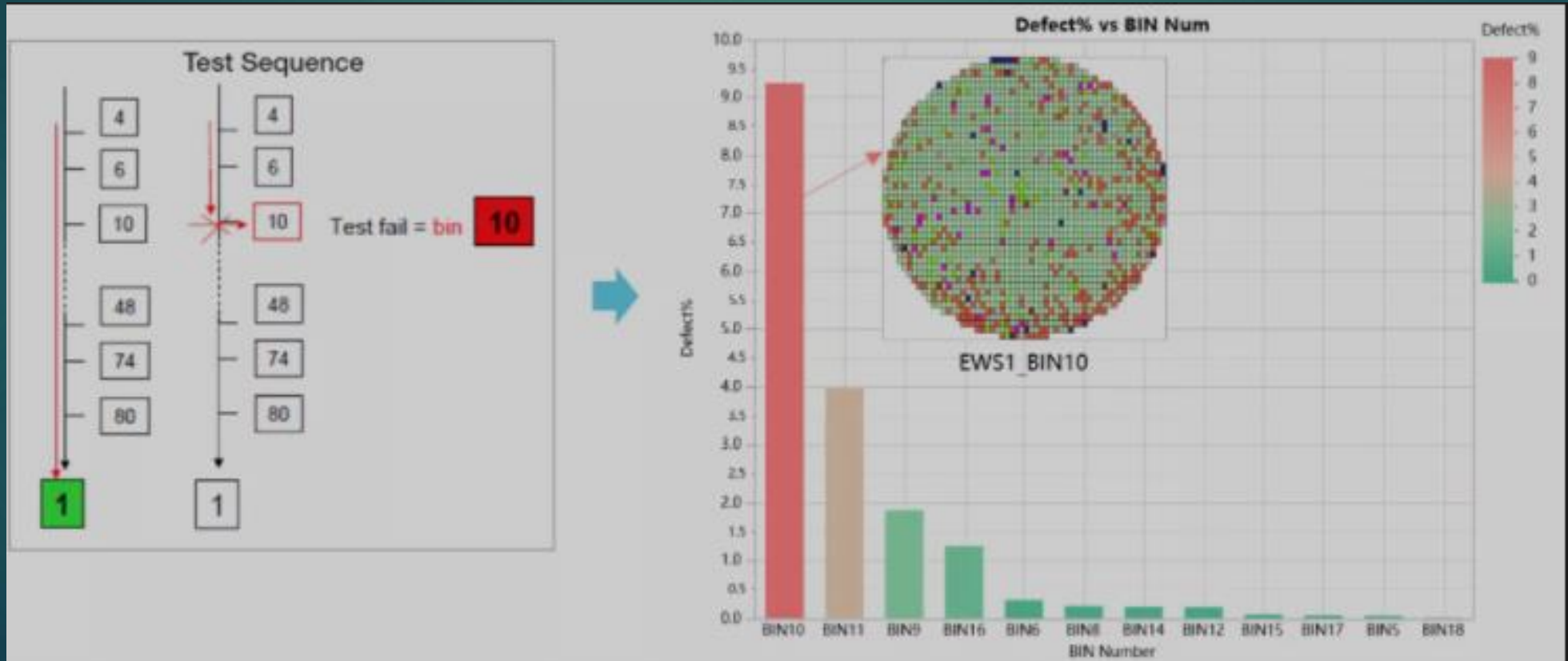
案例背景

半導體終端產品的產生需要經由數百個步驟組成的製造過程，資料量龐大，而且異常複雜；
在製造過程的最後，還有兩個用於對晶片進行分類的補充步驟：
參數測試和電性晶圓分類，這兩個步驟都是電性測試，也是本案例的分析物件。



Step 1 定義業務問題

工程師在某半導體產品 EWS (Electrical Wafer Sort) 測試的 BIN10 上觀測到了良率損失 (在某些晶圓上的良率損失高達 10%)，推測該損失與某個 PT 參數效應有關，因此，一旦確定了 **PT 參數**，就很容易找到**相關的過程參數**，從而形成糾正和預防措施。



Step 2 資料獲取和準備

將用於良率分析的資料進行關聯；比如，按 Lot ID + Wafer ID 關聯 EWS, Parameter Test, Equipment, 甚至 OVL, ADI, THK 等 Inline Test 資料；

