A Multi-step-ahead Markov Conditional Forward Model with Cube Perturbations For Extreme Weather Forecasting

Presenter : Chia-Yuang Chang

Chia-Yuang Chang, Cheng-Wei Lu, Chuan-Ju Wang

February, 2021
AAAI 2021, Virtual

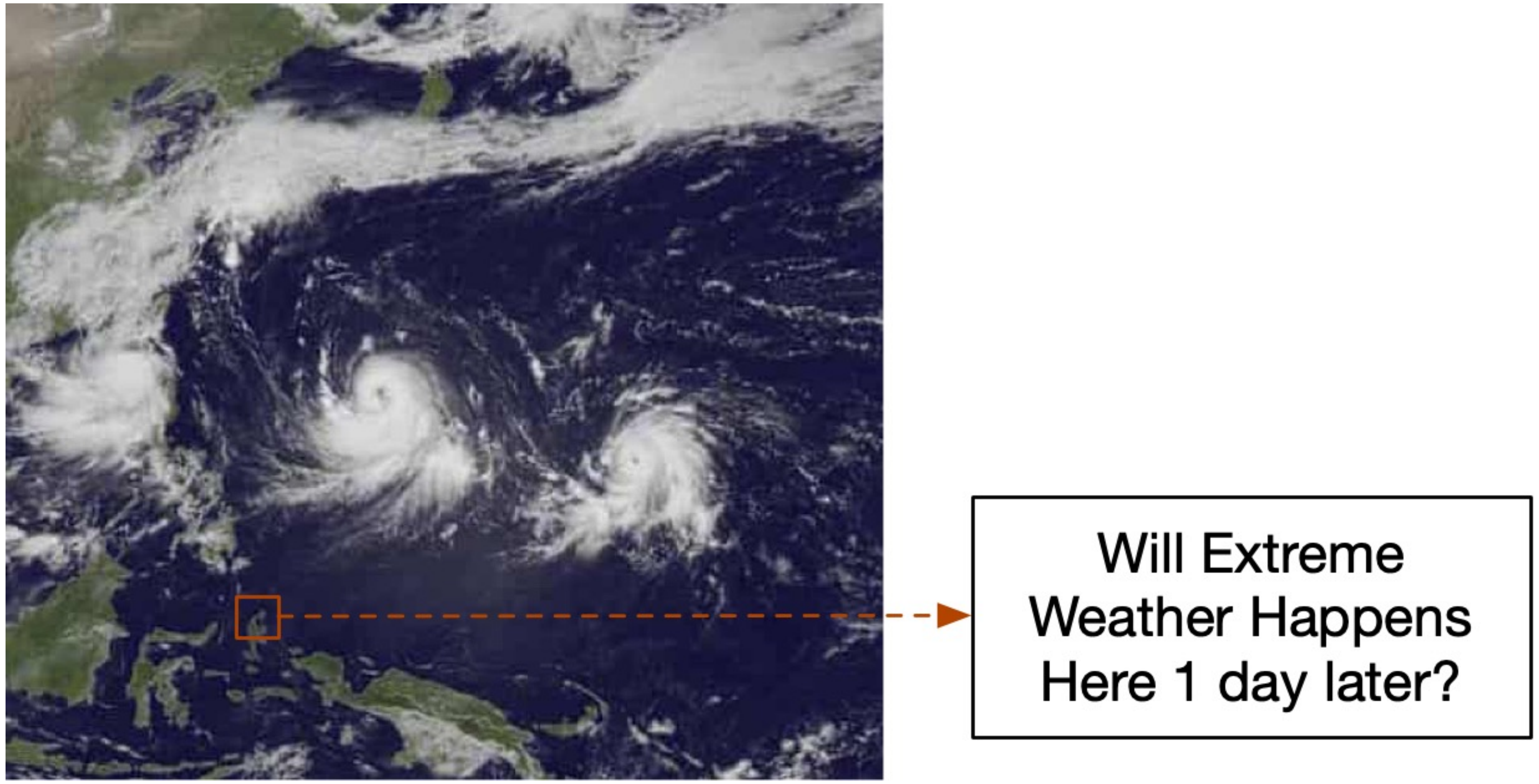
中央研究院
ACADEMIA SINICA

Agenda

- Introduction
- MSA Prediction Model
- Cube Perturbation
- Experiment Results
- Conclusion

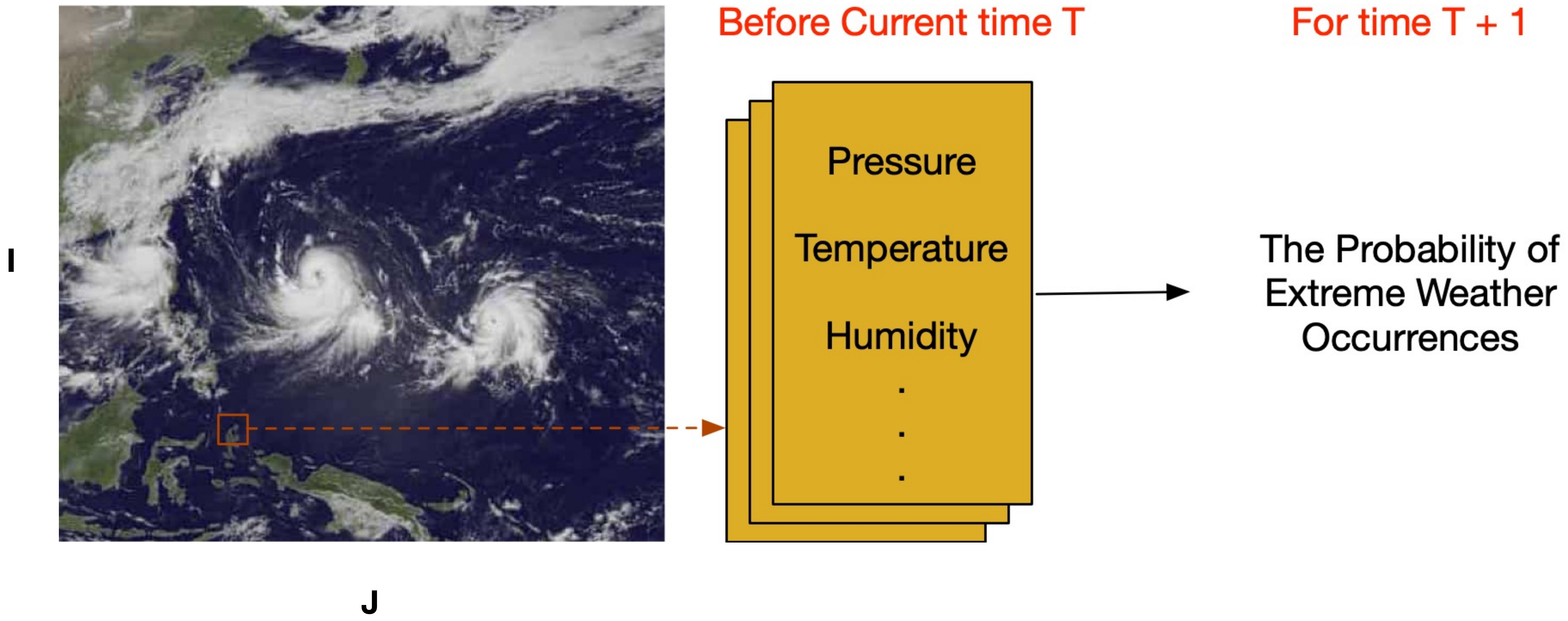
Introduction

- Goal : Predict extreme weather at future horizons (time stamps).
- Multi-step-ahead (MSA) Prediction.



