

環氧氯丙烷(ECH) 反應區收率優化

密



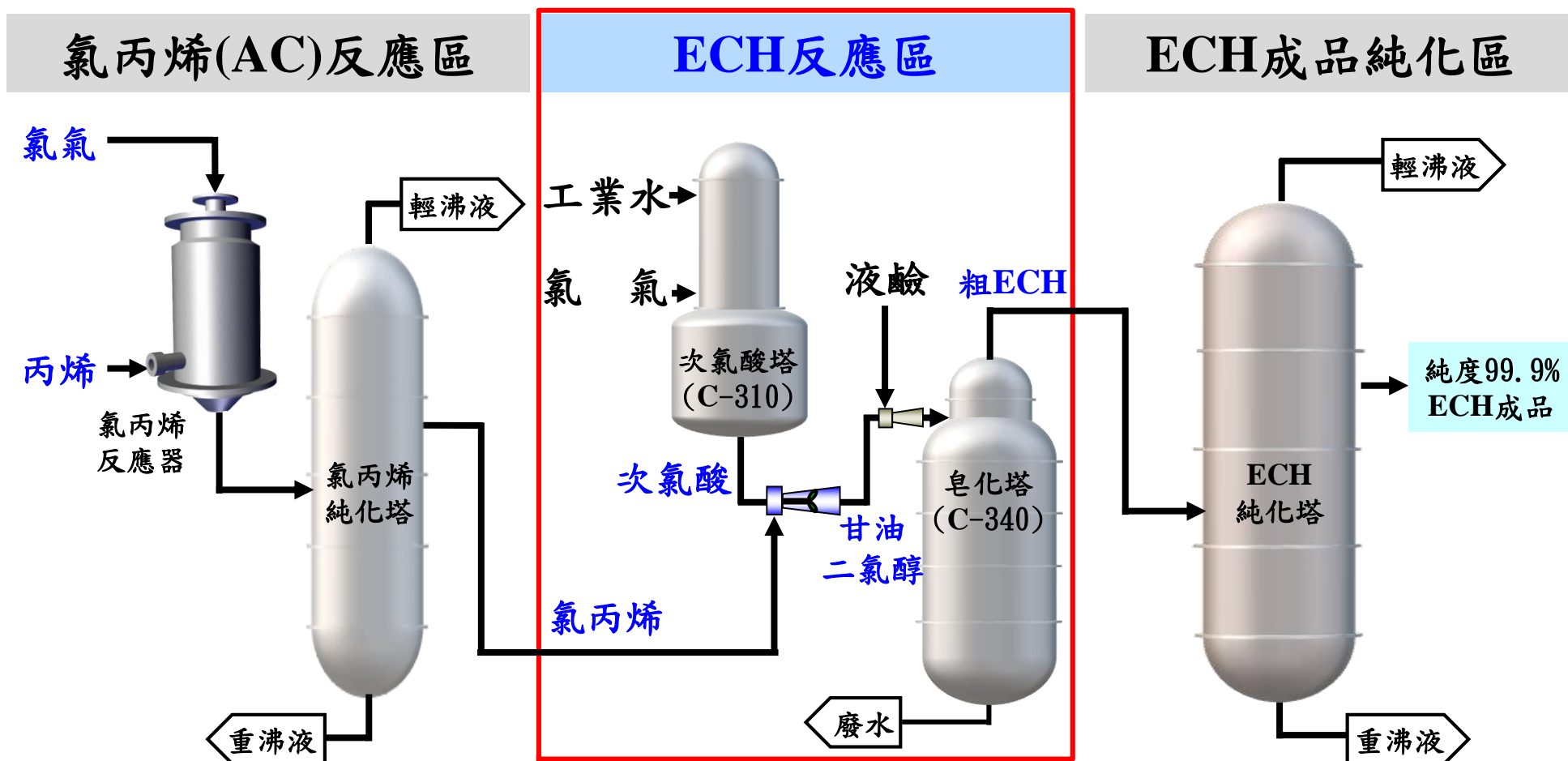
台塑化學品部 ECH廠

報告人：吳承翰 2020.11.13

1. 麥寮ECH廠製程說明
2. 改善動機
3. 解決方案
4. AI模型開發歷程
5. 各階段詳細說明
6. 效益說明
7. 結論及後續推動事項

1. 麥寮ECH廠製程說明(1/2)

- ECH係由氯氣和丙烯生成氯丙烯，與次氯酸(HOCl)生成甘油二氯醇(DCH)，再與液鹼皂化產生粗ECH，再經純化後製成ECH成品。

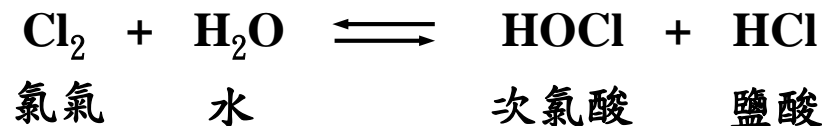


本次AI應用範圍

1. 麥寮ECH廠製程說明(2/2)

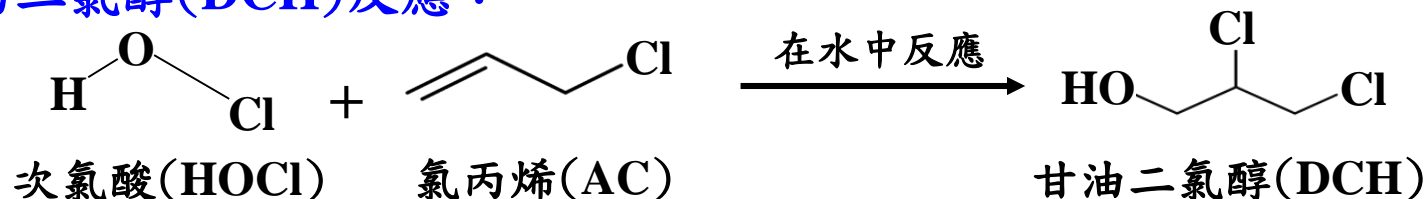
➤ ECH反應區三個主要反應：

● 次氯酸(HOCl)反應：



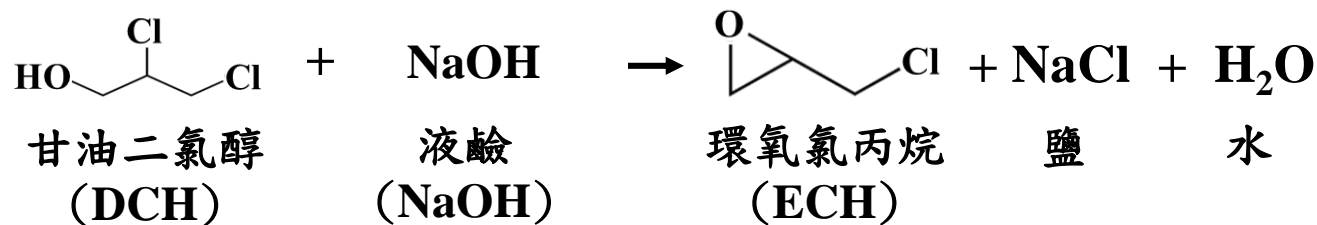
說明：因氯氣不易溶於水，且為放熱反應，故低溫和良好的混合有利次氯酸生成。

● 甘油二氯醇(DCH)反應：



說明：因氯丙烯要溶於水中才可與次氯酸反應，故大量的水有利甘油二氯醇生成。

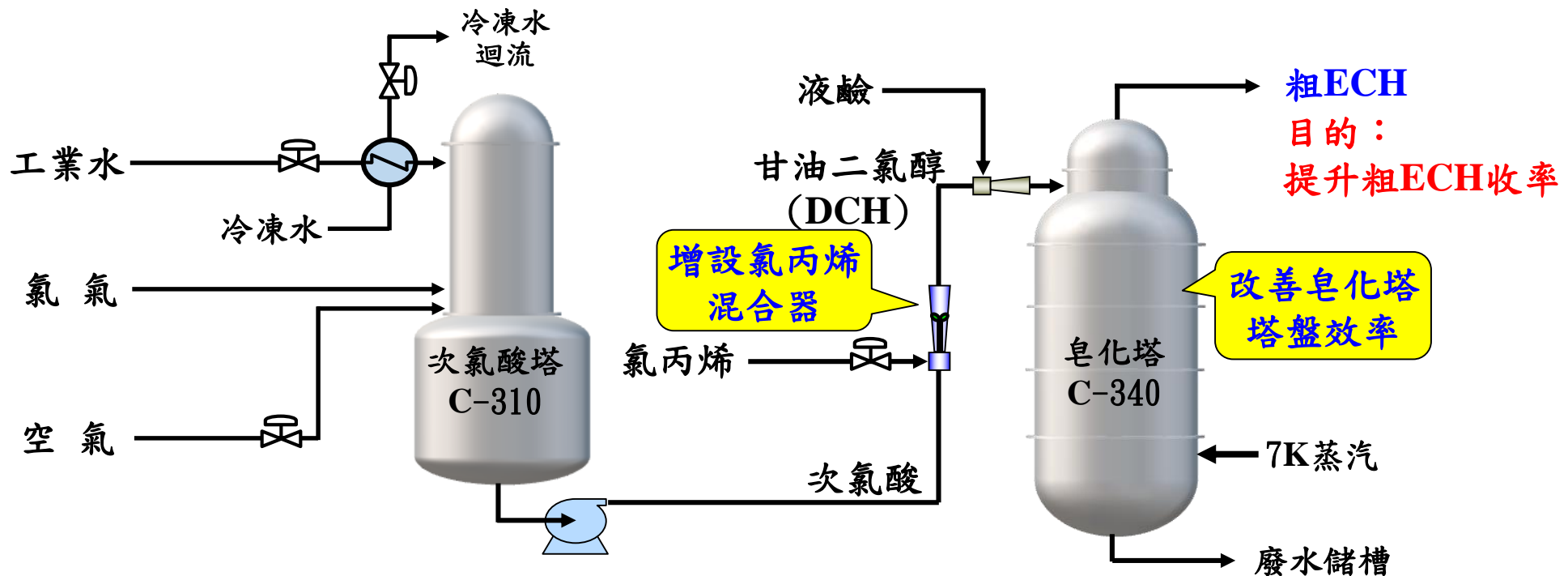
● 環氧氯丙烷(ECH)皂化反應：



說明：甘油二氯醇與液鹼反應非常快速，製得的粗ECH由塔頂汽提出來後冷凝。

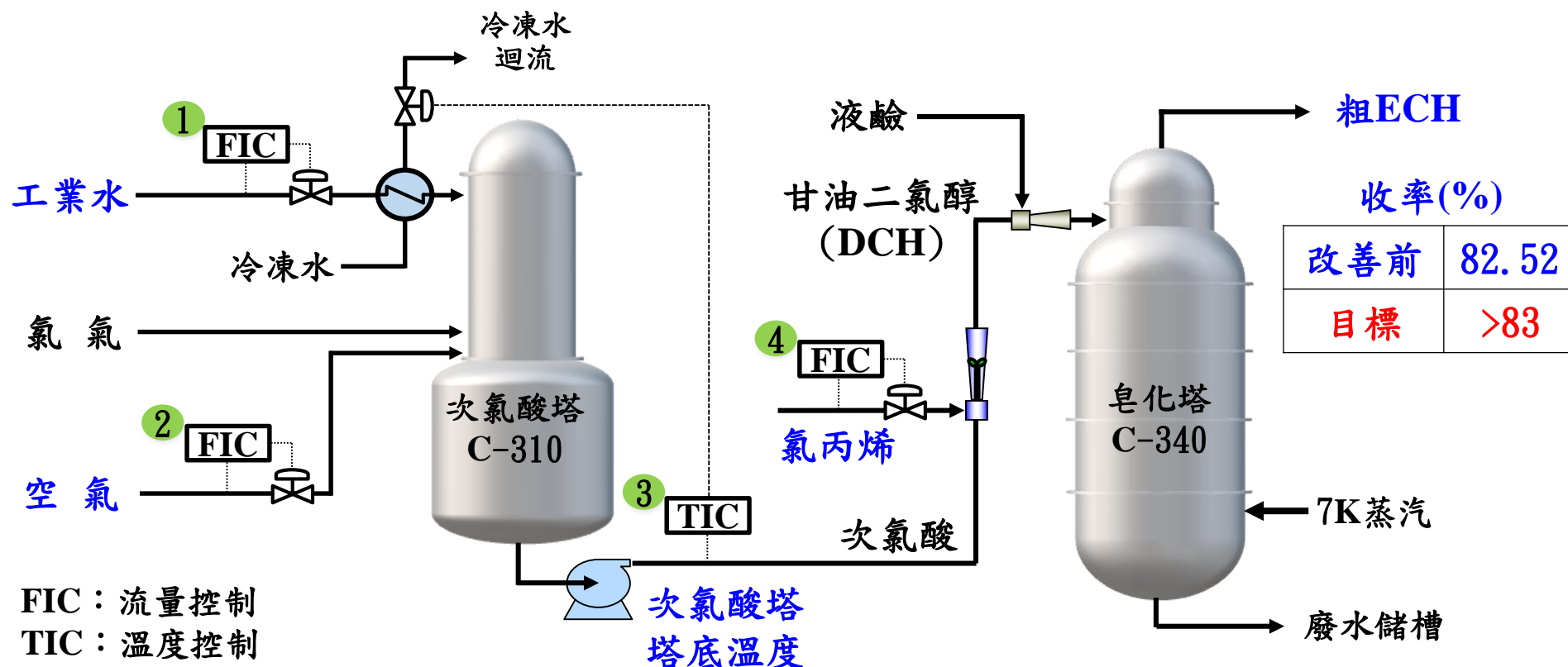
2. 改善動機(1/2)

