

基於類神經網路的 數據劉維方程式與展期預報應用

Construction and Validation of the MJO Tangent Linear Operator
Using the ECMWF S2S Ensemble Forecasts Dataset

ECMWF

QNTD



A7-19

中央氣象署科技發展組 台灣大學大氣科學研究所

陳奕安、曾開治、李明營

大綱

研究背景介紹

Introduction

數學工具與
理論框架

Mathematical Framework

資料介紹、
線性逆模型

Data and linear inverse model

物理資訊神經網
路模型設計-軌跡
逆算子

Physics-Informed
Neural Network,
PINN

模型驗證與非自
治性診斷

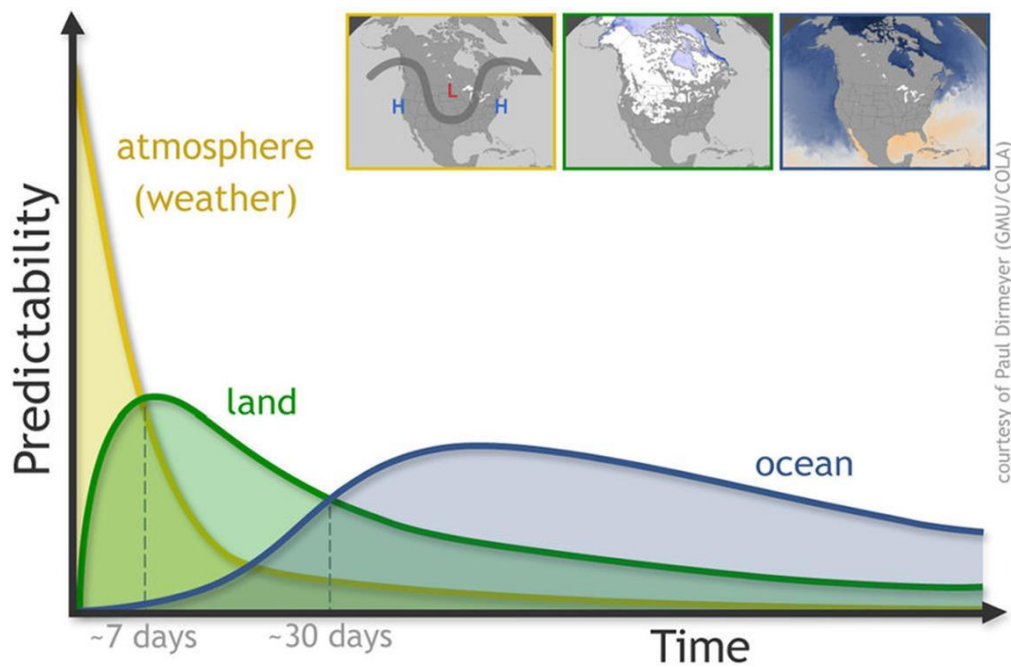
Validation of model

結語與展望

Conclusions

介紹:s2s的預報

S2S 的現實狀況：在次季節至季節（S2S）時間尺度（預測期 2 週至 2 個月），氣象預測會進入「可預報性沙漠」（predictability desert）。S2S 恰好於這交界處。大氣失去對其初始狀態的記憶。S2S 的可預報性主要受慢變化的邊界條件支配：海面溫度（SST）、土壤濕度、積雪、海冰覆蓋率以及平流層相互作用。



This is figure 3.2 in the report "NOAA Subseasonal and Seasonal Forecasting Innovation: Plans for the Twenty-First Century". Credit: NOAA Climate.gov graphic, adapted from original by Paul Dirmeyer.

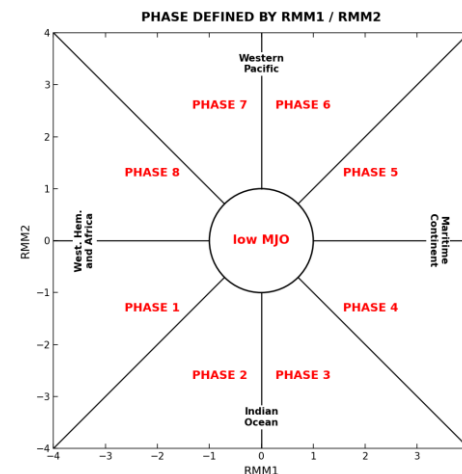
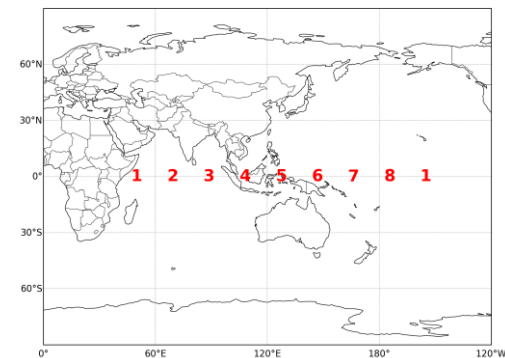
S2S 預報因子：

- 大氣熱帶遙連：MJO（熱帶季內振動）
- 平流層影響：平流層極渦、SSW（平流層突然偏溫）
- 海溫邊界條件：ENSO（聖嬰/反聖嬰）、IOD（印度洋洋熱帶偏振）
- 地表邊界條件：積雪覆蓋率、土壤濕度異常



<https://climate.cwa.gov.tw/ISO>

Approximate locations of the MJO centre of convection RMM index phases 1-8



系集預報

大氣是混沌的

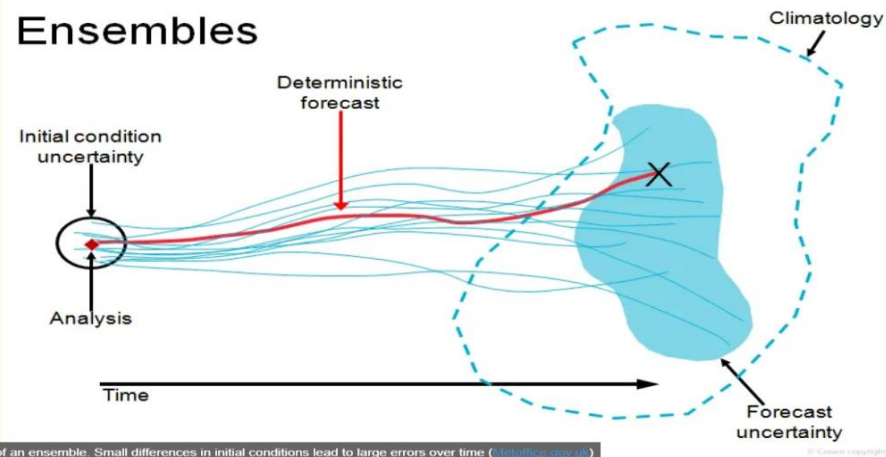
- **蝴蝶效應**：初始觀測資料的極微小誤差，會在非線性系統中指數級放大，最終導致完全不同的天氣結果。

非線性交互作用

- 可預報度極限：由於這種混沌特性，人類對精確天氣預報的理論極限大約是 10 到 15 天。
- 為了克服混沌帶來的預報誤差，現代氣象學使用**系集預報（Ensemble Forecasting）**。來掌握不確定性。



Ensembles



以數值系集預報為基礎，使用**ML-Liouville equation** 計算密度的分配。

比較 **NWP** 的優勢 與 **MLWP** 的局限

1. 耦合地球系統動力：具有大氣、海洋、陸地與海冰反饋機制。多以資料訓練當作下邊界，缺乏耦合反饋。
2. 自迴歸與能量傳遞：流體力學方程，不損失能量在頻譜中。**MSE**損失函數與自迴歸（低通濾波效應），易出現平滑模糊化與小尺度動能耗散。
3. 物理守恆與氣候漂移：數學物理，確保全球質量、濕度與能量的守恆。缺乏物理約束，易累積非物理誤差，出現「氣候漂移、偏差」。
4. **S2S**（次季節至季節）現象模擬：直接模擬平流層—對流層耦合、波力學與雲—輻射反饋等底層因果機制。僅學習歷史資料相關性而非物理因果，易產生無法以物理解釋的現象。
5. 氣候變化與極端事件（OOD）：依賴不變的基礎物理定律（**N-S** 方程等），適用於未知的氣候狀態。本質為基於歷史數據（如 **ERA5**）的統計插值，面對極端氣候或分佈外事件易因假設失效而預測失準。
6. 集合校準與不確定性量化：結合物理擾動與隨機參數化，能提供校準良好、具物理意義的**概率離散度**。雖運算速度快，但隨機加噪容易導致信號衰減或非物理惡化，構建合理的集合離散度仍為瓶頸。