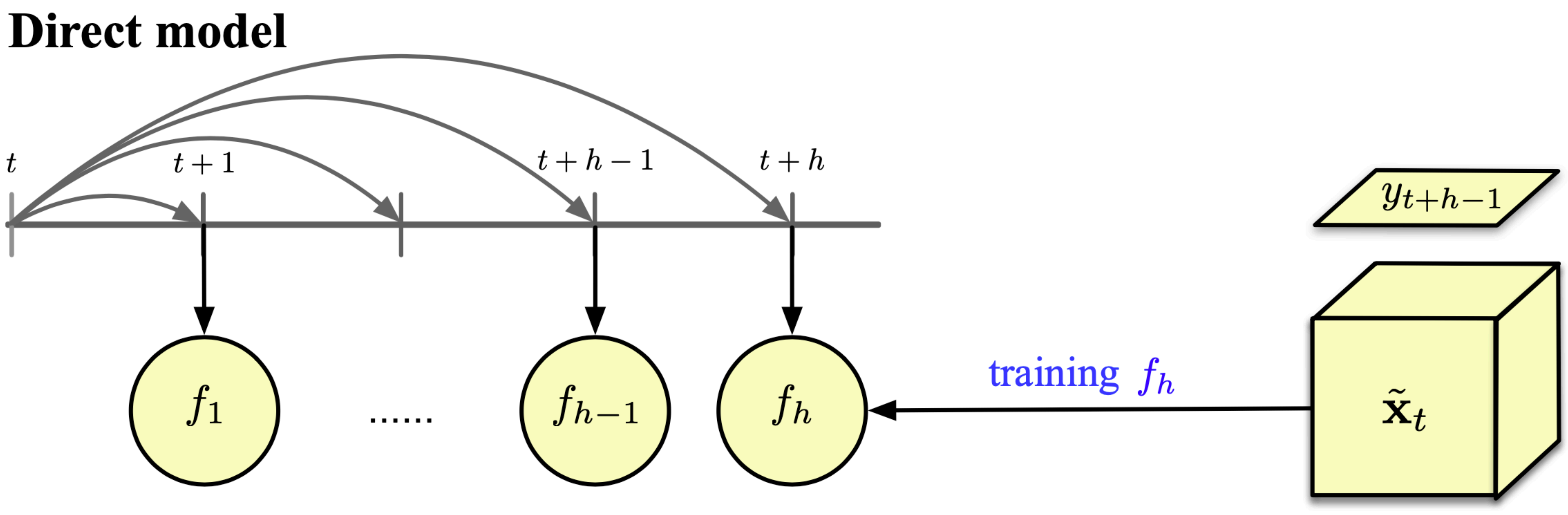
Introduction

- What we have : weather features (channels) in the past.



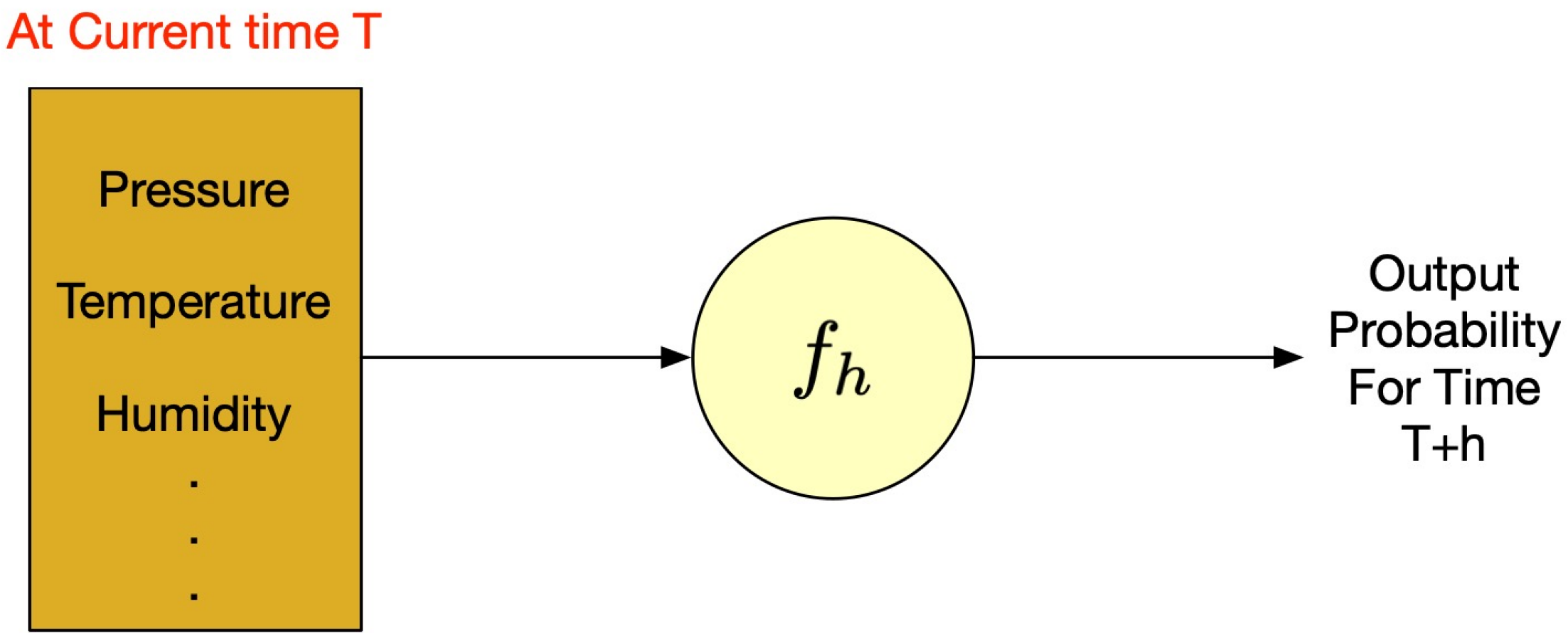
MSA Prediction Model

- Direct Model : it treats each future prediction time point independently and build a model for each of them



