Combining spatio-temporal neural networks with mechanistic interpretability to investigate teleconnections in S2S forecasts

Philine Lou Bommer*

Prof. Marlene Kretschmer

Kirill Bykov

Fiona Spurler

Prof. Marina M.-C. Hoehne

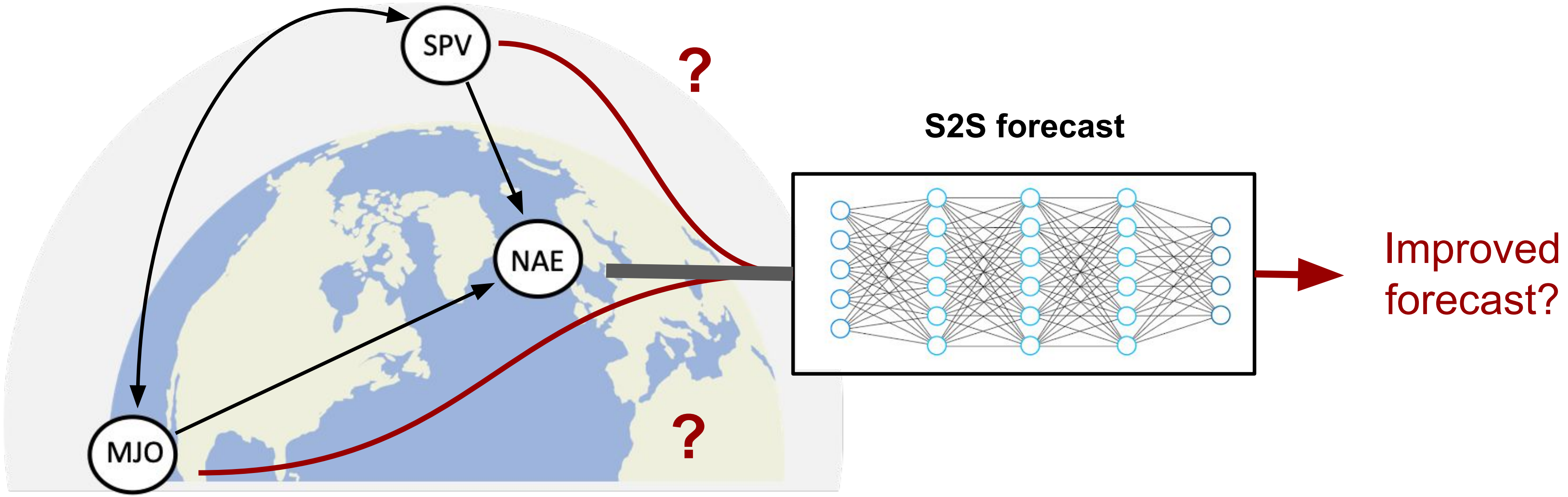
UMI LAB
TU berlin

ATB
Leibniz-Institut für Agrartechnik und Bioökonomie

UNIVERSITÄT LEIPZIG

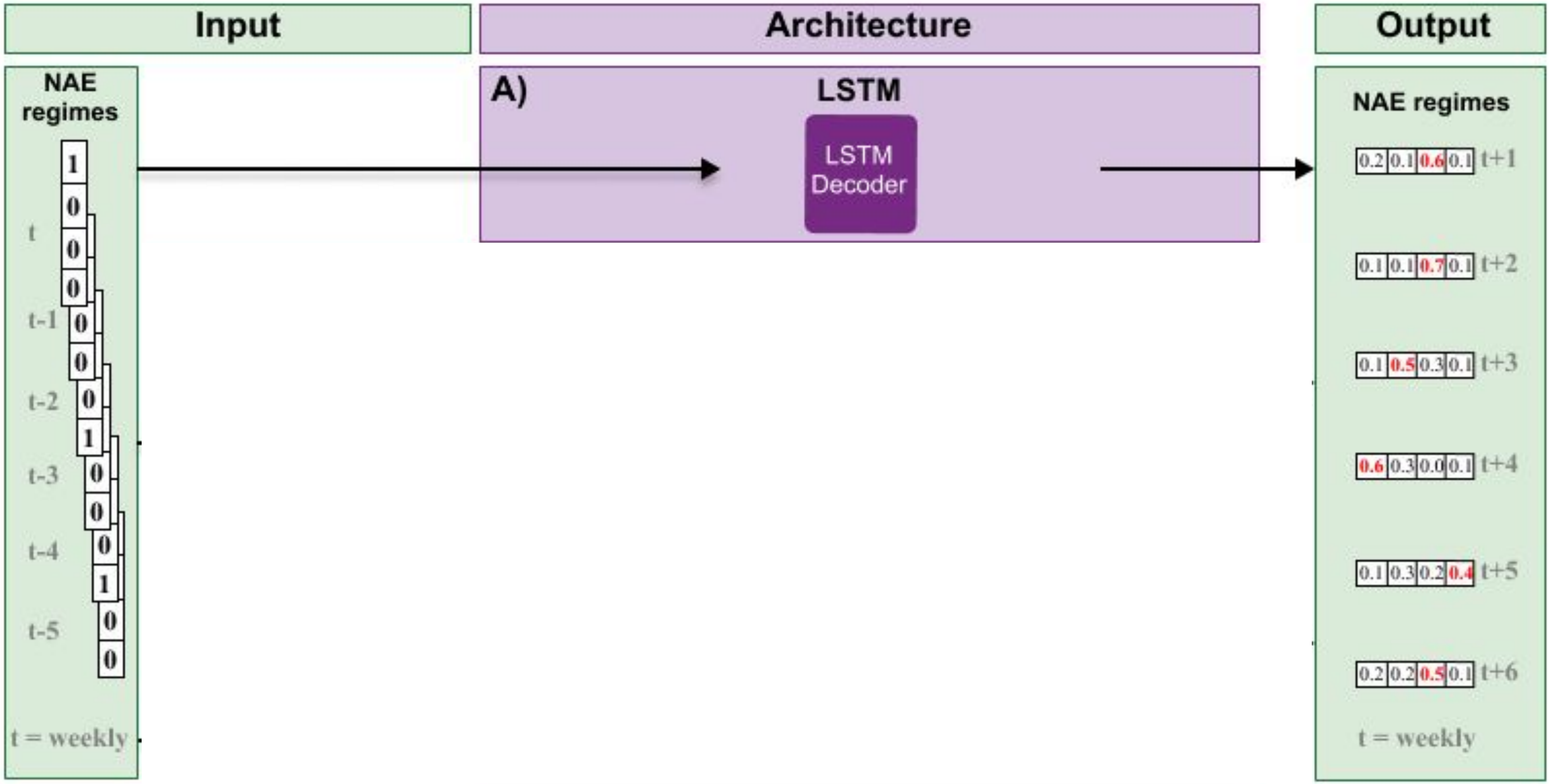
University of Reading