- 本區在過去已歷經多次的設備改善，如增設氯丙烯混合器和皂化塔(C-340)塔盤效率改善等，收率已從80.63%提升至82.52%。
- 考量設備改善的效益已出現瓶頸，故以優化製程操作參數作為現階段改善方向。在操作調整方面，過去都依照SOP、檢驗數據(檢驗頻率：8小時/次)及操作經驗進行調整，可改善的方向與最佳製程操作參數不易找出。



2. 改善動機(2/2)

- 經Aspen模擬，發現ECH反應區可再調整的重要控制參數為以下四個：
 - ① 工業水流量
 - ② 空氣流量
 - ③ 次氯酸塔塔底溫度
 - ④ 氯丙烯流量
- 本案擬根據Aspen模擬的結果，進行AI操作優化，期以再增加中間產物（次氯酸和甘油二氯醇）的產量，達到提升粗ECH收率的目標。



3. 解決方案

- 運用AI對ECH反應區的製程控制點進行優化，依下列4個步驟來實現增加ECH收率之目標。



4. AI模型開發歷程

Start

正式上線

2019年 6/3

11/28

1. Aspen
模擬

找出4個製程
操作參數，
增加中間產物
的產量，提升
粗ECH收率。



2. AI收率
模型

篩選出15個特徵，
選用 XGBoost 為 AI
收率模型演算法。



3. AI模型
驗證

驗證收率模型的準確度，
收率目標 $MAE < 0.7\%$ ，
並運用 SHAP 演算法確保
預測趨勢符合化工原理。



4. AI操作
建議

用基因演算法
(GA)，算出可使
粗ECH獲得最大
收率的控制參數。



5. 正式
上線

粗ECH收率達83.18%
(+0.66%)，成品增加
2.28噸/日。

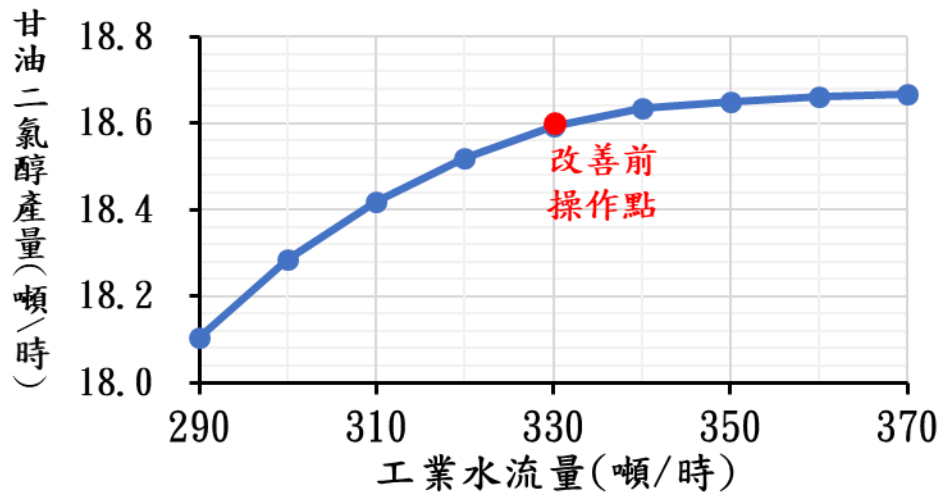
MAE(平均絕對誤差, Mean Absolute Error)

5.1 第一階段：Aspen模擬(1/2)

➤ 為了要找尋最佳的操作條件，用 Aspen 軟體建立 ECH 反應區穩態模型，模擬製程操作參數與中間產物(次氯酸和甘油二氯醇)的關係，發現調整4個製程操作參數，可再增加中間產物的產量，其模擬的趨勢圖及結果如下：

➤ 工業水流量的模擬：

甘油二氯醇產量 vs 工業水流量

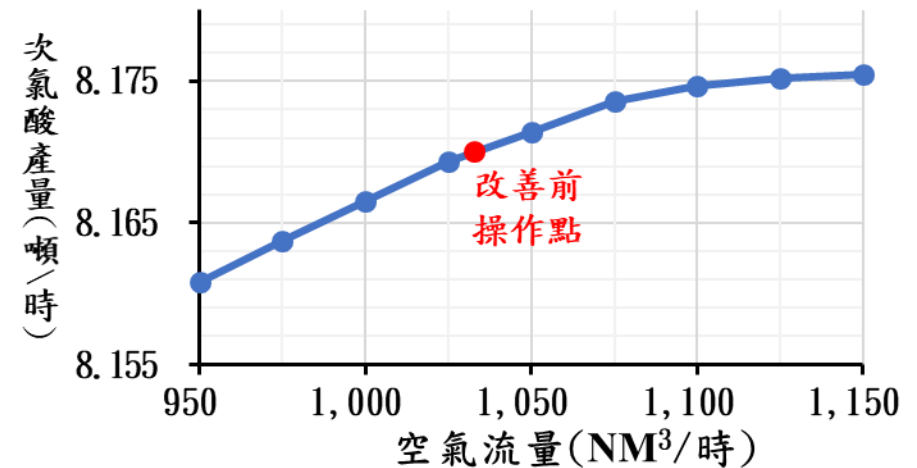


● 結果：

適量提升工業水流量，可再提高中間產物甘油二氯醇產量。

➤ 空氣流量的模擬：

次氯酸產量 vs 空氣流量



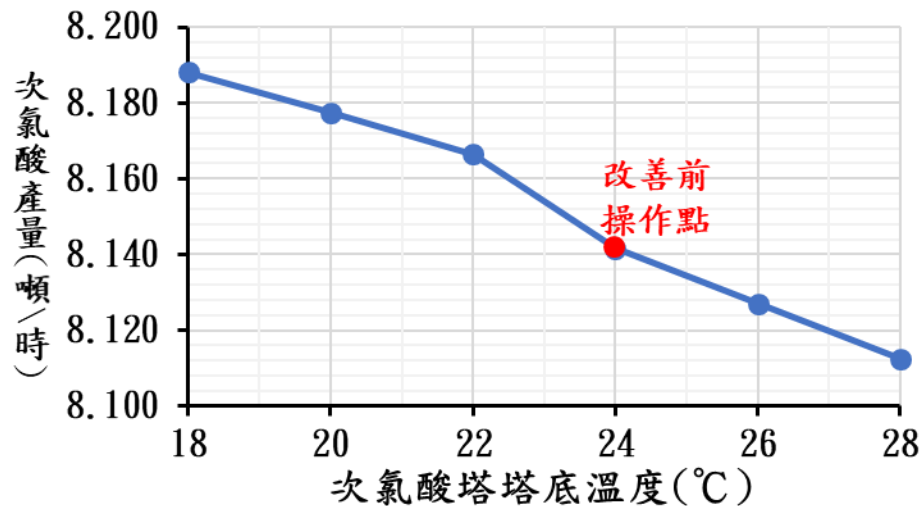
● 結果：

適量提升空氣流量，可提高次氯酸塔氣體負荷量，有助於氯氣和水產生次氯酸。

5.1 第一階段：Aspen模擬(2/2)

➤ 次氯酸塔塔底溫度模擬：

次氯酸產量 vs 次氯酸塔塔底溫度

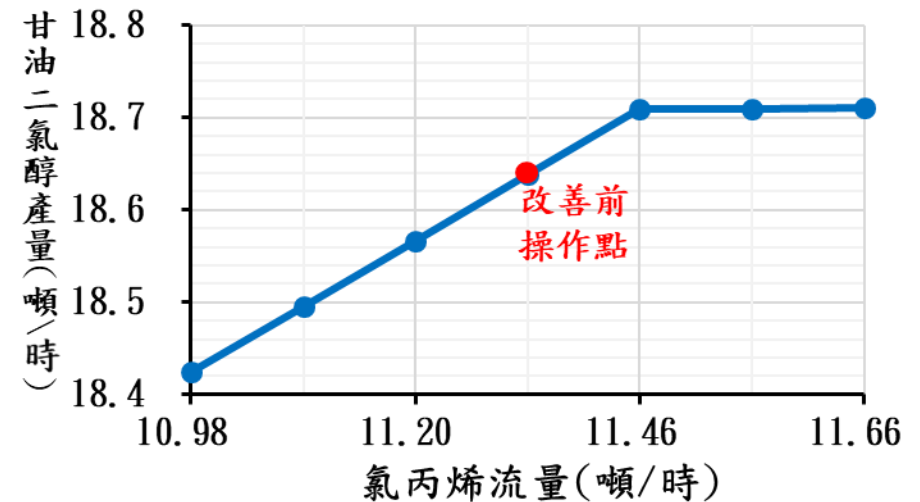


● 結果：

溫度越低，次氯酸產量越高，然過度降低溫度會大幅增加耗用冷凍水的成本。

➤ 氯丙烯流量模擬：

甘油二氯醇產量 vs 氯丙烯流量



● 結果：

適度增加氯丙烯流量，有利中間產物甘油二氯醇產量的提升。

➤ 藉由上述4個控制參數的模擬結果，發現製程還有改善的空間，後續將進行AI模型的建立。

- 本階段目的係建立「粗ECH純度」與「粗ECH產量」的AI預測模型，再藉由粗ECH收率公式算出收率。

$$\text{粗ECH收率} = \frac{\text{粗ECH純度} \times \text{粗ECH產量}}{\text{理論產量}} \times 100\%$$

Step1. 資料收集

(1) 時間：2014年1月～2019年5月

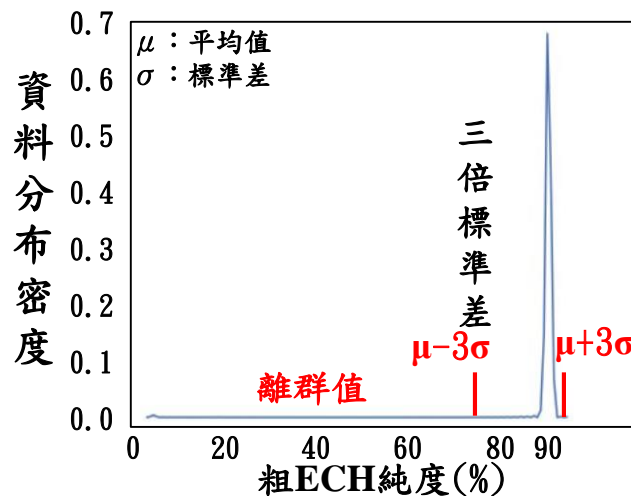
(2) 筆數：60萬筆(包含製程參數和檢驗數據)，其中訓練樣本：48萬筆；
驗證樣本：12萬筆。