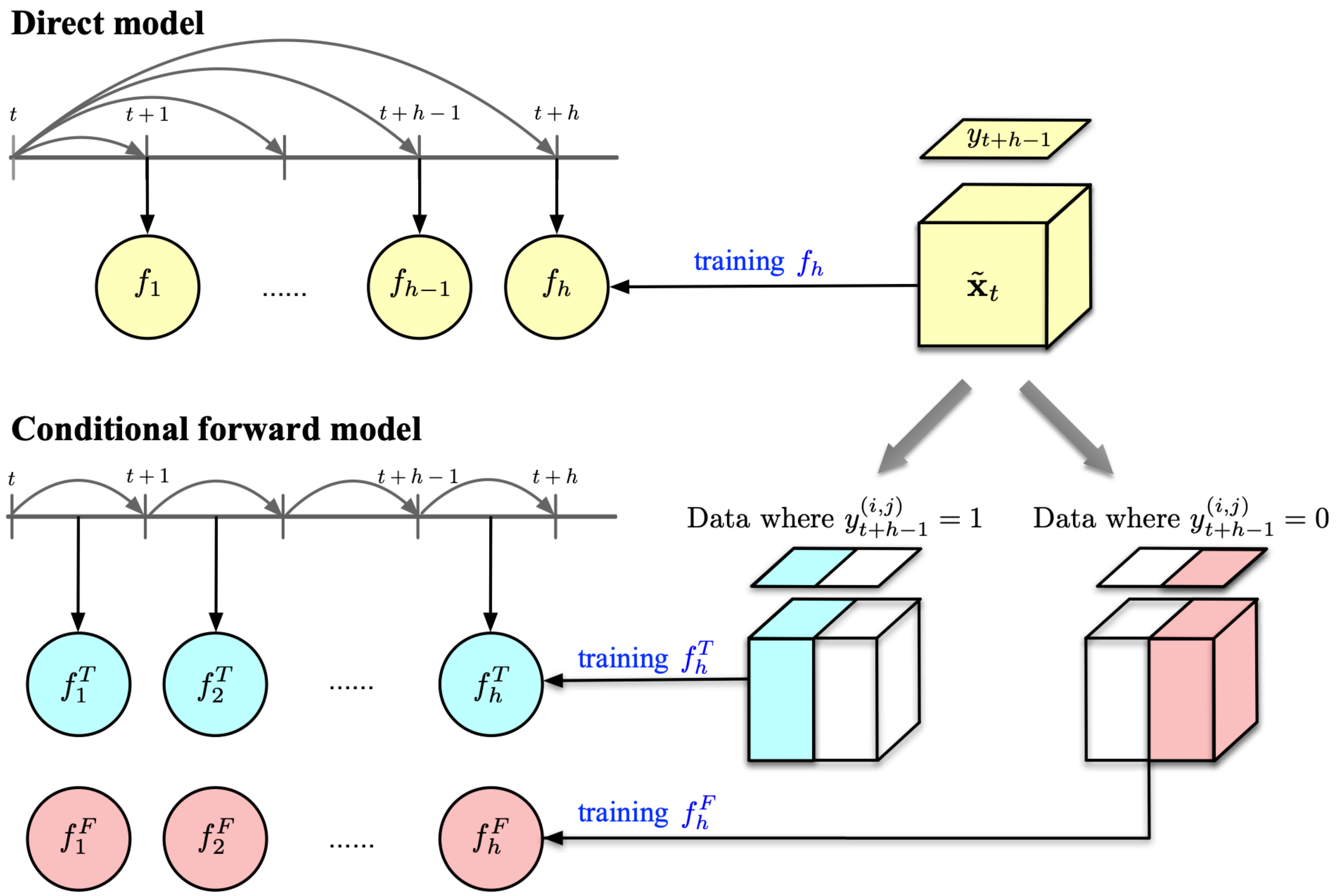
MSA Prediction Model

- Direct Model : if we want to predict the probability of extreme weather for T+h (the current time is T), we just need to do the following :



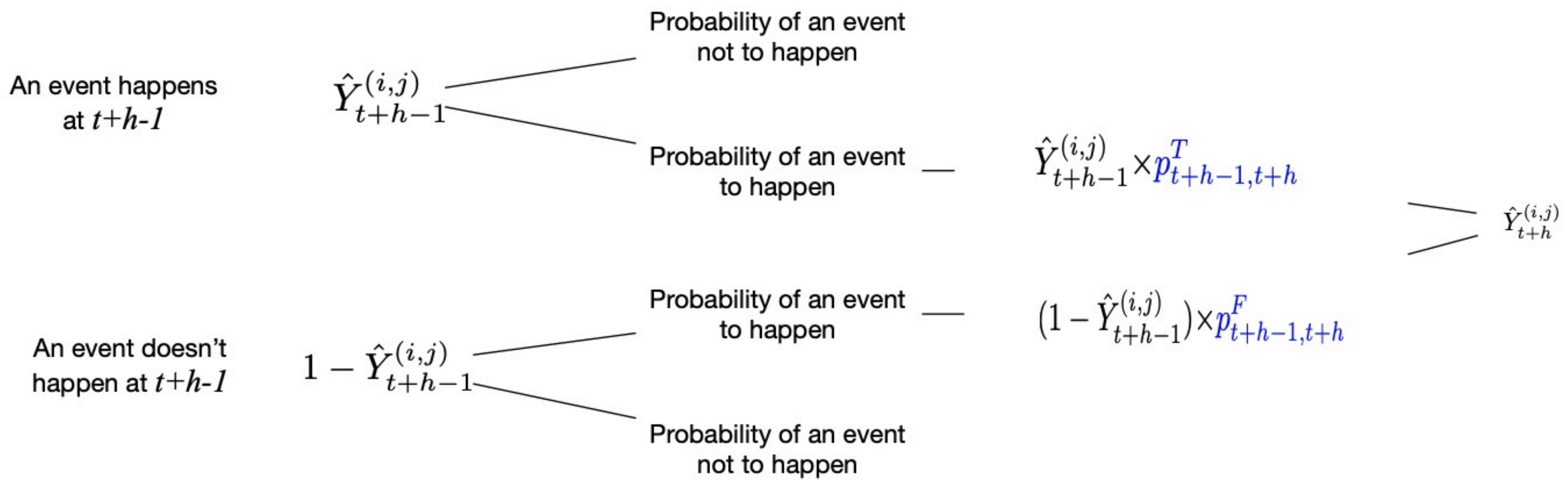
MSA Prediction Model

- Markov Conditional Forward Model (MCF): We build two models to calculate conditional probability between two consecutive prediction time and adopt Markov Probability to do probability inference.



