

### 三、研究計畫內容：

#### 計畫目錄

三、研究計畫內容：	1
計畫目錄	1
計畫之背景與文獻回顧	2
驗室團隊前期相關成果	4
研究目的	4
研究原創性、重要性、預期影響性及國內外有關本計畫之研究情況	5
年度研究進度	5
參考文獻	6
分年列述研究方法	8
第一年：開發利用少量數據重建動態工廠軟儀表，並引入局部性增益增強物理可解釋性能力	8
第二年：開發工廠軟儀表結合穩態理論代理模型，引入全局穩態增益方向性，強化軟儀表物理可解釋性	15
第三年：開發工廠軟儀表結合穩態與動態理論代理模型，引入全局動態與穩態增益方向性，可進行全域動態外插能力	18
初步研究成果	24
1. 建構基於多取樣頻率之少樣數據重建動態資訊軟儀表	24
2. 製程孿生模型	24
預期完成之工作項目	25
對於參與之工作人員，預期可獲之訓練	25
預期完成之研究成果（如實務應用績效、期刊論文、研討會論文、專書、技術報告、專利或技術移轉等質與量之預期成果）	25

(一) 研究計畫之背景。請詳述本研究計畫所要探討或解決的問題、研究原創性、重要性、預期影響性及國內外有關本計畫之研究情況、重要參考文獻之評述等。如為連續性計畫應說明上年度研究進度。

### 計畫之背景與文獻回顧

在當今快速發展的人工智能技術浪潮中，模型預測控制（MPC）結合神經網絡的核心架構，即 AI-MPC，已經取得廣泛的成功應用。AI-MPC 革新了過往的控制策略，使製程變化控制更有效，並提高了製程效益。Kittisupakorn, et al. [1]的研究提供了早期實例，他們將神經網路與 MPC 結合，提出了一種基於神經網路的模型預測控制（NNMPC）來控制銅酸洗流程，其中神經網路的作用是尋找最優控制動作。

進一步地，Shin, et al. [2]將帶有 MPC 的神經網路應用於丙烷蒸發器的控制中。相比傳統的線性 MPC，基於神經網路的非線性 MPC 表現出更優越的控制性能。通過對新控制器和傳統 PI 控制器進行的性能比較，研究發現新控制器的穩定時間提升了 60%。此類型的非線性控制策略對於程序控制尤為有益，如 Ławryńczuk [3]的研究在比較了攪拌反應器和蒸餾塔中的線性與非線性 MPC 時得出的結論。另一方面，Piche, et al. [4]對於聚乙烯(PE)反應器設定點的變化控制引入了基於經驗數據的神經網路。這進一步印證了 AI-MPC 在工業應用中帶來的大幅效率提升，其中在產品轉換過程中實現了 30%的加速以及整體受控變量的降低。

儘管成果卓著，但 AI-MPC 的穩定性和有效性極度依賴於模型的準確表示[5,6]，以及在多變量和複雜約束條件下的決策策略。一個穩定的控制系統，必須具備可行性、穩定性和穩健性這三大特性。這些特性受限於多個關鍵因素，包括如何克服硬性控制變數限制、解決現場量測時延問題以及處理程序中的各種擾動。量測時延是一個特別挑戰性的問題，它可能由於無法及時更新數據讓模型重新預測，導致控制反應過慢，無法有效應對程序擾動。

這些研究結果突顯了 AI-MPC 在工業自動化控制領域帶來的創新和提升。然而，仍需進一步研究來解決當前的挑戰，尤其是在模型準確性和控制策略的適應性方面。未來的工作將集中於如何改進 AI-MPC 架構，使其更加穩健，並能快速適應不斷變化的製程條件。

通常 AI-MPC 模型是否具備水平預測能力被視為控制效能的關鍵[5, 7]，但實際上 AI 模型是否具有物理可解釋性才是製程控制能否在約束條件下提供合理解的關鍵因素[5]。作為 MPC 導航用 AI 目前技術問題在於，無法僅透過真實工廠數據來驅動的模型解決：1)由於數據取樣頻率限制而導致的動態資訊丟失，2)製程動態的時間常數精準識別，3)製程增益方向變化的處理，以及 4)缺少歷史數據情況下的外插預測分析。其關鍵在於真實製程數據存在著操作局部性，低頻取樣離散性。並且數據受限於現場安裝表點位置合理性以及感測器品質，要能從數據中提取資訊並建構合理且健全的孿生模型，困難度非常高。經常有許多物理資訊會因為製程雜訊所掩蓋。

AI-MPC 的模型常採用軟儀錶做為 MPC 模型，通常採用簡單的 ANN 網路進行處理[2, 7]或是全連結層的 RFB(Radial Basis Function)網路[8]，並以單一時間點建模 $y_{t+1} = f(y_t, x_t, y_{t-1}, x_{t-1}, \dots)$ 進行預測，

以及規劃操作策略。主要原因是大部分工廠仍以實驗室測量為主，通常測量頻率為半天到一週一次測量不等。在取樣頻率低情況下，要建立長時間水平模型是有困難的。近年，序列對序列式模型(Sequence-to Sequence, StS)結合遞迴神經網路(recurrent Neural Network, RNN)被廣泛應用於時間序列數據分析及自然語言處理，同時也被使用在具時間序列關係的化工製程中並有準確的預測效果[9]。雖然 LSTM 在獲取長期水平預測性能方面非常有幫助，但它不能在不同的時間步長上預測不同的變量。為了克服這個問題，Bahdanau, et al. [10]中提出了一種基於注意力的編碼器-解碼器網絡。基於 LSTM 單元，可以構建編碼器-解碼器網絡來解決序列到序列問題。有了 encoder 部分，可以將輸入序列轉化為固定長度的向量，然後 decoder 將生成的固定向量轉化為輸出序列。然而，該模型的性能會隨著輸入序列長度的增加而迅速下降。這是工業數據序列中的一個常見問題，因為質量變量通常是根據非常長的滯後輸入序列來預測的。Du, et al. [11] 通過基於注意力的編碼器-解碼器框架進行多變量時間序列應用於多種預測，包括空氣 PM2.5、交通、以及能量消耗等，其驗證了深度學習在多變數水平預測的能力，但是該模型只能建構基於單一取樣頻率的水平預測。也就是預測變數取樣為一個小時，模型僅具有預測間隔一個小時的水平預測能力。這對於需要高頻率預測的系統來說並無法及時提供資訊。

Yuan, et al. [12]文獻使用注意力層配合 StS 開發水平預測的模型。但文章指出，雖模型可建構高頻水平預測能力，但是對於低頻率採樣數據來建構高頻率預測模型仍有待克服。其關鍵在於建構模型的數據一年有效數據僅能約 54~600 筆，對於現有 AI 模型屬於數據量稀少，且取樣頻率過低，通常會通過補值(Imputation)將遺失數據透過內插或者是統計方法進行補值在進行模型訓練。傳統常見方法包括下採樣(Down-sampling)及上採樣(Up-sampling)。下採樣，其目的是找出最低採樣頻率，以最低採樣率建立人工神經網路模型的軟儀錶[13]；相反，上採樣試圖通過補值方法估計相應的低採樣頻率變量，使其具有最高採樣頻率的測量值。上採樣的補值方法有使用均值插補[14, 15]、最後觀察值推估(Last Observation Carried Forward, LOCF)[16]、線性內插法[17, 18]或是使用模型對於狀態進行估計。使用模型進行估計的方法如下：概率主成分分析模型、卡爾曼濾波器、神經網路及自動編碼器。概率主成分分析模型(Probabilistic Principal Component Analysis)適用於處理缺失值，可以作為上採樣方法中的估計量[19]。也有針對不同採樣速率，分別建立兩個卡爾曼濾波器進行估計[20]。Petrozziello, et al. [21]透過無缺失數據集訓練類神經網路並對缺值數據集進行補值並有良好補值結果，此類神經網路可使用小量完整數據集訓練，適用於稀少的品質變量數據的插補。然而，在好的補值方法，終究是利用統計方法產生出的假數據。使用上採樣會丟失一些有用的過程信息，而下採樣會由於可用質量數據稀疏，導致不精確的補值致使模型性能不佳[22]，如嚴重延遲預測問題失去反映製程的能力。因此，補值方式建構軟儀錶並非是理想的方法。

AI 模型物理解釋性的具備是 AI 模型能否套用於 MPC 控制的關鍵。且因為實廠數據的操作侷限性，使得模型難以外插預測歷史操作範圍外的控制變量。但如具備物理可解釋性的 AI 模型，可透過模型對於製程物理性的理解，推估歷史外操作，即便具有誤差，但因基於物理解釋能力做出的決策必能符合製程需求，僅須透過製程誤差修正即可做到穩定的外插控制。對於化工製程，會透過 Aspen Plus 進行理論建模，來避免自行書寫上千萬條的微分方程式。透過 Aspen Plus 擬合實場狀態，再透過 AI 代理模型進行線上控制。雖這樣的模型具備強的物理解釋性，但透過 Aspen Plus 產生的數據為穩態數據不具備動態預測能力，且理論模型需要絕對的操作條件來預測製程控制變量，但實場的感測器並非對正確且經常有偏差，在這樣的情況下，理論模型很難具備控制能力。

最近的研究表明，物理模型與神經網絡模型的結合在各個領域是一種強大的方法。物理知神經網絡（PINN）在各種應用中解決局部微分方程（PDE）顯示出了很大的潛力。他們通過結合自我訓練機制和動態標準化權重[23]來解決現有 PINN 的低精度和聚合問題的問題。此外，還提出了一種稱為 NAS-PINN 的神經架構搜索引導方法，以自動搜索用於解決 PDE 的最佳神經架構[24]此外，有限基礎物理信息神經網絡（FBPINN）通過將它們與重疊的域分解方法結合來改善 PINN 的性能，尤其是對於高頻和多規模解決方案的問題 [25]。PINN 的這些進步在使用自動微分方程的解決方法中提供了實用性，以參數化 PDE 和殘留損失，使其適用於化學工程的研究。

但對於化工製程控制來說，引入微分方程式來實現模型的物理解釋性，成本過於龐大。會使用數據驅動模型取代理論模型來實現控制，主因便是因為理論建模一直存在高成本開發以及費時等問題。但如果為了讓數據驅動模型具備物理意義，便將描述製程中多個單元操作設備的上千條微分方程式引入，會導致 MPC 模型開發曠日廢時。但如果能在建立數據驅動的 AI 模型時，透過將 Aspen Plus 理論模型引入物理解釋性，便可讓 AI 模型同時具備實場預測能力並同具備高物理解釋性的外插能力。這也是目前尚待開發的技術。

### 驗室團隊前期相關成果

為能夠讓化工製程虛實整合控制技術計畫順利推動，本團隊於前期完成初步成果，已具備開發本計畫之能力。

1. 本團隊在前期以開發獨特 Sequence-to-Sequence 模型。該模型具有水平預測能力[9]。

Chou, Cheng-Hung, et al. "Physically consistent soft-sensor development using sequence-to-sequence neural networks." *IEEE Transactions on Industrial Informatics* 16.4 (2019): 2829-2838.

2. 本研究已投入人力與資源開發長時間水平預測數據驅動數位孿生模型，已掌握開發技術並有極大信心可開發成功。目前初步成果已發表有國際會議論文[26]。

Kang, J. L., Mirzaei, S., Lee, Y. C., Chuang, Y. C., Frias, M., Chou, C. H., ... & Jang, S. S. (2021). Digital Twin Model Development for Chemical Plants Using Multiple Time-Steps Prediction Data-Driven Model and Rolling Training. In *Computer Aided Chemical Engineering* (Vol. 50, pp. 567-572). Elsevier.