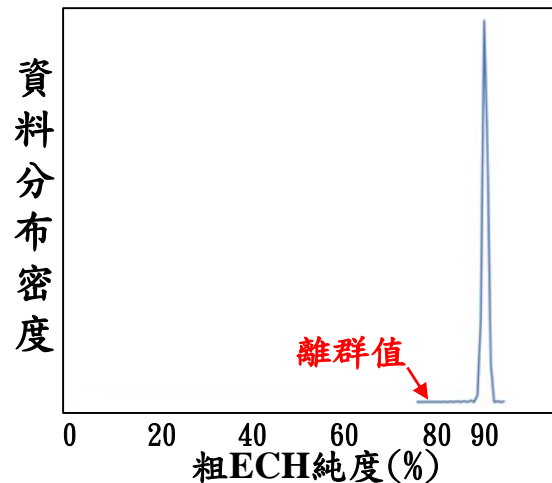
Step2. 資料前處理問題點與改善方案：

- 問題點：原僅使用三倍標準差進行篩選，模型準確度不高，容易導致AI操作建議產生錯誤的建議值。
- 改善方案：經檢討原因，部分數據非常態分布，用三倍標準差篩選後仍有離群值，故再參考現場操作經驗及專業知識進行篩選獲得改善。

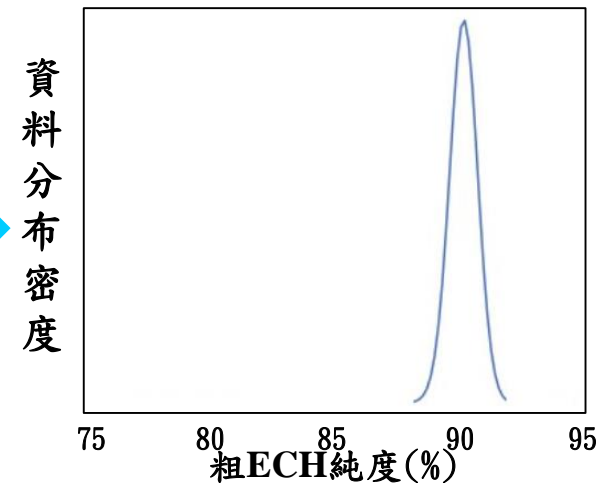
粗ECH純度離群值篩選改善前後分布圖



原始數據非常態分布



篩選後仍有離群值

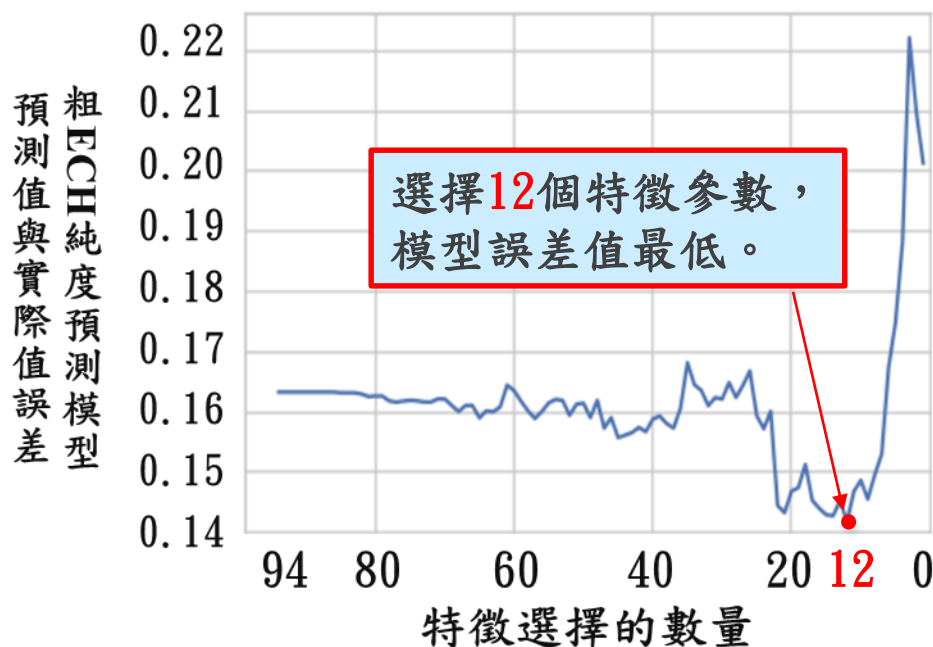


改善後無離群值

Step3.1 特徵篩選

1. 遞迴特徵篩選法(Recursive Feature Elimination, RFE)：

先用94個製程參數，進行各種AI演算法的評估後，選用XGBoost來進行多輪訓練，每輪訓練後，移除權重最低的特徵參數，再重新進行下一輪訓練，找出預測值與實際值誤差最小的特徵參數。

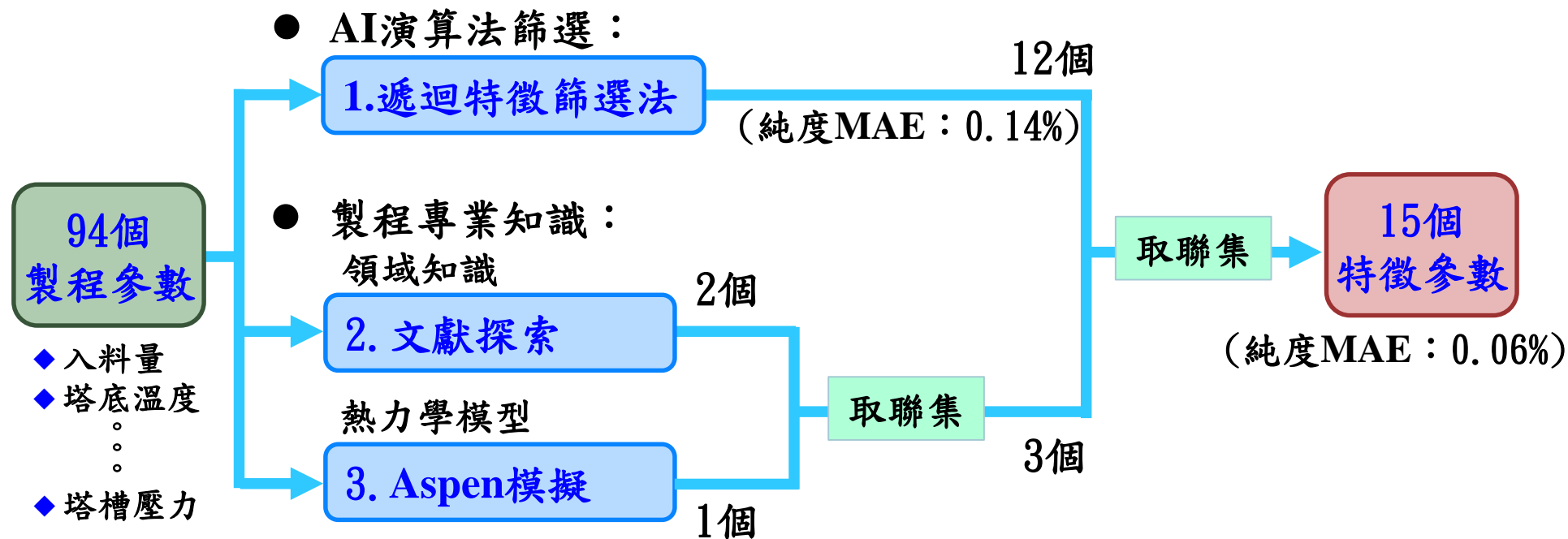


● 篩選出的12個特徵參數：

1. 皂化塔液鹼流量
2. 氯丙烯壓力控制
3. 次氯酸塔液鹼流量
4. 氯丙烯濃度
5. 次氯酸塔氯氣流量
6. 皂化塔入口層與塔底壓差
7. 次氯酸塔塔底溫度
8. ECH反應區次氯酸入口流量
9. 皂化塔塔底溫度
10. 皂化塔入口層壓力
11. 次氯酸塔活性氯濃度
12. 皂化塔塔頂壓力

Step3.2 特徵篩選困難點與改善方案：

- 困難點：原僅採用遞迴特徵篩選法，模型預測準確度的平均絕對誤差(純度MAE：0.14%)雖低，但所篩選的12個特徵參數，只找到由Aspen模擬出來的次氯酸塔塔底溫度，並未包含另外3個重要製程參數(工業水流量、空氣流量及氯丙烯流量)。
- 改善方案：根據文獻探討及Aspen模擬，組合出新的特徵參數加入模型訓練，降低預測誤差(純度MAE：0.14% → 0.06%)，並使模型符合化工原理。

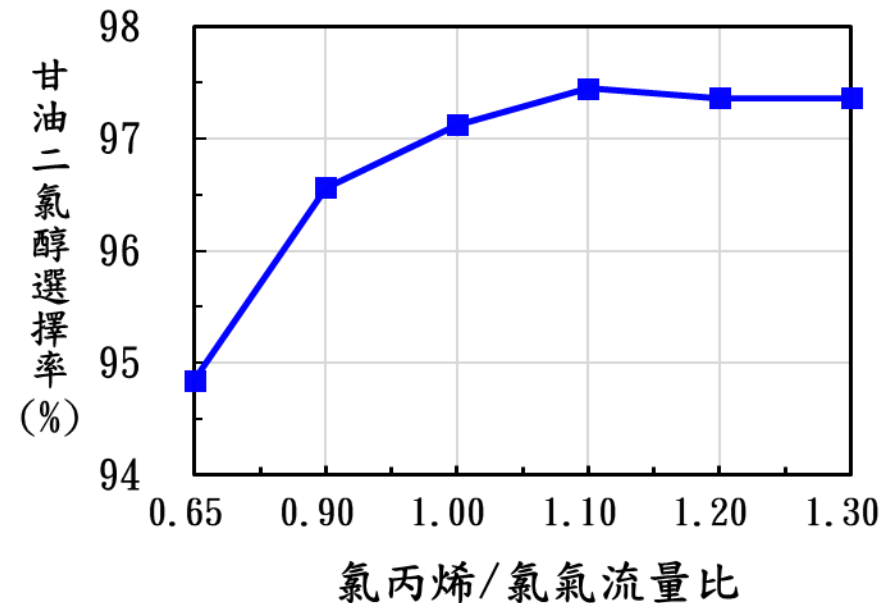
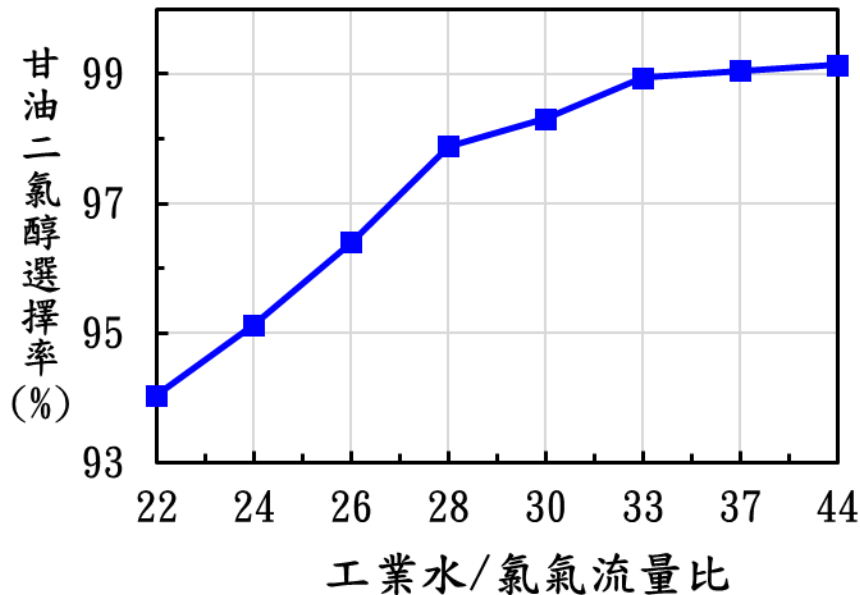


Step3.3 特徵篩選

2. 文獻探索：

I&EC Research於2012年發表的文獻顯示：甘油二氯醇(DCH)的選擇率(Selectivity)，與「工業水/氯氣流量比」和「氯丙烯/氯氣流量比」這2個比值有關係，故再將上述2個特徵，分別加入模型參與訓練。