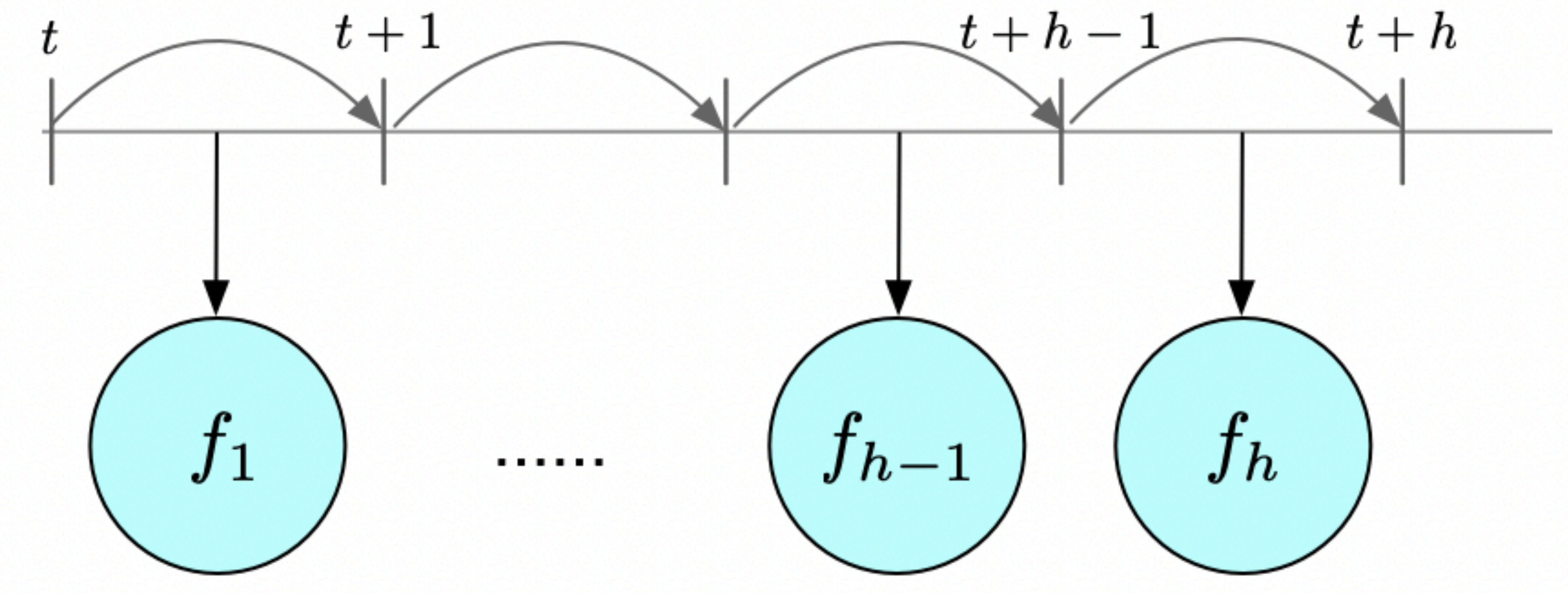
MSA Prediction Model

- Markov Conditional Forward Model (MCF) : if we want to predict the probability of extreme weather for T+h by inferencing using the probability at T+h-1, we need to do the following inference:



Cube Perturbation

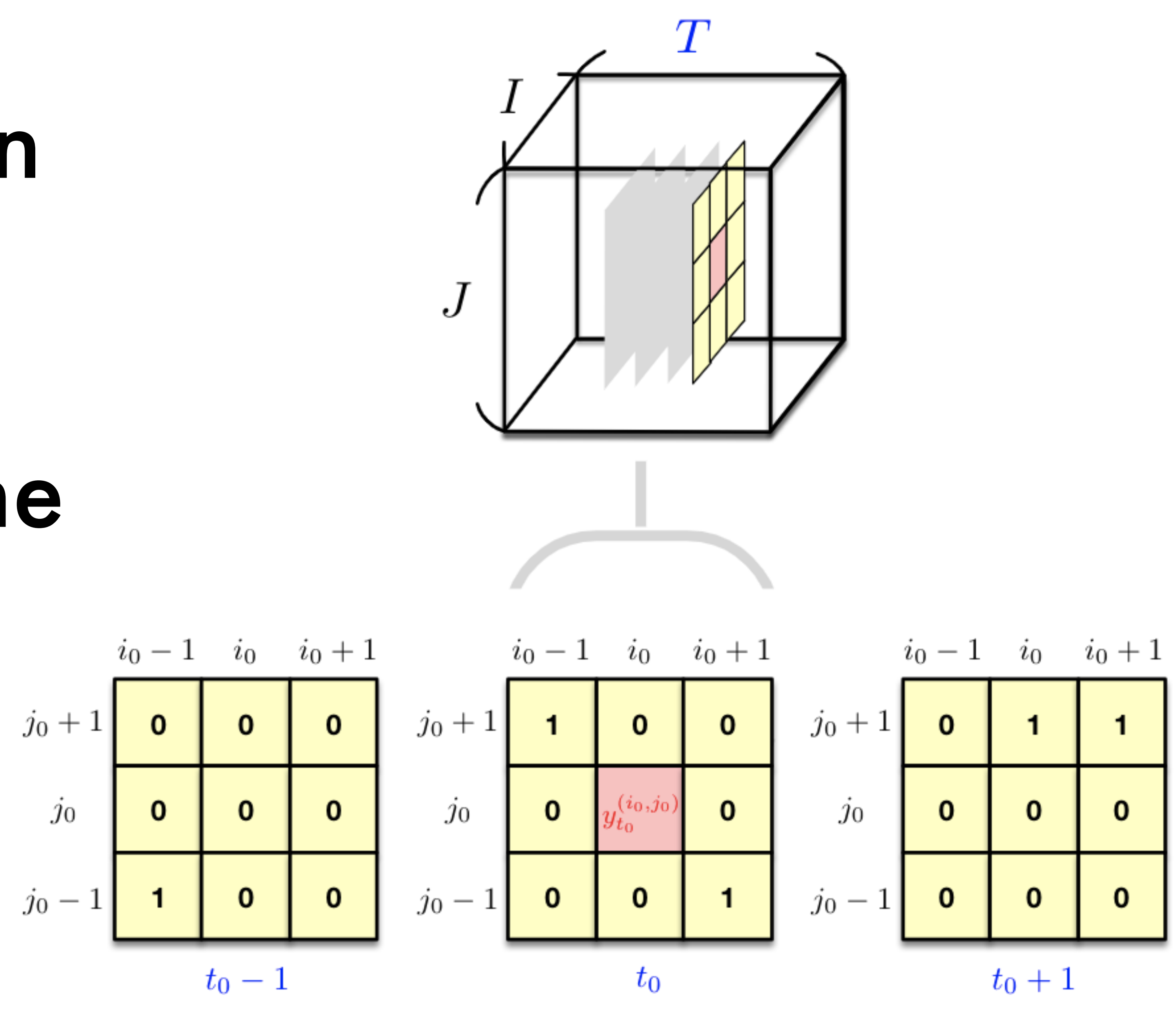
- To address error accumulation
  - MCF model accumulates error because of using predicted values to compute the next prediction.