3. 今年發現數據驅動數位孿生模型於高點數期刊[27]

Kang, Jia-Lin, Somayeh Mirzaei, and Zi Hang Yang. "Sequence-to-sequence digital twin model in chemical plants with internal rolling training algorithm." *Applied Soft Computing* 146 (2023): 110608.

### 研究目的

由上述可知 AI-MPC 的表現效益 AI 模型貢獻比非常高。本研究希望開發目前 AI-MPC 技術待解決的議題：1)由於數據取樣頻率限制而導致的動態資訊丟失，2)製程動態的時間常數精準識別，3)製程增益方向變化的處理，以及 4)缺少歷史數據情況下的外插預測分析。團隊認為預測模型需要具有高頻水平預測能力，並保有明確的製程動態行為與明確的物理意義，來提升 AI-MPC 的控制穩定性，是主要 AI-MPC 領域需要發展的重點項目。

近年 Self-supervised learning 概念[28]提出重要觀點，大部分任務中，標籤數據(y)(等同化工的有取樣時間的數據)都非常稀少，但是無標籤的數據(非取樣時間的輸入數據)卻是無窮無盡。但無標籤的數



據內涵蓋大量資訊量，應充分使用。而此發展重點則必須藉由從低取樣頻率以及大量無標籤的數據中建構才能真正解決因為數據稀疏造成模型會有延遲以及動態不足的問題。(解決第一點與第二點)

現階段 Aspen 理論模型與數據驅動模型仍處於平行狀態，團隊希望透過手動引入或者 AI 代理模型技術來複製 Aspen 理論穩態與動態模型。藉由理論的 AI 代理模型取代 PINNS 中複雜的微分方程式，來結合實廠數據來訓練神經網路模型，來改善現有技術的 AI 模型限制。(解決第三點與第四點)

為達本研究之目的，本計畫需分年執行逐步完成：

第一年：建構基於多取樣頻率品質變數建構可高頻率水平預測之數據驅動軟儀錶，並透過 Aspen 分析製程增益局部性導入軟儀錶增強歷史操作範圍物理可解釋性能力。

第二年：開發工廠軟儀錶結合穩態理論代理模型，引入全局穩態增益方向性，強化軟儀錶物理可解釋性。

第三年：開發工廠軟儀錶結合穩態與動態理論代理模型，引入全局動態與穩態增益方向性，可進行全域動態外插能力。

### 研究原創性、重要性、預期影響性及國內外有關本計畫之研究情況

#### AI-MPC 核心模型開發提升 AI-MPC 控制穩定度

AI-MPC 的表現效益 AI 模型貢獻比非常高。因此，本團隊認為，預測模型需要具有高頻水平預測能力，並保有明確的製程動態行為與明確的物理意義，來提升 AI-MPC 的控制穩定性，是主要 AI-MPC 領域需要發展的重點項目。本團隊提出希望透 Self-supervised learning 概念[28]利用稀少標籤數據(y)(有取樣時間的數據)配合大量無標籤的數據(非取樣時間的輸入數據)建構具有高頻水平預測能力(即一次可預未來 3~5 小時，且時間間隔為 10~15 分鐘一筆預測)。此發展重點則必須藉由從低取樣頻率以及大量無標籤的數據中建構才能真正解決因為數據稀疏造成模型會有延遲以及動態不足的問題。目前化工文獻尚無提出類似概念。

並透過 Aspen 引入增益方向性，使得模型具備物理可解釋性，期望模型具備外插能力。本技術如開發成功，可望有效將理論模型與實場數據緊密結合，透過理論模型的資訊彌補實廠數據的資訊的不足。

### 年度研究進度

為達本研究之目的，本計畫需分年執行逐步完成：

第一年：建構基於多取樣頻率品質變數建構可高頻率水平預測之數據驅動軟儀錶，並透過 Aspen 分析製程增益局部性導入軟儀錶增強歷史操作範圍物理可解釋性能力。

第二年：開發工廠軟儀錶結合穩態理論代理模型，引入全局穩態增益方向性，強化軟儀錶物理可解釋性。

第三年：開發工廠軟儀錶結合穩態與動態理論代理模型，引入全局動態與穩態增益方向性，可進行全域動態外插能力。

## 參考文獻

- [1] P. Kittisupakorn, P. Thitiyasook, M. A. Hussain, and W. Daosud, "Neural network based model predictive control for a steel pickling process," *Journal of Process Control*, vol. 19, no. 4, pp. 579-590, 2009.
- [2] Y. Shin, R. Smith, and S. Hwang, "Development of model predictive control system using an artificial neural network: A case study with a distillation column," *Journal of Cleaner Production*, vol. 277, p. 124124, 2020.
- [3] M. Ławryńczuk, "A family of model predictive control algorithms with artificial neural networks," *International Journal of Applied Mathematics and Computer Science*, vol. 17, no. 2, pp. 217-232, 2007.
- [4] S. Piche, B. Sayyar-Rodsari, D. Johnson, and M. Gerules, "Nonlinear model predictive control using neural networks," *IEEE Control Systems Magazine*, vol. 20, no. 3, pp. 53-62, 2000.
- [5] M. Schwenzer, M. Ay, T. Bergs, and D. Abel, "Review on model predictive control: An engineering perspective," *The International Journal of Advanced Manufacturing Technology*, vol. 117, no. 5, pp. 1327-1349, 2021.
- [6] J. B. Rawlings, "Tutorial overview of model predictive control," *IEEE control systems magazine*, vol. 20, no. 3, pp. 38-52, 2000.
- [7] D. Lee, R. Ooka, Y. Matsuda, S. Ikeda, and W. Choi, "Experimental analysis of artificial intelligence-based model predictive control for thermal energy storage under different cooling load conditions," *Sustainable Cities and Society*, vol. 79, p. 103700, 2022.
- [8] S. Li, P. Jiang, and K. Han, "RBF Neural network based model predictive control algorithm and its application to a CSTR process," in *2019 Chinese control conference (CCC)*, 2019: IEEE, pp. 2948-2952.
- [9] C.-H. Chou *et al.*, "Physically consistent soft-sensor development using sequence-to-sequence neural networks," vol. 16, no. 4, pp. 2829-2838, 2019.
- [10] D. Bahdanau, K. Cho, and Y. Bengio, "Neural machine translation by jointly learning to align and translate," *arXiv preprint arXiv:1409.0473*, 2014.
- [11] S. Du, T. Li, Y. Yang, and S.-J. Horng, "Multivariate time series forecasting via attention-based encoder-decoder framework," *Neurocomputing*, vol. 388, pp. 269-279, 2020.
- [12] X. Yuan, L. Li, Y. A. Shardt, Y. Wang, and C. Yang, "Deep learning with spatiotemporal attention-based LSTM for industrial soft sensor model development," *IEEE Transactions on Industrial Electronics*, vol. 68, no. 5, pp. 4404-4414, 2020.
- [13] M. T. Tham, G. A. Montague, J. Glassey, M. J. J. M. Willis, and Control, "Practical inferential estimation using artificial neural networks," vol. 35, no. 1, pp. 5-9, 2002.
- [14] J. Osborne and A. J. B. p. i. q. m. Overbay, "Best practices in data cleaning," vol. 1, no. 1, pp. 205-213, 2008.
- [15] A. C. J. J. o. M. Acock and family, "Working with missing values," vol. 67, no. 4, pp. 1012-1028, 2005.
- [16] H. Zhou, K.-M. Yu, M.-G. Lee, and C.-C. Han, "The application of last observation carried forward method for missing data estimation in the context of industrial wireless sensor networks," in *2018 IEEE*

*Asia-Pacific Conference on Antennas and Propagation (APCAP)*, 2018: IEEE, pp. 1-2.

- [17] M. Kano, Y. J. C. Nakagawa, and C. Engineering, "Data-based process monitoring, process control, and quality improvement: Recent developments and applications in steel industry," vol. 32, no. 1-2, pp. 12-24, 2008.
- [18] D. Wang, J. Liu, and R. J. I. T. o. I. I. Srinivasan, "Data-driven soft sensor approach for quality prediction in a refining process," vol. 6, no. 1, pp. 11-17, 2009.
- [19] M. E. Tipping and C. M. J. J. o. t. R. S. S. S. B. Bishop, "Probabilistic principal component analysis," vol. 61, no. 3, pp. 611-622, 1999.
- [20] A. Fatehi and B. J. J. o. P. C. Huang, "Kalman filtering approach to multi-rate information fusion in the presence of irregular sampling rate and variable measurement delay," vol. 53, pp. 15-25, 2017.
- [21] A. Petrozziello, I. Jordanov, and C. Sommeregger, "Distributed neural networks for missing big data imputation," in *2018 International Joint Conference on Neural Networks (IJCNN)*, 2018: IEEE, pp. 1-8.
- [22] Y. Jiang, S. Yin, J. Dong, and O. J. I. S. J. Kaynak, "A review on soft sensors for monitoring, control and optimization of industrial processes," 2020.
- [23] J. Yan, X. Chen, Z. Wang, E. Zhoui, and J. Liu, "ST-PINN: A Self-Training Physics-Informed Neural Network for Partial Differential Equations," *arXiv preprint arXiv:2306.09389*, 2023.
- [24] S. Deguchi and M. Asai, "Dynamic & norm-based weights to normalize imbalance in back-propagated gradients of physics-informed neural networks," *Journal of Physics Communications*, vol. 7, no. 7, p. 075005, 2023.
- [25] Y. Wang and L. Zhong, "NAS-PINN: Neural architecture search-guided physics-informed neural network for solving PDEs," *Journal of Computational Physics*, vol. 496, p. 112603, 2024.
- [26] J.-L. Kang *et al.*, "Digital Twin Model Development for Chemical Plants Using Multiple Time-Steps Prediction Data-Driven Model and Rolling Training," in *Computer Aided Chemical Engineering*, vol. 50: Elsevier, 2021, pp. 567-572.
- [27] J.-L. Kang, S. Mirzaei, and Z. H. Yang, "Sequence-to-sequence digital twin model in chemical plants with internal rolling training algorithm," *Applied Soft Computing*, vol. 146, p. 110608, 2023.
- [28] Y. Wang, C. M. Albrecht, N. A. A. Braham, L. Mou, and X. X. Zhu, "Self-supervised learning in remote sensing: A review," *arXiv preprint arXiv:2206.13188*, 2022.
- [29] M. T. Ribeiro, S. Singh, and C. Guestrin, "" Why should i trust you?" Explaining the predictions of any classifier," in *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining*, 2016, pp. 1135-1144.
- [30] Y.-D. Hsiao, J.-L. Kang, and D. S.-H. J. P. Wong, "Development of Robust and Physically Interpretable Soft Sensor for Industrial Distillation Column Using Transfer Learning with Small Datasets," vol. 9, no. 4, p. 667, 2021.