Dr. CJ Huang Edited, copyrights reserved
12/27/2021

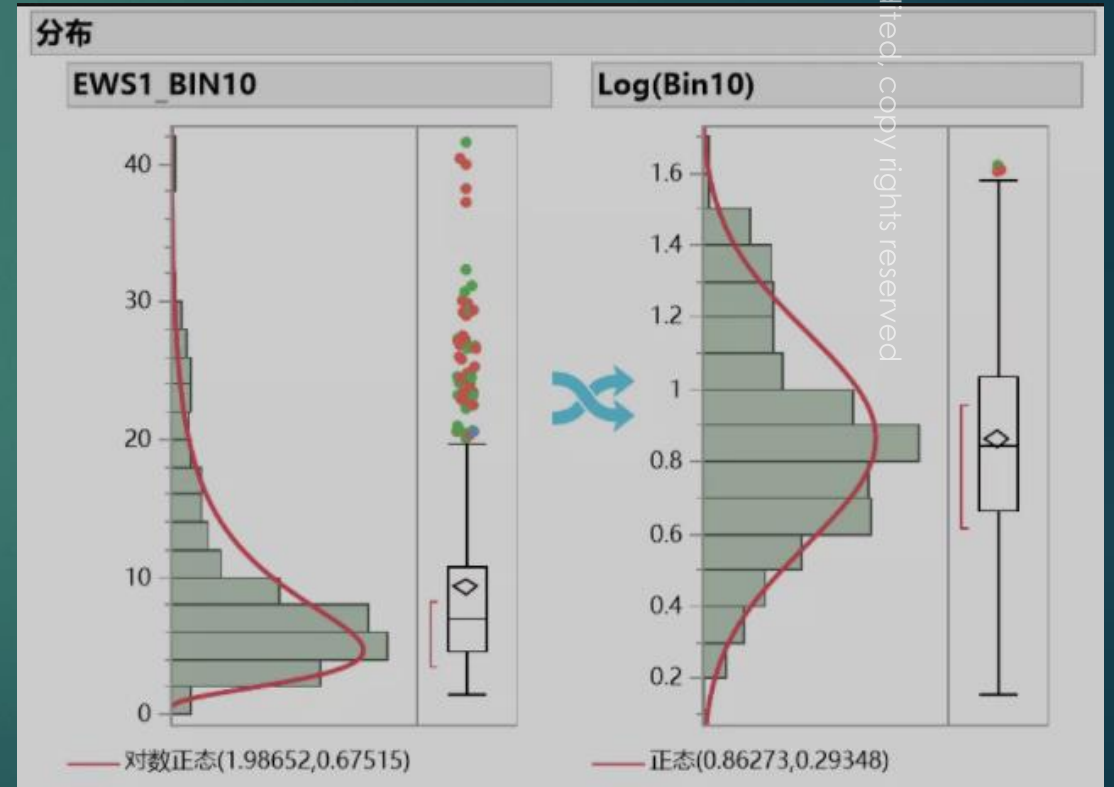
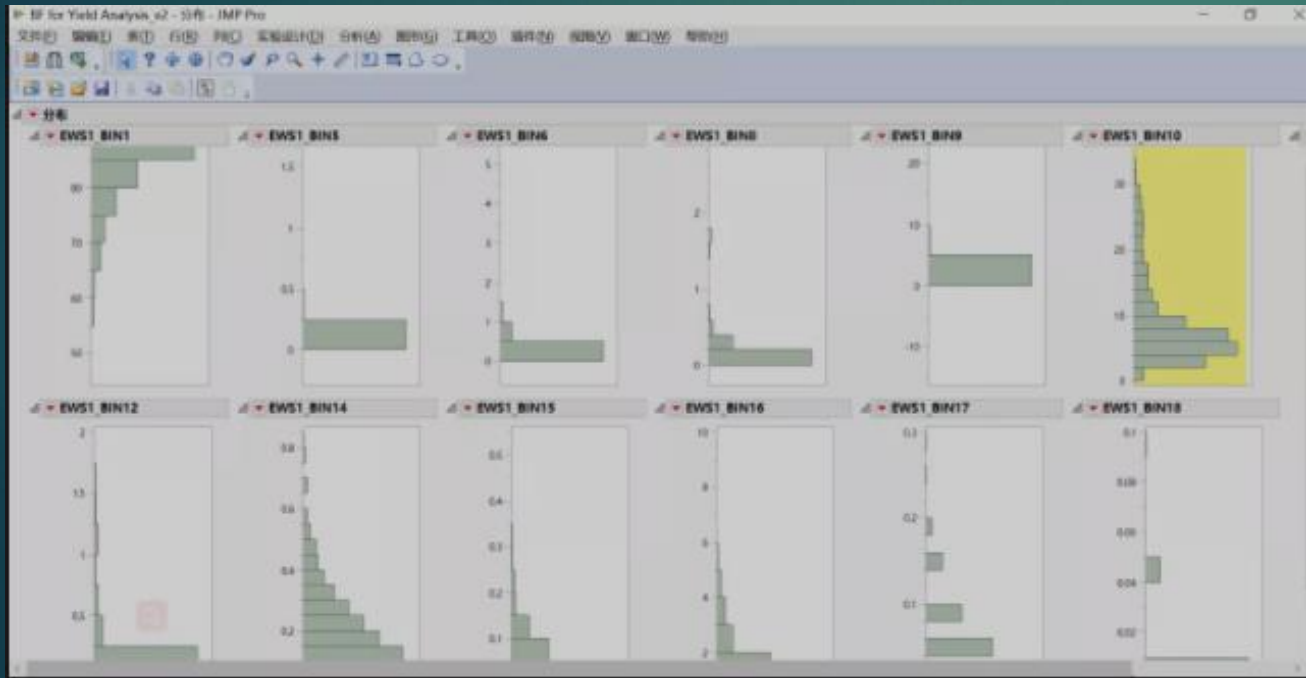
Lot	Wafer	#EWS1_BIN1 (%)	#EWS1_BIN10 (%)
LOT_01	1	88.7	2.0
LOT_01	2	81.7	8.4
LOT_01	3	86.7	4.1
LOT_01	4	90.9	2.2
LOT_01	5	87.3	7.7
LOT_02	1	84.9	6.9
LOT_02	2	76.6	11.6
LOT_02	3	77.1	9.3
LOT_02	4	63.3	30.0
LOT_02	5	81.2	8.8
...