Research questions

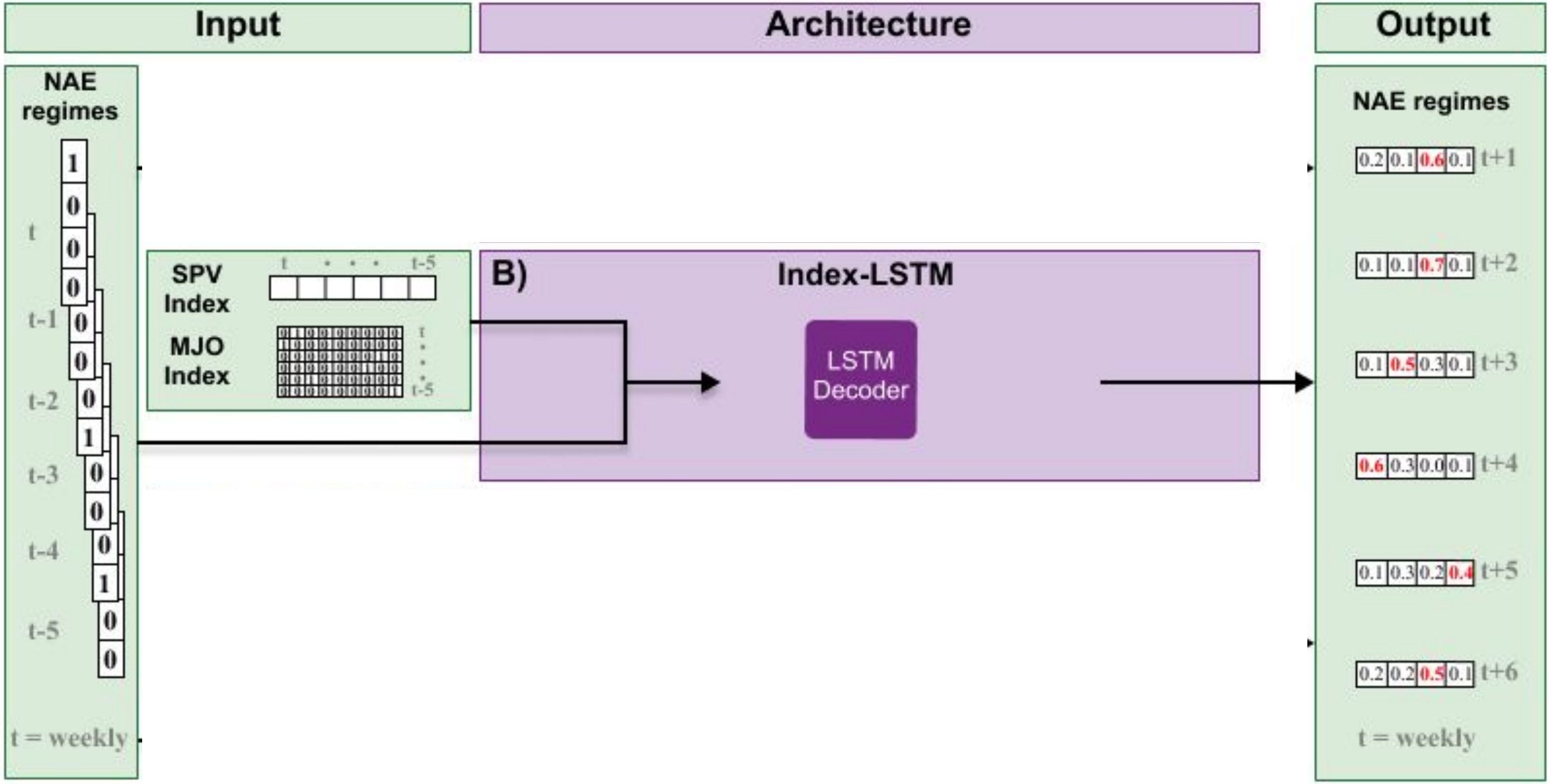


Task & Architecture

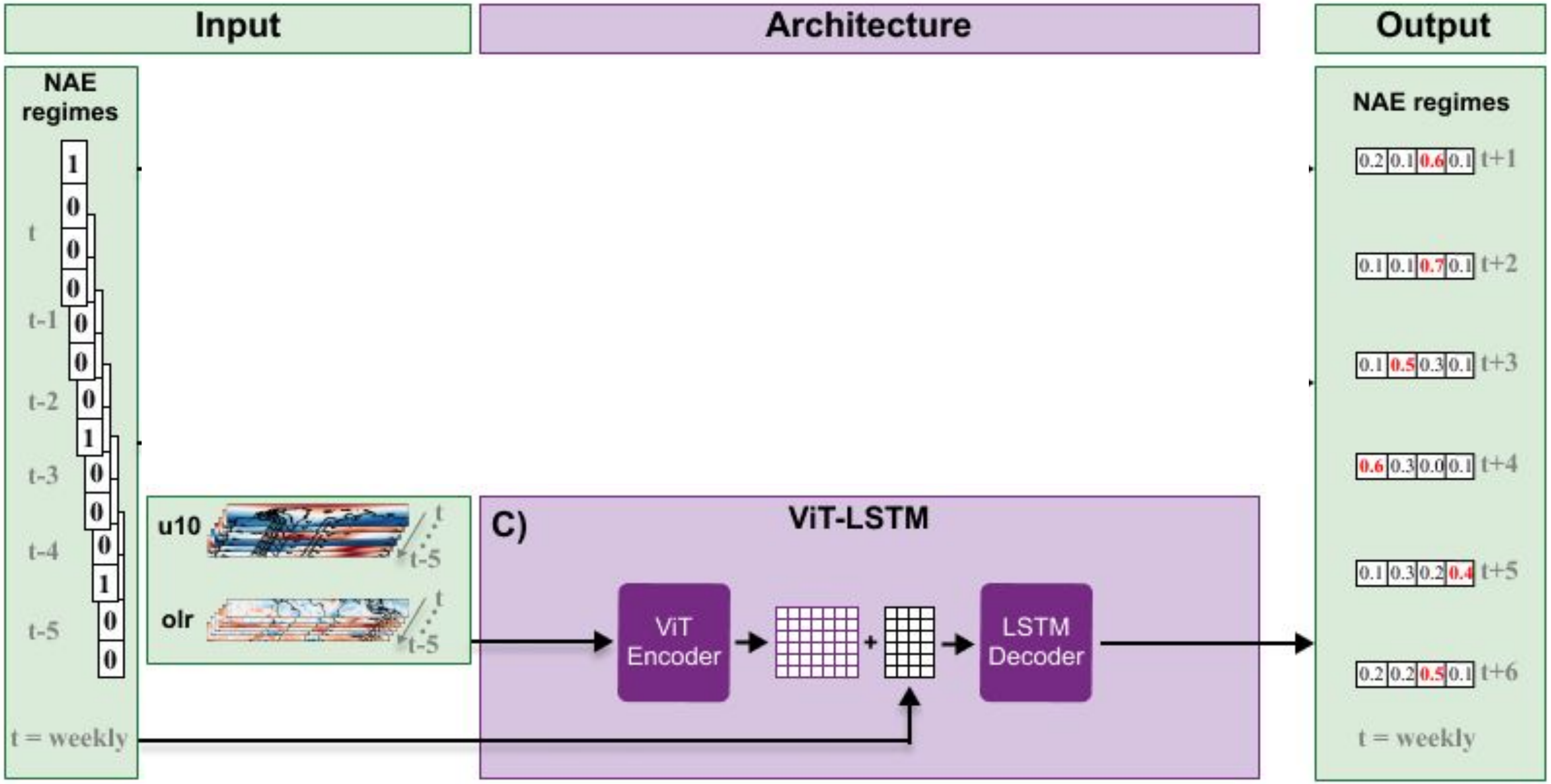
Architecture



Architecture



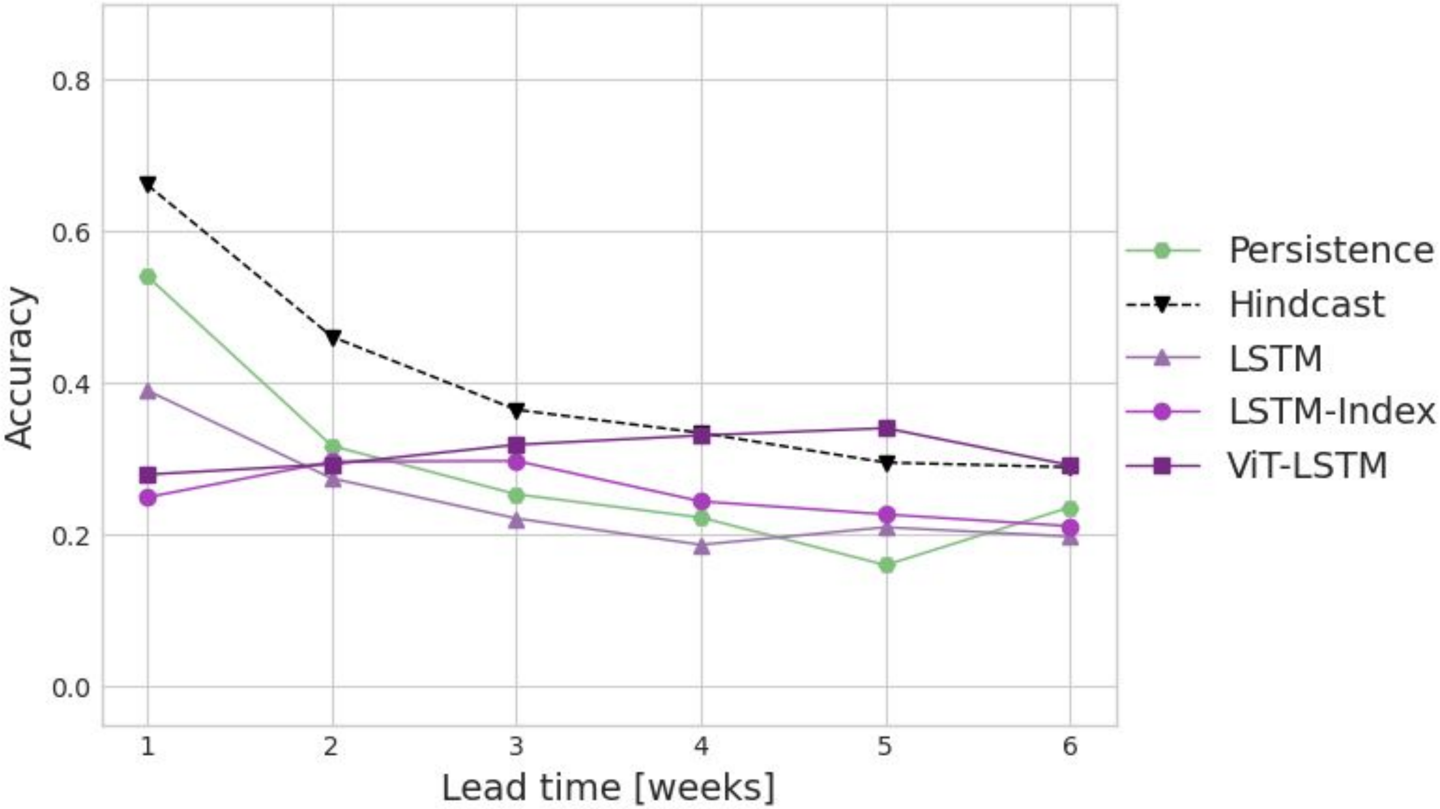
Architecture



Results

I) Forecast Skill

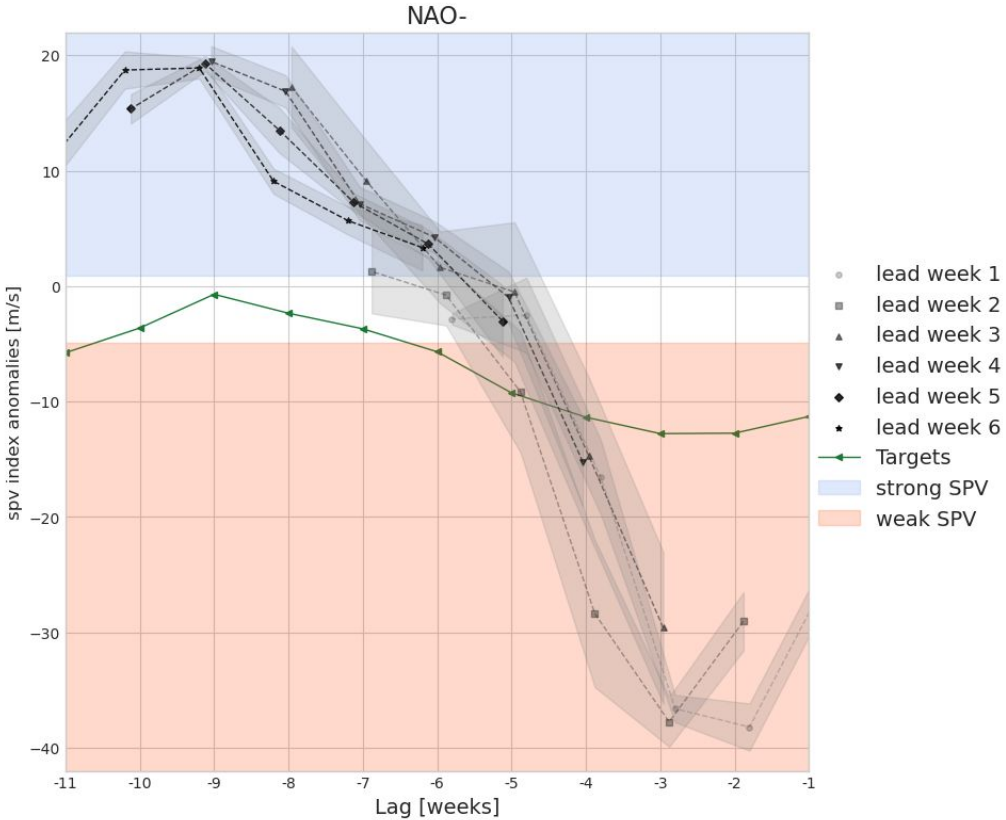
Performance



Bommer et al. 2025: Deep Learning Meets Teleconnections: Improving S2S Predictions for European Winter Weather