- Naïve Perturbation
  - Randomly switch the label from 0 to 1 (or from 1 to 0) based on a predefined probability.

Cube Perturbation

- Two assumptions of cube perturbations
  - Positive data are more critical for model training than negative data because of their sparsity.
  - Spatial and temporal correlations among labels in the obtained data exist.



- The definition of "Cube" (or "Neighbor")

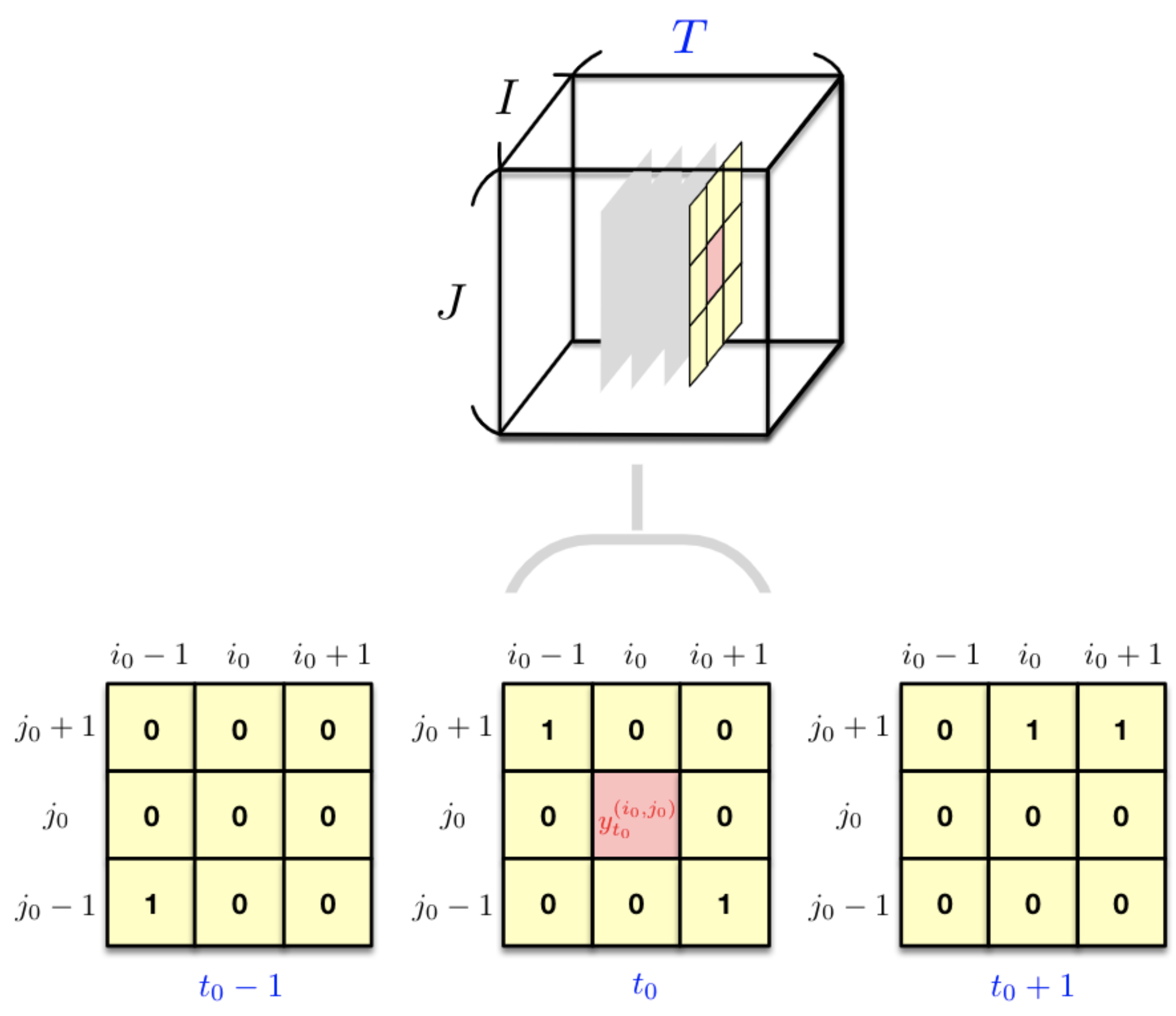
$$\mathcal{N}_{t_0}^{(i_0,j_0)} = \{(i,j,t) \mid i \in \mathcal{I} \wedge j \in \mathcal{J} \wedge t \in \mathcal{T} \wedge |i-i_0| \leq \kappa \wedge |j-j_0| \leq \kappa \wedge |t-t_0| \leq \kappa\} \setminus \{(i_0,j_0,t_0)\}$$

  - E.g., κ = 1 results in a 3 × 3 × 3 cube perturbation.

Cube Perturbation

- Simple Cube Perturbation (SCP)
  - Based on the aforementioned two assumptions:

$$y_{t_0}^{(i_0,j_0)} \neq 1 \wedge \exists (i,j,t) \in \mathcal{N}_{t_0}^{(i_0,j_0)} \left(y_t^{(i,j)} = 1 \right)$$



Cube Perturbation

- Neighborhood Probability Cube Perturbation (NPCP)
  - The perturbation probability is proportional to how many of its neighbors have positive labels

$$\xi^{\mathrm{NPCP}}_{(i_0,j_0,t_0)} = \frac{\left|\left\{(i,j,t) \,\middle|\, (i,j,t) \in \mathcal{N}^{(i_0,j_0)}_{t_0} \wedge y^{(i,j)}_t = 1\right\}\right|}{\left|N^{(i_0,j_0)}_{t_0}\right|}$$

Cube Perturbation

- Multinomial Cube Perturbation (MCP)
  - Models the probability of counts for each side of a k-sided die rolled n times.