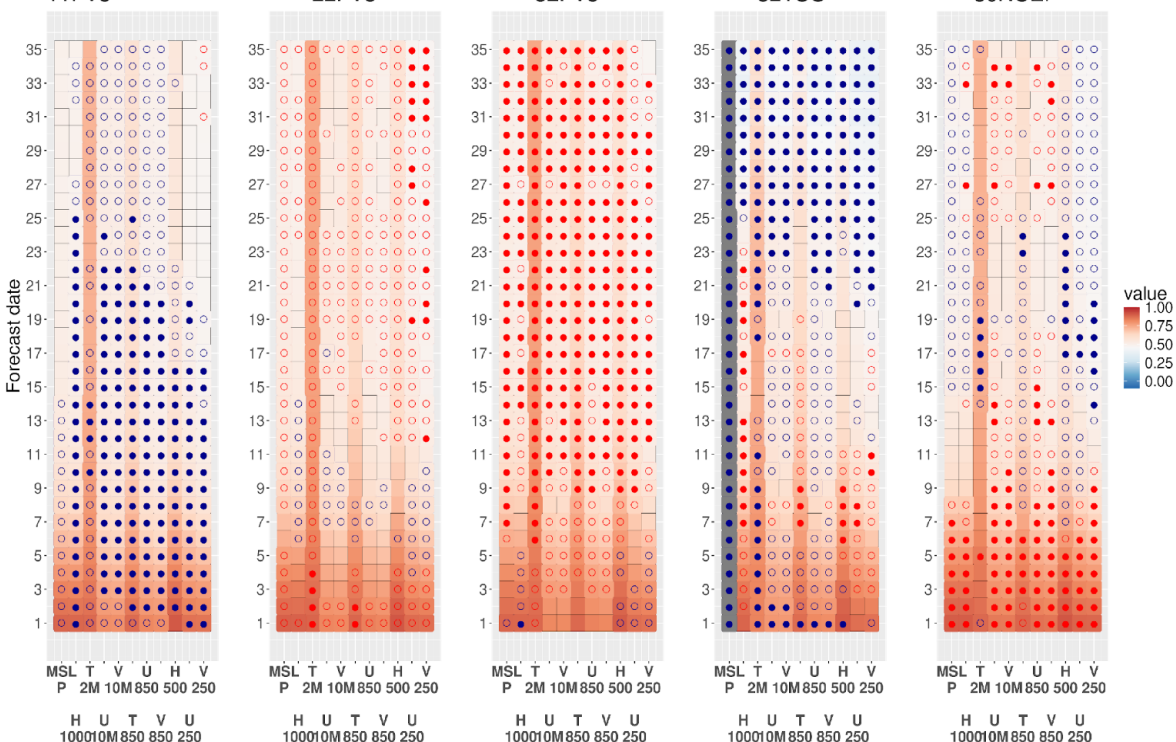
NH the area of the ROC curve 202310
11FV3 22FV3

113年A2-12

32FV3

32TCO

30NCEP



本研究的數學工具

- 線性逆模型 (**Linear Inverse Model**) 「給定觀測數據，反過來推導 (逆向求解) 出最符合數據的線性系統算子」

$$\frac{d\mathbf{x}}{dt} = \mathbf{L}\mathbf{x} + \xi$$

- 劉維方程式 (**Liouville equation**) 描述了系統機率密度函數 (**PDF**) 在相空間 (**Phase space**) 中的演變，具有密度守恆的特性。

$$\mathbf{J} = [\dot{q}, \dot{p}] \rho \quad \text{where} \quad \rho = \rho(q, p, t)$$

○ Coordinates (q, p) :

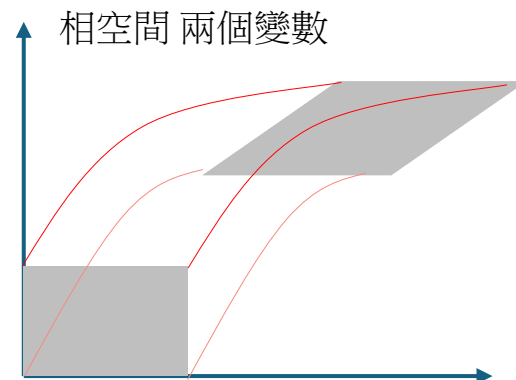
- $q = \text{PC1}$
- $p = \text{PC2}$

$$0 = \frac{\partial \rho}{\partial t} + \nabla \cdot \mathbf{J} \quad \text{where} \quad \nabla \cdot \mathbf{J} = \frac{\partial(\rho \dot{q})}{\partial q} + \frac{\partial(\rho \dot{p})}{\partial p}$$

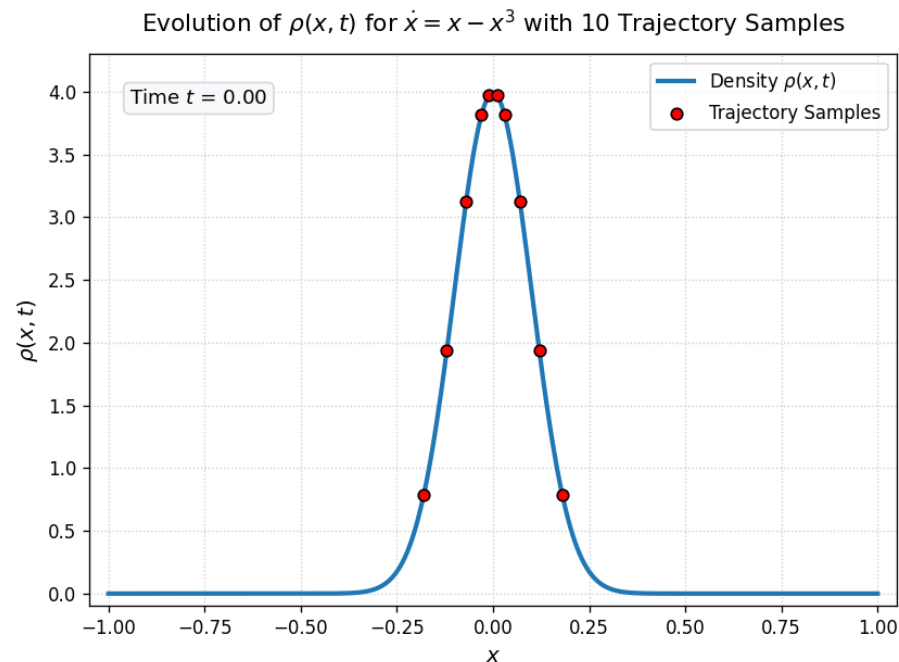
舉例

$$\frac{d}{dt}x = x - x^3$$

$$\frac{\partial \rho(x, t)}{\partial t} + (x - x^3) \frac{\partial \rho(x, t)}{\partial x} = -(1 - 3x^2) \rho(x, t).$$
$$\rho(x, t) = \rho_0(\xi[x, t]) \exp \left[- \int_0^t \underbrace{\{1 - 3(x\{\xi[x, t], t'\})^2\}}_{=\psi(x)} dt' \right].$$



Joseph Liouville published his foundational work in 1838 detailing what is now known as **Liouville's Theorem** in Hamiltonian mechanics and statistical physics.