(二) 研究方法、進行步驟及執行進度。請分年列述：1.本計畫採用之研究方法與原因及其創新性。2. 預計可能遭遇之困難及解決途徑。3.重要儀器之配合使用情形。4.如為須赴國外或大陸地區研究，請詳述其必要性以及預期效益等。

## 分年列述研究方法

為達本研究之目的，本計畫需分年執行逐步完成，並分年承述研究方法。

### 第一年：開發利用少量數據重建動態工廠軟儀表，並引入局部性增益增強物理可解釋性能力

計畫以前期開發之序列對序列式模型作為基礎，進行開發，該模型本身已具有短時間水平預測能力，故應能在此基礎延長水平預測能力，希望能開發出長時間預測(超過 3 小時以上，每 10 分鐘間格預測)。此序列對序列模型(Sequence to Sequence, StS)與一般文獻 StS 輸入資訊有所差異。圖 2 為本專案序列對序列式神經網路(StS)結構示意圖[9]，由編碼器(Encoder)與解碼器(Decoder)組成，使用當前觀察到的數據  $o_1 \sim o_m$  和最後一步的隱藏狀態，輸出當前隱藏狀態  $h_m$ ，其包含了所有過去歷史觀察到的信息並傳遞至解碼器，再透過給定當前輸入  $x_1 \sim x_t$  至解碼器進行預測  $y_1 \sim y_t$ 。於製程取樣頻率過低的數據，因為數據間格過長，一般模型無法建立水平預測模型，或需藉由滾動方式進行水平預測，且此類型模型經常發生嚴重延遲預測以及反應不靈敏的問題。StS 亦面臨相同問題。為解決 StS 可適用於任何取樣頻率數據，並能重構製程動態，本研究提出數據遮罩損失函數(Masking loss)，透過高取樣頻率的  $x$  輸入數據配合低頻取樣數據  $y$  在不同時間點的預測，可增加 1. 數據訓練量以及 2. 製程動態資訊重構。並且為了模型具有合理控制方向性(gain-informed loss)，我們引入增益一致性損失函數來確保模型訓練具有合理物理意義。訓練損失函數定義如下：

$$\mathcal{L}_{\text{total}} = W_{\text{Mask}} \times \mathcal{L}_{\text{Mask}} + W_{\text{Gain}} \times \mathcal{L}_{\text{Gain}} \quad (1)$$

而使用兩者損失函數模型合稱為(M-GI-StS)

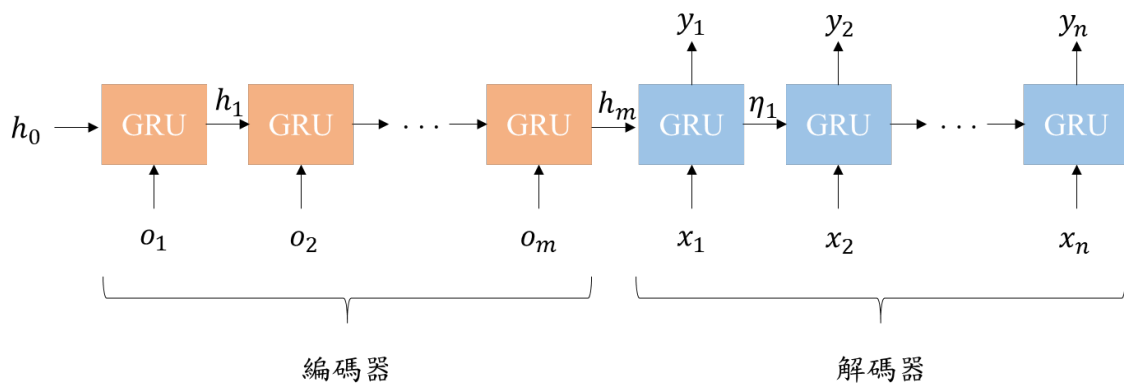


圖 1 序列對序列式神經網路

### 遮罩損失函數

遮罩損失函數藉由訓練過程中將，不屬於實際取樣數據  $y$ (資料庫自動補值)進行遮蔽不計算。在模型的損失函數中會加入遮罩變數來遮蔽非採樣點的數據如下列公式所示：



$$S_{Mask_{i,t}} = \begin{cases} 1, & t = \text{sampling point} \\ 0, & t = \text{unsampling point} \end{cases} \quad (2)$$

$$\mathcal{L}_{Mask} = \frac{\sum_{t=0}^T \sum_{i=0}^n |\tilde{y}_{i,t} - y_{i,t}| \times S_{Mask_{i,t}}}{\sum_{i=0}^n N_{Mask_i}} \quad (3)$$

$S_{Mask_{i,t}}$  為遮罩變數，品質變量為  $i \in 0, 1, 2, \dots, n$ ，預測時間步長為  $t \in 0, 1, 2, \dots, T$ 。在採樣點時，遮罩變數為 1；反之，遮罩變數為 0。透過在平均絕對誤差中乘上遮罩變數， $\tilde{y}_{i,t}$  為模型預測值， $y_{i,t}$  為實際值， $N_{Mask_i}$  為品質變量有採樣點的總數。遮罩遮蔽概念示意圖如圖 2，在非採樣點的 loss 會因為乘上遮罩變數而變為 0，進而不會影響模型訓練。但卻因為遮罩方式，讓取樣點  $y$  可以放在不同位置的水平預測時間點上進行訓練而使得產生大量訓練數據可使用，並讓模型學習在不同長度的輸入  $x_{t \sim t-W}$  資訊下，與取樣  $y_{i,t}$  之間的關係。並且可以同時訓練多種不同取樣頻率之變量，使模型可同時將多種變量進行高頻預測。

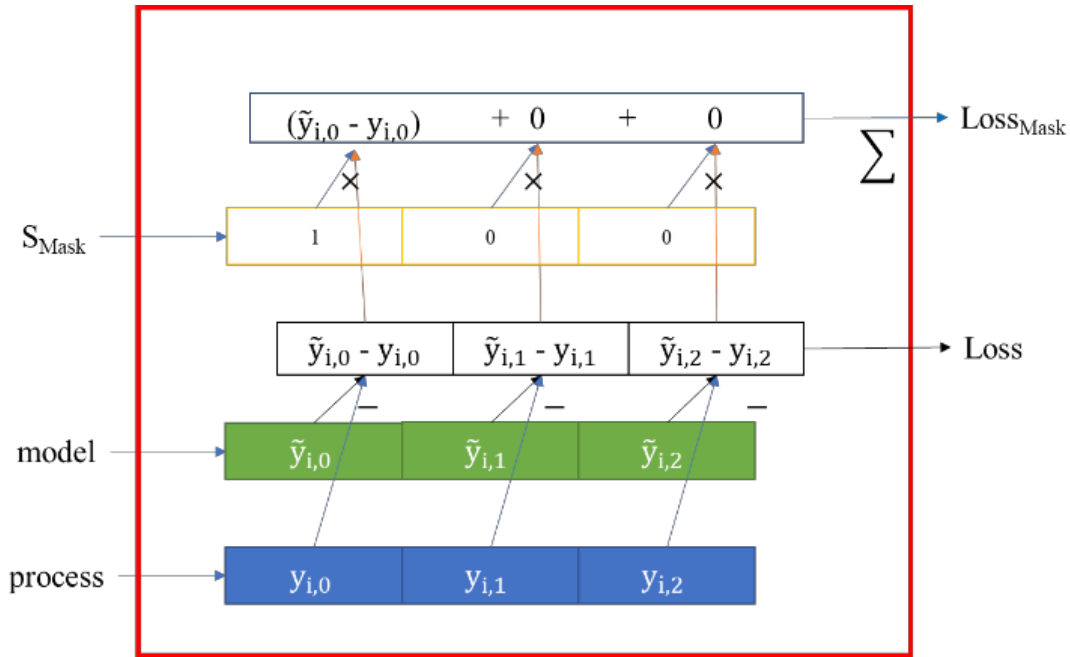


圖 2 遮罩遮蔽原理

### 模型物理解釋性

Ribeiro, et al. [29]指出儘管深度學習模型被廣泛應用在軟儀錶，但其仍然屬於黑盒子模型，了解背後的物理意義對於模型使用上非常重要。當模型用於決策時，必須取得使用者信任。目前模型可使用準確性指標評估，除此之外，檢查模型解釋也是其中一個選擇。使用者通常對應用領域有先驗知識，如果能使其了解模型預測背後的推理，便能接受或否決模型。

Hsiao, et al. [30]使用先驗分析來解釋模型。透過對輸入中的特定特徵進行擾動，而輸出對擾動輸入的變化，稱為過程增益，其應與化工領域知識在物理上一致。而過程增益( $K_{ij}^{dyn}$ )如公式 4，其中 $\Delta qv_{i,t}$ 為品質變量的變化， $\Delta u_{j,t}$ 為操作條件的擾動， $i$ 為品質變量， $j$ 為操作條件：

$$K_{ij}^{dyn} = \left( \frac{\Delta qv_{i,t}}{\Delta u_{j,t}} \right) \quad (4)$$

將動態和穩態過程( $K_{ij}^{ss}$ )增益方向一致的測試樣本的百分比定義為增益一致性( $Con_{ij}$ )，如下公式：

$$Con_{ij} = \frac{\sum_{n=1}^N \left( Hv \left( sign(K_{ij}^{dyn}) sign(K_{ij}^{ss}) \right) \right)}{N} \times 100\% \quad (5)$$

$Hv$  為 Heaviside 函數，可解釋軟儀錶應具備高增益一致性，在應用於 MPC 中，才得以提出合理的最佳操作條件。

### 增益一致性損失函數

模型在物理解釋方面，本研究透過增益方向，驗證穩態增益方向與動態增益方向是否為同向，以確認模型與製程相同。圖 6 為增益方向示意圖，通過計算改變操作後，預測目標的變化除上操作條件的變化去計算增益方向。

採用增益方向的概念[37]，穩態增益方向計算方式如公式 6，穩態增益方向為  $K_{ss}$ ， $\hat{y}_{i,j,t}$  為原操作穩態模型預測值， $\hat{y}'_{i,j,t}$  為操作變化後穩態模型預測值， $mv_{i,j,t}$  為原操作條件， $mv'_{i,j,t}$  為變化後操作條件， $i$  為品質變量編號， $j$  為操作變量編號。模型動態方向計算方式如公式 7，動態增益方向為  $K_{dyn}$ ， $\hat{y}_{i,j,t}$  為原操作動態模型預測值， $\hat{y}'_{i,j,t}$  為操作變化後動態模型預測值。公式 8 為增益損失函數， $H$  為邏輯判斷函數，其受模型增益方向  $K_{dyn}$  與穩態增益方向  $K_{ss}$  影響，當兩者相乘大於 0 時，代表兩者方向相同， $H$  會等於 1；當兩者相乘小於 0， $H$  會等於 0。當兩者同向時，增益損失函數為 0；當兩者反向時，損失函數為  $K_{dyn}$  的絕對值平均，當反向的程度越高，會給出更高的損失以校正模型方向。

$$K_{ss_{i,j,t}} = \begin{cases} 1, & \text{if } \frac{\hat{y}_{i,j,t} - \hat{y}'_{i,j,t}}{mv_{i,j,t} - mv'_{i,j,t}} > 0 \\ -1, & \text{if } \frac{\hat{y}_{i,j,t} - \hat{y}'_{i,j,t}}{mv_{i,j,t} - mv'_{i,j,t}} < 0 \end{cases} \quad (6)$$

$$K_{dyn_{i,j,t}} = \frac{\hat{y}_{i,j,t} - \hat{y}'_{i,j,t}}{mv_{i,j,t} - mv'_{i,j,t}} \quad (7)$$

$$\mathcal{L}_{Gain} = \frac{\sum_{t=0}^L \sum_{j=0}^m \sum_{i=0}^n \left| K_{dyn_{i,j,t}} \times [H(K_{dyn_{i,j,t}}, K_{ss_{i,j,t}}) - 1] - 0 \right|}{\sum_{j=0}^m \sum_{i=0}^n N_{Gain_{i,j}}} \quad (8)$$

$$\text{where } \begin{cases} H(K_{dyn_{i,j,t}}, K_{ss_{i,j,t}}) = 1, & \text{if } K_{dyn_{i,j,t}} \times K_{ss_{i,j,t}} > 0 \\ H(K_{dyn_{i,j,t}}, K_{ss_{i,j,t}}) = 0, & \text{if } K_{dyn_{i,j,t}} \times K_{ss_{i,j,t}} < 0 \end{cases}$$