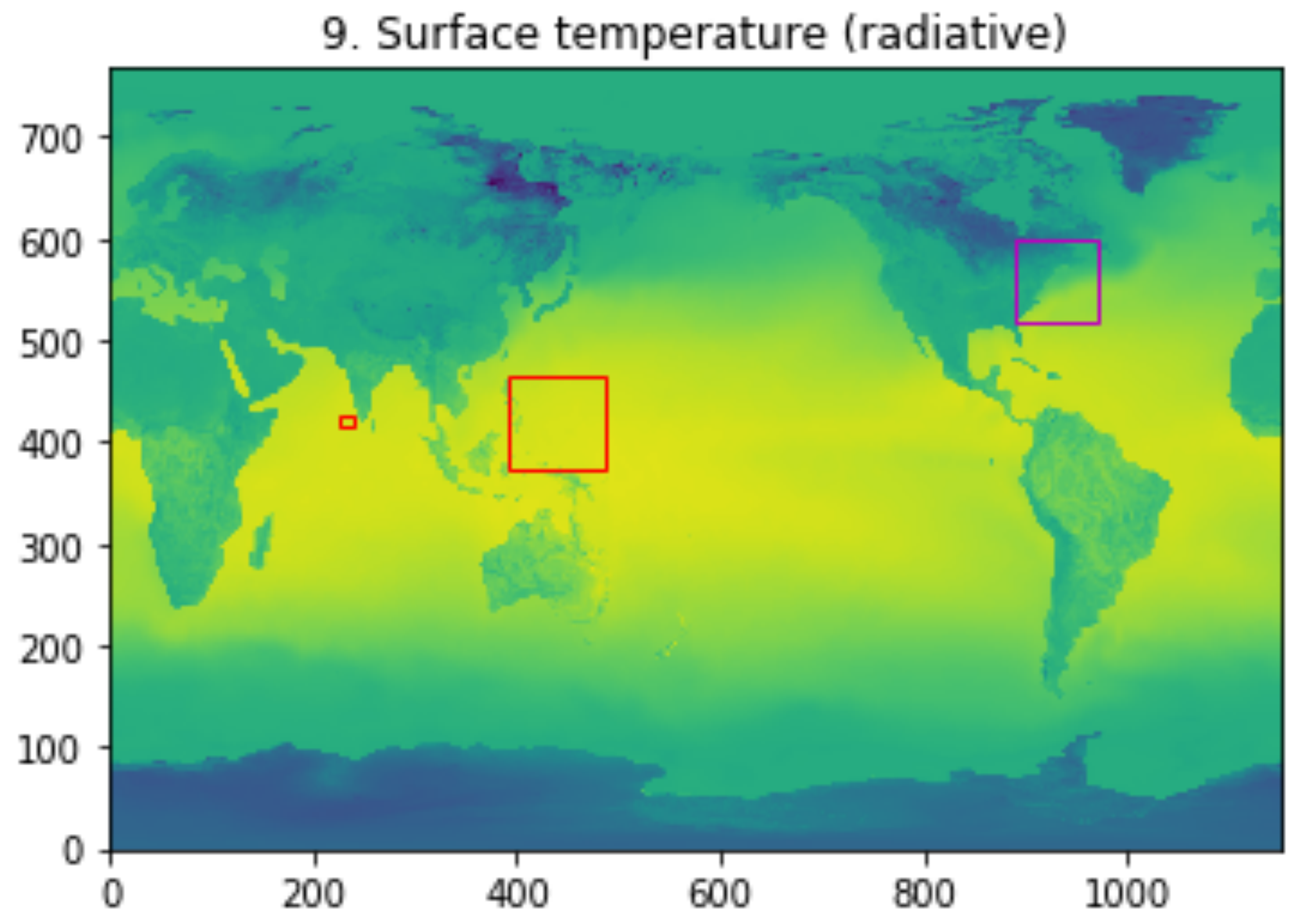
$$\xi^{\mathrm{MCP}}_{(i_0,j_0,t_0)} = \frac{\left|\left\{(i,j,t) \,\middle|\, (i,j,t) \in \mathcal{N}^{(i_0,j_0)}_{t_0} \wedge y^{(i,j)}_t = 1\right\}\right|}{\sum_{\forall \hat{i},\hat{j},\hat{t}} \left|\left\{(i,j,t) \,\middle|\, (i,j,t) \in \mathcal{N}^{(\hat{i},\hat{j})}_{\hat{t}} \wedge y^{(i,j)}_t = 1\right\}\right|}$$

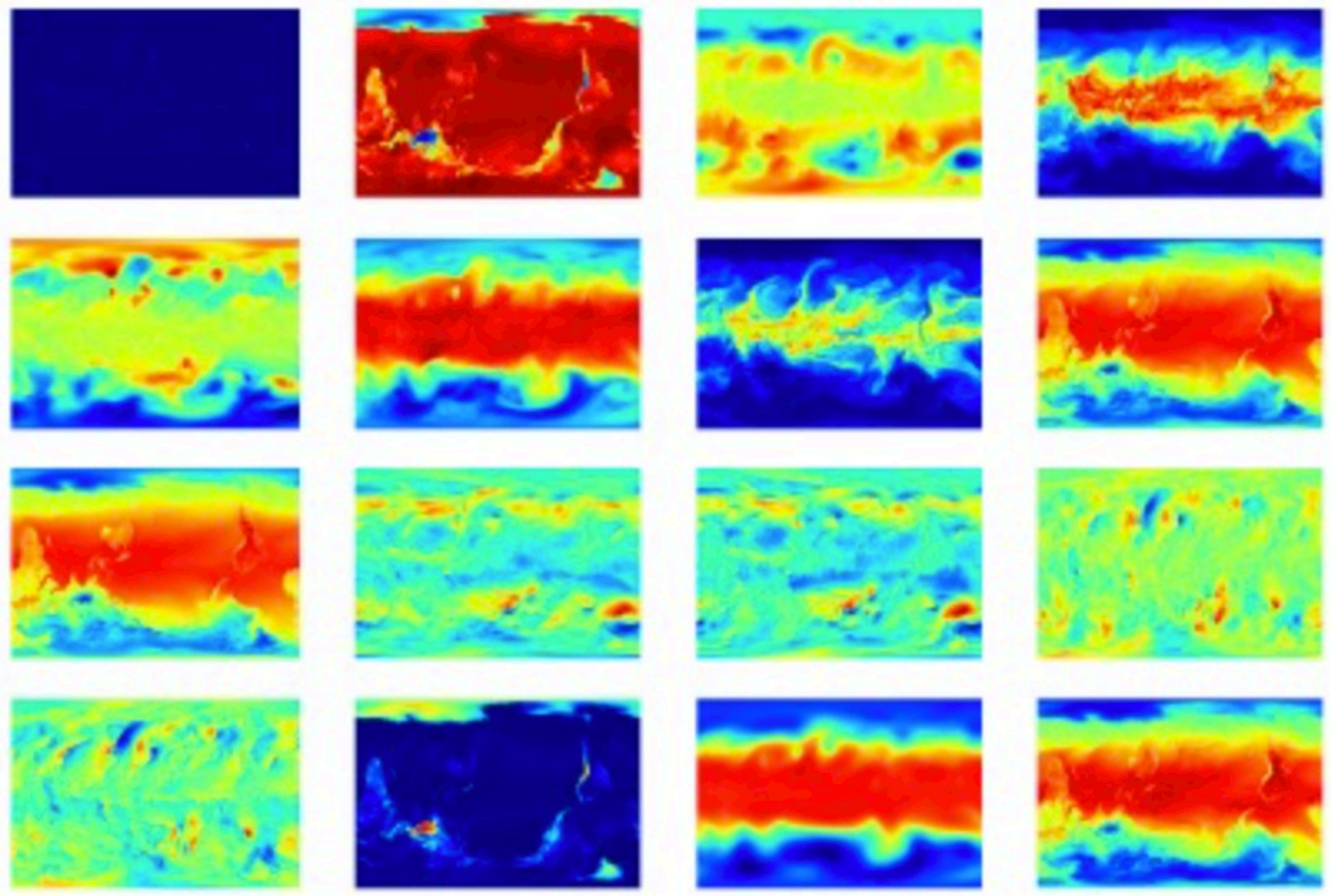
  - We then roll the k-sided die n times to obtain n data points for perturbation, where n is associated with:

$$r^{\mathrm{MCP}} = n/k \qquad \text{with } k = I \times J \times T$$

Experiment Results

- Datasets: ExtremeWeather
  - A physical variables image dataset containing extreme weather labels in each temporal resolution
    - 16 channels (e.g. pressure, temperature, humidity…)
    - Each horizon is equal to 6 hours
    - 4 labels of extreme weather