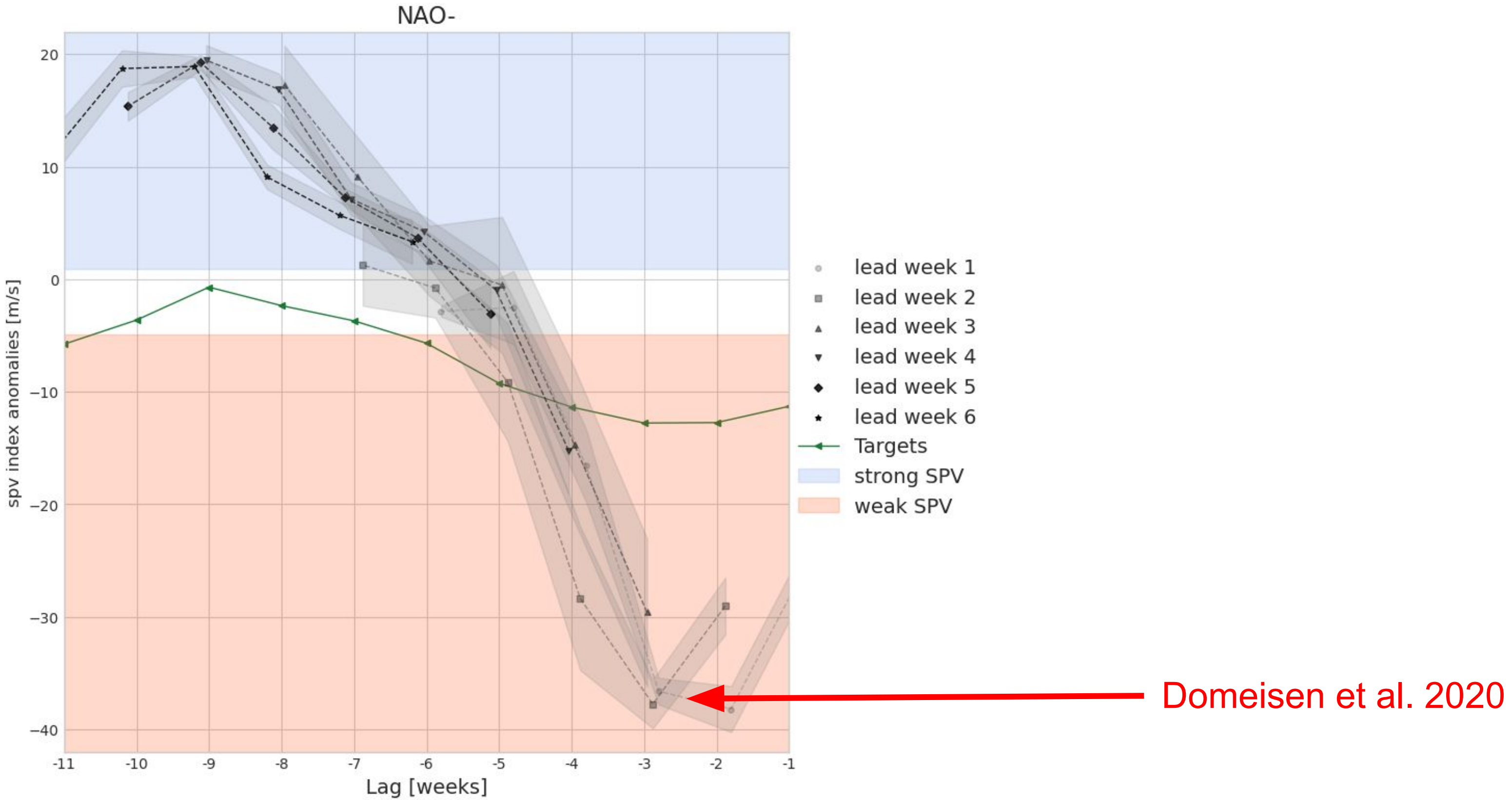
II) Impact of teleconnections

SPV Precursor



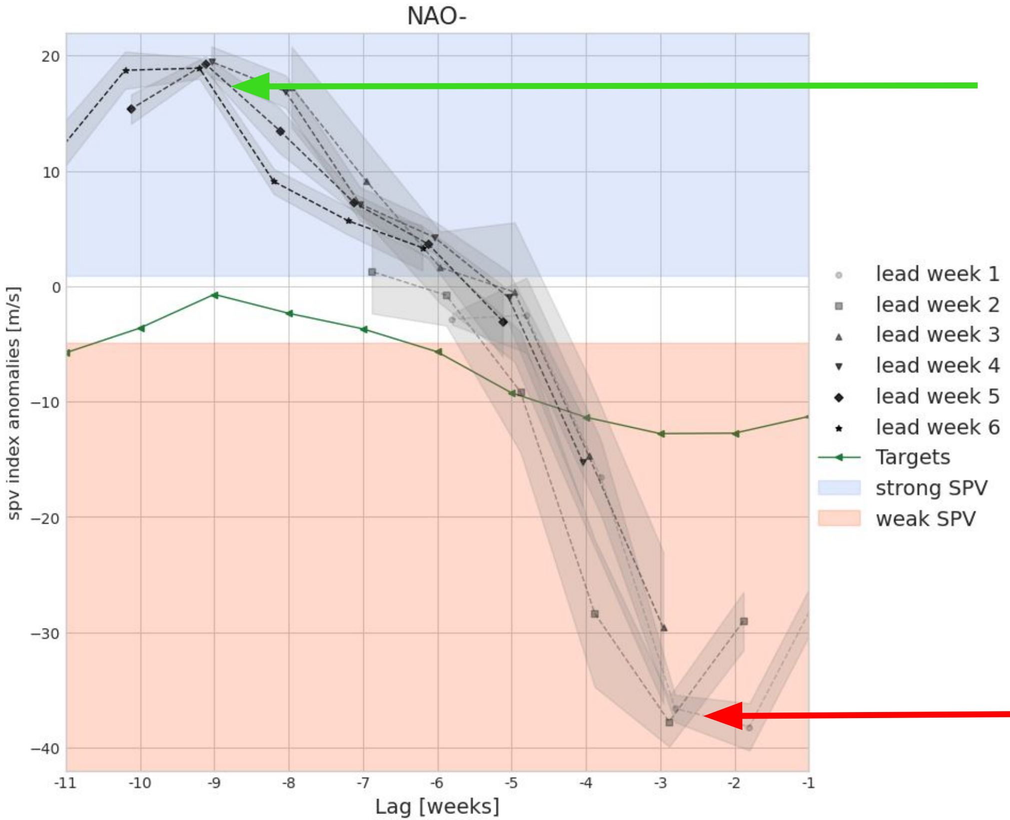
Bommer et al. 2025: Deep Learning Meets Teleconnections: Improving S2S Predictions for European Winter Weather

SPV Precursor



Bommer et al. 2025: Deep Learning Meets Teleconnections: Improving S2S Predictions for European Winter Weather

SPV Precursor



Strong SPV improves predictability on longer timescales?

Domeisen et al. 2020