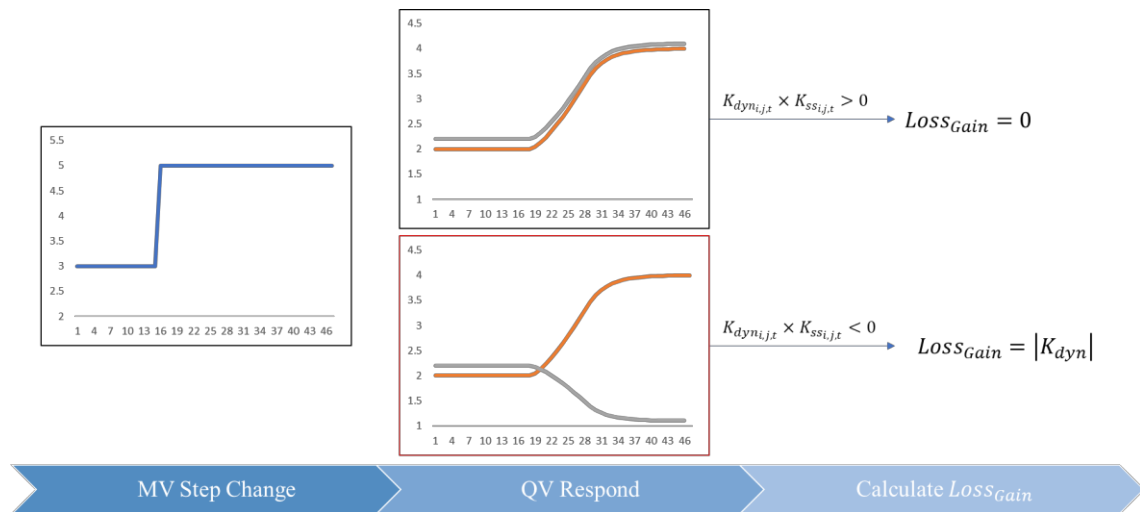


圖 3 增益方向示意圖

### 探討製程案例

本研究目前以針對選擇性氫化單元(Selective Hydrogenation Unit, SHU)中的碳四分離塔為建模對象，製程如圖 4 所示。該製程的主要產品品質目標是要將上塔的碳四含量純度提高至 99.9%，且下塔<sub>的</sub>出料碳四的純度不能高過 17.5%。製程可控的變數有上塔的回流閥及蒸餾塔的鍋爐溫度。下表顯示 C4/C5 分離製程變數清單。未來會再增加更多製程探討以確認模型適用性。

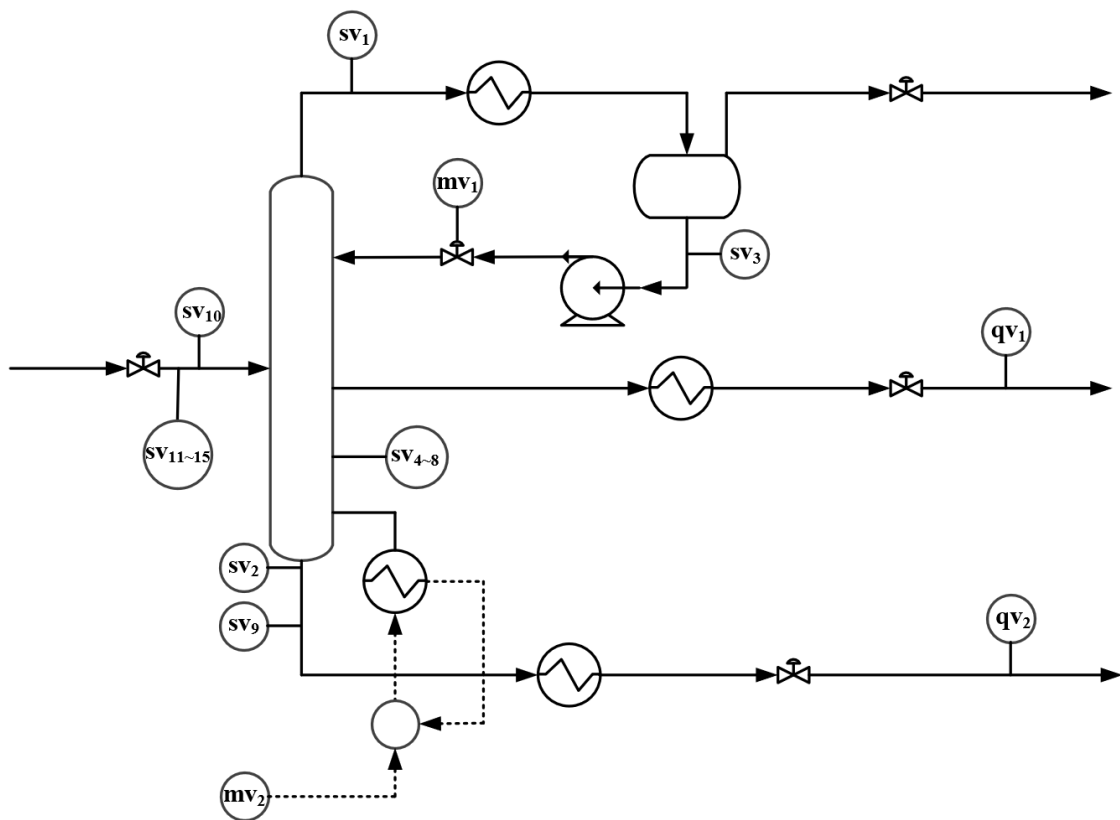


圖 4 製程架構圖

表 1 製程變數標籤名稱

變數種類	變數名稱	變數種類	變數名稱
QV <sub>1</sub>	Top C5 Product concentration	SV <sub>9</sub>	Reboiler heat
QV <sub>2</sub>	Bottom C4 Product concentration	SV <sub>10</sub>	Inlet flowrate density
SV <sub>1</sub>	Top pressure	SV <sub>11</sub>	Inlet flowrate
SV <sub>2</sub>	Bottom pressure	SV <sub>12</sub>	Inlet flowrate temperature
SV <sub>3</sub>	Condenser temperature	SV <sub>13</sub>	Inlet C4 composition concentration
SV <sub>4</sub>	1st stage temperature	SV <sub>14</sub>	Inlet C5 composition concentration
SV <sub>5</sub>	2nd stage temperature	SV <sub>15</sub>	Inlet impurity composition concentration
SV <sub>6</sub>	5th stage temperature	MV <sub>1</sub>	Reflux flowrate
SV <sub>7</sub>	35th stage temperature	MV <sub>2</sub>	Reboiler temperature
SV <sub>8</sub>	49th stage temperature		

在此研究中，將比較常見 AI 預測模型包含 DNN 團隊開發的 StS、以及文獻的 att-StS[10]以及本研究提出之 M-GI-StS 模型分析其製程動態預測行為、預測準確度、以及物理意義合理性。未來將針對更多製程進行分析，其中包含批次反應與更多蒸餾製程，如萃取蒸餾，並應用於後續各年計畫中進行探討適用性。

## 驗證環境

但目前 AI 建模都基於 python 環境建模，因此所建構的 AI-MPC 架構以 python 環境編寫較為方便，但 python 要與 Aspen Dynamics 動態模型互動，需要額外架構互動平台。下圖為第二年進行研究示意圖。首先，需建構 AI-MPC 以及 Aspen Dynamics 互動介面，確定資訊傳輸，接著將前期建構的軟儀錶以滾動或以直接水平預測方式結合 MPC 架構變成 AI-MPC，來與虛擬工廠進行互動，確認各種 AI-MPC 控制效能。分析任務包含單純品質控制效能比較以及引入節能目標下的控制表現，來確認何種 AI 模型更適合製程動態控制。

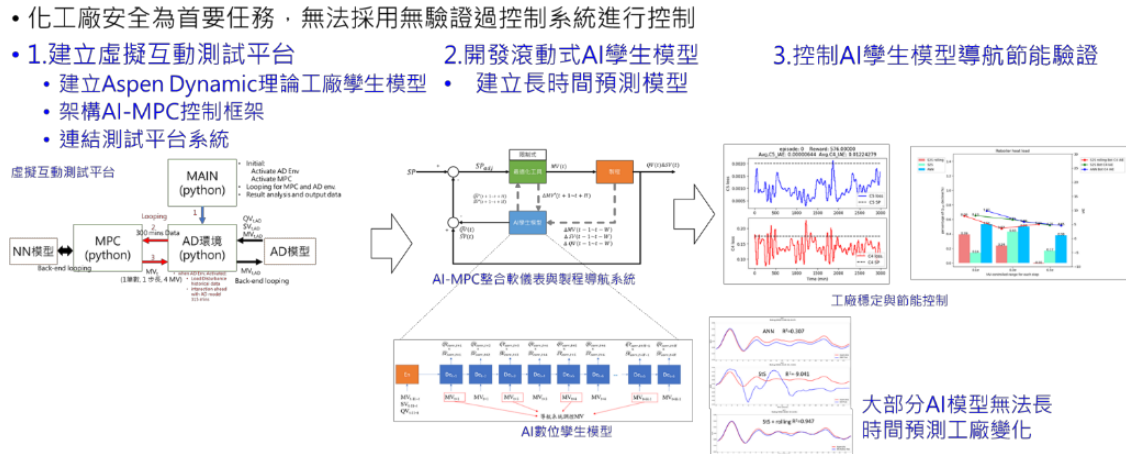


圖 5 AI-MPC 控制表現流程

## AI-MPC 與 Aspen Dynamics 互動平台開發

Aspen Dynamics 並無介面可以與 python 開發之 AI-MPC 進行互動，因此必須開發 AI-MPC 與 Aspen Dynamics 共同平台，傳遞資訊來模擬分析 AI-MPC 控制表現。圖 6 為 MPC 與 AD 的互動架構圖。首先，由紅色虛線處開始，此部分為預跑階段只執行一次，將歷史數據從數據集給入 AD 模型，同時對每步的資料進行紀錄，待 AD 模型輸入足夠的數據供神經網路做預測輸入。本研究預測模型需要 18 步，但預測控制系統取 19 步，其中 1~18 步用作誤差計算用，2~19 步做預測控制用，透過主程式將資料傳遞至 MPC 系統，逐步透過基於限制式的最適化方程式進行最佳操作搜索，搜索到的  $\Delta mv$  會與原始值做計算得到最佳操作動作，並回傳第一個時間點的操作動作至 AD 模型進行下一個時間步的規劃，依黑線處步驟 1 至 6 反覆迭代。以 C4/C5 製程為例 Aspen Dynamics 可要使用得 tags 如圖 7(輸出接口)與圖 8(輸出接口)所示。