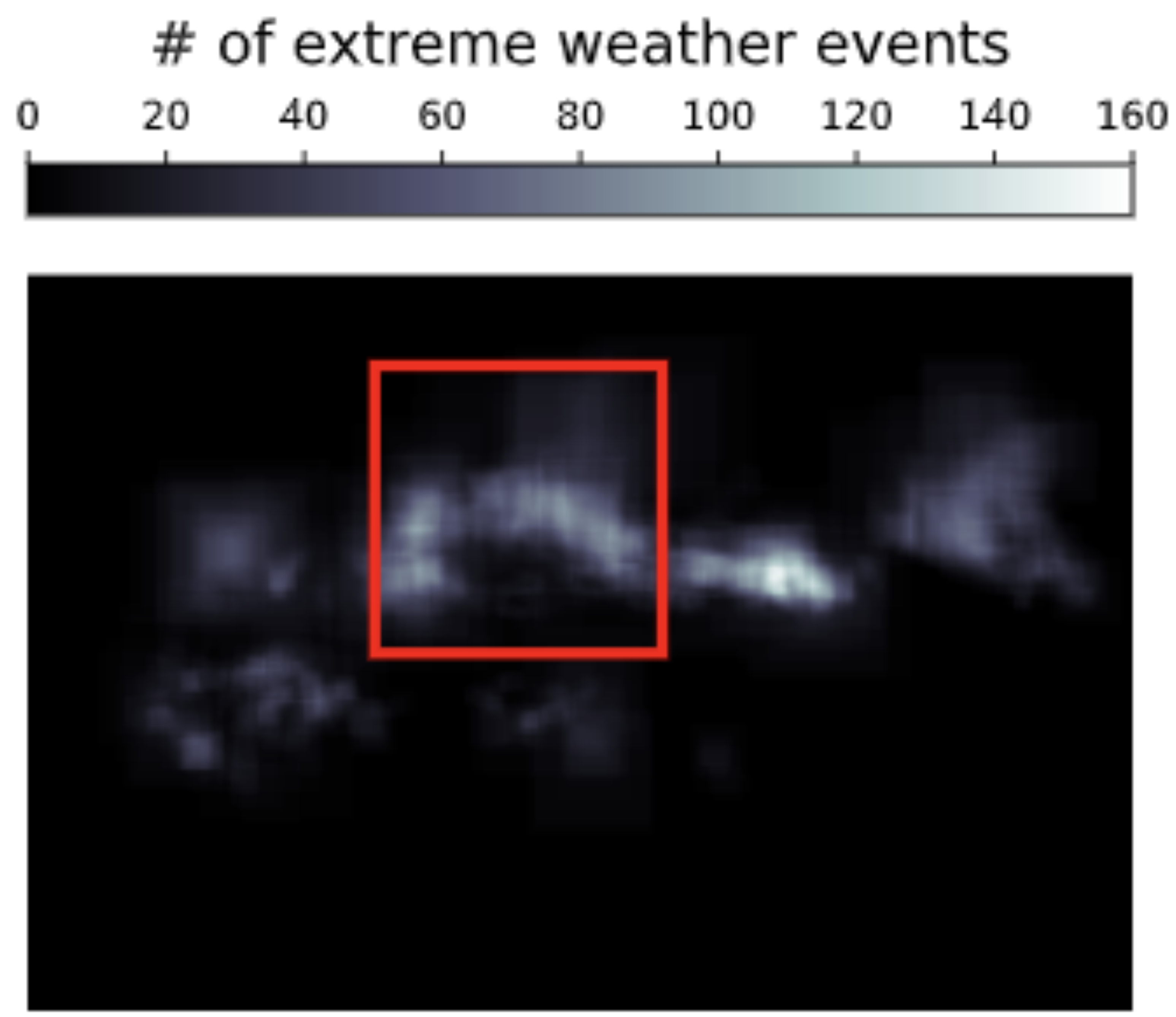




CFDA and CLIP LABS

Experiment Results

- Experiment Setting
  - Two extreme weather phenomena
    - Tropical cyclone
    - Extratropical cyclone
  - Conducted experiments in a certain area
    - 150 x 150 most active area
    - 50 - km resolution



Experiment Results

- Direct and MCF model performance (AUC)

Tropical cyclones									
Horizons (hours)	6	12	18	24	30	36	42	48	Average
LR (direct)	0.6682	0.6664	0.6641	0.6490	0.6472	0.6502	0.6576	0.6428	0.6557
LR (MCF)	0.7376	0.7213	0.7081	0.6747	0.6598	0.6634	0.6567	0.6428	0.6830
LR improvement (%)	**10.39****	**8.25****	**6.63****	**3.96****	**1.93****	**2.04****	**−0.14***	0.00	4.16
CNN (direct)	0.6760	0.6760	0.6773	0.6629	0.6577	0.6601	0.6674	0.6518	0.6662
CNN (MCF)	0.7441	0.7309	0.7188	0.6864	0.6739	0.6756	0.6689	0.6506	0.6937
CNN improvement (%)	**10.07****	**8.12****	**6.13****	**3.55****	**2.46****	**2.35****	0.22	−0.18	4.13

Extratropical cyclones									
Horizons (hours)	6	12	18	24	30	36	42	48	Average
LR (direct)	0.8243	0.8113	0.7976	0.7831	0.7628	0.7429	0.7303	0.7242	0.7721
LR (MCF)	0.9361	0.8858	0.8426	0.8104	0.7821	0.7564	0.7376	0.7275	0.8098
LR improvement (%)	**13.56****	**9.19****	**5.64****	**3.48****	**2.53****	**1.81****	**1.00****	**0.46****	4.88
CNN (direct)	0.8244	0.8119	0.7979	0.7843	0.7734	0.7614	0.7489	0.7396	0.7802
CNN (MCF)	0.9360	0.8851	0.8421	0.8103	0.7852	0.7651	0.7532	0.7413	0.8148
CNN improvement (%)	**13.54****	**9.02****	**5.54****	**3.32****	**1.53****	**0.49****	**0.57****	**0.23****	4.43