文獻趨勢圖



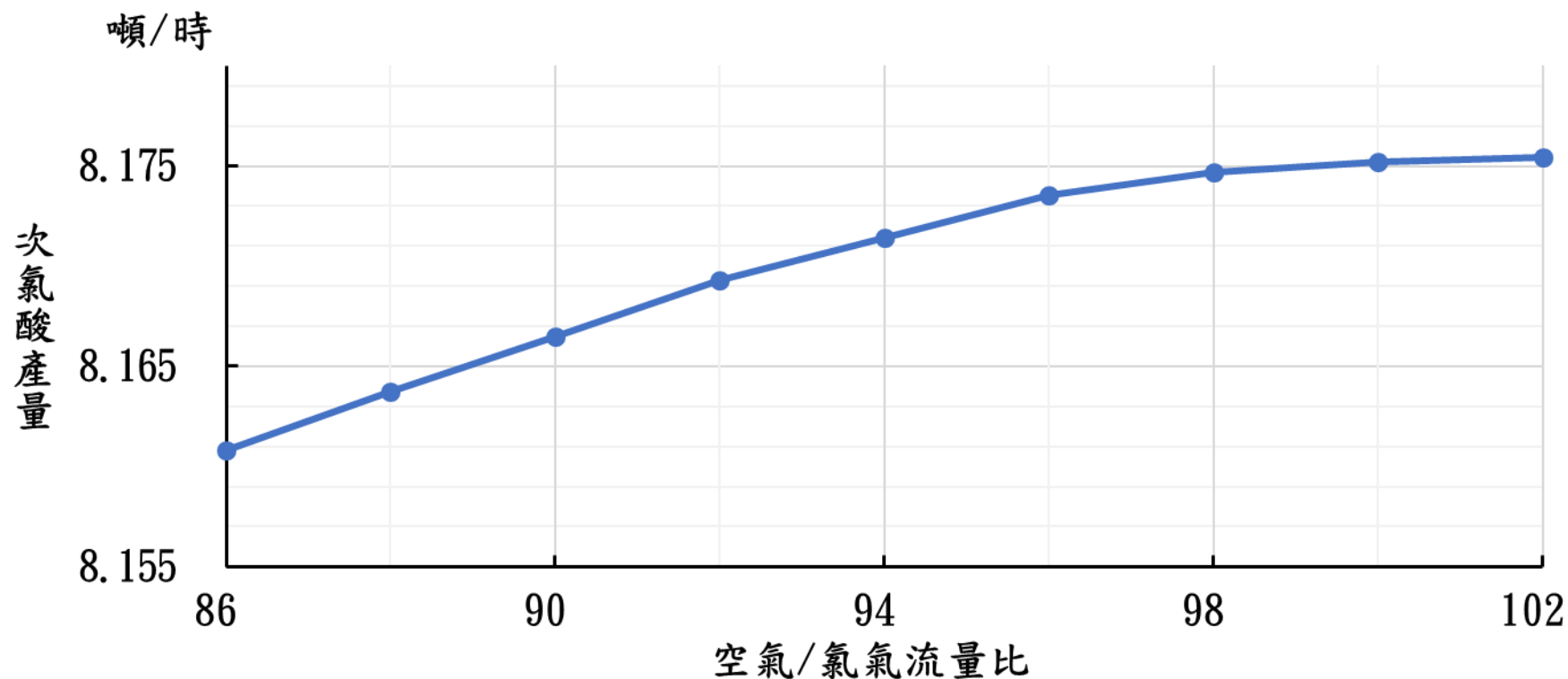
$$\text{選擇率(\%)} = \frac{\text{甘油二氯醇 wt\%}}{\text{甘油二氯醇 wt\%} + \text{副產物 wt\%}}$$

I&EC Research : Industrial & Engineering Chemistry Research

Step3.4 特徵篩選

3. Aspen模擬：

根據Aspen模擬結果，發現次氯酸塔(C-310)空氣流量與次氯酸產量有部分正相關，故增加「空氣/氯氣流量比」作為特徵參數。



Step4. 模型建立(演算法選定)

- 本階段分別訓練「粗ECH純度預測模型」和「粗ECH產量預測模型」，目的係計算粗ECH收率。

$$\text{粗ECH收率} = \frac{\text{粗ECH純度} \times \text{粗ECH產量}}{\text{理論產量}} \times 100\%$$

- 根據製程的反應特性，對2個預測模型各測試6種演算法，最終選擇XGBoost作為預測模型的演算法，並用SHAP演算法評估是否符合化工原理。

粗ECH純度預測模型

項次	演算法	MAE (註1)	是否符合 化工原理
1	XGBoost	0.0585	是
2	KNN	0.0603	否
3	隨機森林	0.0629	是
4	SVR	0.0652	否
5	決策樹	0.0702	否
6	DNN	0.0819	是

粗ECH產量預測模型

項次	演算法	MAE (註1)	是否符合 化工原理
1	Ridge	2.6712	否
2	Lasso	2.6784	否
3	SVR	2.7720	否
4	XGBoost	2.8320	是
5	隨機森林	2.9928	是
6	決策樹	3.1848	否

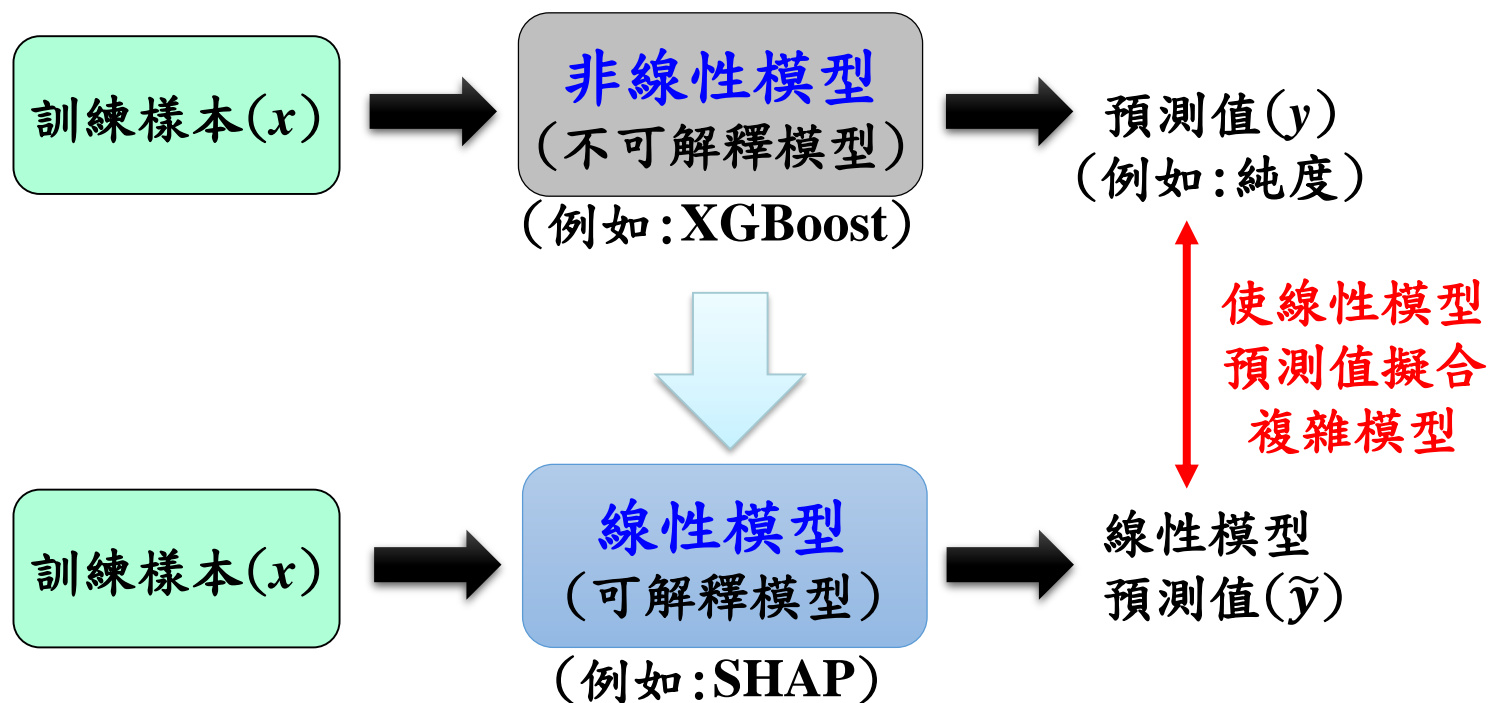
註1：MAE(平均絕對誤差, Mean Absolute Error)

$$MAE(y, \hat{y}) = \frac{1}{n} \sum_{i=1}^n |\hat{y}_i - y_i|$$

\hat{y}_i ：預測值 n ：樣本數 y_i ：實際值

► 模型預測趨勢解釋：

- (1) 目前大多數的AI演算法難以解釋模型的輸入跟預測值之間的對應關係，應用於化工製程會有疑慮，而SHAP演算法可以幫助我們解釋模型運作方式。
- (2) SHAP演算法的理念：是將一個任意的複雜模型，藉由具有解釋能力的線性模型去擬合，再由線性模型解釋模型輸入值(x)與預測值(y)的關係。

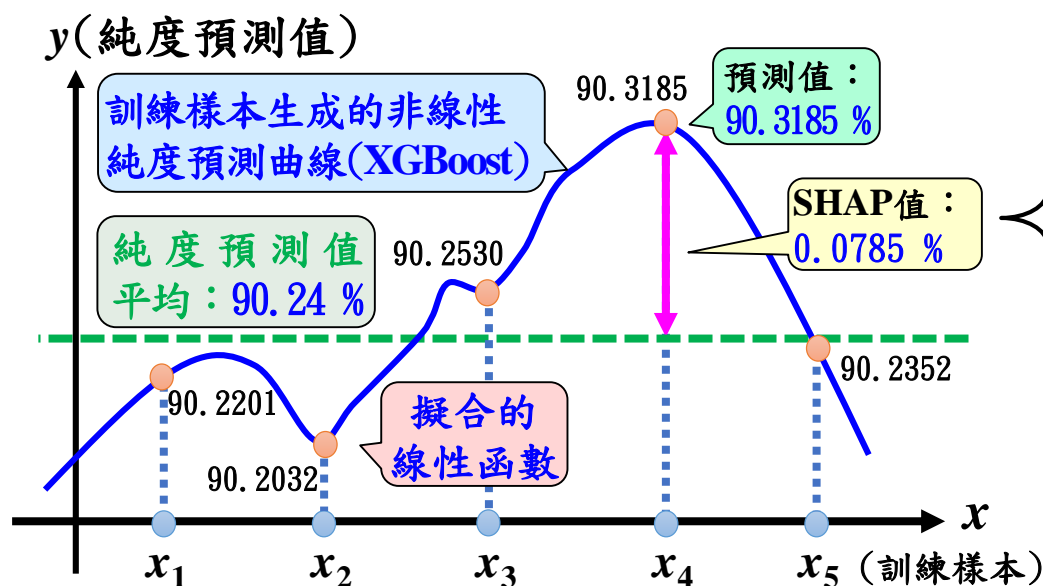


► 模型預測趨勢解釋：

(3) 藉由無數個線性函數來擬合出複雜的非線性模型，針對每個局部區域，轉換成輸入與輸出的線性函數。

(4) 將模型的預測結果轉換為：「預測平均值」+「15個特徵參數的SHAP值總和」。

● 純度模型預測趨勢圖：



項次	特徵參數	操作條件	SHAP值(%)
1	氯丙烯/氯氣流量比	1.015	0.0502
2	工業水/氯氣流量比	30.05	0.0415
3	次氯酸塔塔底溫度	23.50	0.0231
4	空氣/氯氣流量比	87.68	-0.0018