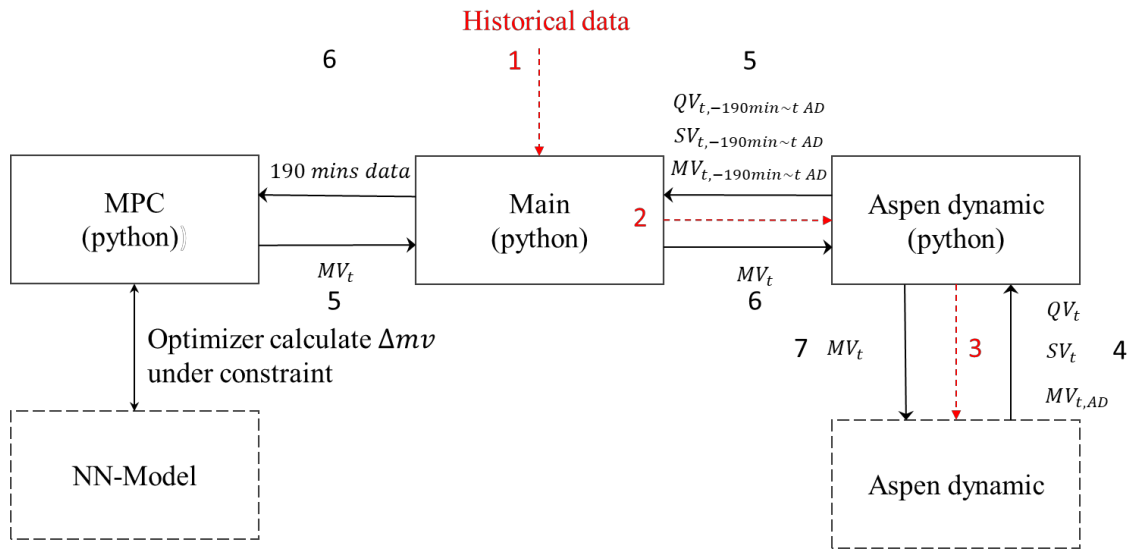


圖 6 AI-MPC 與 Aspen Dynamics 互動控制平台流程

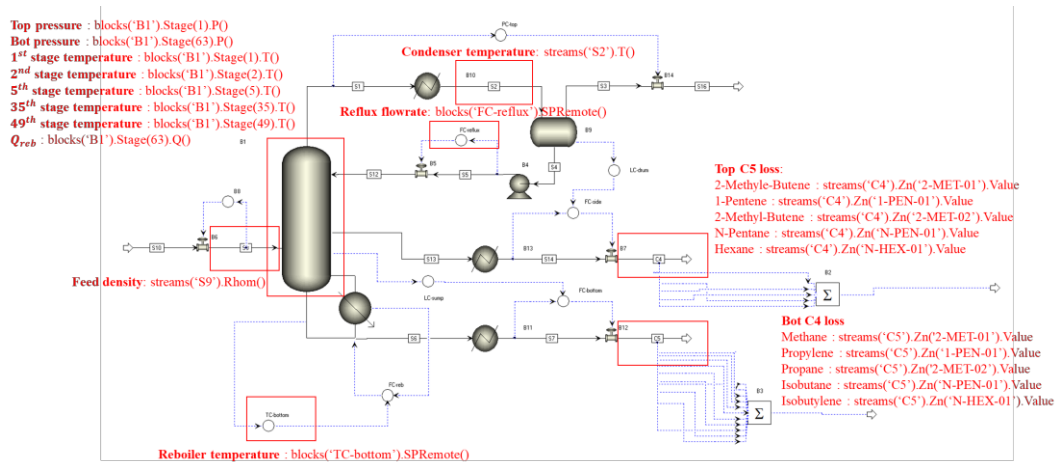


圖 7 Aspen Dynamic 輸出數據標籤與製程對應接口

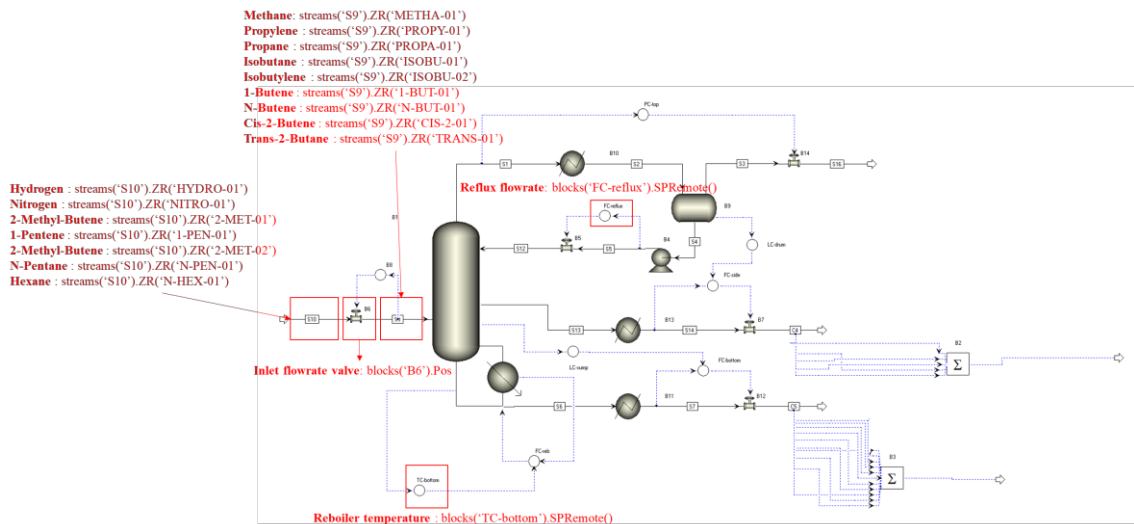


圖 8 Aspen Dynamic 輸入數據標籤與製程對應接口

## AI-MPC 架構

圖 9 為本研究模型預測控制系統流程圖，模型預測控制以 Aspen dynamics 建立的製程做為與模型對接的窗口，提供預測模型所需的輸入資料，做 100 個時間步的動作控制。在製程預跑 19 個時間點後，如圖 9 神經網路模型處分別使用三種模型：DNN, StS, StS rolling 作為比較模型，提供  $t-18 \sim t$  時間點的資料。其中  $t-18 \sim t-1$  時間點的資料作為誤差修正用，將預測結果與現有數據依據 EMWA 方法進行誤差修正。對  $t-17 \sim t$  的結果進行修正。透過差分演算法作為最適化演算法搜尋最佳操作條件後，供製程調控，藉此得到最佳的品質變量。限制式會根據製程需求寫入 Optimizer 中進行決策。

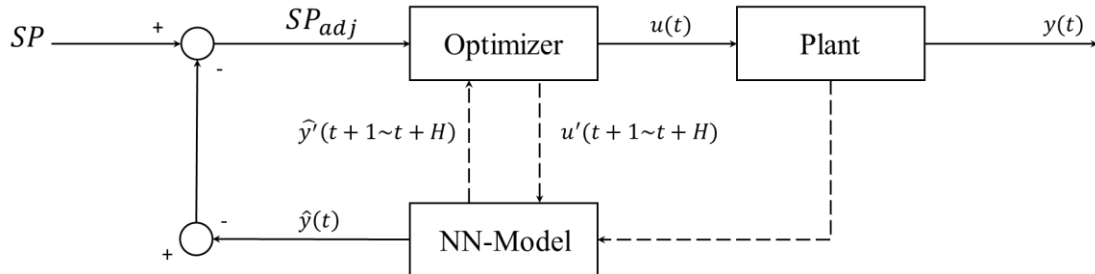


圖 9 模型預測控制系統示意圖

## 研究內容

1. 建構 M-GI-StS 模型
2. 分析模型對於多品質變數、多取樣頻率動態資訊重構能力
3. 分析模型各品質變量對於各調控變量的物理解釋性
4. 分析模型對於數值延遲的預測反應動態
5. 各種取樣頻率驗證模型極限
6. 3~5%製程雜訊挑戰
7. 建構 M-GI-StS 為基礎的 AI-MPC
8. 加入取樣延時劇本，分析 AI-MPC 控制能力
9. 套用於新式 Transformer 與 Mamba 結構，確認本研究提出套件的通用性。

**第二年：開發工廠軟儀表結合穩態理論代理模型，引入全局穩態增益方向性，強化軟儀表物理可解釋性**

基礎概念為使用前期開發的 M-GStS，針對於 Gain informed 部分進行改良。透過結合全局 Aspen Plus 穩態數據獲得正確穩態增益方向性。圖 10 為開發流程。總共分為兩步驟。第一步利用 NN 網路建構穩態 AI 代理模型，接著在第二步驟訓練工廠軟儀表並利用穩態 AI 代理模型引入穩態模擬數據與穩態增益方向性。

## 開發步驟

前期開發指引入在製程局部範圍內的方向一致性，透過損失函數方式帶入到 AI 軟儀錶中。但這有一缺點，此假設  $K_{ij}^{ss} = K_{ij}^{dyn}$ ，但現實上可能  $K_{ij}^{ss} \neq K_{ij}^{dyn}$ 。因此，在進行增益方向性訓練時，需要確保模型已達穩態，再要求穩態下的增益方向性才能合理的讓模型學會實質的物理意義。為了使 Aspen Plus 可

以有效的與 AI 軟儀表結合，我們需要透過 AI 代理模型來複製 Aspen Plus 的理論模型，稱之為理論 AI 代理模型，來加速理論模型與 AI 軟儀表訓練時間，如圖 11 所示。

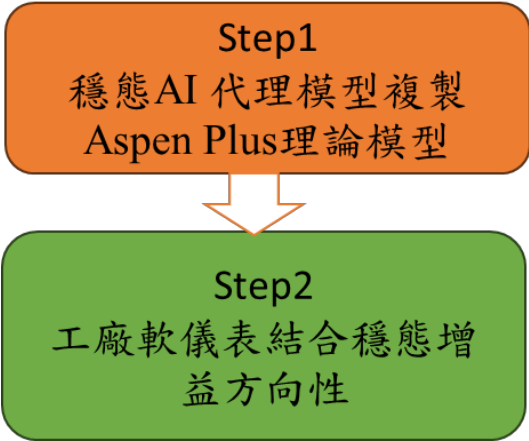


圖 10 工廠軟儀表結合穩態增益方向開發步驟說明

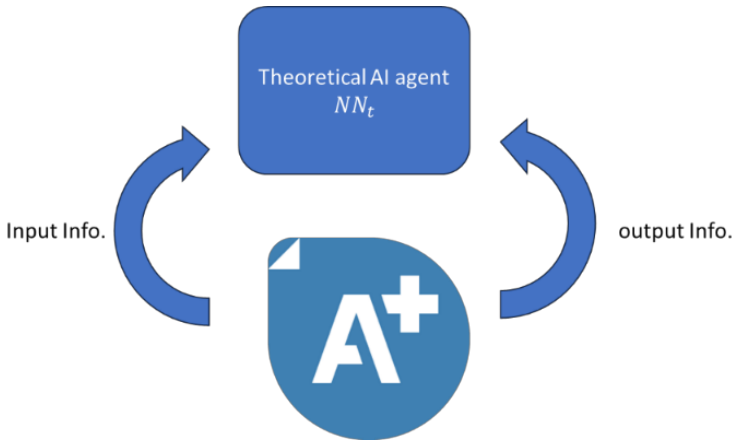


圖 11 利用 AI 代理模型複製 Aspen Plus 理論模型(Step 1)