Experiment Results

- Direct and MCF model performance (AUC)

Tropical cyclones									
Horizons (hours)	6	12	18	24	30	36	42	48	Average
LR (direct)	0.6682	0.6664	0.6641	0.6490	0.6472	0.6502	0.6576	0.6428	0.6557
LR (MCF)	0.7376	0.7213	0.7081	0.6747	0.6598	0.6634	0.6567	0.6428	0.6830
LR improvement (%)	**10.39****	**8.25****	**6.63****	**3.96****	**1.93****	**2.04****	**−0.14***	0.00	4.16
CNN (direct)	0.6760	0.6760	0.6773	0.6629	0.6577	0.6601	0.6674	0.6518	0.6662
CNN (MCF)	0.7441	0.7309	0.7188	0.6864	0.6739	0.6756	0.6689	0.6506	0.6937
CNN improvement (%)	**10.07****	**8.12****	**6.13****	**3.55****	**2.46****	**2.35****	0.22	−0.18	4.13
Extratropical cyclones									
Horizons (hours)	6	12	18	24	30	36	42	48	Average
LR (direct)	0.8243	0.8113	0.7976	0.7831	0.7628	0.7429	0.7303	0.7242	0.7721
LR (MCF)	0.9361	0.8858	0.8426	0.8104	0.7821	0.7564	0.7376	0.7275	0.8098
LR improvement (%)	**13.56****	**9.19****	**5.64****	**3.48****	**2.53****	**1.81****	**1.00****	**0.46****	4.88
CNN (direct)	0.8244	0.8119	0.7979	0.7843	0.7734	0.7614	0.7489	0.7396	0.7802
CNN (MCF)	0.9360	0.8851	0.8421	0.8103	0.7852	0.7651	0.7532	0.7413	0.8148
CNN improvement (%)	**13.54****	**9.02****	**5.54****	**3.32****	**1.53****	**0.49****	**0.57****	**0.23****	4.43

Experiment Results

- Performance of various perturbation methods (AUC)

Tropical cyclones									
Horizons (hours)	6	12	18	24	30	36	42	48	Average
CNN (MCF)	0.7441	0.7309	0.7188	0.6864	0.6739	0.6756	0.6689	0.6506	0.6937
+Naive ($\xi^{\text{Naive}} = 5\%$)	0.7405	0.7144	0.6898	0.6614	0.6551	0.6555	0.6602	0.6438	0.6776
+Multinomial cube ($r^{\text{MCP}} = 10\%$)	0.7395	0.7266	0.7182	0.6863	0.6749	0.6791	0.6678	0.6481	0.6926
+Simple Cube ($\xi^{\text{SCP}} = 47.5\%$)	**0.7458**	**0.7317**	0.7266	0.6999	0.6908	0.7045	0.6820	0.6633	0.7056
+Neighborhood prob cube	0.7445	0.7308	**0.7272**	**0.7002**	**0.6922**	**0.7077**	**0.6846**	**0.6668**	**0.7068**
Improvement (%)	0.23	0.11	**1.16***	**2.01***	**2.71****	**4.76****	**2.35***	**2.48***	1.89

Extratropical cyclones									
Horizons (hours)	6	12	18	24	30	36	42	48	Average
CNN (MCF)	0.9360	0.8851	0.8421	0.8103	0.7852	0.7651	0.7532	0.7413	0.8148
+Naive ($\xi^{\text{Naive}} = 2.5\%$)	**0.9371**	0.8808	0.8288	0.7957	0.7716	0.7566	0.7470	0.7385	0.8070
+Multinomial cube ($r^{\text{MCP}} = 5\%$)	0.9365	0.8861	0.8425	0.8105	**0.7857**	0.7661	0.7545	0.7435	0.8157
+Simple Cube ($\xi^{\text{SCP}} = 47.5\%$)	0.9356	0.8855	0.8428	0.8101	0.7845	**0.7663**	**0.7558**	**0.7442**	0.8156
+Neighborhood prob cube	0.9368	**0.8873**	**0.8444**	**0.8112**	0.7849	0.7650	0.7538	0.7428	**0.8158**
Improvement (%)	0.12	0.24	0.27	0.12	0.06	0.16	0.35	0.39	0.12

Experiment Results

- Performance of various perturbation methods (AUC)

Tropical cyclones									
Horizons (hours)	6	12	18	24	30	36	42	48	Average
CNN (MCF)	0.7441	0.7309	0.7188	0.6864	0.6739	0.6756	0.6689	0.6506	0.6937
+Naive ($\xi^{\text{Naive}} = 5\%$)	0.7405	0.7144	0.6898	0.6614	0.6551	0.6555	0.6602	0.6438	0.6776
+Multinomial cube ($r^{\text{MCP}} = 10\%$)	0.7395	0.7266	0.7182	0.6863	0.6749	0.6791	0.6678	0.6481	0.6926
+Simple Cube ($\xi^{\text{SCP}} = 47.5\%$)	**0.7458**	**0.7317**	0.7266	0.6999	0.6908	0.7045	0.6820	0.6633	0.7056
+Neighborhood prob cube	0.7445	0.7308	**0.7272**	**0.7002**	**0.6922**	**0.7077**	**0.6846**	**0.6668**	**0.7068**
Improvement (%)	0.23	0.11	**1.16***	**2.01***	**2.71****	**4.76****	**2.35***	**2.48***	1.89

Extratropical cyclones									
Horizons (hours)	6	12	18	24	30	36	42	48	Average
CNN (MCF)	0.9360	0.8851	0.8421	0.8103	0.7852	0.7651	0.7532	0.7413	0.8148
+Naive ($\xi^{\text{Naive}} = 2.5\%$)	**0.9371**	0.8808	0.8288	0.7957	0.7716	0.7566	0.7470	0.7385	0.8070
+Multinomial cube ($r^{\text{MCP}} = 5\%$)	0.9365	0.8861	0.8425	0.8105	**0.7857**	0.7661	0.7545	0.7435	0.8157
+Simple Cube ($\xi^{\text{SCP}} = 47.5\%$)	0.9356	0.8855	0.8428	0.8101	0.7845	**0.7663**	**0.7558**	**0.7442**	0.8156
+Neighborhood prob cube	0.9368	**0.8873**	**0.8444**	**0.8112**	0.7849	0.7650	0.7538	0.7428	**0.8158**
Improvement (%)	0.12	0.24	0.27	0.12	0.06	0.16	0.35	0.39	0.12

Conclusion

- MCF model provides better performance than traditional direct model on MSA extreme weather prediction.
- MCF model learned with the use of CNN yields prominent results for both short-term and long-term predictions.
- Cube perturbation methods successfully enhance the fault tolerance of the MCF model by addressing error accumulation.

Thanks For Your Listening! Any Question?