Bommer et al. 2025: Deep Learning Meets Teleconnections: Improving S2S Predictions for European Winter Weather

Conclusion

Take home!

1. ***External drivers*** *and DL* ***can improve S2S predictability***

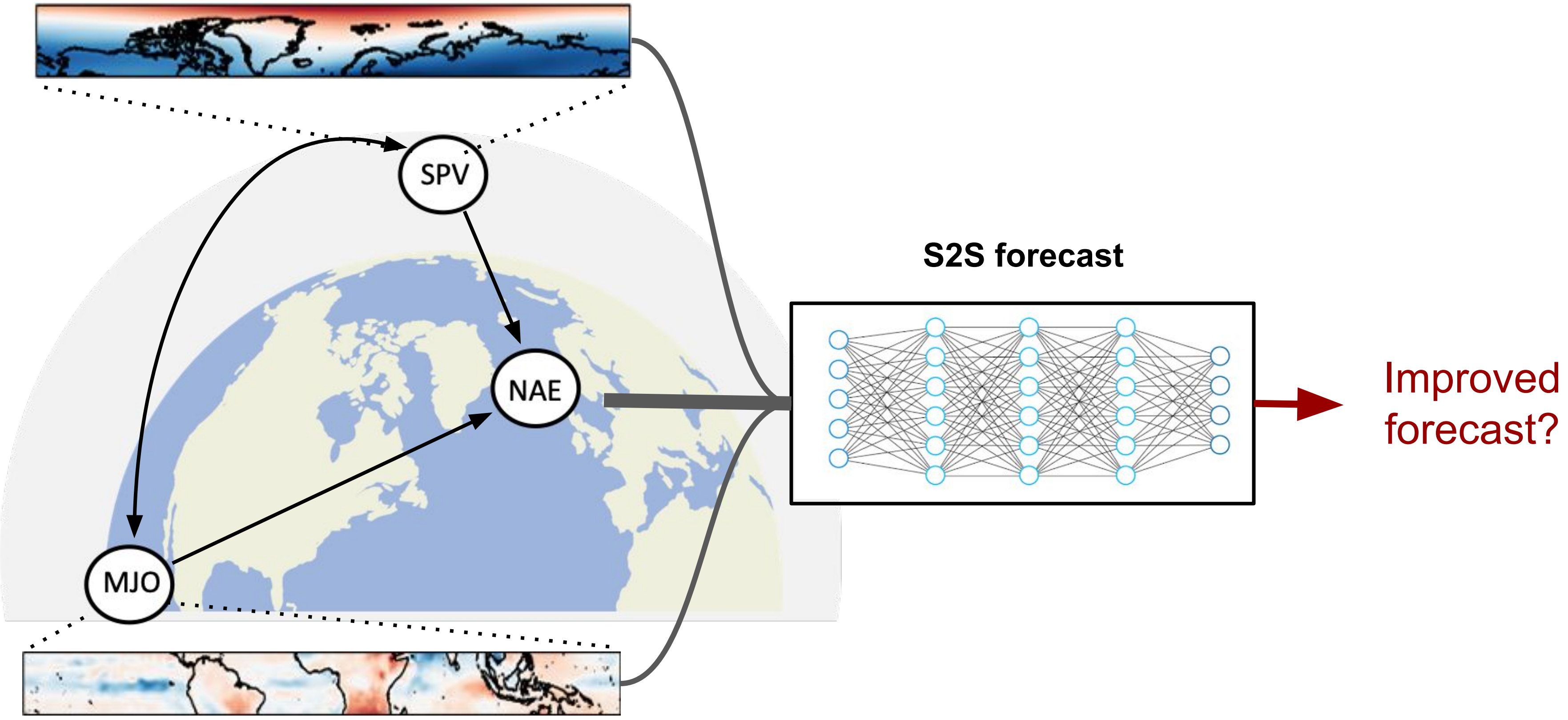
2. **Purely data-driven embeddings** potentially complementary to established physics

3. **ML as an investigative tool** for unknown atmospheric relationships

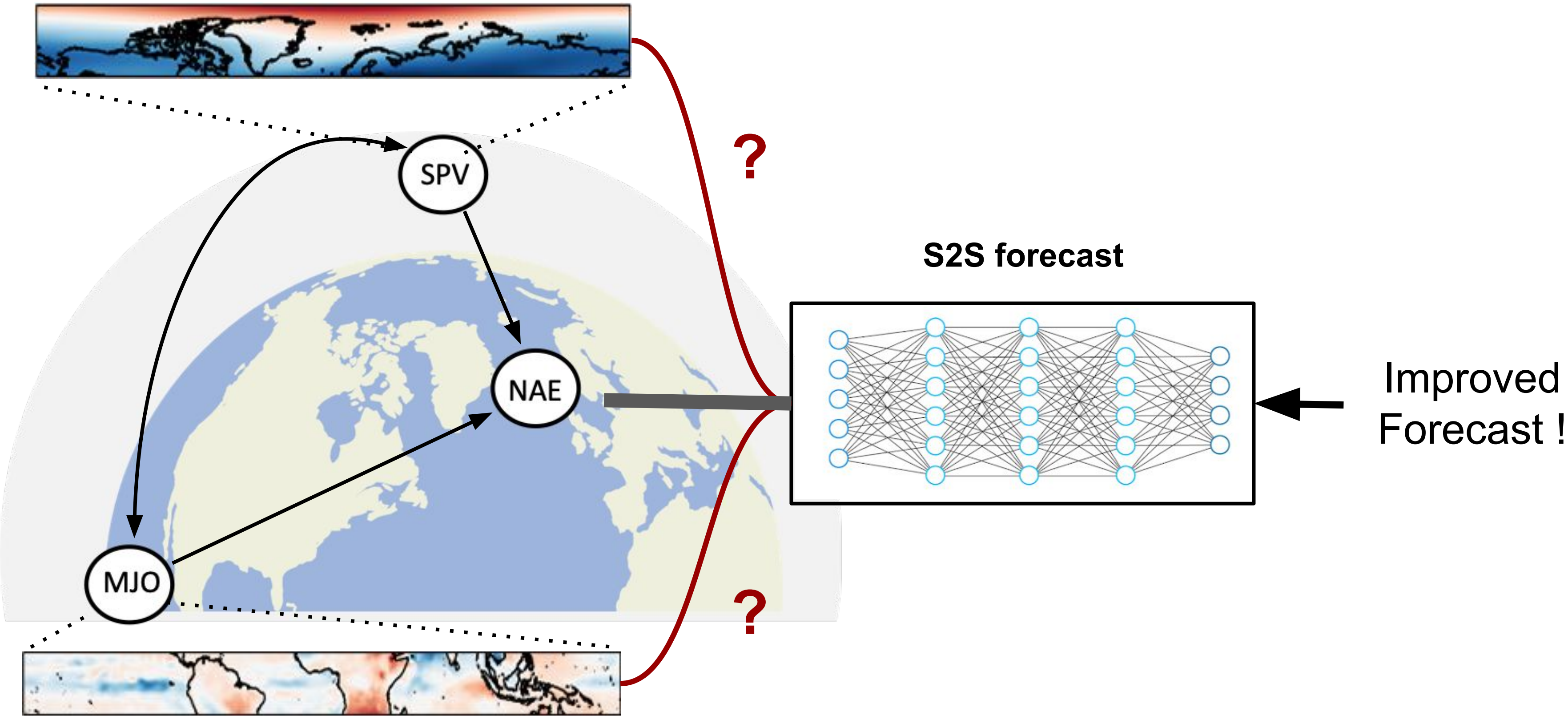
Check out the paper and github!