資料和前處理

ERA5 daily qv

Year: 43 (1979-2021)

horizontal grid spacing : 0.5 degree * 0.625 degree

S2S ensemble forecasts data

Dimensions:

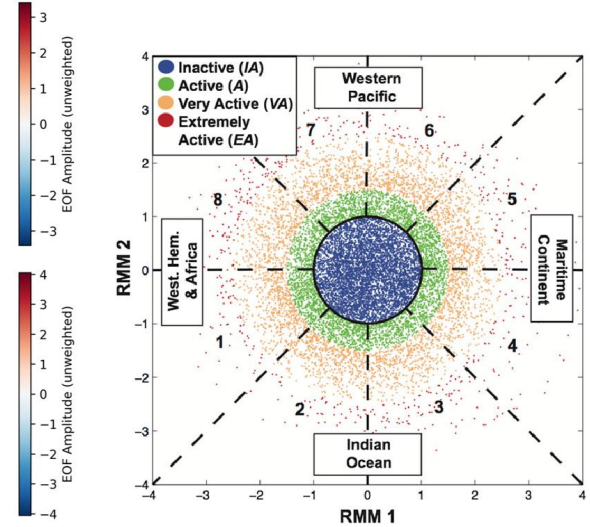
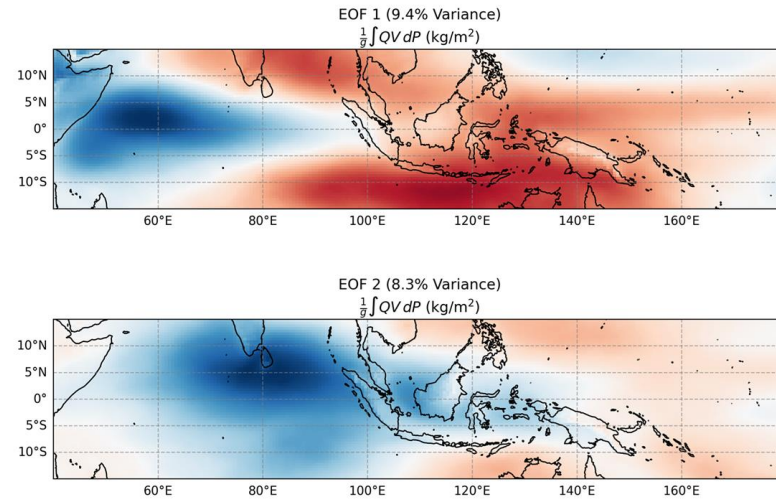
Ensemble members (file): 10

Time: 8 Year: 19 (2001-2019) Month: 12

Step: 47 (forecast days)

horizontal grid spacing : 1.5 degree * 1.5 degree

ERA5



https://www.researchgate.net/figure/Phase-space-diagram-of-the-RMM-index-WH04-showing-daily-phase-quadrant-and-magnitude_fig1_276474259

$\int (ERA5 \text{ qv}) dp$

Remove first 3 harmonic

EOF1 EOF2

$\int (S2S \text{ qv}) dp$

Subtract climatology

Anomaly

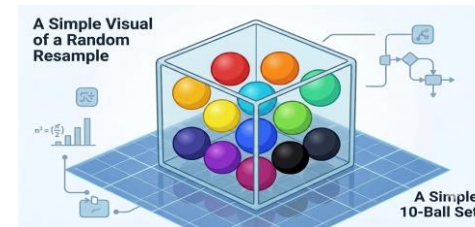
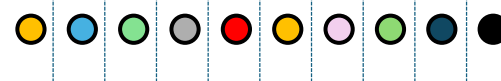
PC1 PC2



線性逆模型、線性算子G

1,024,000 bootstrap iterations , then averaged all iterations

10^{10}



10 members

$\begin{pmatrix} pc1 \\ pc2 \dots \end{pmatrix}$

$_{step(t)=2}$

$= \begin{pmatrix} G11 & G12 \\ G21 & G22 \end{pmatrix}$

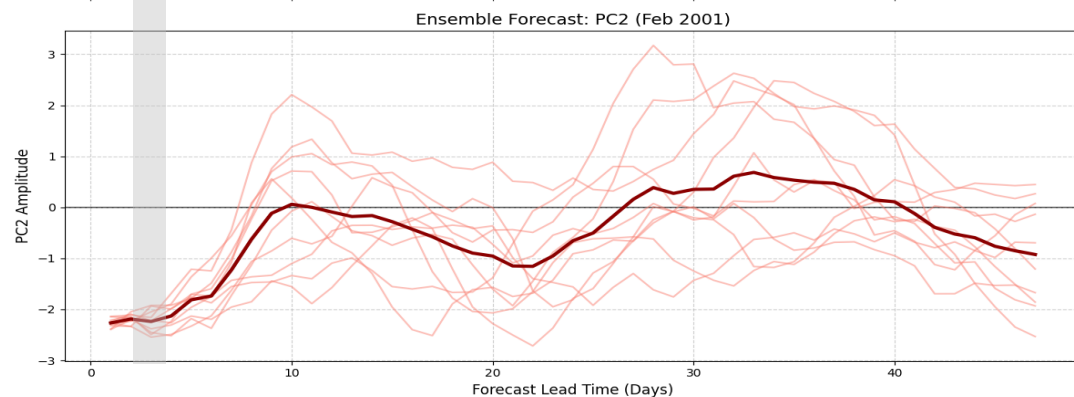
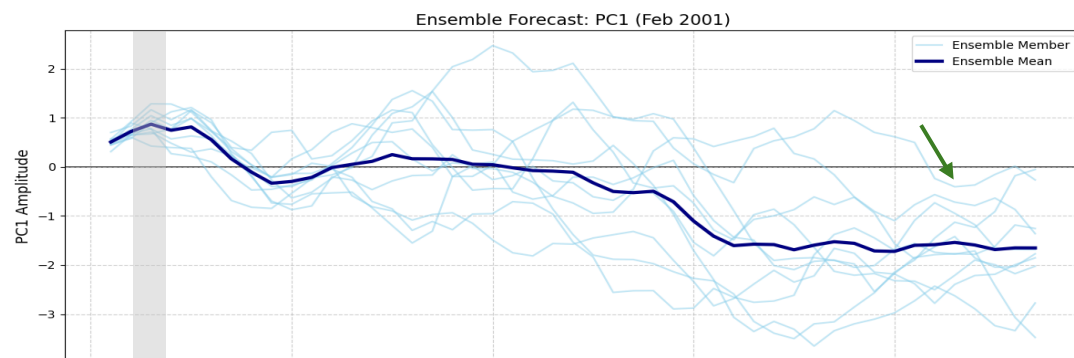
10 members

$\begin{pmatrix} pc1 \\ pc2 \dots \end{pmatrix}$

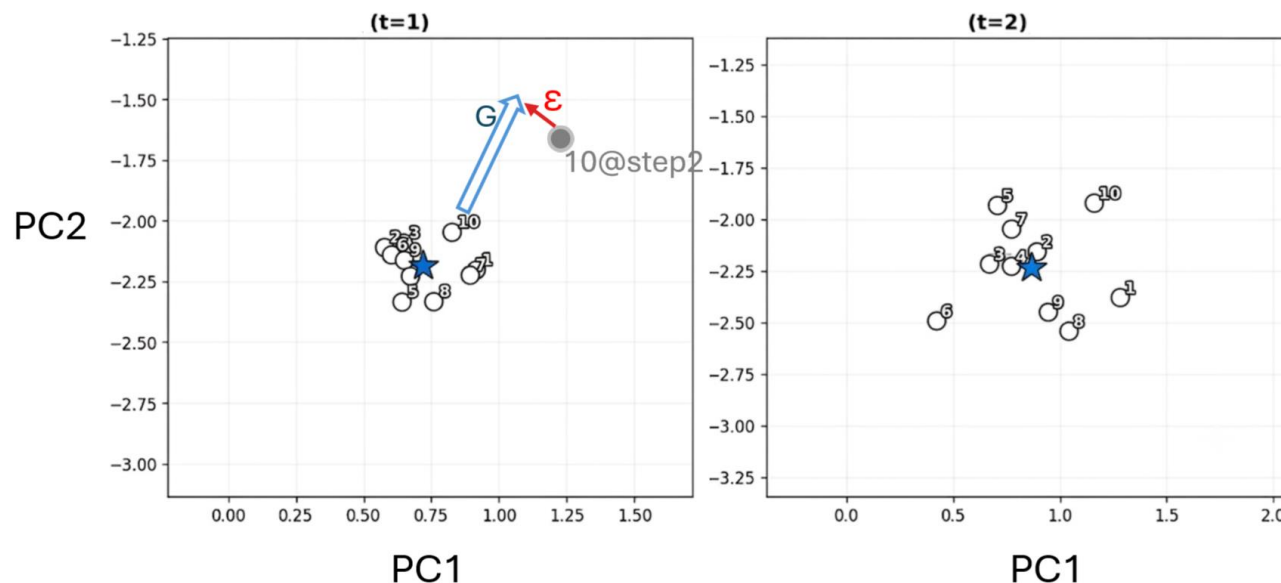
$_{step(t)=1}$

4 unknown variables need at least 4 equations to solve 10 ensemble members

$$\underbrace{G}_{2 \times 2} = \underbrace{PC_{t=2}}_{2 \times 10} \times \underbrace{PC_{t=1}^T}_{10 \times 2} \times \underbrace{(PC_{t=1} PC_{t=1}^T)^{-1}}_{2 \times 2}$$