真實數據平行帶入工廠軟儀表模型進行訓練，如圖 12 所示。在訓練過程，透過理論代理模型亂數生成全域操作範圍穩態數據，並計算理論的穩態增益，如圖 13 所示。圖 14 為真實數據帶入軟儀表訓練的水平輸入與輸出，並要求輸出的 y 需跟實場測量相同。

圖 15 為將理論的穩態數據 $mv_{ss}$ 與 $sv_{ss}$ 帶入軟儀表並期望模型輸出要擬合 $y_{ss}$ 。考量相同操作 $mv_{ss}$ 下，真實狀況的 $sv_{ss}'$ 與 $QV_{ss}'$ 的理論模型的 $sv_{ss}$ 與 $QV_{ss}$ 有偏差，故需透過 $\delta_{sv}$ 與 $\delta_v$ 進行修正，此兩參數以透過權重或非線性 NN 模型取代 $(\delta_{sv} = NN_{sv}(SV_{ss})$  與  $\delta_{qv} = NN_{qv}(QV_{ss}))$ 並且與工廠軟儀表一同訓練。

透過理論 AI 代理模型引入工廠軟儀表模型訓練中透過神經網路的自動微分計算當下操作的增益，取代前期的手動帶入的增益方向性方法來計算 $K_{ij}^{ss} = \left( \frac{\Delta qv_{i,t}}{\Delta u_{j,t}} \right)_{ss}$ 。理論 AI 代理模型複製理論模型的正確製程增益方向性，因此增益方向性訓練可以不侷限於歷史數據範圍，可透過擴大亂數取樣範圍來增加軟儀表的增益方向性合理操作範圍。

模型損失函數可改寫為：

$$\mathcal{L} = W_{QV}\mathcal{L}_{QV} + W_{SS}\mathcal{L}_{SS} + W_{gain}\mathcal{L}_{gain} \quad (9)$$

$$\mathcal{L}_{QV} = \text{MSE}(\widehat{QV}_{t \sim t+H}, QV_{t \sim t+H}) \quad (10)$$

$$\mathcal{L}_{SS} = \text{MSE}(\widehat{QV}'_{t \sim t+H}, \delta_{qv}\widehat{QV}_{SS}) \quad (11)$$

$$\mathcal{L}_{gain} = \text{MSE}(K_{SS}, K_{SS}') \quad (12)$$

損失函數考慮工廠軟儀表預測損失函數( $\mathcal{L}_{QV}$ )、穩態模擬數據預測損失函數( $\mathcal{L}_{SS}$ )、以及穩態增益方向性損失函數( $\mathcal{L}_{gain}$ )。工廠軟儀表預測損失函數( $\mathcal{L}_{QV}$ )是要求預測品質變量( $\widehat{QV}_{t \sim t+H}$ )要與實場測量變量相同( $QV_{t \sim t+H}$ )。穩態模擬數據預測損失函數( $\mathcal{L}_{SS}$ )則要求模擬數據帶入後的預測( $\widehat{QV}'_{t \sim t+H}$ )要與修正後的理論 AI 代理模型的預測( $\delta_{qv}\widehat{QV}_{SS}$ )相同。穩態增益方向性損失函數( $\mathcal{L}_{gain}$ )則要求工廠軟儀表的穩態增益方向( $K_{SS}$ )與 $K_{SS}'$ 與理論 AI 代理模型的增益方向( $K_{SS}'$ )要一致。

透過此訓練方式，可有效地將理論模型的增益方向性透過 AI 代理模型與工廠軟儀表模型耦合，並讓工廠軟儀表學習全域的增益方向性來增強模型的物理解釋性。

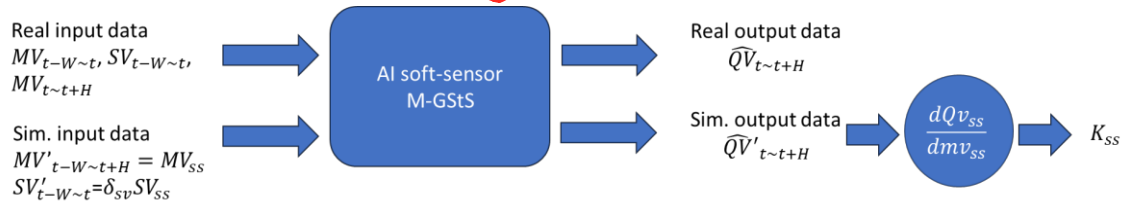


圖 12 將真實數據與生成的穩態數據帶入 M-GStS 進行訓練(Step 1)

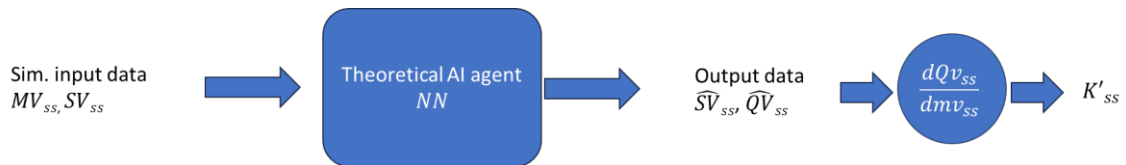


圖 13 利用 AI 代理模型在訓練過程中，亂數生成小批量穩態數據 (Step 2)

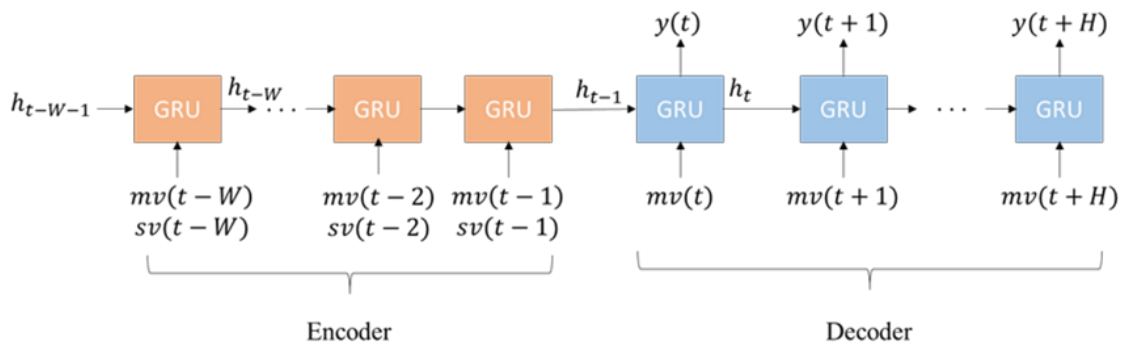


圖 14 真實數據帶入 M-GStS 模型的示意圖

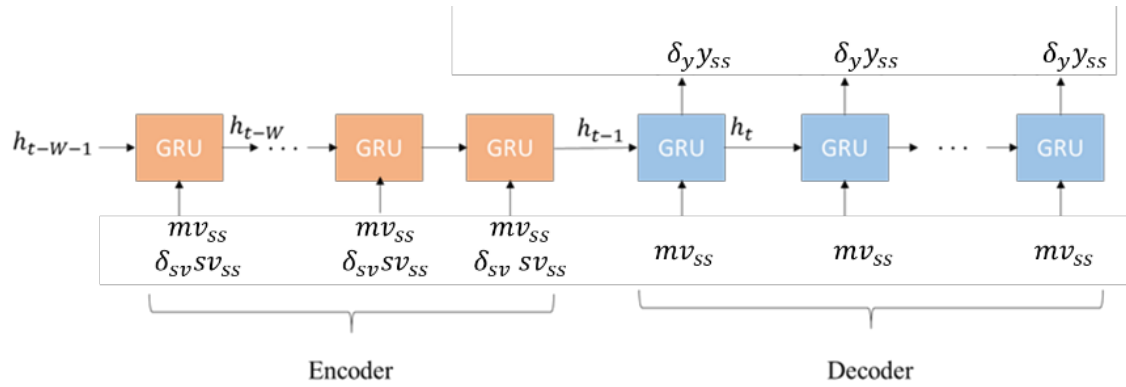


圖 15 生成的穩態數據帶入 M-GStS 模型的示意圖

## 研究內容

1. 建構 M-GI-StS 結合全域穩態增益物理解釋性
2. 調整損失函數權重，找到最佳理論與時廠數據的平衡模型
3. 套用於 AI-MPC 進行多情境測試確認模型增益方向性
4. 套用於 AI-MPC 進行多情境測試確認模型外插能力表現
5. 套用於新式 Transformer 與 Mamba 結構，確認本研究提出套件的通用性。

## 第三年：開發工廠軟儀表結合穩態與動態理論代理模型，引入全局動態與穩態增益方向性，可進行全域動態外插能力

然而前一年的開發是以引入全局穩態增益方向性，強化軟儀表物理可解釋性。但本研究的工廠軟儀表是為動態軟儀表，單純引入穩態增益方向性，僅能確定模型在用於 AI-MPC 控制時最終能帶向合理的穩態控制，但無法確保過程的增益方向性合理性。因此第三年開發將基於本團隊開發的 StS-rolling 數據驅動學生模型作為理論動態 AI 代理模型來複製 Aspen Dynamics 動態行為。該模型以證明可以有效地複製 Aspen Dynamics 蒸餾塔動態行為。透過理論動態 AI 代理模型將理論的動態增益方向性加入到工廠的軟儀表模型中。

## 列對序列滾動模型(Sequence to Sequence Rolling, Seq2Seq Rolling)介紹

本研究在序列對序列神經網路模型(Seq2Seq)[27]的基礎上加入了滾動訓練，將解碼器預測出的製程狀況及當前的控制動作，放入到下一次編碼器的輸入，並將最舊的歷史製程狀況進行移除，維持編碼器的輸入為只有 18 個時間點的資料。如圖 16 所示。

序列對序列滾動模型使用之神經網路結構和序列對序列神經網路模型相同，區別在滾動模型在訓練時，多了滾動的步驟，藉由滾動的方式將模型之預測結果迭代至模型編碼器的輸入值，並且模型的更新方式也有所更改，如圖 17 所示。模型在訓練時，除了使預測值跟實際值之間的誤差盡量最小化，式(13)，還會去擬合編碼器與解碼器的權重式(14)以提升模型的穩定性。因此模型的總損失函數，式(15)為式(13)和式(14)的加總。