Appendix

Research question A

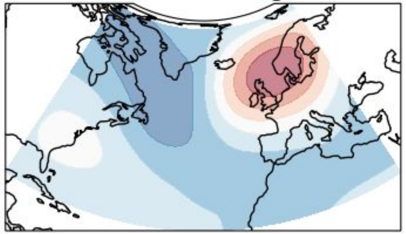


Research question B



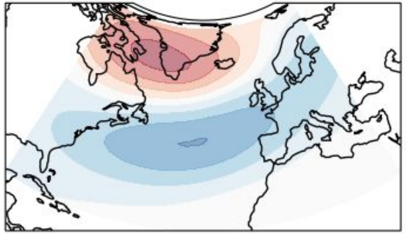
Prediction Task

Scandinavian blocking (24.50%)



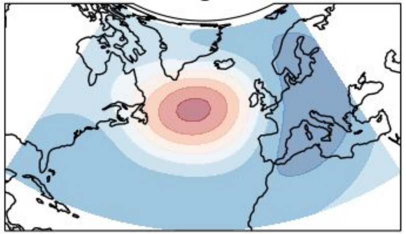
Class 1

NAO - (19.03%)



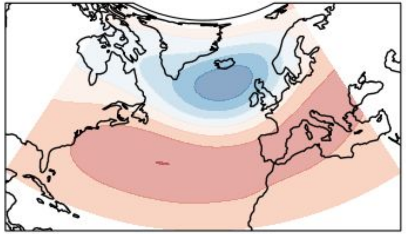
Class 2

Atlantic Ridge (24.50%)



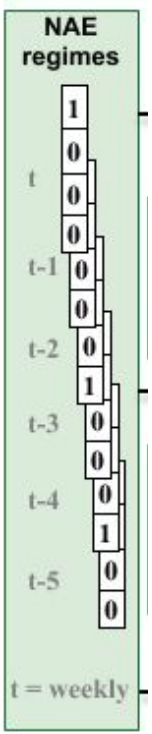
Class 3

NAO+ (31.96%)



Class 4

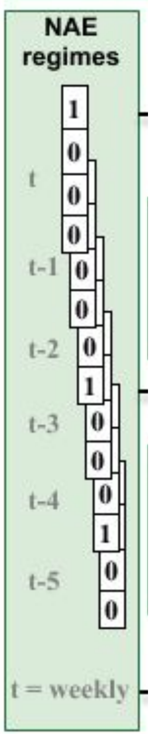
Prediction Tasks



Inputs:

- 6 weekly averages
- Time series:
  1. 4 **NAE regimes** as classes

Prediction Tasks

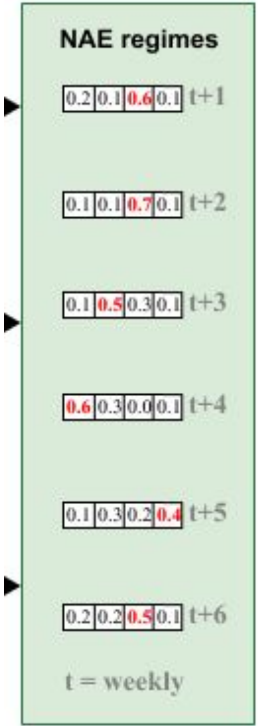


Inputs:

- 6 weekly averages
- Time series:
    1. 4 **NAE regimes** as classes

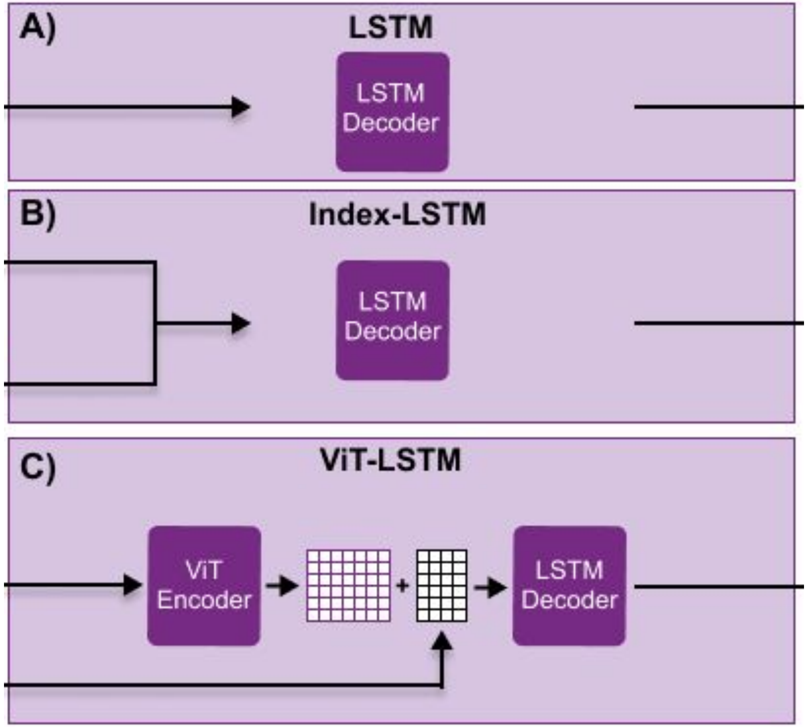
Outputs:

- weekly regime probabilities for t +1 to t+ 6 weeks
- predicted class: regime of highest probability



Training

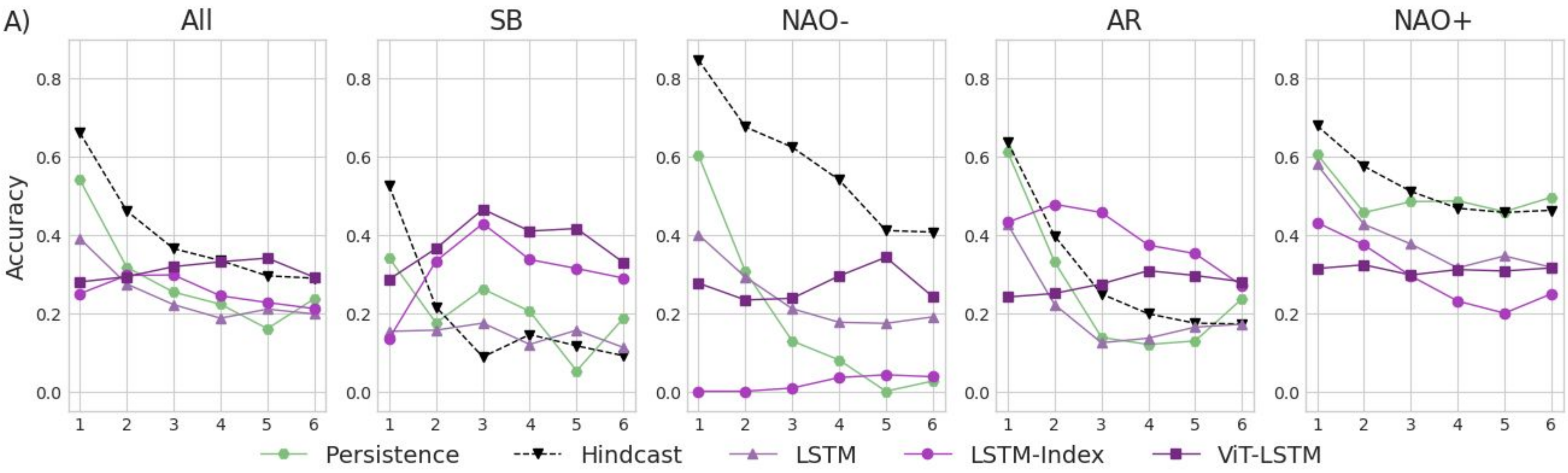
1. **Pre-training (MAE-based)**
   - 20CRv3 (1836 - 1979)
   - ERA5 (1980 - 2023)
2. **Fine-tuning (Classification)**
   - ERA5 (1980 - 2023)