⋮

14	皂化塔入口層壓力	0.05	-0.0127
15	皂化塔塔頂壓力	0.00	-0.0029
15個特徵SHAP值總和			0.0785

模型預測值

90.3185 %

SHAP轉換

=

模型訓練資料的預測平均值

90.24 %

+ 15個特徵參數SHAP值總和

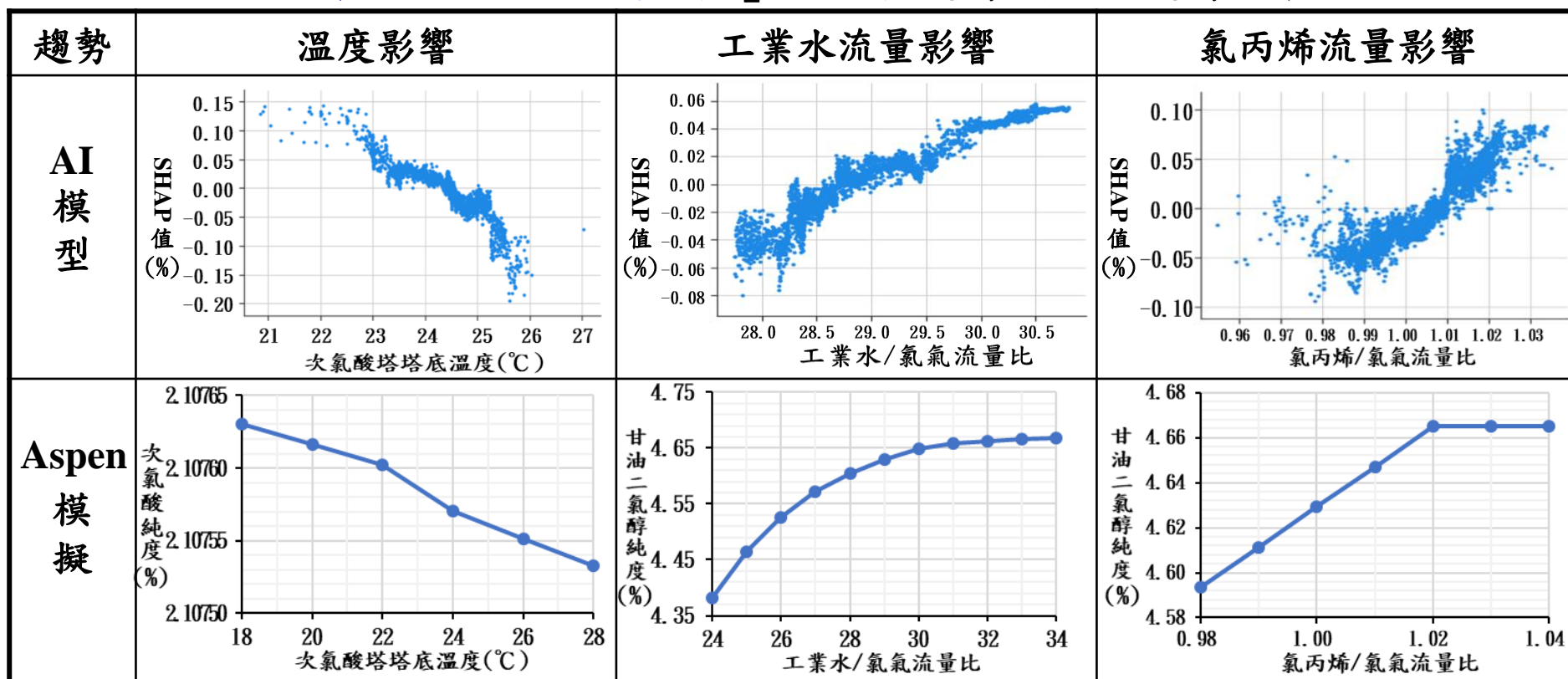
+

0.0785 %

► 化工原理驗證

(1) 將每個特徵參數在不同的操作條件下，所對應SHAP值，畫成AI模型趨勢圖（如下圖），再與Aspen模擬趨勢比較，發現AI模型的預測方式與模擬趨勢相符。

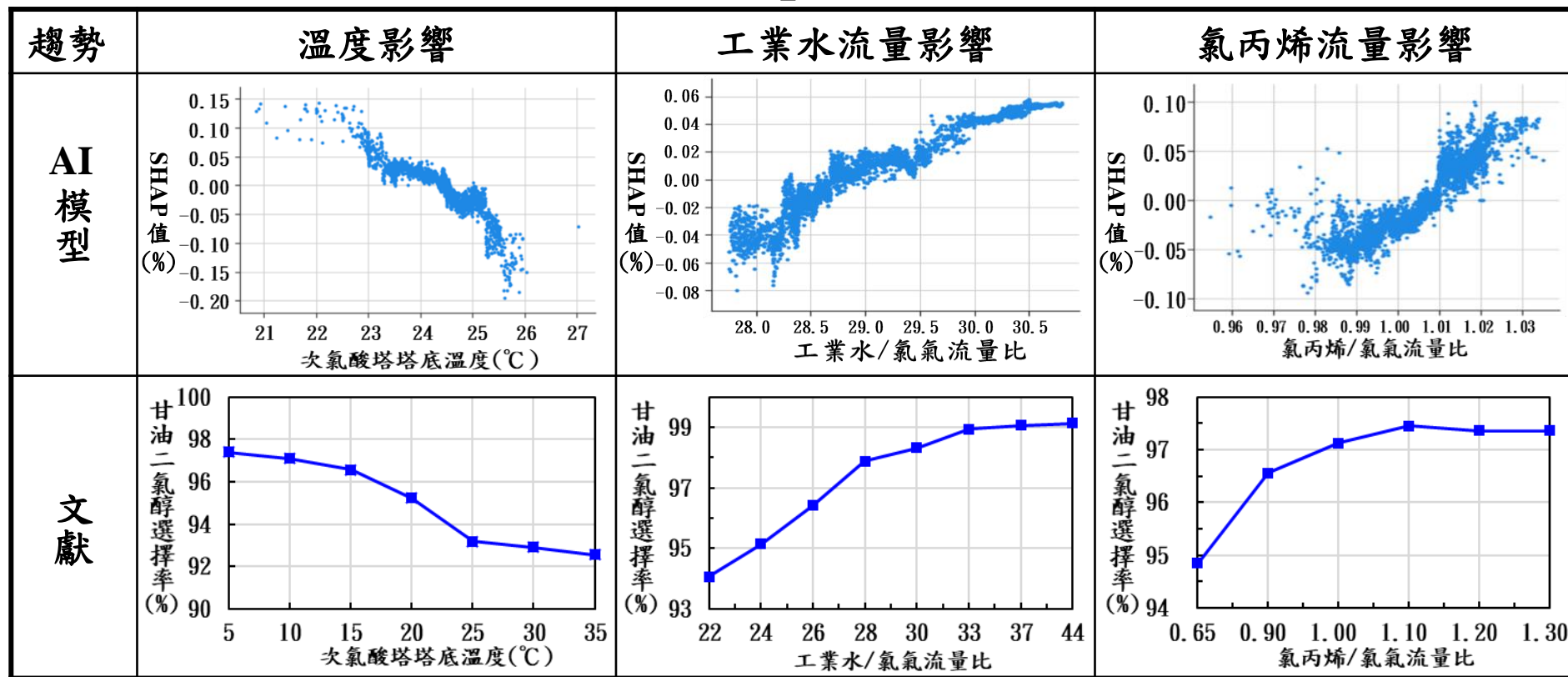
「粗ECH純度預測模型」AI模型趨勢與模擬趨勢比較



➤ 化工原理驗證

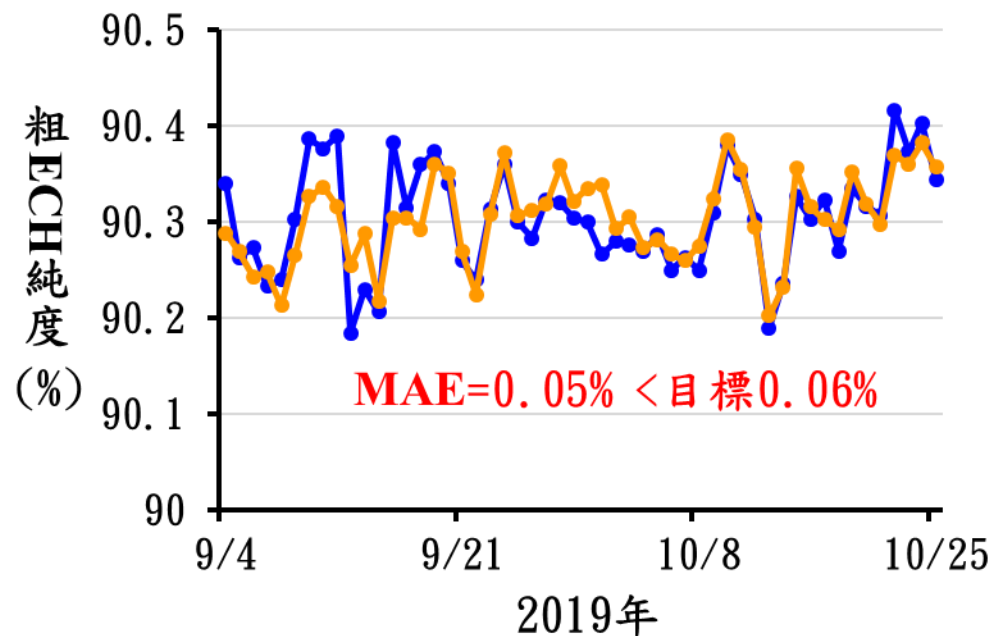
(2)比較I&EC Research 2012年的文獻，AI模型預測趨勢與文獻的實驗數據趨勢相符，證明本案訓練出的預測模型符合化工原理。

「粗ECH純度預測模型」預測趨勢與文獻趨勢比較



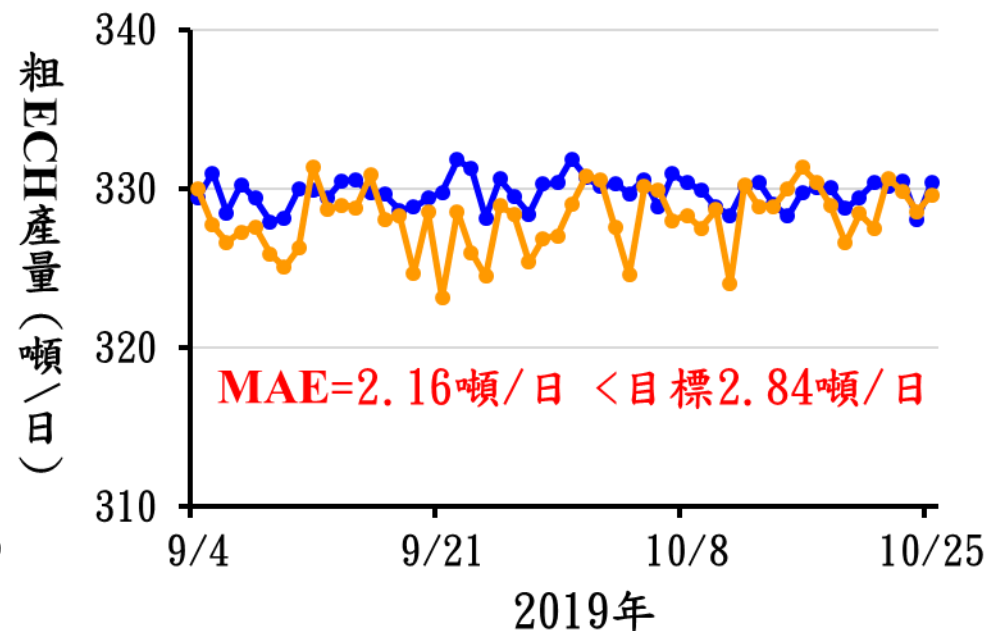
- 模型驗證時間為2019/9/4~2019/10/25。
- 驗證結果：純度與產量變化趨勢相近。