#EQP	#Parameter1 Test (Value)	#Parameter(X) Test (N Cols)
EQP1	29.3	...
EQP1	25.2	...
EQP1	24.1	...
EQP1	24.6	...
EQP1	27.5	...
EQP2	26.6	...
EQP2	27.3	...
EQP2	28.4	...
EQP2	24.9	...
EQP2	27.4	...
...

	9 PT	590 PT	593 PT	594 PT	595 PT	596 PT	597 PT	598 PT	599 PT	600 PT	601 PT	602 PT	603 PT	604 PT	605 PT	606 PT	
*	1	211.2	215.3	39.7	40.79	143.3	80.1	4104	4021	3443	3431	2764	23.24	1.57	1.406	0.4294	
*	2	210.8	214.9	39.69	40.79	143.2	80.09	4035	4131	3369	3570	2792	16.41	1.68	1.426	0.3998	
*	3	211.3	215.2	39.71	40.81	142.8	80.11	4134	3993	3462	3566	2771	21.43	1.555	1.375	0.4026	
*	4	210.9	214.9	39.66	40.78	142.7	80.1	4125	4115	3401	3486	2785	18.12	1.617	1.414	0.3938	
*	5	210.7	214.2	39.77	40.88	141.7	80.15	109.9	4316	3976	3270	3265	2869	14.24	1.656	1.375	0.3772
*	6	210.9	215	39.72	40.82	143.1	80.06	4088	4101	3382	3518	2806	14.14	1.586	1.402	0.3873	
*	7	211.1	215	39.71	40.81	143	80.07	4211	3984	3412	3511	2773	18.97	1.584	1.371	0.3752	
*	8	210.9	214.9	39.7	40.8	143.3	80.05	4098	4095	3346	3451	2786	20.45	1.648	1.437	0.4075	
*	9	211.1	214.9	39.71	40.82	142.6	80.06	109.9	4201	3971	3520	3575	2770	16.7	1.609	1.391	0.4178
*	10	211.2	215	39.68	40.8	143.2	80.08	4311	4099	3492	3342	2785	21.37	1.602	1.438	0.4662	
*	11	211.3	214.9	39.73	40.84	143.3	79.98	4359	3990	3525	3587	2770	26.82	1.562	1.426	0.4	
*	12	211.1	214.9	39.74	40.84	143.5	79.98	4262	4109	3348	3519	2795	22.94	1.586	1.434	0.3993	
*	13	211.3	215	39.75	40.87	143.4	80.05	4449	3965	3441	3385	2779	17.46	1.617	1.383	0.4081	
*	14	211.7	215.2	39.8	40.92	143.3	80.27	4448	4111	3341	3392	2948	15.24	1.719	1.453	0.4416	
*	15	211.3	215	39.73	40.83	143	80.06	4238	3991	3425	3498	2763	25.23	1.531	1.379	0.4009	
*	16	210.9	214.7	39.71	40.82	144.3	79.94	110	4309	4082	3381	3330	2795	21.13	1.57	1.371	0.4145
*	17	211.2	214.9	39.76	40.87	143.7	80	4418	3934	3406	3518	2785	21.61	1.57	1.371	0.3784	
*	18	211.1	215	39.76	40.86	144	80	4283	4093	3387	3504	2786	19.78	1.656	1.434	0.4501	
*	19	211	214.9	39.77	40.88	144	80	4222	3973	3388	3367	2772	20.35	1.57	1.359	0.4196	
*	20	211.3	215	39.75	40.85	143.9	79.99	110	4361	4096	3436	3408	2812	17.55	1.594	1.383	0.3962
*	21	211.3	215	39.83	40.92	144.5	79.92	110	4460	3950	3384	3451	2780	23.41	1.57	1.391	0.4367
*	22	210.7	214.8	39.79	40.89	144.8	79.92	110	4095	4096	3424	3341	2768	21.21	1.625	1.422	0.4155
*	23	210.9	214.9	39.8	40.89	144.3	79.95	110	4223	3968	3410	3432	2754	19.07	1.672	1.445	0.4279
*	24	211.2	214.8	39.8	40.88	144.7	79.93	110	4462	4096	3379	3495	2804	25.24	1.57	1.418	0.4118
*	25	211.1	215.1	39.8	40.9	145	79.93	110	4370	3958	3426	3450	2790	21.24	1.586	1.391	0.4484
*	26	209.9	215.5	40.05	41.18	145	79.54	109.6	4121	4028	3266	3209	2899	13.02	1.711	1.402	2.185
*	27	209.2	214.7	40.06	41.18	144.6	79.53	109.6	3794	4040	3431	3430	2851	12.23	1.773	1.445	1.493
*	28	209.4	214.4	40.07	41.19	144.6	79.45	109.6	3956	4051	3438	3410	2883	12.86	1.711	1.398	1.629
*	29	209.2	214.6	40.11	41.22	143.8	79.48	109.6	3948	4036	3461	3563	2857	12.82	1.703	1.391	1.322
*	30	209.7	214.7	39.88	41	145.3	79.72	109.9	4157	4090	3316	3460	2807	12.52	1.766	1.445	1.717
*	31	210.1	215.2	40.05	41.15	144.4	79.71	109.9	3942	4060	3461	3405	2873	12.5	1.734	1.414	1.222
*	32	210.2	215.1	40.05	41.17	144.3	79.66	109.8	4056	4081	3440	3422	2859	12.35	1.672	1.348	1.35
*	33	209.7	214.5	40.06	41.18	144.7	79.68	109.8	3947	4049	3474	3506	2843	12.21	1.703	1.374	1.327