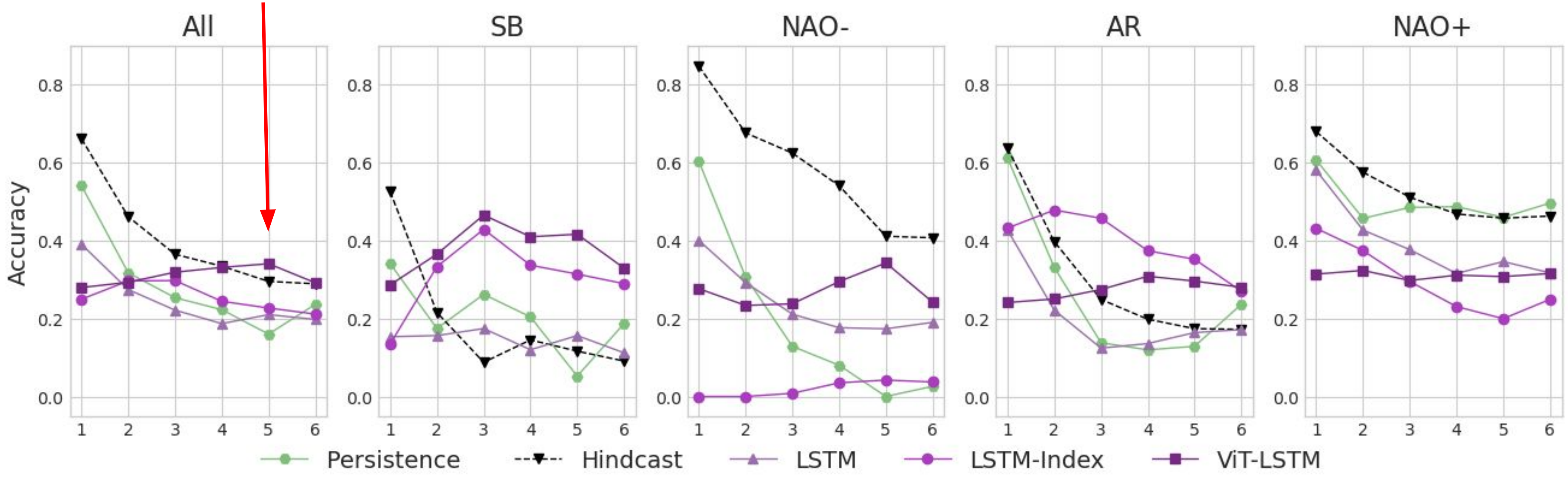
Performance

		lead week 1	lead week 2	lead week 3	lead week 4	lead week 5	lead week 6
Baseline	Persistence	54.1%	31.7%	25.3%	22.3%	16.0%	23.6%
	Climatology	24.8%	24.6%	24.3%	23.6%	23.6%	22.6%
	Hindcast	**66%**	**46.1%**	**36.5%**	**33.4%**	29.5%	**28.9%**
ML	LR	$35.65 \pm 0.09\%$	$32.53 \pm 0.08\%$	$26.89 \pm 0.09\%$	$20.46 \pm 0.09\%$	$16.1 \pm 0.1\%$	$23.30 \pm 0.08\%$
	LSTM	$39 \pm 3\%$	$28 \pm 1\%$	$22.1 \pm 0.7\%$	$18 \pm 1\%$	$21 \pm 1\%$	$21 \pm 1\%$
	Index - LSTM	$25 \pm 1\%$	$30 \pm 1\%$	$30 \pm 1\%$	$24 \pm 1\%$	$23 \pm 1\%$	$21 \pm 2\%$
	ViT - LSTM	$28 \pm 2\%$	$30 \pm 2\%$	$31 \pm 2\%$	$\mathbf{33} \pm \mathbf{2}\%$	$\mathbf{33} \pm \mathbf{2}\%$	$\mathbf{30} \pm \mathbf{2}\%$
	Aurora-T	$37 \pm 1\%$	$26 \pm 1\%$	$26 \pm 2\%$	$22 \pm 2\%$	$22 \pm 1\%$	$21 \pm 2\%$