粗ECH純度預測圖

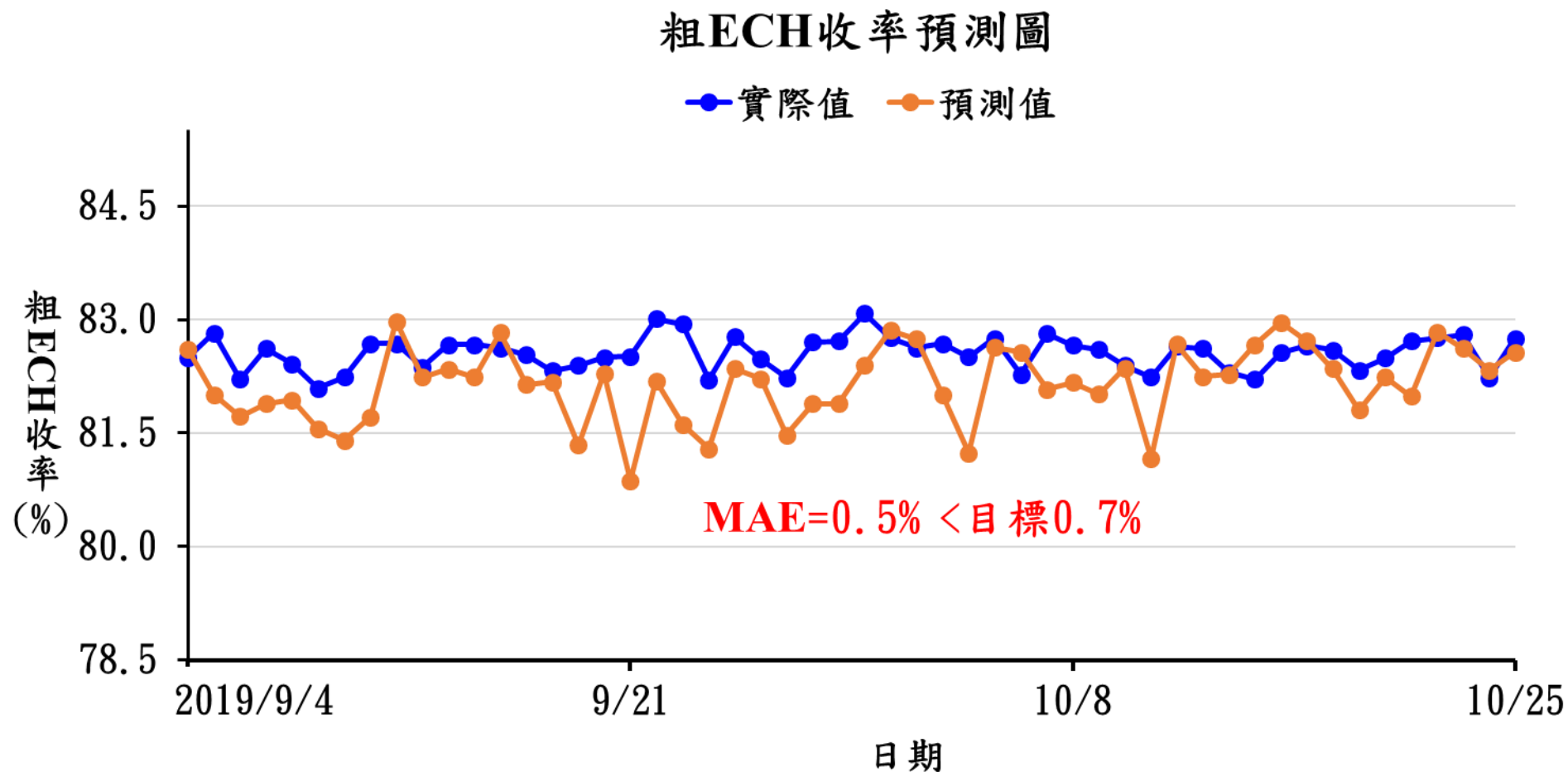


—●— 實際值
—●— 預測值

粗ECH產量預測圖

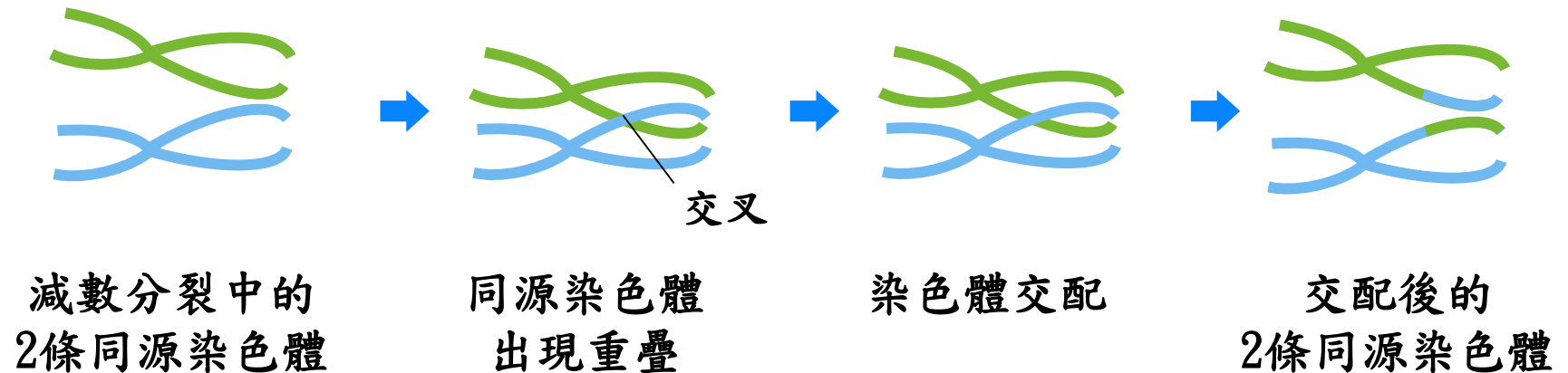


- 模型驗證時間為2019/9/4~2019/10/25。
- 驗證結果：粗ECH收率的變化趨勢相近。



5.4 第四階段：AI操作建議(1/5)

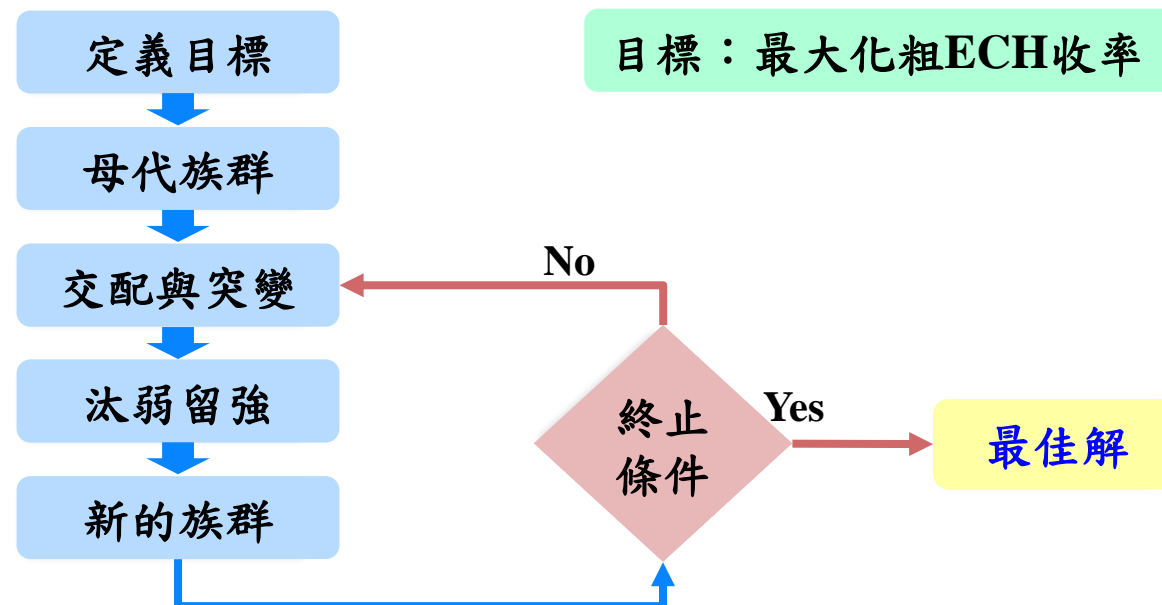
- 本階段的目的，係使用AI收率模型，找到一組控制參數，使粗ECH收率最大化。
- 基因演算法原理簡述：
 - (1) 基因演算法 (Genetic Algorithm) 運用達爾文「適者生存」的概念，模擬自然界透過交配使基因重組與排列，保留優良基因淘汰不良基因。因具有搜尋最佳解的能力，目前已被工業界廣泛使用。



5.4 第四階段：AI操作建議(2/5)

(2)從15個特徵參數中，針對ECH反應區4個重要的可控制程參數(工業水流量、空氣流量、次氯酸塔塔底溫度及氯丙烯流量，其餘11個製程參數視為常數，不做變動)，藉由基因演算法汰弱留強的法則，找出可使粗ECH收率最大化的控制參數。

➤ 基因演算法運作流程：



5.4 第四階段：AI操作建議(3/5)

➤ 基因演算法運作流程：

(1) 定義目標：

目標：粗ECH收率最大化

$$\text{粗ECH收率} = \frac{\text{粗ECH純度} \times \text{粗ECH產量}}{\text{理論產量}} \times 100\%$$

(2) 以ECH反應區四個重要控制參數，產生母代族群：

染色體	基因A	基因B	基因C	基因D
	次氯酸塔 塔底溫度	工業水/氯氣 流量比	空氣/氯氣 流量比	氯丙烯/氯氣 流量比



產生
母代族群

23.00	27.76	79.77	0.981	染色體1
24.10	28.86	80.87	0.987	染色體2

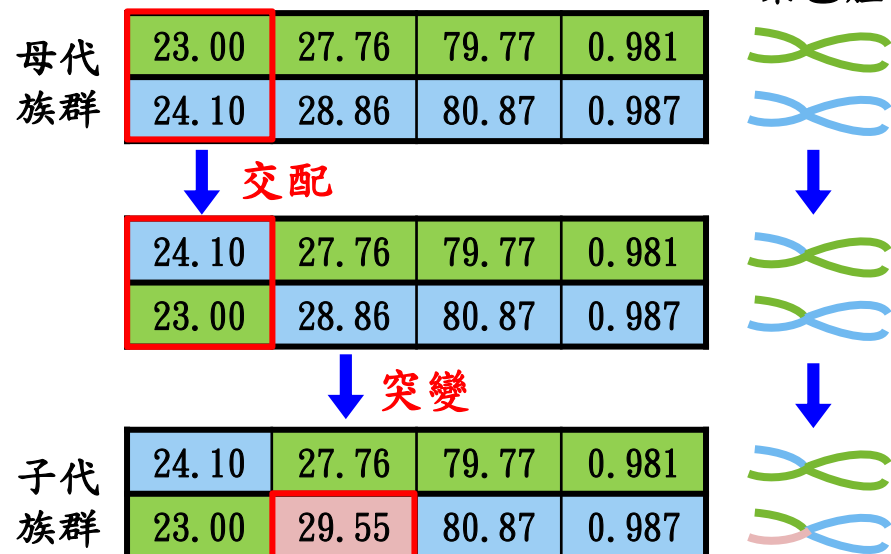
⋮

500組不同的參數組合

5.4 第四階段：AI操作建議(4/5)

➤ 基因演算法運作流程：

(3) 遺傳操作(交配和突變)：



(4) 預測收率和汰弱留強：

		粗ECH 收率	是否 保留
母代 族群	23.00 27.76 79.77 0.981	81.5%	✓
	24.10 28.86 80.87 0.987	80.3%	✗
子代 族群	24.10 27.76 79.77 0.981	81.3%	✗
	23.00 29.55 80.87 0.987	82.3%	✓

汰弱
留強

(5) 產生新族群，持續演化：

				粗ECH 收率
23.00	27.76	79.77	0.981	81.5%
23.00	29.55	80.87	0.987	82.3%

否

終止
條件

是

(6) 最佳操作條件

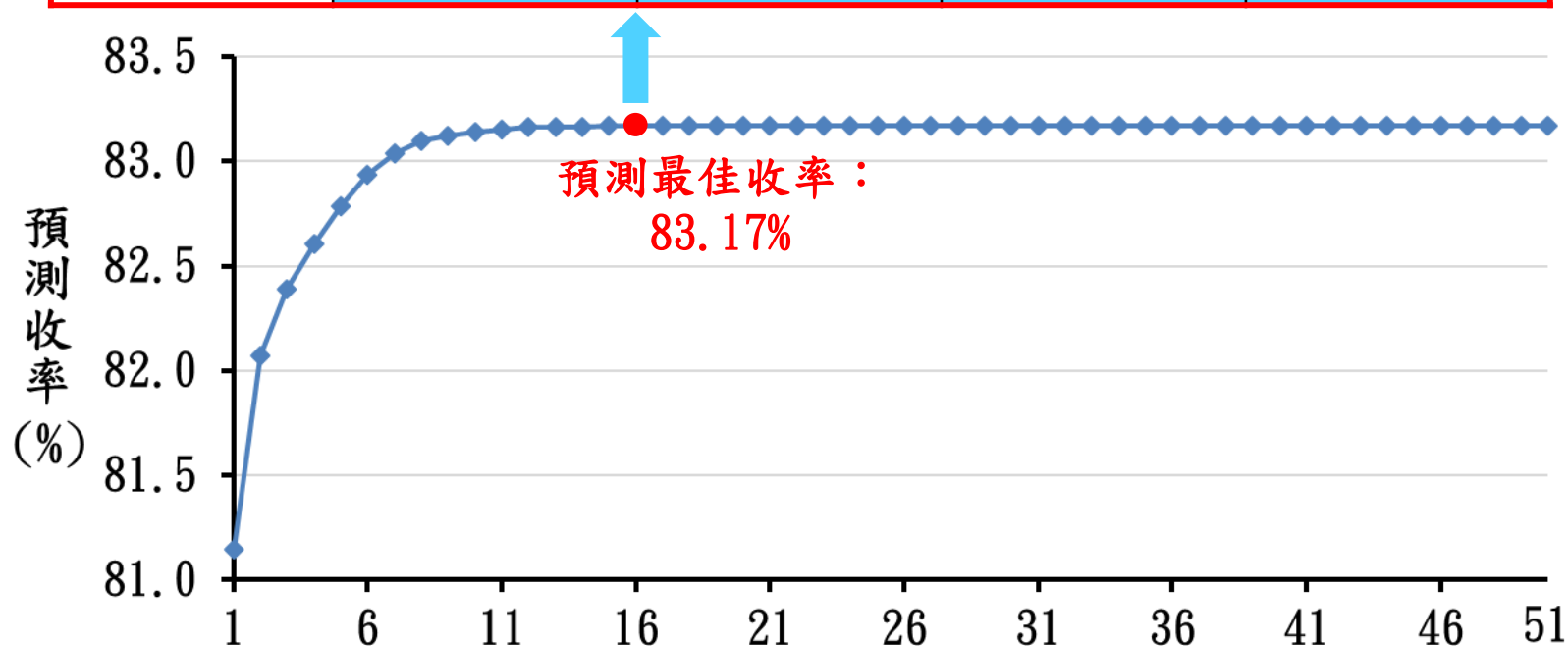
5.4 第四階段：AI操作建議(5/5)