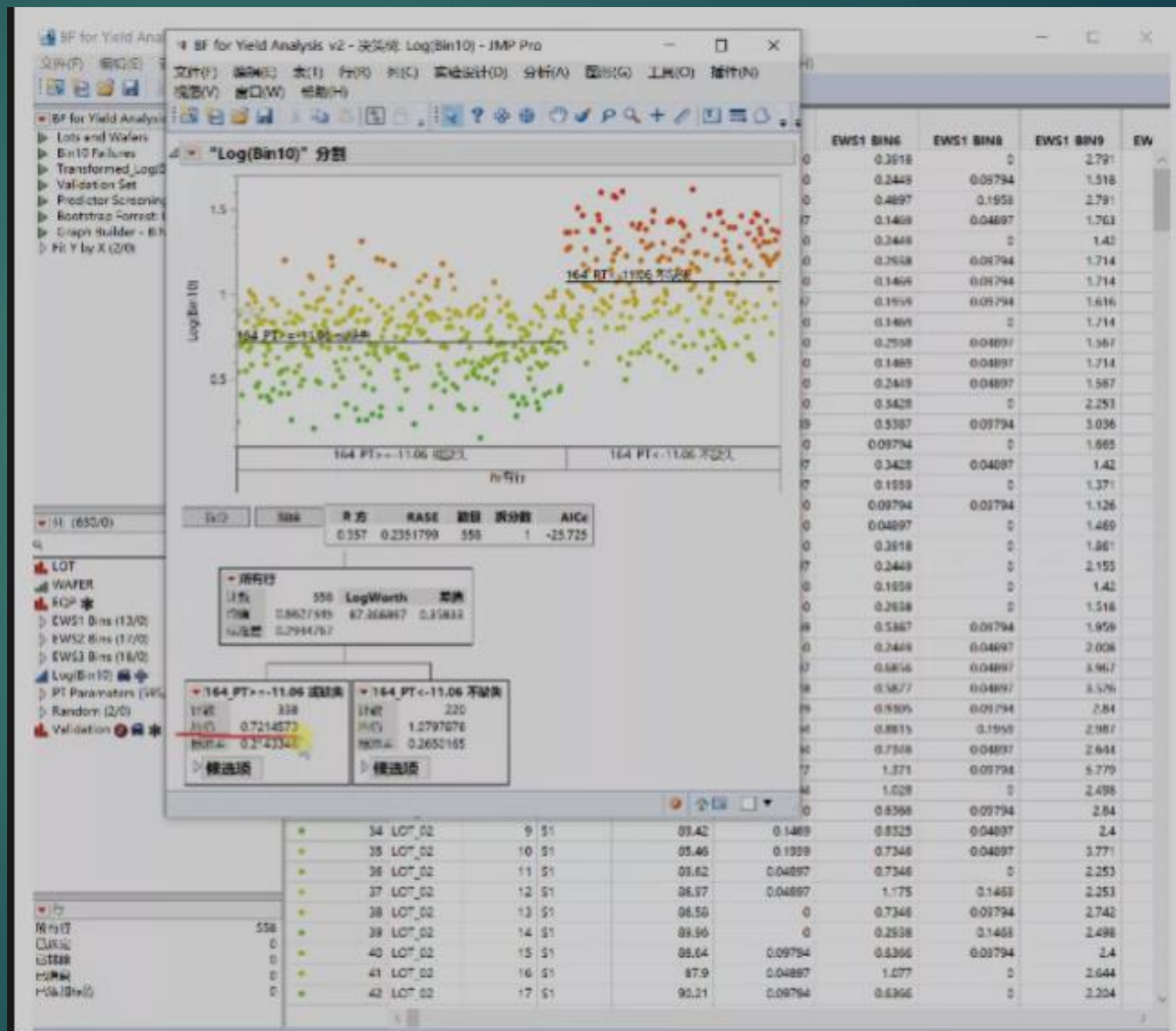
Step 3 數據變換

由JMP直觀的長條圖，可以看到，由於 EWS1_BIN10 資料近似服從對數正態分佈 (偏態分佈)，因此工程師對其進行對數變換，從而轉換為近似正態分佈 (對稱分佈)