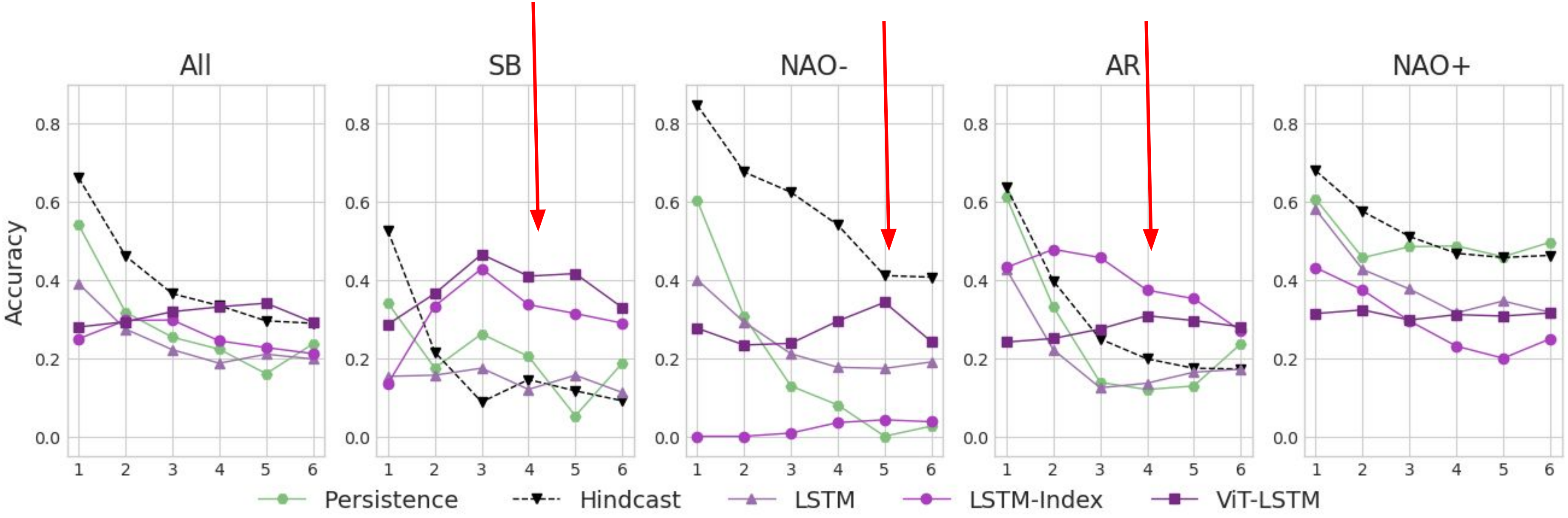
Performance



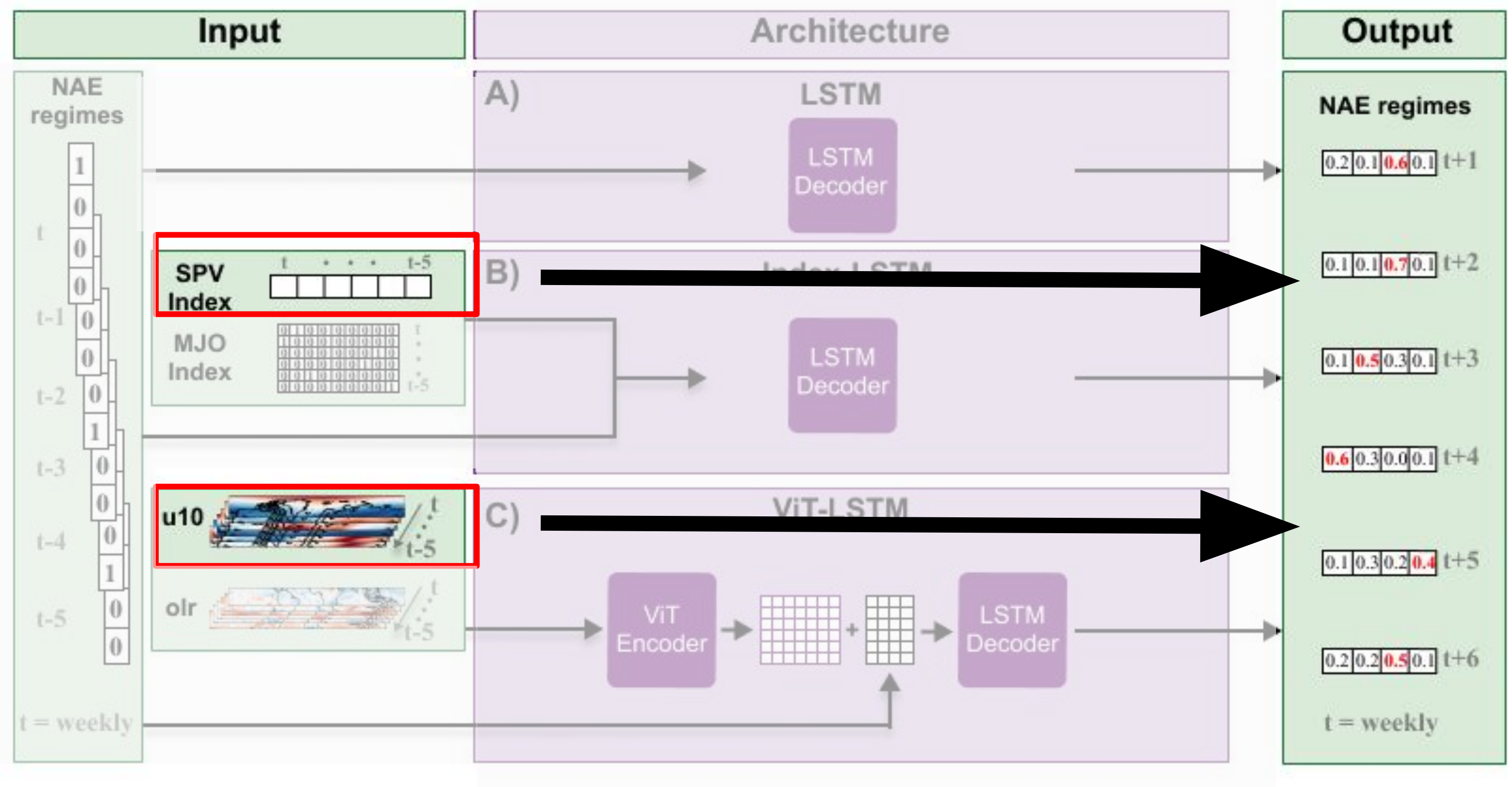
Performance



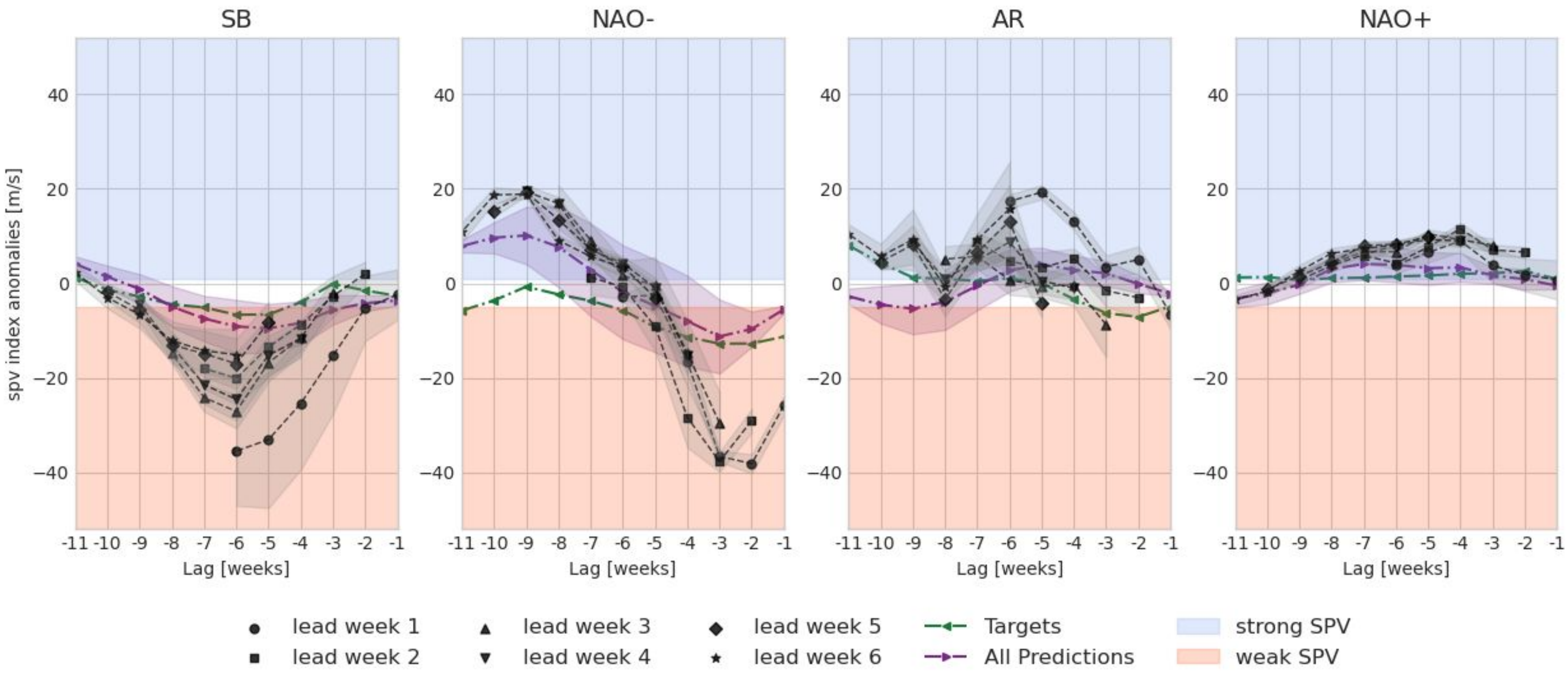
Performance



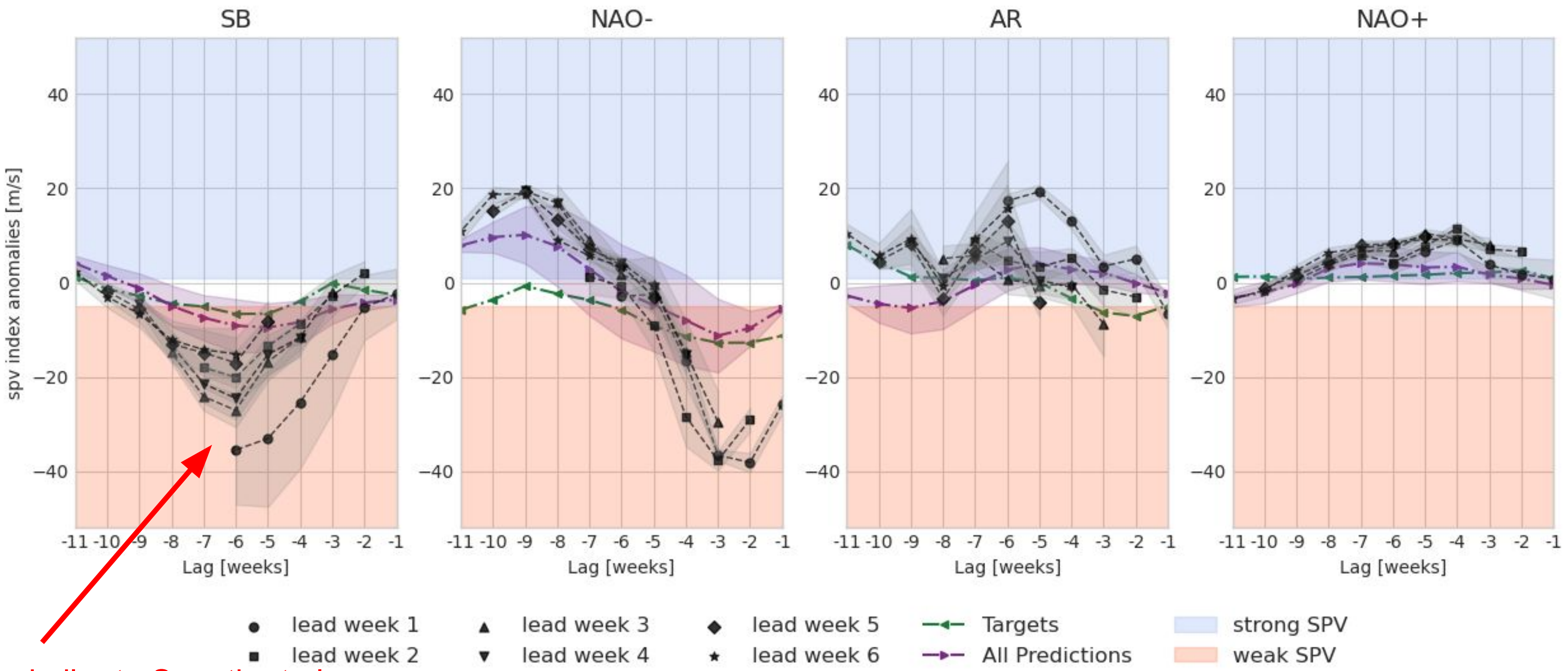
Example - SPV Precursor