星星是系集平均



10 members

$(\epsilon \dots)$

$= \begin{pmatrix} G11 & G12 \\ G21 & G22 \end{pmatrix}$

10 members

$\begin{pmatrix} pc1 \\ pc2 \dots \end{pmatrix}$

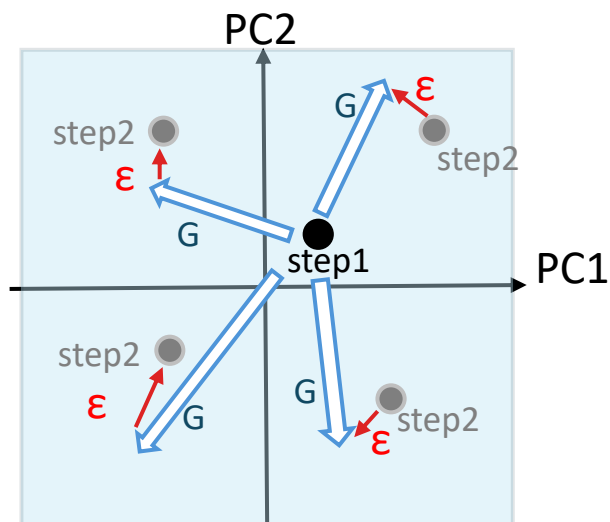
$_{step(t)=1}$

10 members

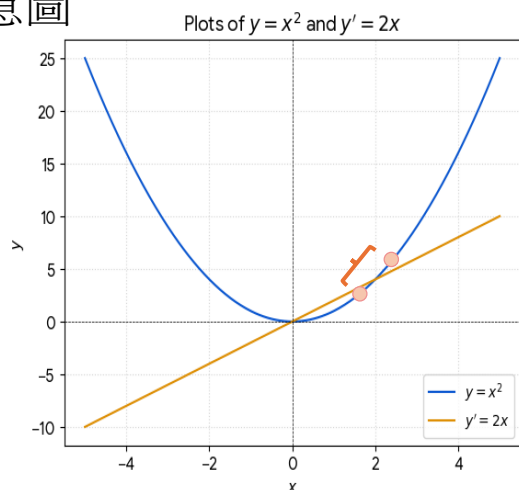
$- \begin{pmatrix} pc1 \\ pc2 \dots \end{pmatrix}$

$_{step(t)=2}$

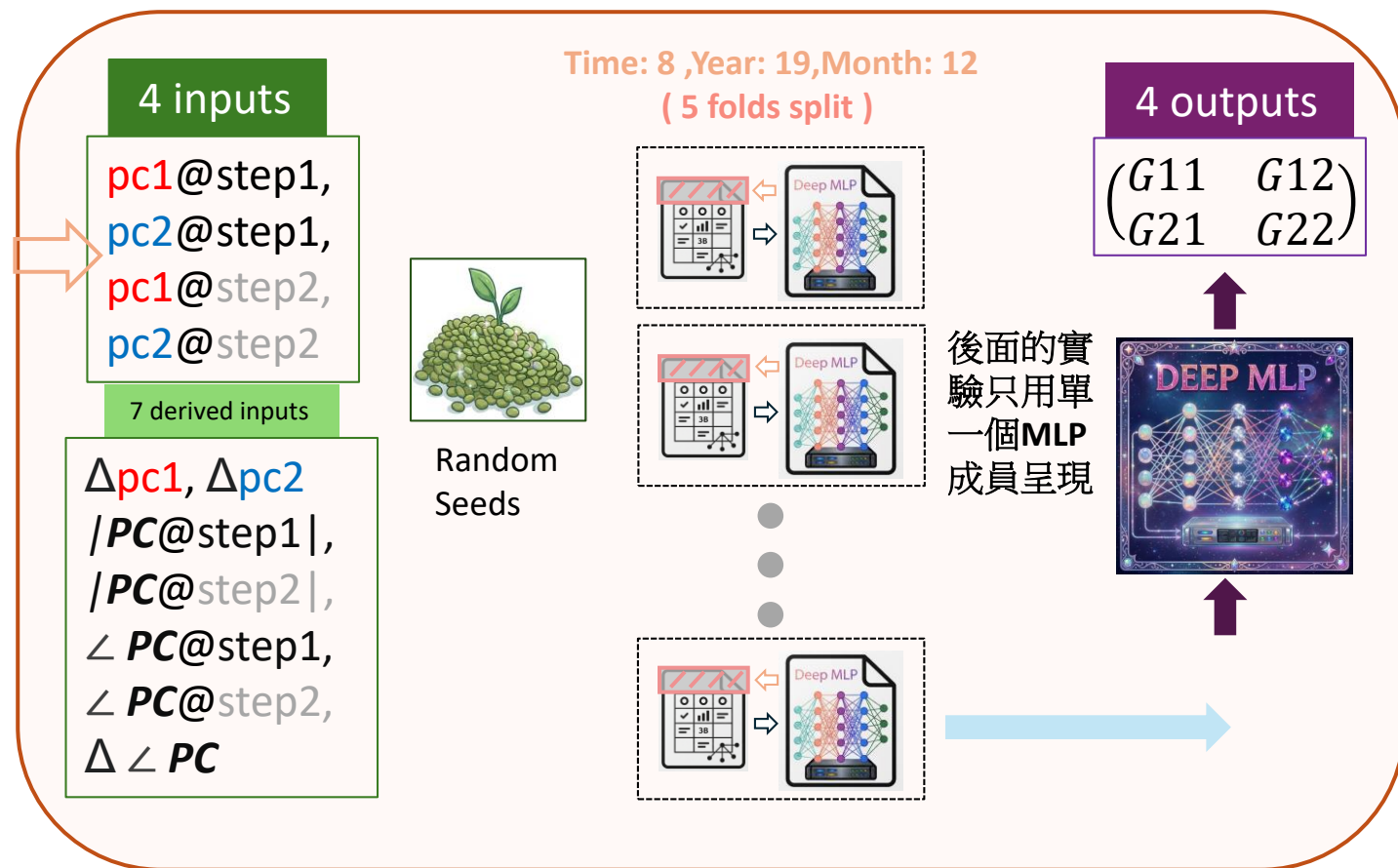
軌跡逆算子 類神經網路MLP



示意圖



都是
系集
平均



Input(11) -> Stem Layer -> Residual Blocks(4)
-> Neck Layer -> Head Layer(4)

Hidden layers : 256 layers
Lamba_phys : 0.2
Activation Function : GELU

Huber($G_{predicted}$, G_{target})

$$\mathcal{L}_{total} = \mathcal{L}_{supervised} + \lambda_{phys} \cdot \mathcal{L}_{physics}$$

$$\{ \text{Residual}(\epsilon) - (G * PC@step1 - PC@step2) \}$$

PINN

(Physics-Informed Neural Networks)
(物理資訊神經網路)

類神經網路MLP det|G|

ML model	R ² g11	R ² g12	R ² g21	R ² g22	R ² det G
Random Forest Step1	-0.123	-0.096	-0.162	-0.175	-0.065
Rank cdf Step1	-0.012	0.010	0.044	0.028	0.014
deep MLP Step1+2	0.302	0.252	0.246	0.226	0.389

$$\begin{pmatrix} \text{pc1@step2} \\ \text{pc2@step2} \end{pmatrix} = \begin{pmatrix} G_{11} & G_{12} \\ G_{21} & G_{22} \end{pmatrix} \begin{pmatrix} \text{pc1@step1 as } x \\ \text{pc2@step1 as } y \end{pmatrix}$$