➤ 基因演算法運作流程：

(6) 滿足終止條件，找到最佳解：

一直重複演化，直到預測的粗ECH收率不再增加，即可找到最佳的控制參數（最佳解）。

控制參數 名稱	次氯酸塔 塔底溫度	工業水/氯氣 流量比	空氣/氯氣 流量比	氯丙烯/氯氣 流量比
最佳控制值	23	30.36	97.23	1.018



5.5 第五階段：正式上線(1/3)

- 2019/11/28正式上線至今，粗ECH平均收率上升0.66%(82.52%→83.18%)，ECH成品可增加2.28噸/日(全載情況下)。

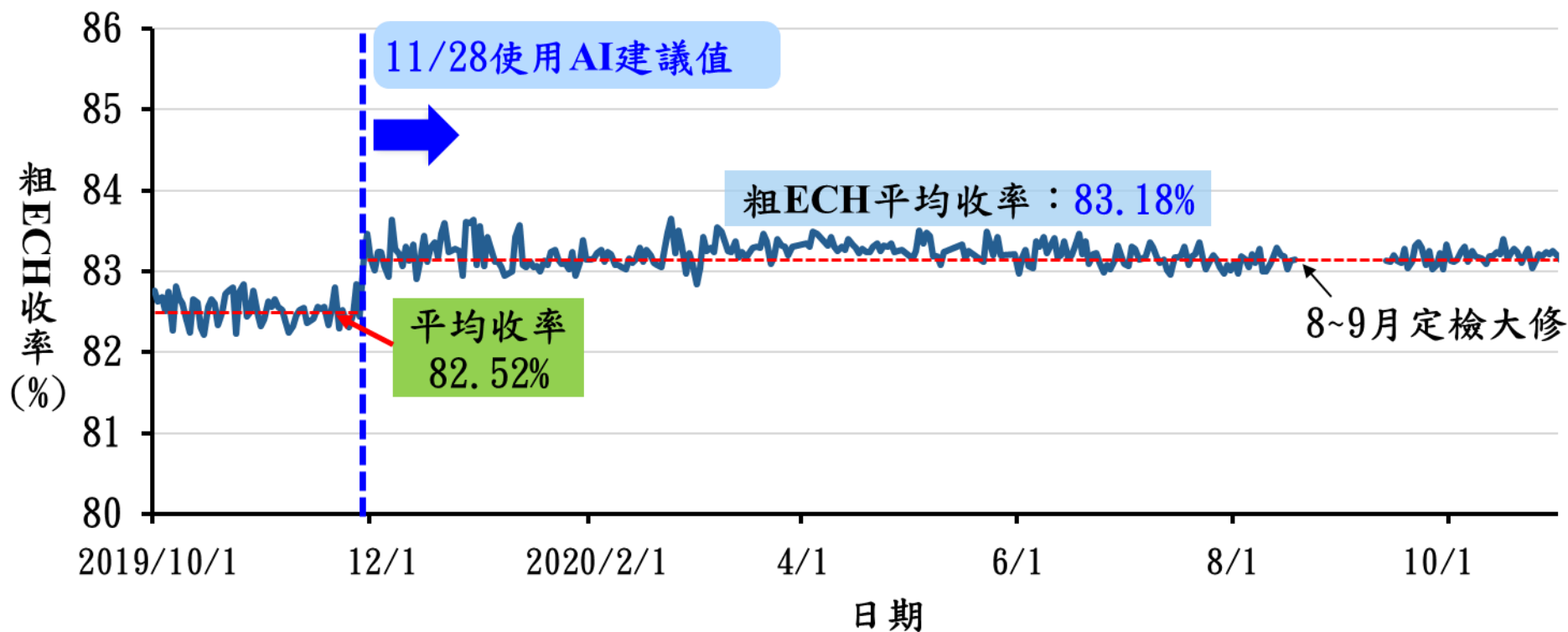
● 製程條件調整：

工業水流量(噸/時)： 330 → 340

空氣流量(NM³/時)： 1,031 → 1,089

次氯酸塔塔底溫度(℃)： 24.2 → 23.0

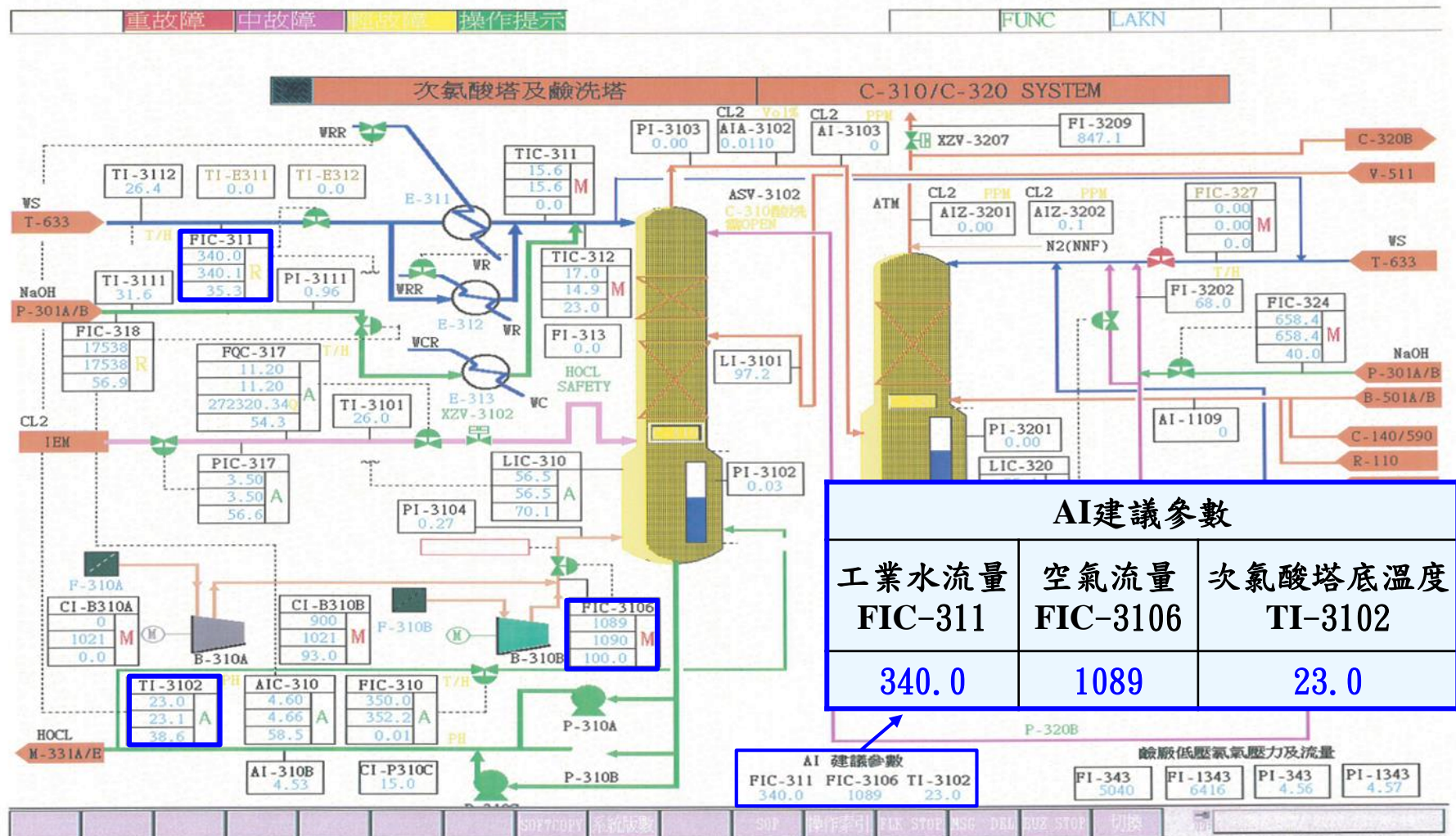
氯丙烯流量(噸/時)： 11.3 → 11.4



5.5 第五階段：正式上線(2/3)

➤ DCS整合：

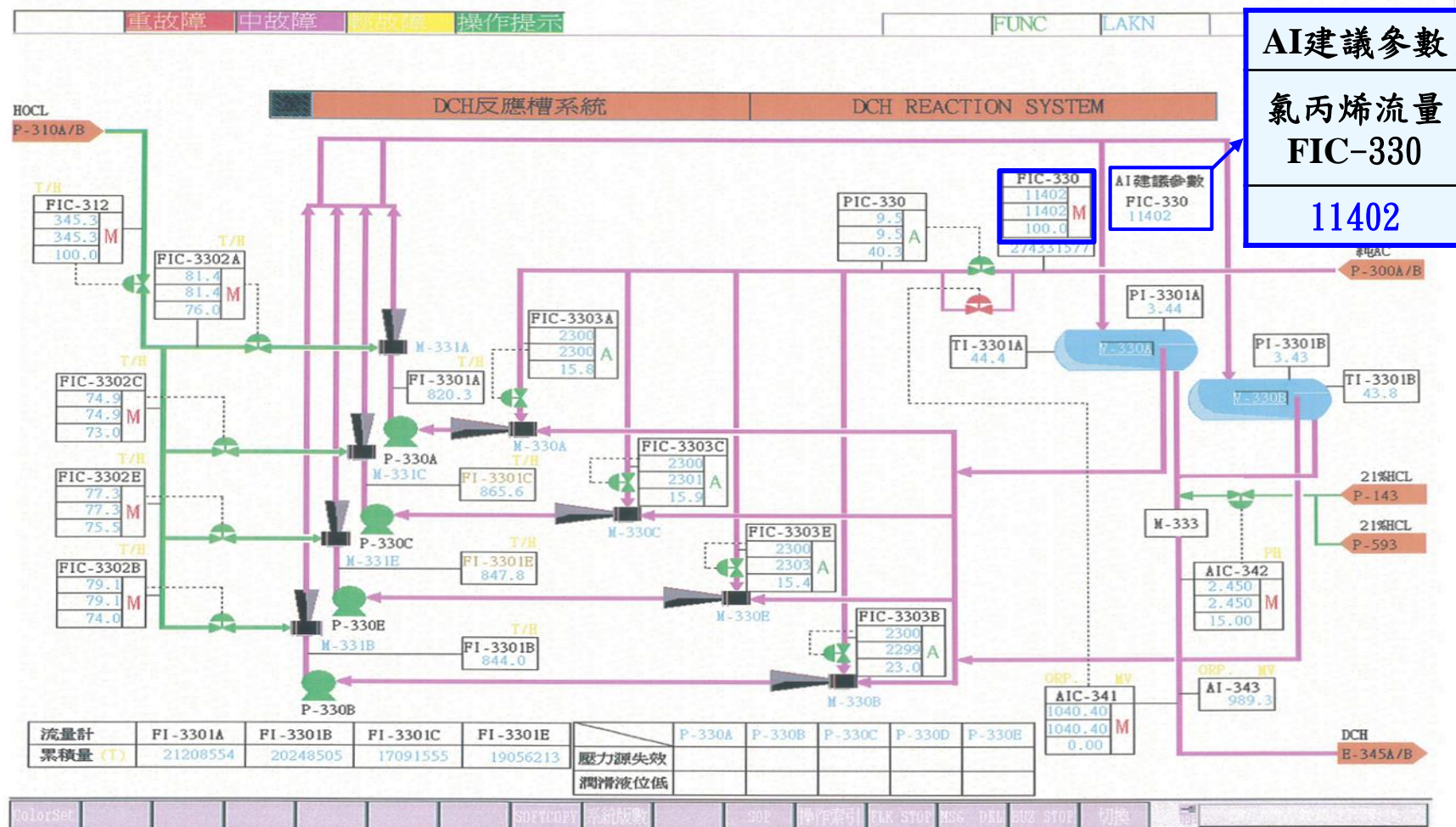
AI模型建議操作值和品質預測值已在DCS盤控上顯示



5.5 第五階段：正式上線(3/3)