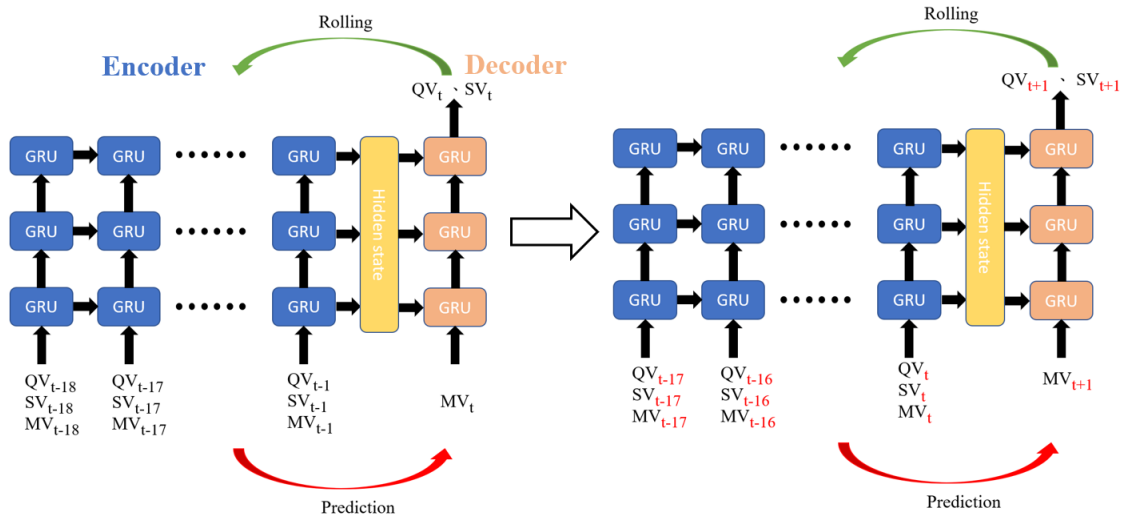


圖 16 序列對序列式滾動模型架構圖

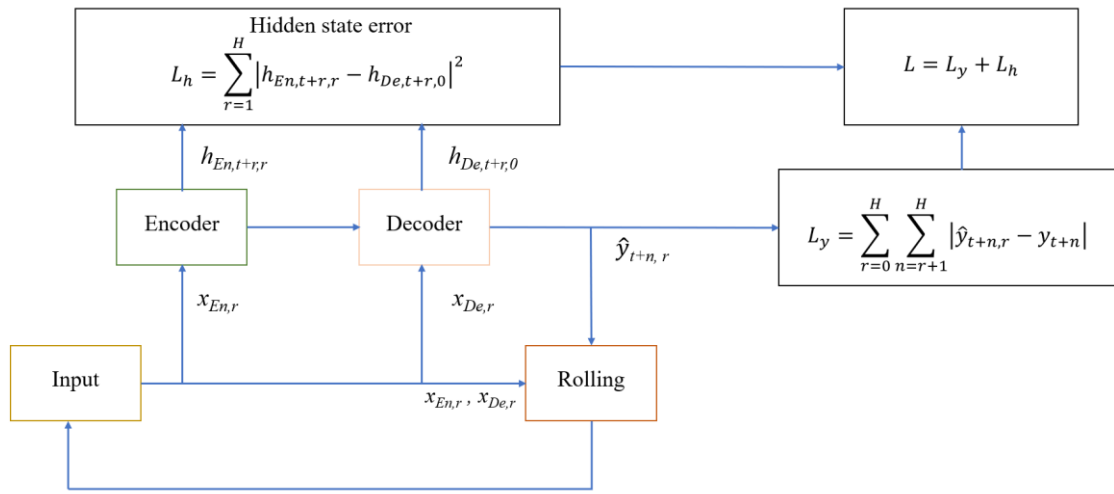


圖 17 序列對序列式滾動模型訓練架構圖

$$\mathcal{L}_y = \sum_{r=0}^H \sum_{n=r+1}^H |\hat{y}_{t+n,r} - y_{t+n}| \quad (13)$$

$$\mathcal{L}_h = \sum_{r=1}^H |h_{En,t+r,r} - h_{De,t+r,0}|^2 \quad (14)$$

$$\mathcal{L}_{rolling} = \mathcal{L}_y + \mathcal{L}_h \quad (15)$$

圖 18 為滾動預測之預測方式。在一開始編碼器接收到過去時間點的製程資訊，並將其轉化成特徵向量後( $h_{t+r+1}$ )，再將特徵向量及當前時間的操作及擾動傳遞給編碼器進行預測。在進行下一步預測前，將模型預測出的當前時間的製程狀態傳入編碼器的輸入資訊進行合併，並移除最舊的一筆歷史資訊，維持輸入資訊的時間長度，最後再將數據輸入編碼器進行特徵向量的轉換，在傳遞至解碼器進行預測，以此概念進行迭代。其中  $r$  和  $w$  分別表示數據滾動的次數以及輸入給編碼器多少時間步的製程資訊。

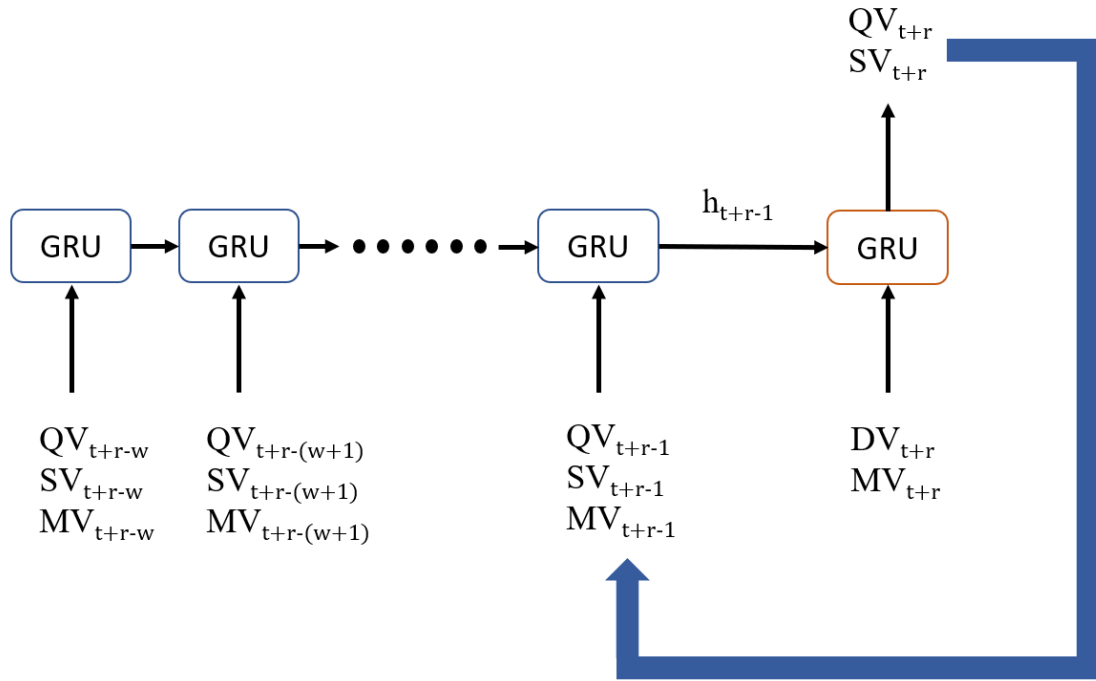


圖 18 滾動預測之測試方式

### 開發步驟

圖 19 為第三年計畫工廠軟儀表結合穩態、動態增益方向開發步驟說明。總共分為三步驟，第一步驟仍是建構穩態 AI 代理模型，其方法與前期一致。接著利用上述的 StS-rolling 方法建構動態 AI 代理模型，配合前期方法將穩態增益帶入動態 AI 代理模型中，而動態增益則透過 StS-rolling 的內部滾動機制學習。接著在第三步驟，將穩態與動態代理模型與工廠軟儀表結合，將穩態與動態增益引入到工廠軟儀表使模

型即便在訓練過程數據品質不穩定都能具有完整的動態與穩態物理意義。

圖 20 為動態 AI 代理模型的開發方法。採用第二年的方法應用於動態 AI 代理模型開發，確保動態 AI 代理模型不只能複製 Aspen Dynamis 的動態行為外，同時可具備正確的穩態增益方向性，減少動態 AI 代理模型與 Aspen Dynamis 的偏差。因為動態 AI 代理模型是以 StS-rolling 為基礎，所以其訓練的損失函數應改寫為如下。

$$\mathcal{L} = W_{rolling}\mathcal{L}_{rolling} + W_{ss}\mathcal{L}_{ss} + W_{gain,ss}\mathcal{L}_{gain,ss} \quad (16)$$

整體損失函數應包含原本 StS-rolling 損失函數外，額外要加入穩態預測( $\mathcal{L}_{ss}$ )以及穩態增益方向( $\mathcal{L}_{gain,ss}$ )的損失函數。

圖 21 為工廠軟儀表結合穩態、動態增益方向開訓練方法。訓練過程中，會透過輸入實場的數據與穩態模擬操作數據進行擬合實場與模擬的品質變量 QV，並透過自動微分方法計算每一個步長的穩態( $K_{ss}$ )與動態( $K_{dyn,t \sim t+H}$ )增益值。而穩態與動態 AI 代理模型也會分別計算理論的穩態與動態 QV 並計算理論的穩態( $K'_{ss}$ )與動態( $K'_{dyn,t \sim t+H}$ )增益值。透過訓練的損失函數將所有性質整合在一起，訓練損失函數可改寫如下：

$$\delta_{sv} = NN_{sv}(SV_{ss}) \quad (17)$$

$$\delta_{qv} = NN_{qv}(QV_{ss}) \quad (18)$$

$$\mathcal{L} = W_{qv}\mathcal{L}_{qv} + W_{dyn}\mathcal{L}_{dyn} + W_{ss}\mathcal{L}_{ss} + W_{gain,ss}\mathcal{L}_{gain,ss} + W_{gain,dyn}\mathcal{L}_{gain,dyn} \quad (19)$$

$$\mathcal{L}_{qv} = MSE(\widehat{QV}_{t \sim t+H}, QV_{t \sim t+H}) \quad (20)$$

$$\mathcal{L}_{dyn} = MSE(\widehat{QV}'_{t \sim t+H}, \delta_{qv}QV''_{t \sim t+H}) \quad (21)$$

$$\mathcal{L}_{ss} = MSE(\widehat{QV}'_{t \sim t+H}, \delta_{qv}QV_{ss}) \quad (22)$$

$$\mathcal{L}_{gain,ss} = MSE(K_{ss}, K_{ss}') \quad (23)$$

$$\mathcal{L}_{gain,dyn} = MSE(K_{dyn}, K_{dyn}') \quad (24)$$

這裡假設 $\delta_{sv}$ 與 $\delta_{qv}$ 理論跟工廠的修正因子在動態與穩態均為相同，因此應可使用同一種修正參數或模型。整體的損失函數包含原本 M-GStS 的損失函數( $\mathcal{L}_{qv}$ )、理論動態預測損失函數( $\mathcal{L}_{dyn}$ )、理論穩態預測損失函數( $\mathcal{L}_{ss}$ )、以及穩態( $\mathcal{L}_{gain,ss}$ )與動態增益損失( $\mathcal{L}_{gain,dyn}$ )函數。

依此方法建構的軟儀表，應可同時具備準確預測以及完整的動態、穩態增益方向性能力。但需要透過權重的調整讓模型找到理論與工廠數據的平衡，可能是本研究需要耗費時間的地方。