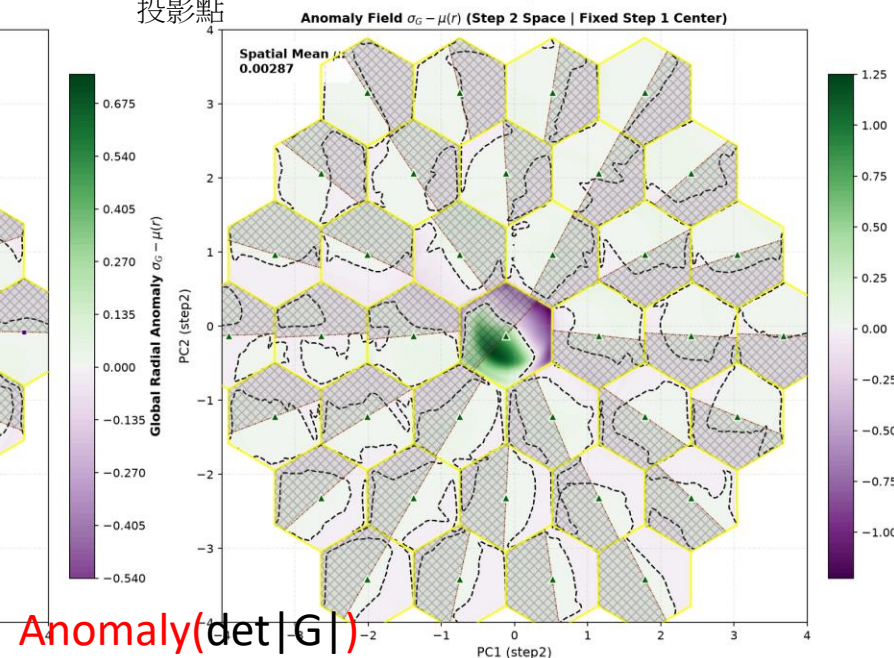
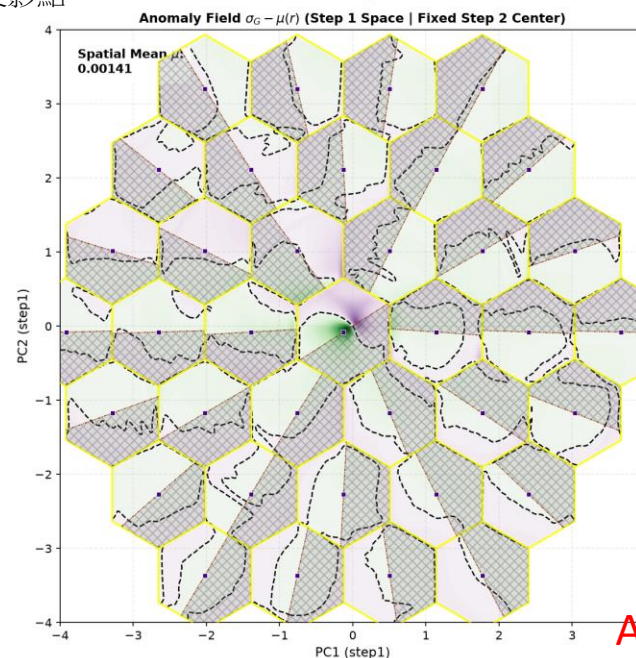
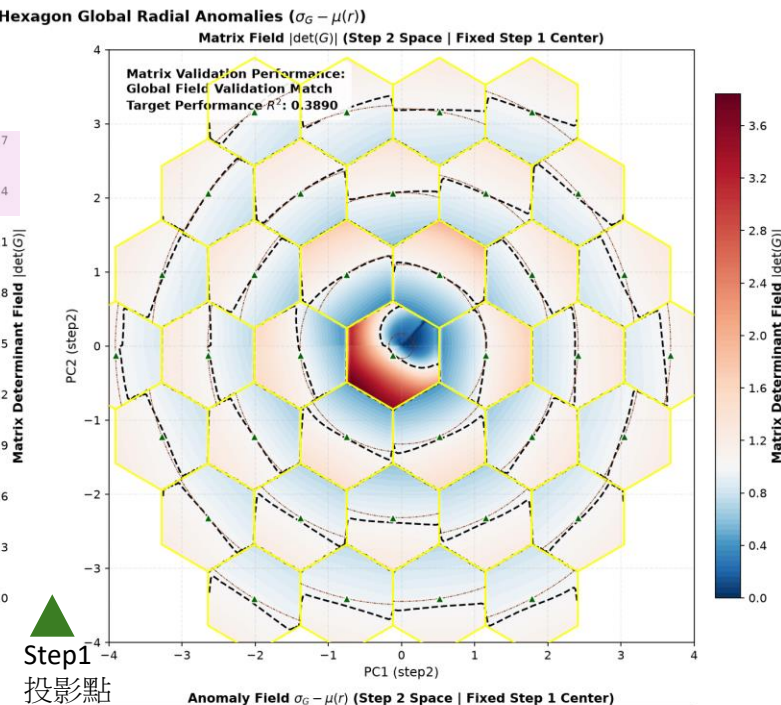
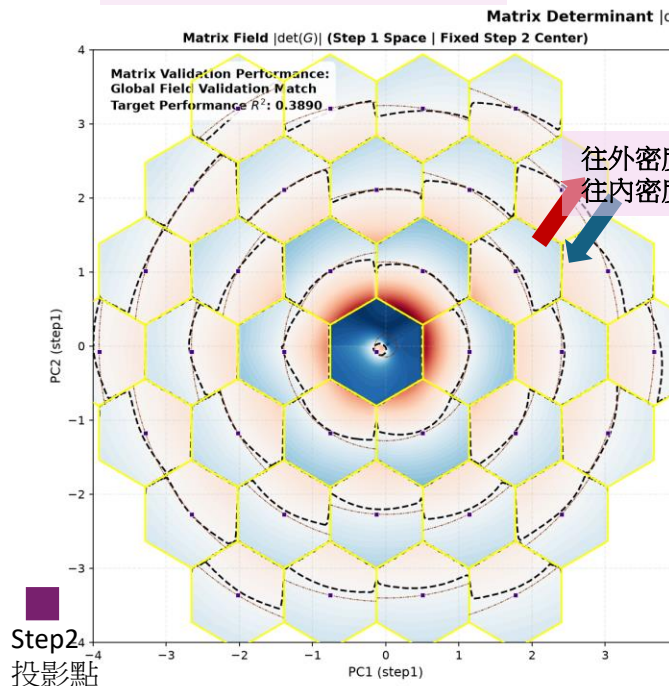
$$\text{Jacobian} = \begin{pmatrix} \frac{\partial \text{pc1@step2}}{\partial x} & \frac{\partial \text{pc1@step2}}{\partial y} \\ \frac{\partial \text{pc2@step2}}{\partial x} & \frac{\partial \text{pc2@step2}}{\partial y} \end{pmatrix}$$

$$\begin{pmatrix} G_{11} + \frac{\partial G_{11}}{\partial x}x + \frac{\partial G_{12}}{\partial x}y & G_{12} + \frac{\partial G_{11}}{\partial y}x + \frac{\partial G_{12}}{\partial y}y \\ G_{21} + \frac{\partial G_{21}}{\partial x}x + \frac{\partial G_{22}}{\partial x}y & G_{22} + \frac{\partial G_{21}}{\partial y}x + \frac{\partial G_{22}}{\partial y}y \end{pmatrix}$$

If $\nabla G=0$, $\det|G|=\det|J|$

$$\rho_{t+1} = \rho_t / \det|G|$$

$$\det|G|$$



$$\rho_{t+1} = \frac{\rho_t}{\det |G|}$$

從step0向前沿者s2s資料
PC1,PC2的軌跡線。計算每個系
集的密度

INPUT: previous density $\rho[t-1]$ & coordinates
(pc1,pc2) at t-1 and t
Output : $\det |G|$

BETA(0.5) relaxation smoothing

- $\det |G| = \det |G| - \text{BETA} \cdot (\det |G| - \text{mean}(\det |G|))$

Step transport (pure advection)

- $\rho_{\text{new}} = \rho[t-1] / \det |G|$

Spatial kernel diffusion (0.15)

- Gaussian RBF kernel over pairwise member
- $\rho_{\text{new}} = (1-D) \cdot \rho_{\text{new}} + D \cdot (\text{kernel} \cdot \rho_{\text{new}})$

Mass conservation restriction

- $\rho_{\text{final}} = \rho_{\text{new}} \cdot (100 / \text{sum}(\rho_{\text{new}}))$