Step 4 預測建模方法 選擇

工程師最終發現基於決策樹的集成演算法**隨機森林 (Bootstrap Forest)** 等機器學習演算法更適合應對此類問題

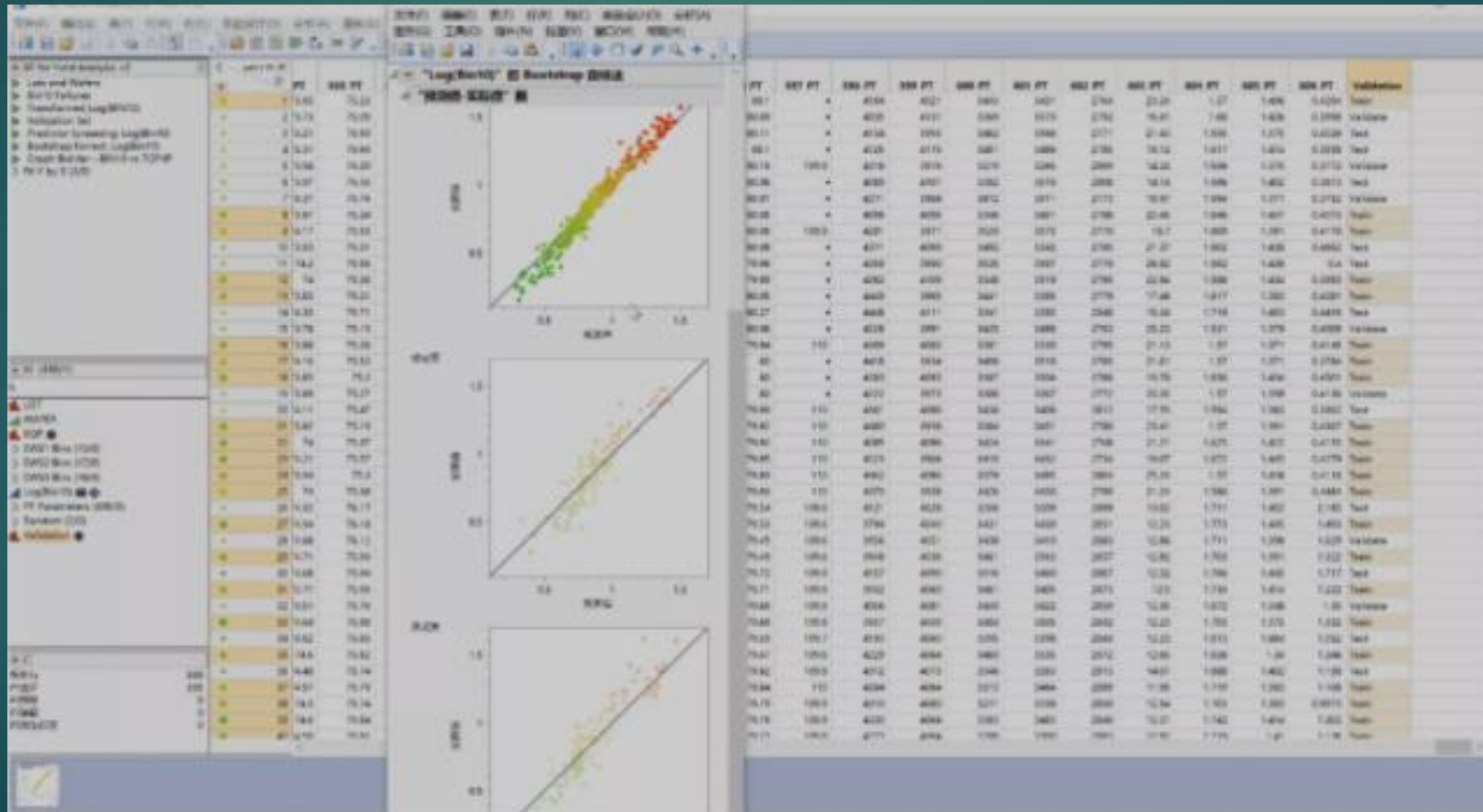


隨機森林是基於決策樹的一種集成機器學習方法

其差異性在於：

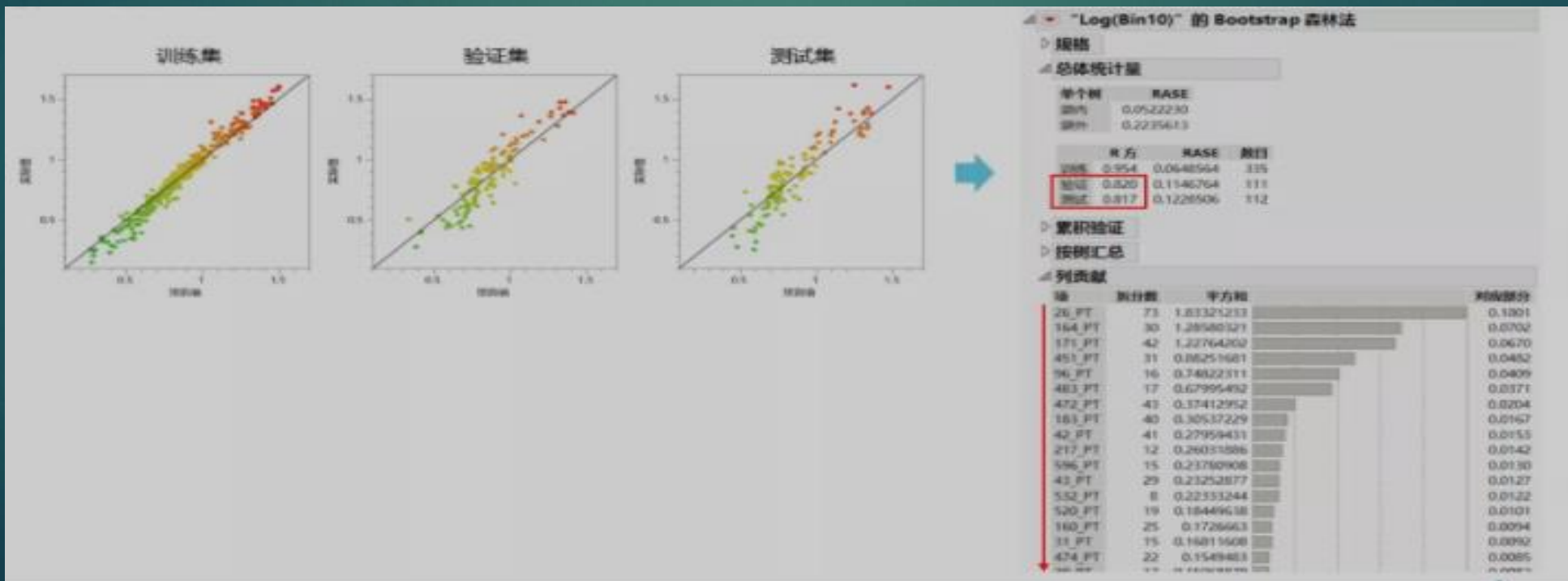
- **隨機**：樣本的隨機性 / 預測變數的隨機性
- **森林**：構建多棵決策樹共同參與決策

從改進單棵決策樹的準確性，並減少方差，讓預測建模效果更好



Step 5 模型診斷及結果分析

從模型的輸出報表，可以看到：
訓練集獲得了較高的 R2，驗證集和測試集的 R2 損失較小，並且兩者的偏差較小，因此，**所建立的模型可以很好地解釋過程變異**
此外，通過查看對森林有非零貢獻的列變數，識別可能影響 Y 的因數，**而根因將是那些貢獻最大的因數，因為特定變數與回應的相關性越大，被選擇的頻率就越高**