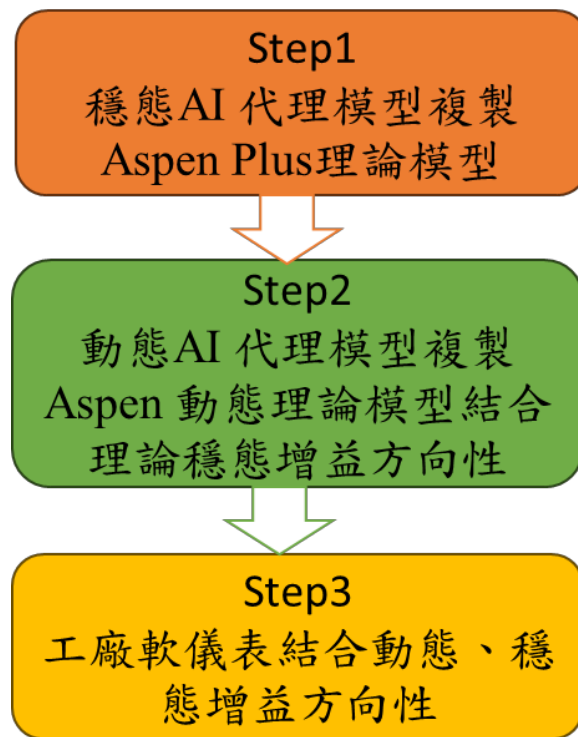


圖 19 工廠軟儀表結合穩態、動態增益方向開發步驟說明

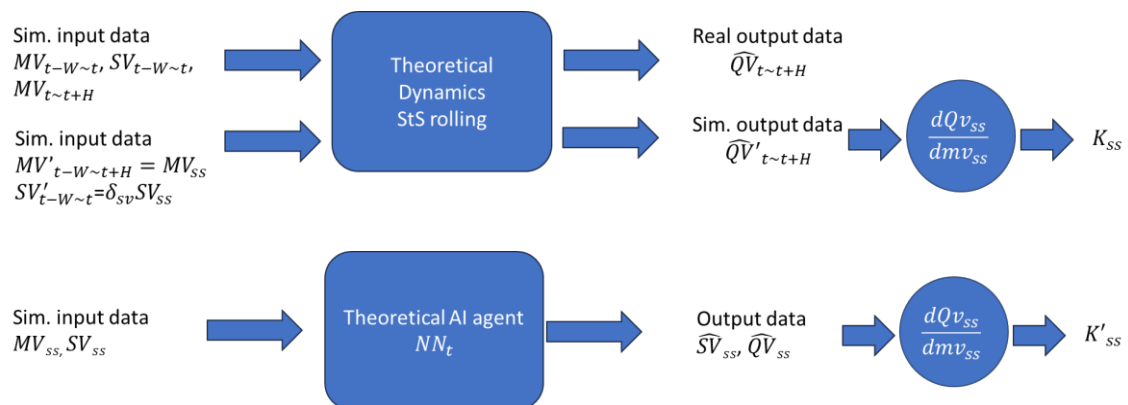


圖 20 動態代理模型 StS-rolling 建構方法

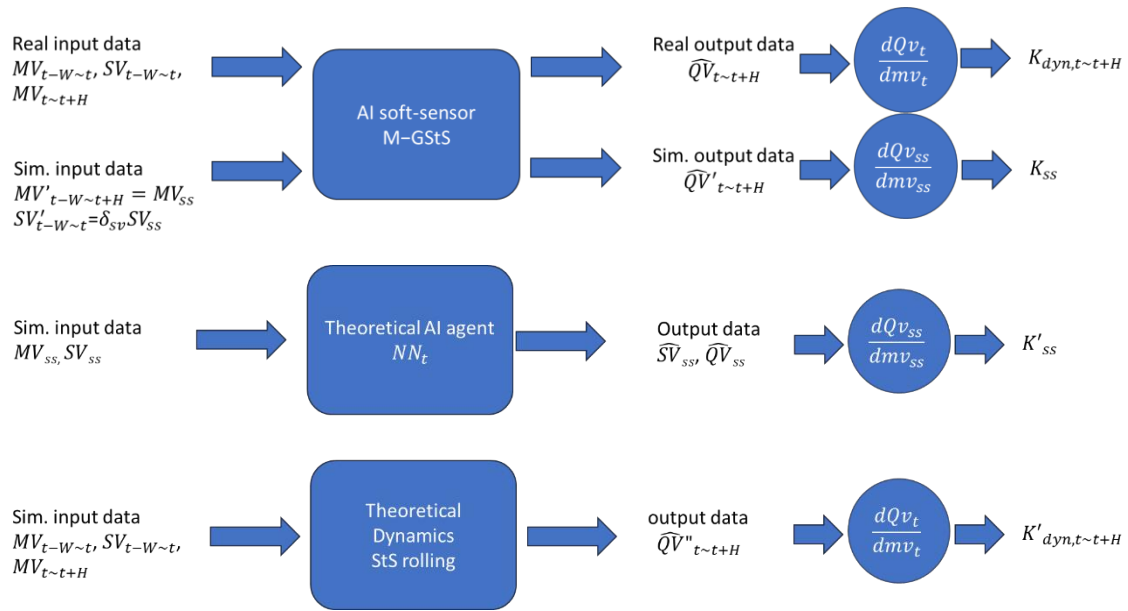


圖 21 工廠軟儀表結合穩態、動態增益方向開訓練方法

### 研究內容

1. 建構 M-GI-StS 結合全域穩態、動態增益物理解釋性
2. 調整損失函數權重，找到最佳理論與時廠數據的平衡模型
3. 套用於 AI-MPC 進行多情境測試確認模型增益方向性
4. 套用於 AI-MPC 進行多情境測試確認模型外插能力表現
5. 套用於新式 Transformer 與 Mamba 結構，確認本研究提出套件的通用性。



## 初步研究成果

### 1. 建構基於多取樣頻率之少樣數據重建動態資訊軟儀錶

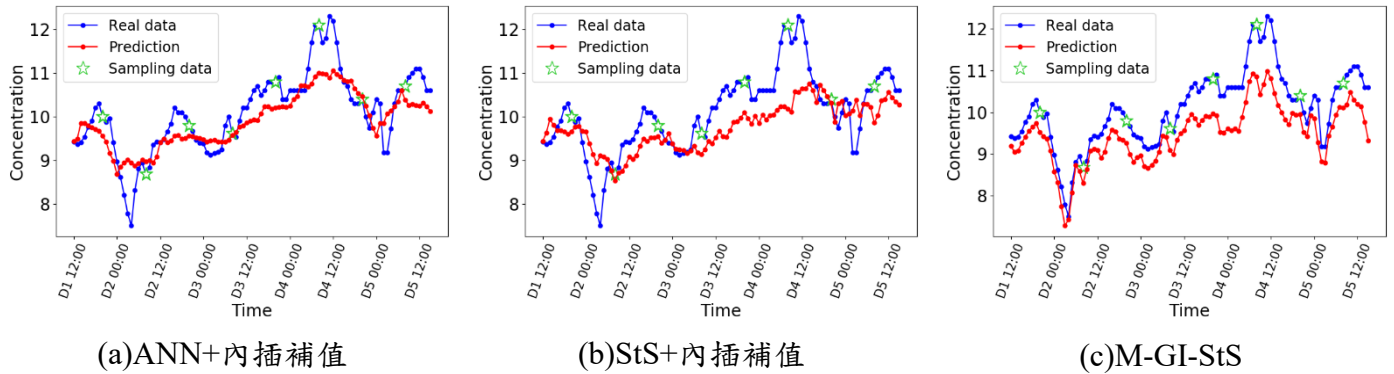


圖 22 三種軟儀錶模型對於 12h 取樣頻率的重建動態能力

圖 22 為(a)ANN 模型配合線性內插數據、(b)StS 模型配合限性內插數據、以及(c)本研究提出 M-GI-StS 直接訓練在 t 時間點製程濃度的區段預測結果，藍點為實際值，紅點為模型預測值，綠色星號為採樣點，橫軸日期 D1 說明：Day 1。前種模型的預測趨勢不明顯，並且發現高點與低點位置不明確，且會有數據延遲預測的問題，約延遲 12 小時。而本研究提出 M-GI-StS 初步發現模型僅使用 12 小時取樣數據，配合高頻率的輸入數據，可以有效重構製程上上下下的動態行為，並且沒有延遲問題。證明提出之方法具有可行性。

### 2. 製程學生模型

圖 23 為使用 ANN、StS 以及所提出的 StS+rolling 方法預測結果。圖中紅線為 Aspen Dynamics 產生的數據結果，而藍色為各模型滾動預測結果。初步我們發現使用傳統 ANN 進行滾動預測以及使用 StS 進行滾動預測都發現隨著預測時間拉長，預測效果就隨之變差，兩種模型 R2 都低於 0.3 以下。但所提出的 StS+rolling 卻可以在滾動預測中仍維持很好的預測能力  $R^2 > 0.9$ ，這證明本研究所提出的方法，非常有潛力可以成功開發出更長的預測能力之數據驅動數位孿生模型。

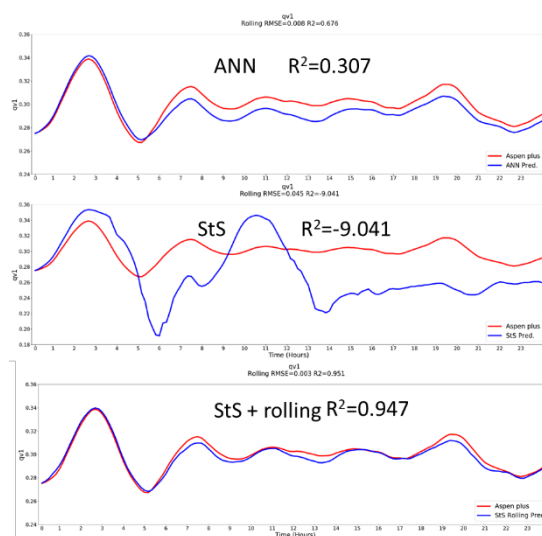


圖 23 ANN、StS、StS+rolling 預測製程一天的結果差異

(三) 期完成之工作項目及成果。請分年列述：1.預期完成之工作項目。2.對於參與之工作人員，預期可獲之訓練。3.預期完成之研究成果（如實務應用績效、期刊論文、研討會論文、專書、技術報告、專利或技術移轉等質與量之預期成果）。4.學術研究、國家發展及其他應用方面預期之貢獻。

#### 預期完成之工作項目

第一年計畫：

1. 建立可應用於多取樣頻率之數據建構高頻率水平預測模型，並具有高物理解釋性
2. 分析模型預測準確度、動態行為、以及初步物理解釋性分析
3. AI-MPC 測試控制效益

第二年計畫

1. 建構 M-GI-StS 結合全域穩態增益物理解釋性
2. 調整損失函數權重，找到最佳理論與時廠數據的平衡模型
3. 套用於 AI-MPC 進行多情境測試確認模型增益方向性
4. 套用於 AI-MPC 進行多情境測試確認模型外插能力表現

第三年計畫

1. 建構 M-GI-StS 結合全域穩態、動態增益物理解釋性
2. 調整損失函數權重，找到最佳理論與時廠數據的平衡模型
3. 套用於 AI-MPC 進行多情境測試確認模型增益方向性
4. 套用於 AI-MPC 進行多情境測試確認模型外插能力表現

#### 對於參與之工作人員，預期可獲之訓練

- (1)學習如何編寫程式，建立數學模型
- (3)熟悉人工智慧與深度學習理論與應用
- (4)學習如何操作與建立深度強化學習
- (5)學習實驗操作及數據整理分析之能力
- (6)瞭解如何撰寫報告及發表學術論文以及申請專利
- (7)熟悉 Aspen Dynamics 操作以及與 Python 程式的連結
- (7)產業合作經驗

#### 預期完成之研究成果（如實務應用績效、期刊論文、研討會論文、專書、技術報告、專利或技術移轉等質與量之預期成果）

- (1)目前學術界上無此類似論文發表，本技術研發成功預期可於國際期刊發表論文。
- (2)預計可發表國際期刊數量：共計 3 篇文章
  1. 多取樣頻率之少樣數據重建動態資訊軟儀錶
  2. M-GI-StS 結合全域穩態增益物理解釋性
  3. M-GI-StS 結合全域穩態、動態增益物理解釋性
- (3)與產業簽訂產學合作，推廣此一技術發展