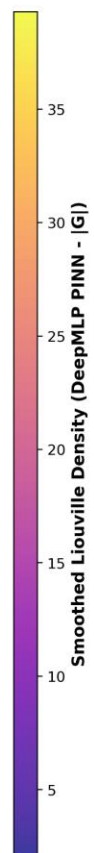
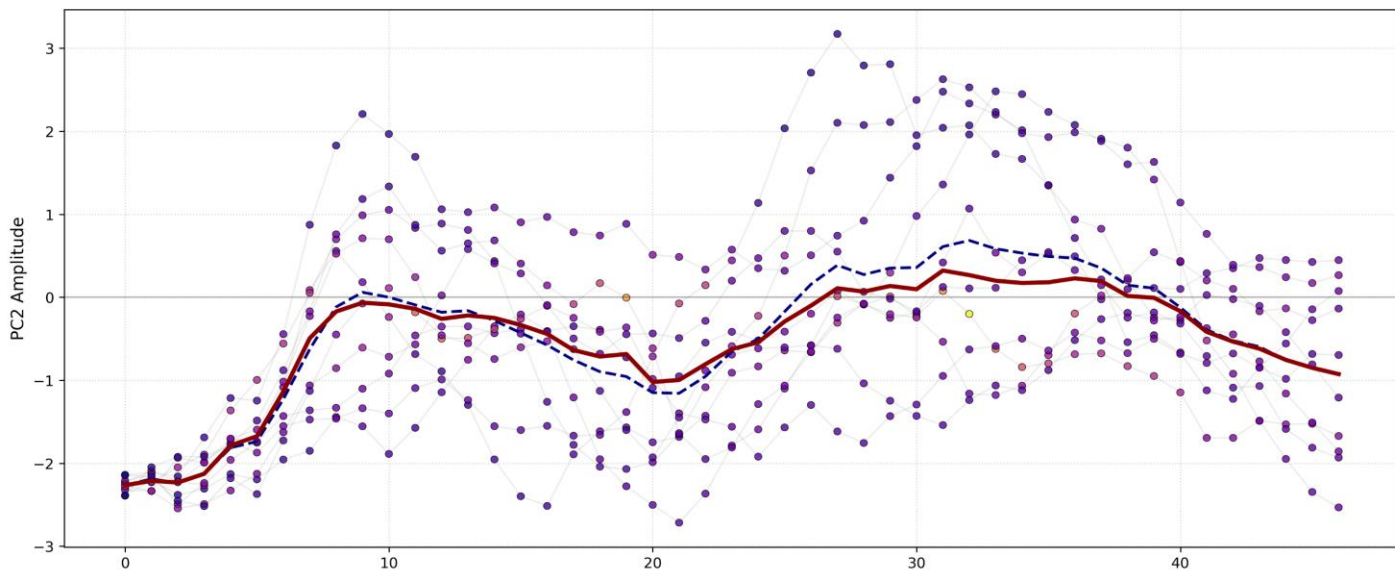
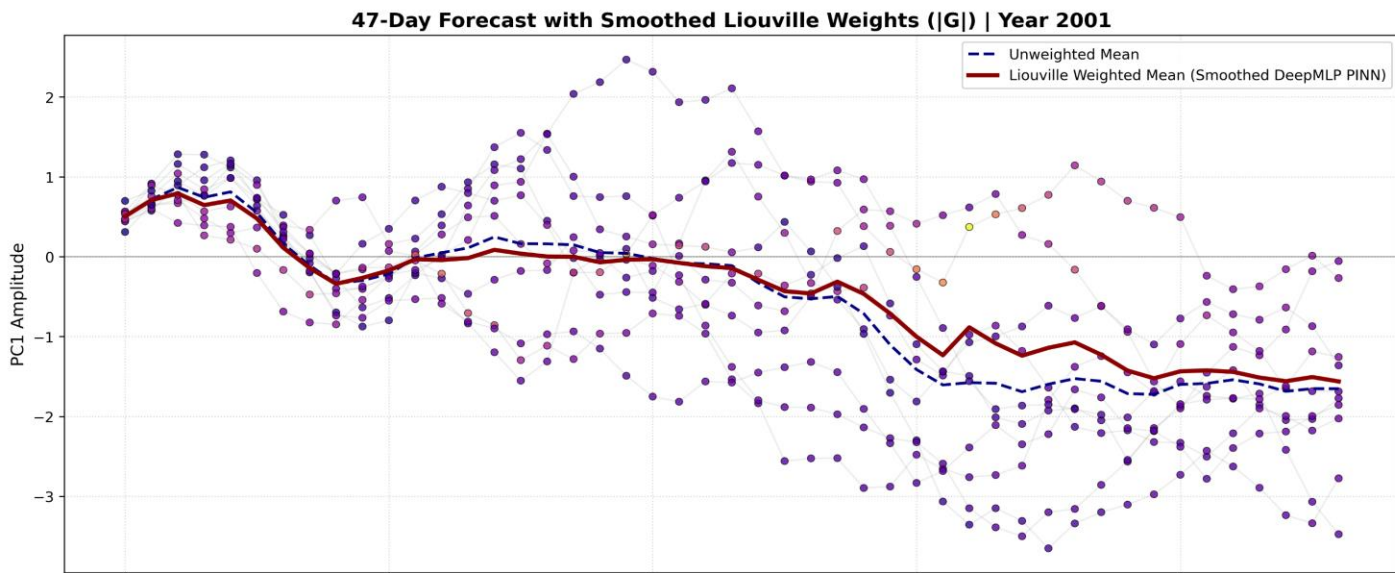
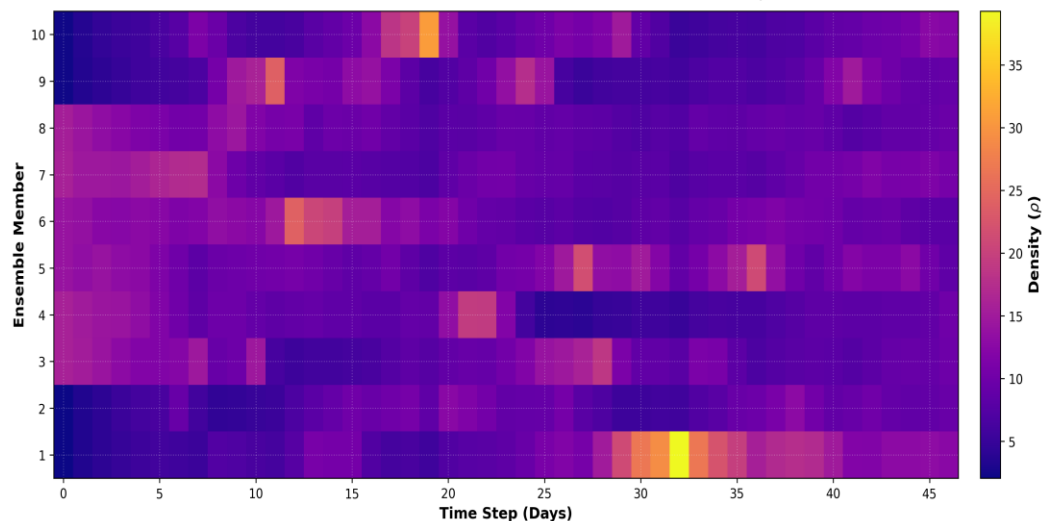


Figure 2: ρ Time Evolution Heatmap across 10 Ensemble Members | Year 2001



Deep MLP增加額外外在的變數



$$f\left\{\begin{pmatrix} pc1 \\ pc2 \end{pmatrix}_{step(t)=1,2}\right\} = \begin{pmatrix} G11 & G12 \\ G21 & G22 \end{pmatrix}$$

4 additional inputs

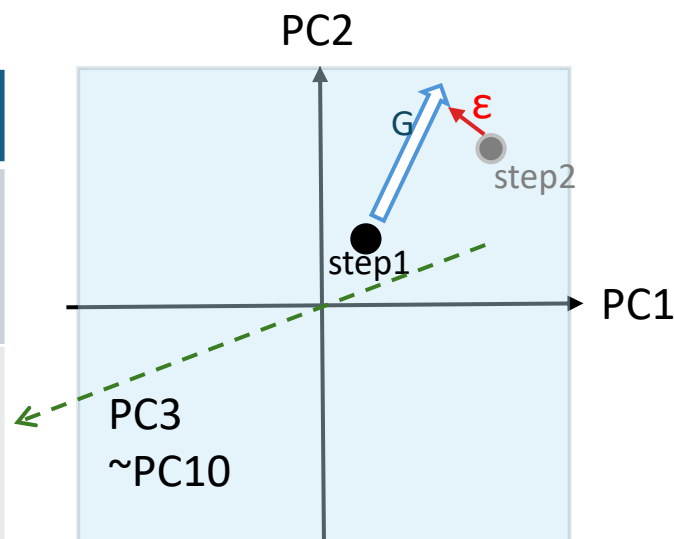
$$\begin{pmatrix} pc3 \\ pc4 \\ \vdots \\ pc10 \end{pmatrix}_{step(t)=1,2}$$