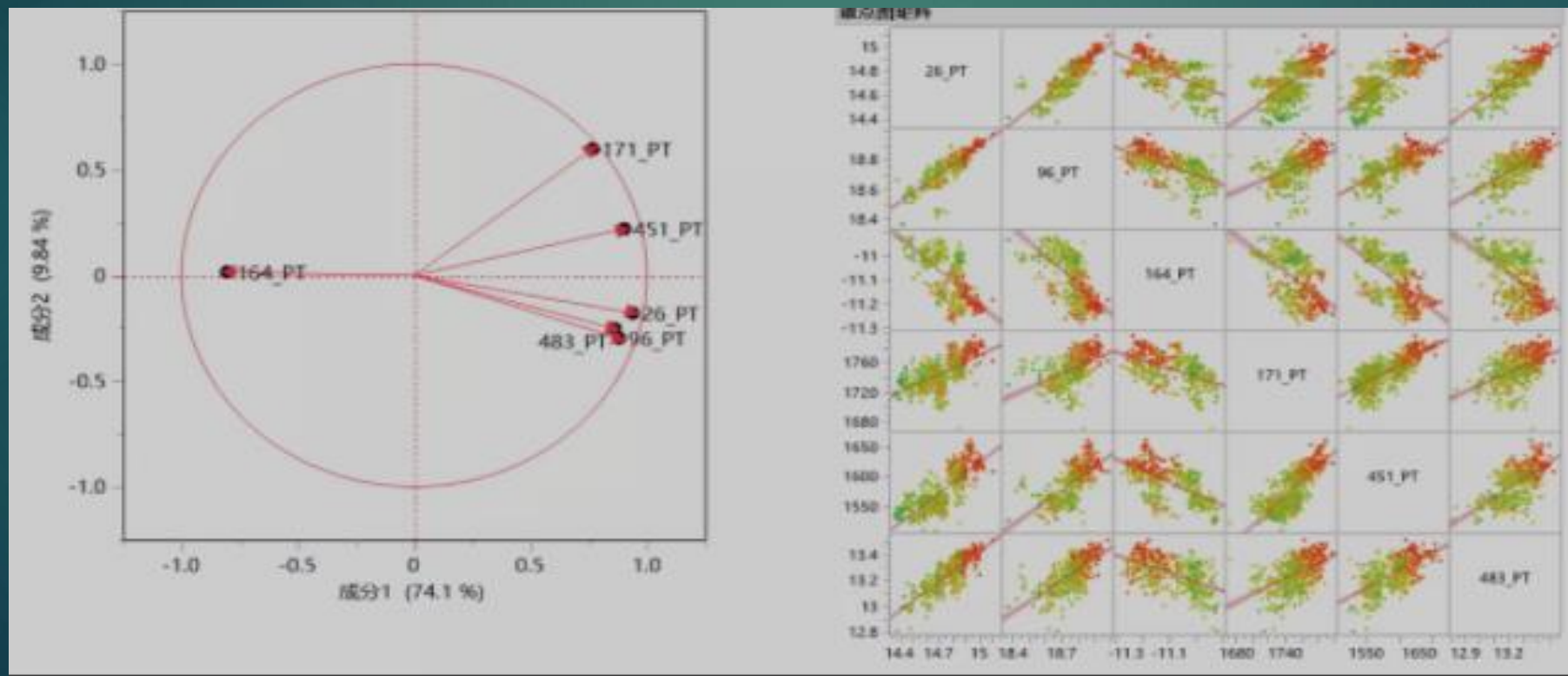


Step 6 根因識別

工程師將列貢獻最大的變數提取出來後發現：

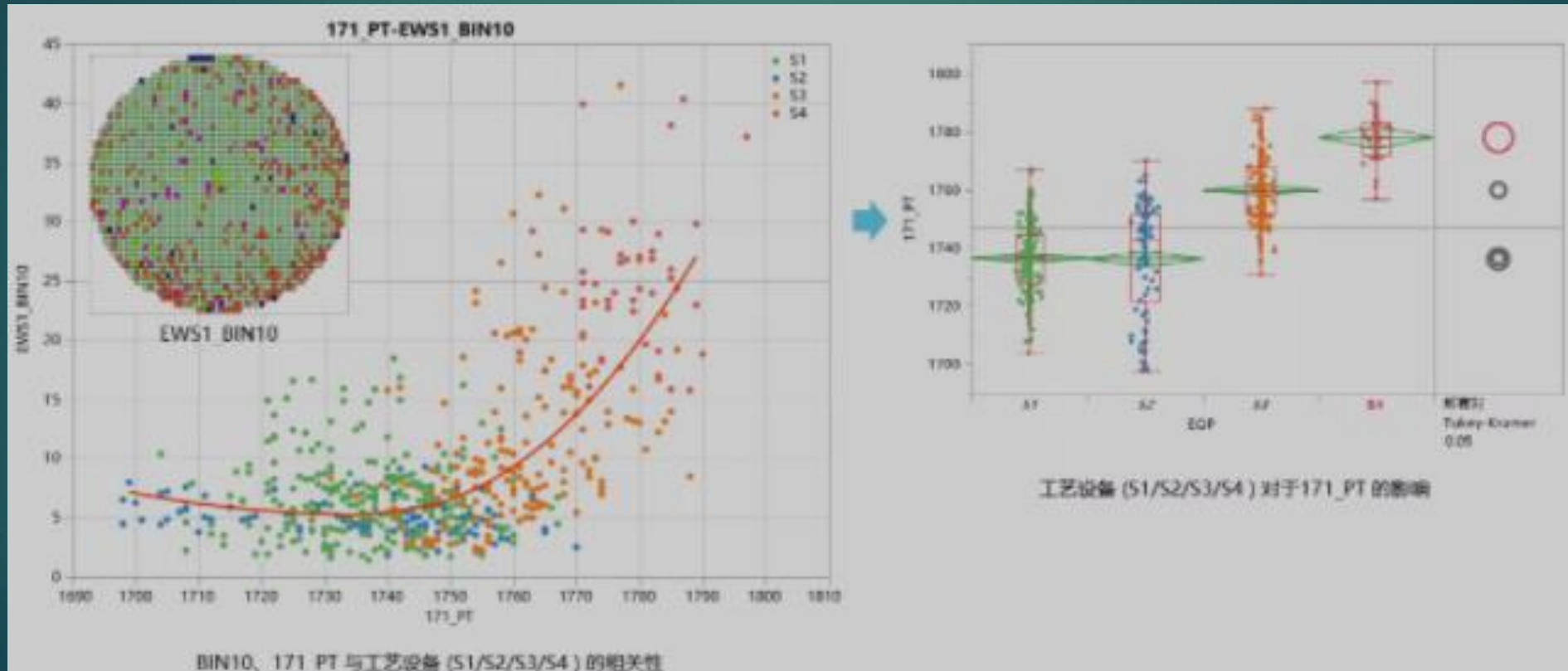
- 列貢獻大 (本例選取Top6) 的預測變數之間可能存在共線性，因此，可以結合主成分或多元相關分析和工程經驗進一步甄別；