➤ DCS整合：

AI模型建議操作值和品質預測值已在DCS盤控上顯示



6. 效益說明(1/2)

➤ 投資費用：4,110千元

(1) Aspen模擬費用：610千元

(2) AI模型部署：3,500千元

➤ 效益計算：

(1) ECH成品產量增加2.28噸/日：可增加收益34,322千元/年
(以過去5年ECH平均售價計算)

(2) 優化調整後，共增加成本25,753千元/年：

a. 次氯酸塔公用流體的用量調整：增加成本3,633千元/年

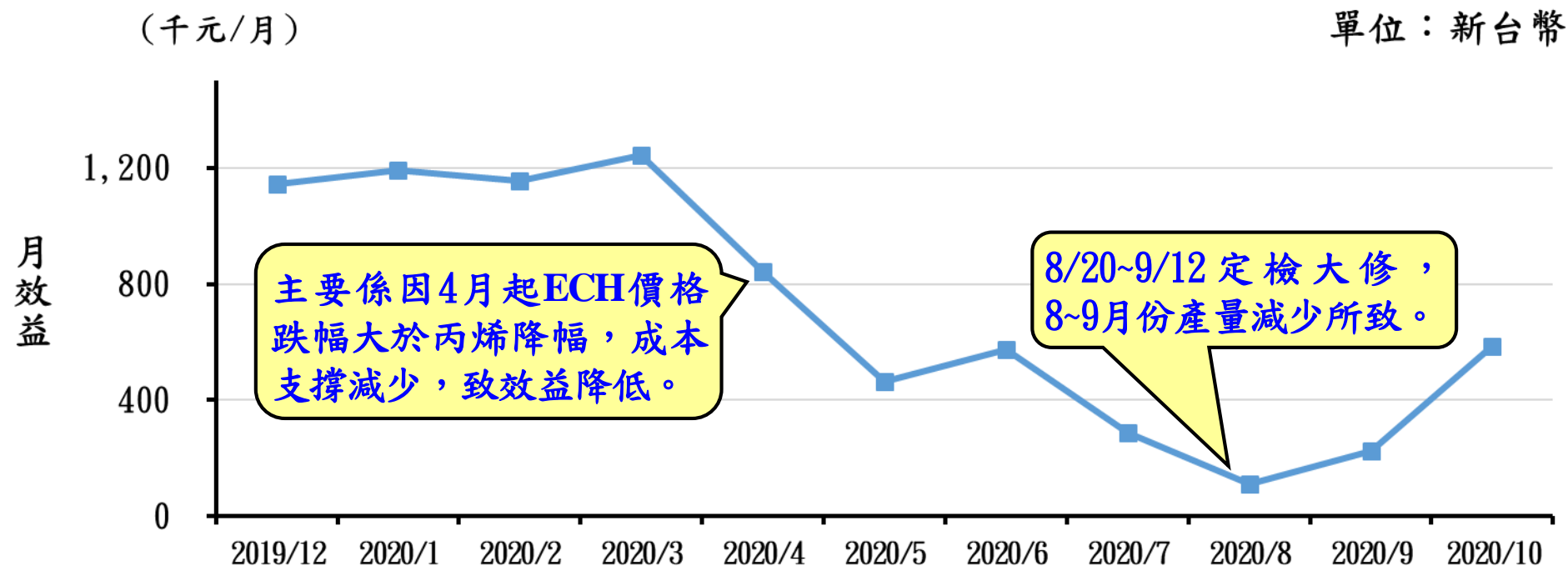
b. 氯丙烯流量調升：增加成本22,120千元/年

(3) 淨年效益： $34,322 - 25,753 = 8,569$ 千元/年

➤ 回收年限：0.48年

6. 效益說明(2/2)

實際效益波動圖



ECH月產量(噸)	9,013	8,945	8,385	7,453	7,220	8,001	8,558	8,872	5,496	5,265	8,904
ECH月增加產量(噸)	73.43	74.60	68.47	64.94	58.06	58.62	65.03	64.50	39.80	40.67	73.32
產量增加收益 (千元/月)	3,471	3,524	3,250	2,968	2,328	2,215	2,541	2,395	1,440	1,511	2,724
增加成本(千元/月)	2,329	2,333	2,095	1,724	1,485	1,749	1,967	2,109	1,328	1,287	2,154
淨月效益(千元/月)	1,143	1,191	1,155	1,244	842	466	574	286	112	224	570

7. 結論及後續推動事項(1/3)

一、運用 Aspen 模擬尋找收率提升的空間：

本案利用 Aspen 模擬 ECH 反應區，找出可再增加粗 ECH 收率的控制條件。

二、利用 SHAP 演算法探究 AI 模型：

大多數 AI 演算法為複雜的非線性模型，故無法提供各特徵的權重並了解其重要性。在導入 SHAP 演算法後，除了評估模型準確性外，亦可確認各特徵的權重，以研判是否與化工原理相符。

三、後續推動事項：

本案將持續精進，將 SHAP 演算法和基因演算法，平行展開至其它類似的 AI 專案，擴大效益。

7. 結論及後續推動事項(2/3)

- 本公司各事業部AI技術交流後，已陸續使用SHAP演算法開發AI專案(合計13案)，現已完成8案，年效益27,326千元，彙總如下表：

部門	單位	主題	應用技術	效益 (千元/年)
台麗朗 事業部	台麗朗廠	脫泡塔品質預測及智慧控制	XGBoost、LSTM、SHAP	1,966
	麥寮AE廠	丙烯酸反應器操作條件智慧監控	Lasso、隨機森林、XGBoost、SHAP	7,290
	麥寮SAP廠	SAP成品品質預測率提升	隨機森林、DNN、SHAP	1,843
	新港SAP廠	SAP改質智慧監控系統	DNN、Lasso、隨機森林、SHAP	973
聚烯 事業部	麥寮LLDPE廠	優化品別轉換過程	XGBoost、SHAP	1,300
化學品 事業部	麥寮MMA廠	甲基丙烯醯胺硫酸鹽(MAAS)轉化率 優化控制	隨機森林、XGBoost、SHAP	1,280
	麥寮ECH廠	皂化塔塔底廢水COD軟儀錶建立	XGBoost、SHAP	4,105
		環氧氯丙烷(ECH)反應區收率優化 (本案)	XGBoost、GA、SHAP	8,569
合 計				27,326

7. 結論及後續推動事項(3/3)

➤ 進行中5案彙總如下表：

部門	單位	主題	應用技術	預完日	預估效益 (千元/年)
塑膠 事業部	麥寮 VCM廠	廢水祛除塔A列智慧效能改善	迴歸分析、 SHAP	2021/12/E	2,800
台麗朗 事業部	林園 AE廠	ZW06煙道排放尾氣品質智能 預測	Lasso、 SHAP	2020/11/E	403 無形效益： 避免環保問題
	麥寮 AE廠	C402共沸蒸餾塔操作最佳化	增強式學習、 SHAP	2020/12/E	1,000
	麥寮 SAP廠	SAP乾燥機蒸汽控制最佳化	Lasso、KNN 隨機森林、 DNN、SHAP	2021/2/E	2,904
福欣	軋鋼 保養廠	鋼卷鑿痕缺陷預警及成因分析	XGBoost、 GA、SHAP	2020/12/E	874
合 計					7,981



報告完畢
恭請指導

附件：英文專有名詞資料表

英文名詞	英文全名	中文名稱	說明
ECH	<u>Ep</u> <u>ic</u> <u>h</u> <u>l</u> <u>o</u> <u>r</u> <u>o</u> <u>h</u> <u>y</u> <u>d</u> <u>r</u> <u>i</u> <u>n</u>	環氧氯丙烷	供應南亞EPOXY廠
MAE	<u>M</u> <u>e</u> <u>a</u> <u>n</u> <u>A</u> <u>b</u> <u>s</u> <u>o</u> <u>l</u> <u>u</u> <u>t</u> <u>e</u> <u>E</u> <u>r</u> <u>r</u> <u>o</u> <u>r</u>	平均絕對誤差	多筆實際值與預測值誤差的絕對值平均
XGBoost	<u>E</u> <u>X</u> <u>t</u> <u>r</u> <u>e</u> m <u>e</u> <u>G</u> <u>r</u> <u>a</u> <u>d</u> <u>i</u> <u>e</u> n <u>t</u> <u>B</u> <u>o</u> <u>o</u> <u>s</u> <u>t</u> <u>i</u> <u>n</u> g	極限梯度提升	將多個弱分類器組合後，形成一個強大的分類器的機器學習演算法。
SHAP	<u>S</u> <u>H</u> <u>a</u> <u>p</u> <u>l</u> <u>e</u> y <u>A</u> <u>d</u> <u>d</u> <u>i</u> <u>t</u> <u>i</u> <u>v</u> <u>e</u> <u>e</u> <u>x</u> <u>P</u> <u>l</u> <u>a</u> <u>n</u> <u>a</u> <u>t</u> <u>i</u> <u>o</u> <u>n</u> s	-	一種可以解釋模型的輸入跟輸出之間的對應關係的演算法
GA	<u>G</u> <u>e</u> <u>n</u> <u>e</u> <u>t</u> <u>i</u> <u>c</u> <u>A</u> <u>l</u> <u>g</u> <u>o</u> <u>r</u> <u>i</u> <u>t</u> <u>h</u> <u>m</u>	基因演算法	一種借鑑生物演化機制中的最佳化搜尋演算法。
RFE	<u>R</u> <u>e</u> <u>c</u> <u>u</u> <u>r</u> <u>s</u> <u>i</u> <u>v</u> <u>e</u> <u>F</u> <u>e</u> <u>a</u> <u>t</u> <u>u</u> <u>r</u> <u>e</u> <u>E</u> <u>l</u> <u>i</u> <u>m</u> <u>i</u> <u>n</u> <u>a</u> <u>t</u> <u>i</u> <u>o</u> <u>n</u>	遞迴特徵篩選法	一種藉由逐步剔除訓練數據的特徵，找出可降低模型誤差的特徵篩選方法。
KNN	<u>K</u> <u>N</u> <u>e</u> <u>a</u> <u>r</u> <u>e</u> <u>s</u> <u>t</u> <u>N</u> <u>e</u> <u>i</u> <u>g</u> <u>h</u> <u>b</u> <u>o</u> <u>r</u>	最近鄰居法	計算與已知類別案例之相似度，來評估未知類別案例可能的分類或平均值
SVR	<u>S</u> <u>u</u> <u>p</u> <u>p</u> <u>o</u> <u>r</u> <u>t</u> <u>V</u> <u>e</u> <u>c</u> <u>t</u> <u>o</u> <u>r</u> <u>R</u> <u>e</u> <u>g</u> <u>r</u> <u>e</u> <u>s</u> <u>s</u> <u>i</u> <u>o</u> <u>n</u>	支持向量迴歸	一種藉由找出預測值與實際值偏離最小的迴歸方法，在SVR中，偏差不大的數據會被視為正確預測。
DNN	<u>D</u> <u>e</u> <u>e</u> <u>p</u> <u>N</u> <u>e</u> <u>u</u> <u>r</u> <u>a</u> <u>l</u> <u>N</u> <u>e</u> <u>t</u> <u>w</u> <u>o</u> <u>r</u> <u>k</u> <u>s</u>	深度神經網路	以神經網路為基礎架構，藉由加深隱藏層來取代傳統特徵工程的演算法。
RIDGE/LASSO	RIDGE/LASSO	嶺迴歸/套索算法	一種可以避免模型過擬合的線性迴歸方法。