$$\frac{(PC3 \sim 10)}{(PC1 * PC2)} @ step1, 2$$

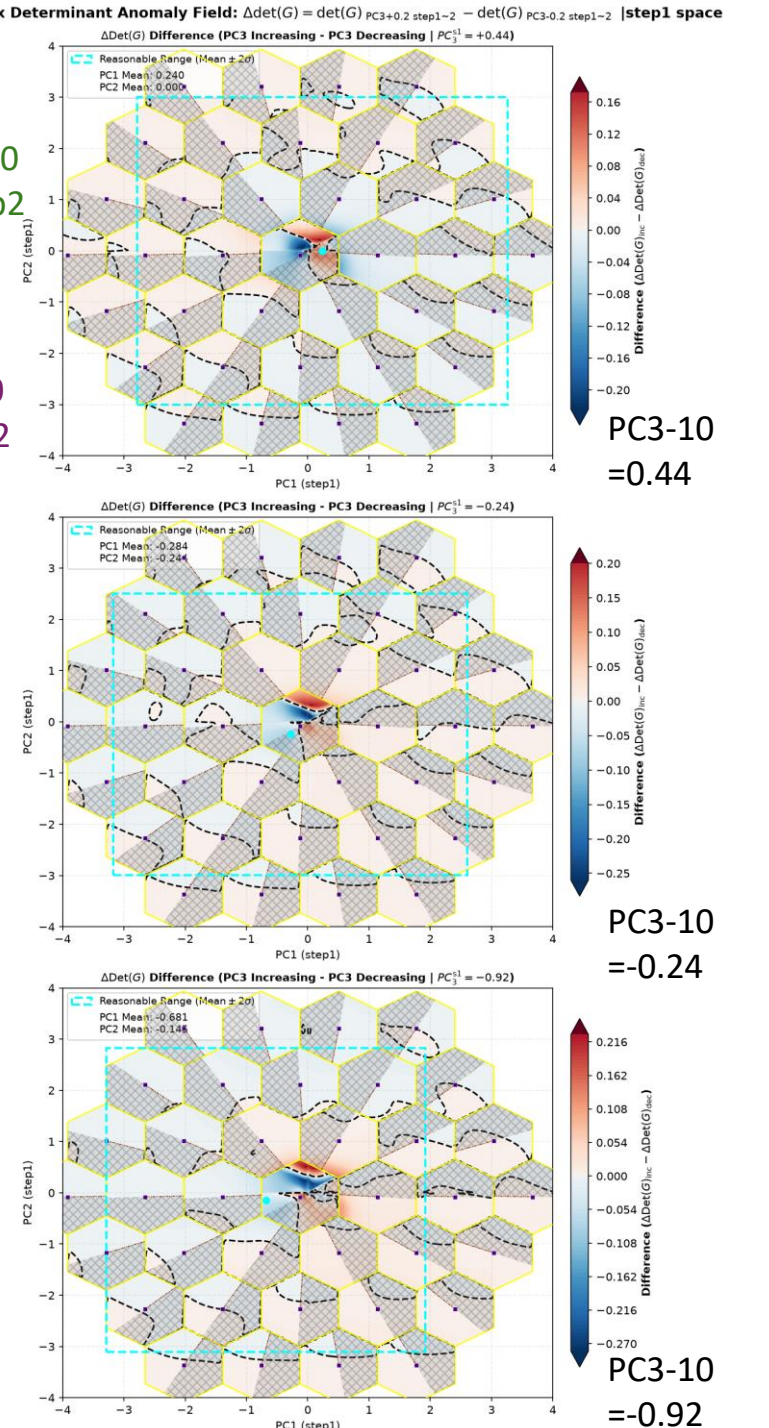
$$\left(\frac{\sqrt{(PC3 \sim 10)^2}}{\sqrt{PC1^2 + PC2^2 + (PC3 \sim 10)^2}}\right) @ step1, 2$$

交互作用項 (scaling)

ML model	R ² G11	R ² g12	R ² g21	R ² g22	R ² det G
deep MLP Step1+2	0.302	0.252	0.246	0.226	0.389
deep MLP Step1+2 +pc3~pc10	0.308	0.257	0.279	0.273	0.417



pc3-10
@step1
pc3-10
@step2

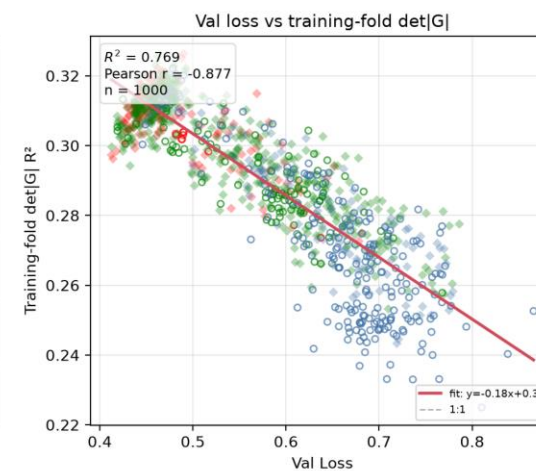
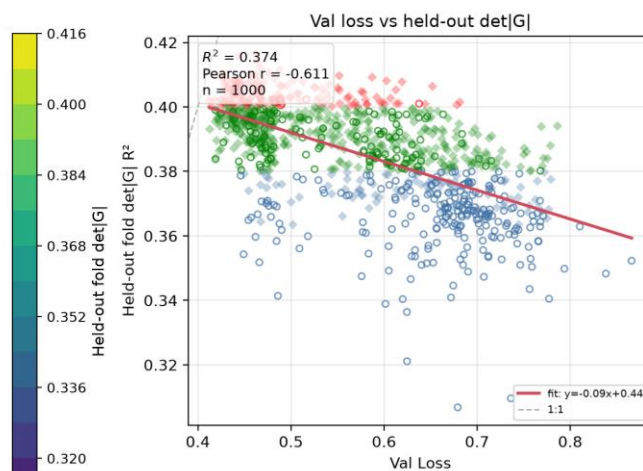
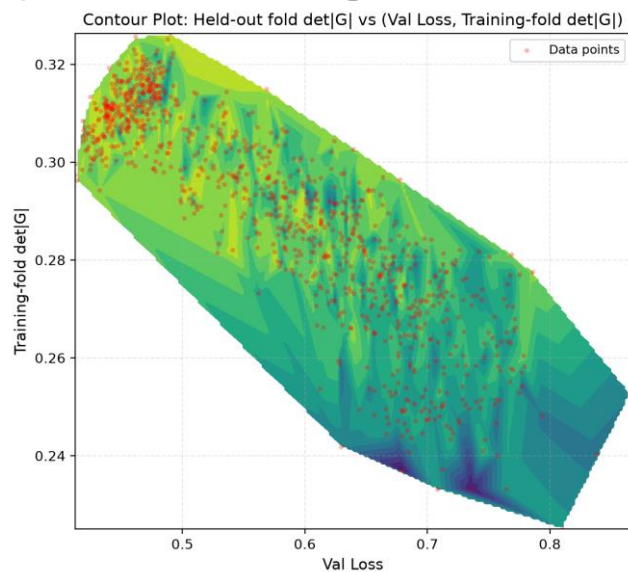


$$G(\text{pc}, y) = f(\text{pc}) + h(y)$$

where $y \notin \text{pc}$

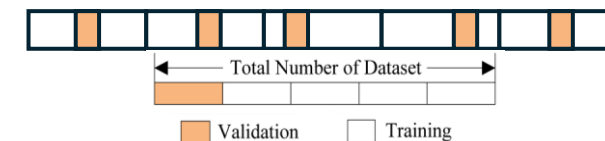
Training fold $G = \bar{G} + G'$
Held-out fold $G = \bar{G} + G'$