對於本例，工程師最終將 171_PT 識別為最重要的影響因數



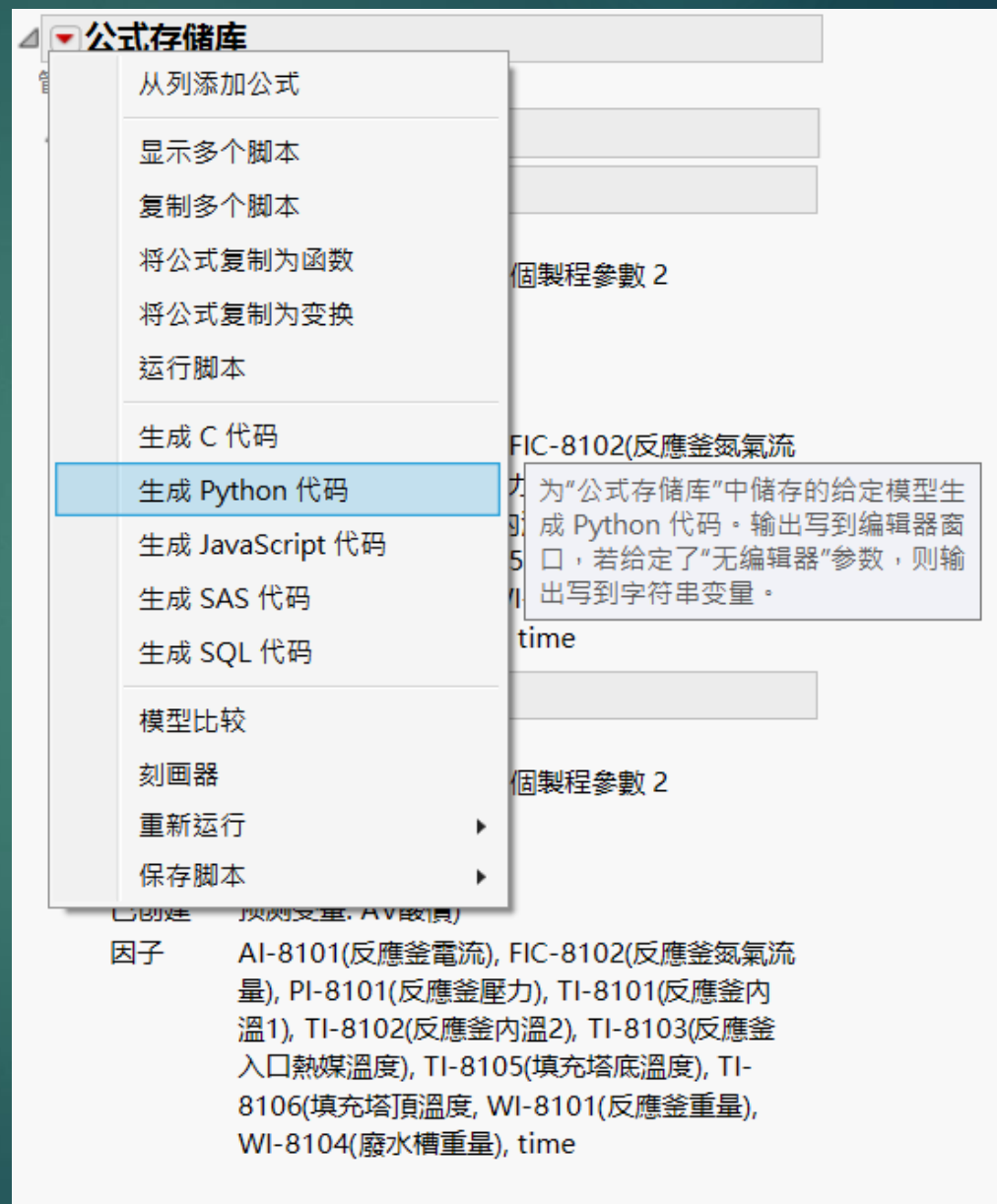
Step 7 結果驗證

工程師通過進一步追溯變異源，發現 171_PT 和某關鍵工藝設備之間存在明顯的相關性，Tukey-Kramer 多重比較確認了該差異的顯著性；因此，工程師可以展開更進一步的排查設備問題，找到導致過程輸出變異和波動的原因，從而形成有針對性的預防措施和改進措施。



Step 8 機器學習

將預測模型運算公式轉換為
與機械溝通語言，例如 C
Code, Python Code, Java
Script Code, SQL Code 或
SAS Code 讓機器學習



模型部署及應用

- ▶ 越來越多的高端製造企業期望借助不同的工具和方法搭建大資料分析平臺，打造智慧工廠，因此，我們將**類似本案例所獲得模型以通用代碼形式部署於企業級的大資料分析平臺**，這是實現**預測性品質控制和預測性維護**的有效手段之一，也是被證明行之有效的科學方法

敬請指教

黃崇仁 博士
黃譯賢 博士

凱律科技股份有限公司

Phoenix-iptech. Corp.

jen@phoenix-tech.com.tw

www.phoenix-
tech.com.tw

886-7-2851008