(1) Random shuffling

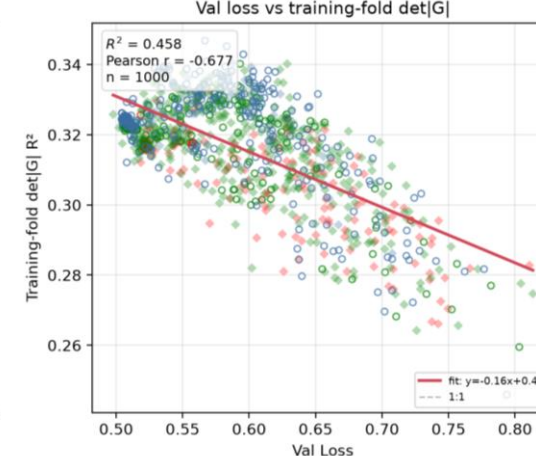
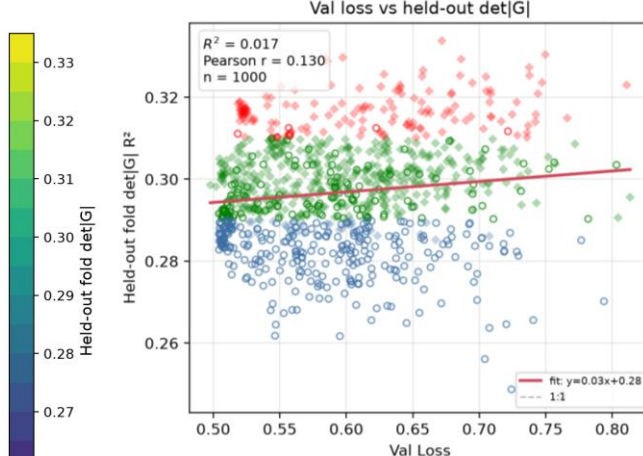
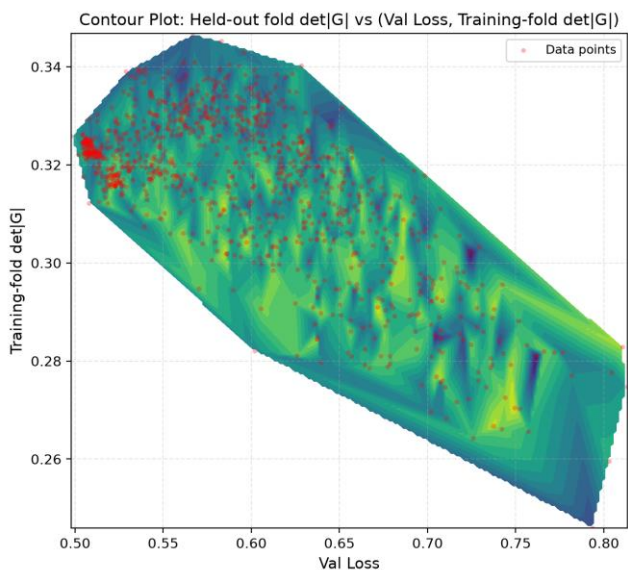


Scatter Plot Legend & Style Explanation:

- Color Scheme (Held-out det[G]):
 - Red: det[G] > 0.40
 - Green: 0.38 < det[G] ≤ 0.40
 - Blue: det[G] ≤ 0.38
- Marker Shape & Transparency:
 - Solid Diamond (alpha = 0.3): RF > 0.00 AND held-out det[G] > RF=0.00 baseline.
 - Hollow Circle: All other data samples.



(2) Chronological split



Scatter Plot Legend & Style Explanation:

- Color Scheme (Held-out det[G]):
 - Red: det[G] > 0.31
 - Green: 0.29 < det[G] ≤ 0.31
 - Blue: det[G] ≤ 0.29
- Marker Shape & Transparency:
 - Solid Diamond (alpha = 0.3): RF > 0.00 AND held-out det[G] > RF=0.00 baseline.
 - Hollow Circle: All other data samples.



結語

$$\frac{d\mathbf{X}}{dt} = A\{\mathbf{X}_{x1,x2}, \mathbf{f}(t)\} \mathbf{X} + \boldsymbol{\varepsilon}$$

$$\rho_{t+1} = \rho_t / \det |G|$$

「非自治軌跡逆算子」

(Non-autonomous Trajectory Inverse Operator)

MJO 的動力學不能簡化為靜態的線性系統，受時間影響的複雜非自治過程。

• 為什麼加入 x_1, x_2 ? (軌跡算子的優勢)

傳統的 SDLIM 只考慮 $L(x_1)$ ，也就是「我在哪裡，物理規則就是什麼」。但 MLP 模型輸入同時包含起點 (Step 1) 和終點 (Step 2)。

從「點」到「線」：神經網路不再只是判斷某個位置的性質，它是在診斷「完成這段特定位移」所需的物理橋樑 G 。

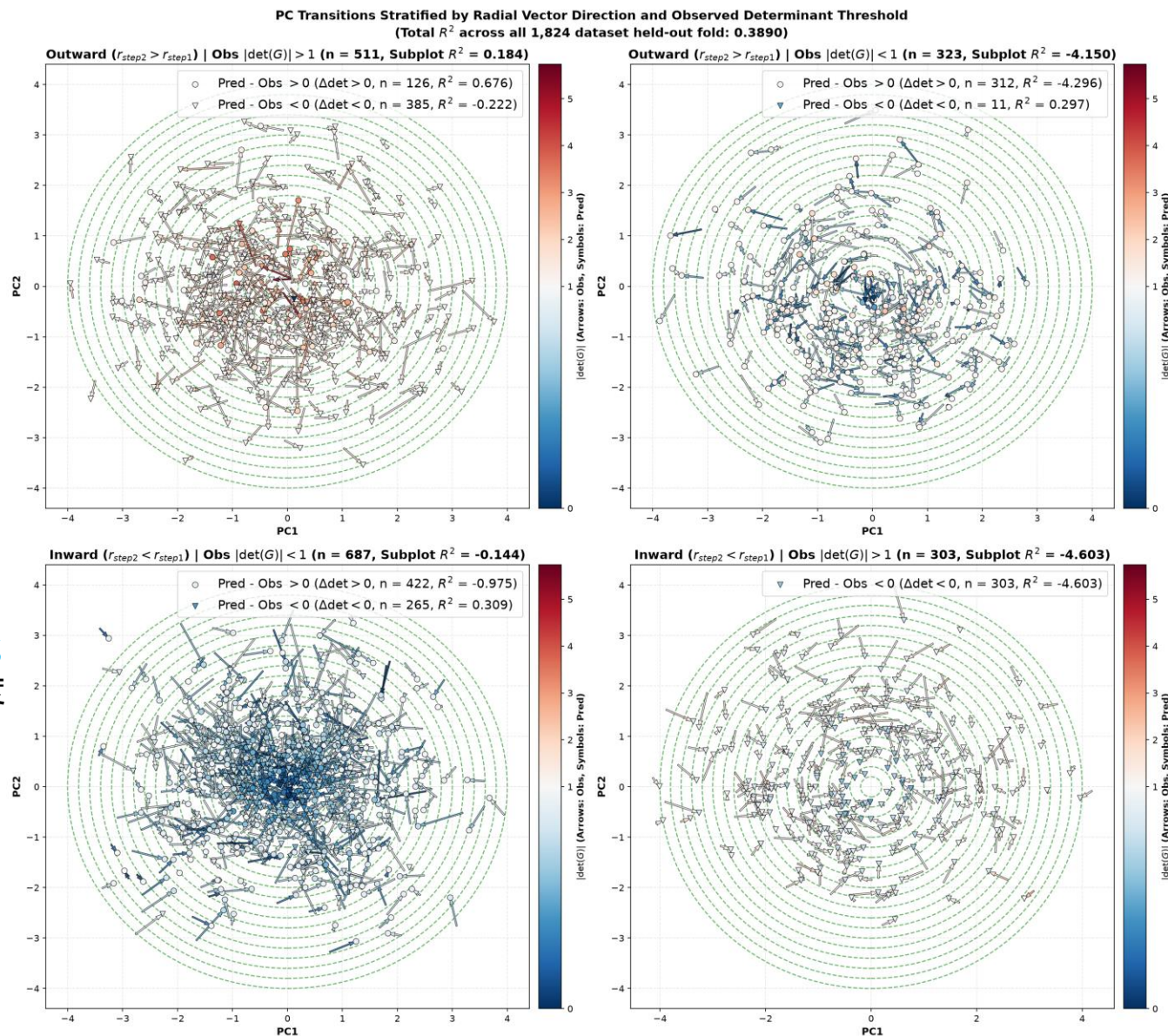
逆算子：這是一個「上帝視角」的設定。模型透過觀察 MJO 實際上是如何從 x_1 走到 x_2 的，進而反推這段路徑上的線性算子。

• 即便知道軌跡，線性矩陣 G 也只能解釋 39%：

這證明了 MJO 的演化絕非單純的「座標映射」，而是受到強大的外部時間強迫 $\mathbf{f}(t)$ 及殘差項 $\boldsymbol{\varepsilon}$ (如季節循環、背景海溫 SST、或是更小尺度的天氣擾動) 的影響，但該模型對於這兩項的掌握度可能不佳。

• 非自治性 (Non-autonomous)： $\mathbf{f}(t)$ 的存在意味著，同樣從 x_1 出發想去 x_2 ，在 1 月和 7 月所需的「物理推力」是不一樣的。這就是為什麼算子必須包含 $\mathbf{f}(t)$ 。

• 殘差項 $\boldsymbol{\varepsilon}$ (線性模型的預測誤差) 推測主要反映隨機強迫，但因目前模型未顯式納入 $\mathbf{f}(t)$ ， $\boldsymbol{\varepsilon}$ 中可能仍摻雜了 $\mathbf{f}(t)$ 相關的系統性誤差，兩者尚無法完全分離。





謝謝聆聽

Thank you for listening