SPV Precursor



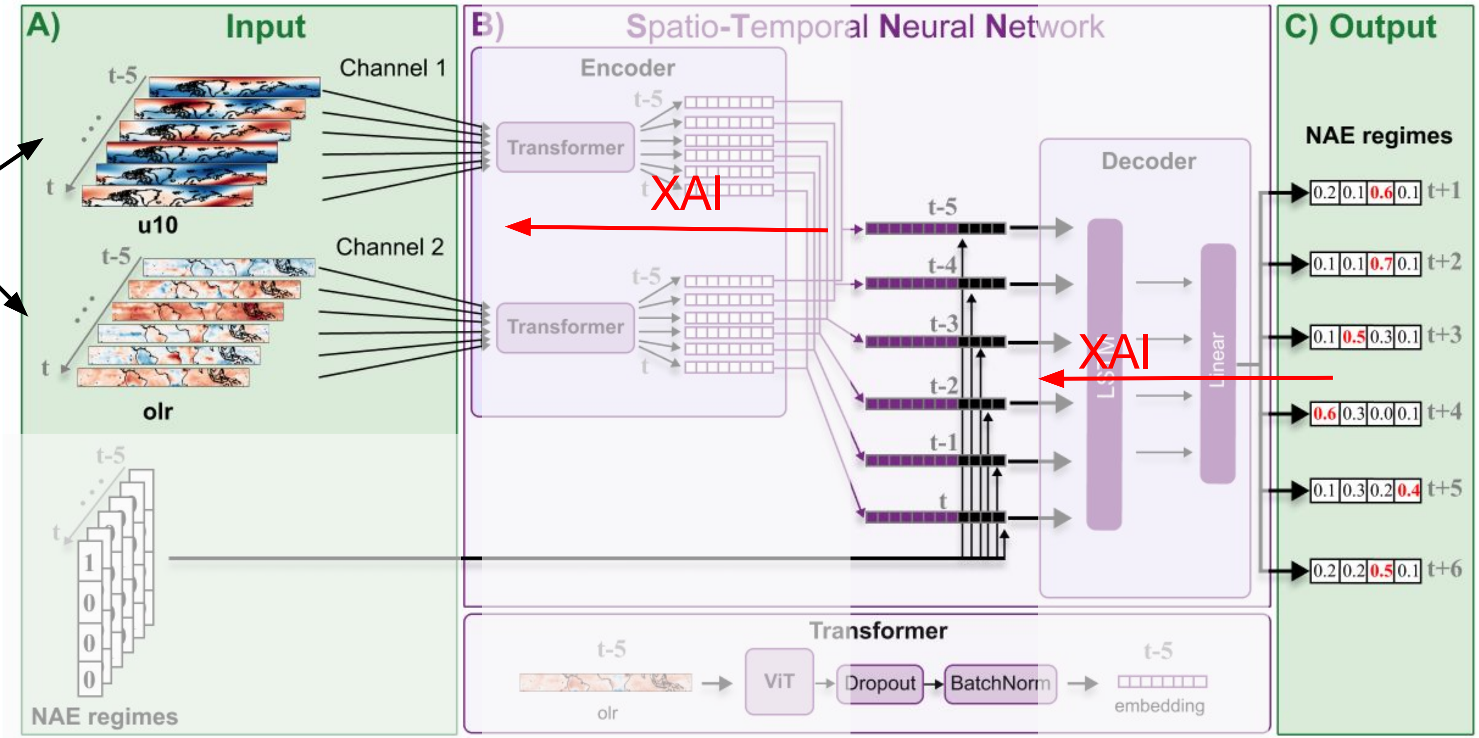
SPV Precursor



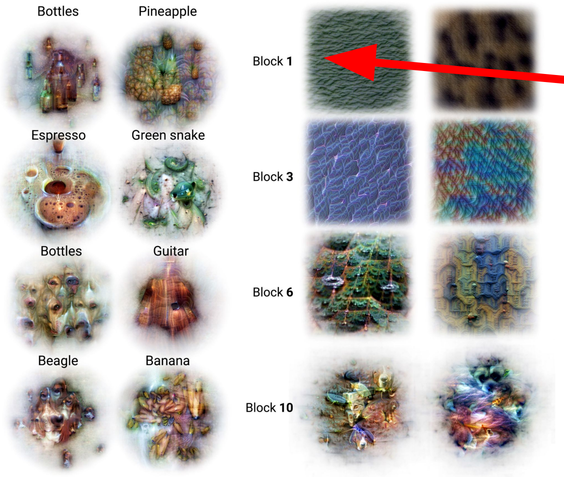
similar to Spaeth et al. 2024

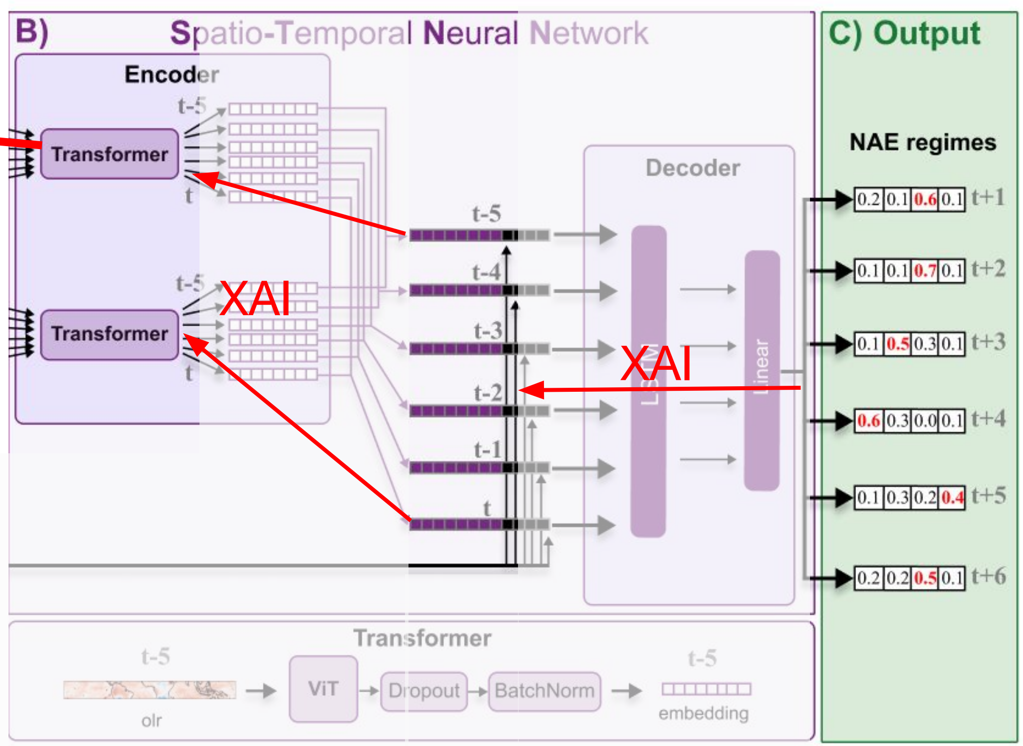
Forecasts of Opportunity?

Statistics



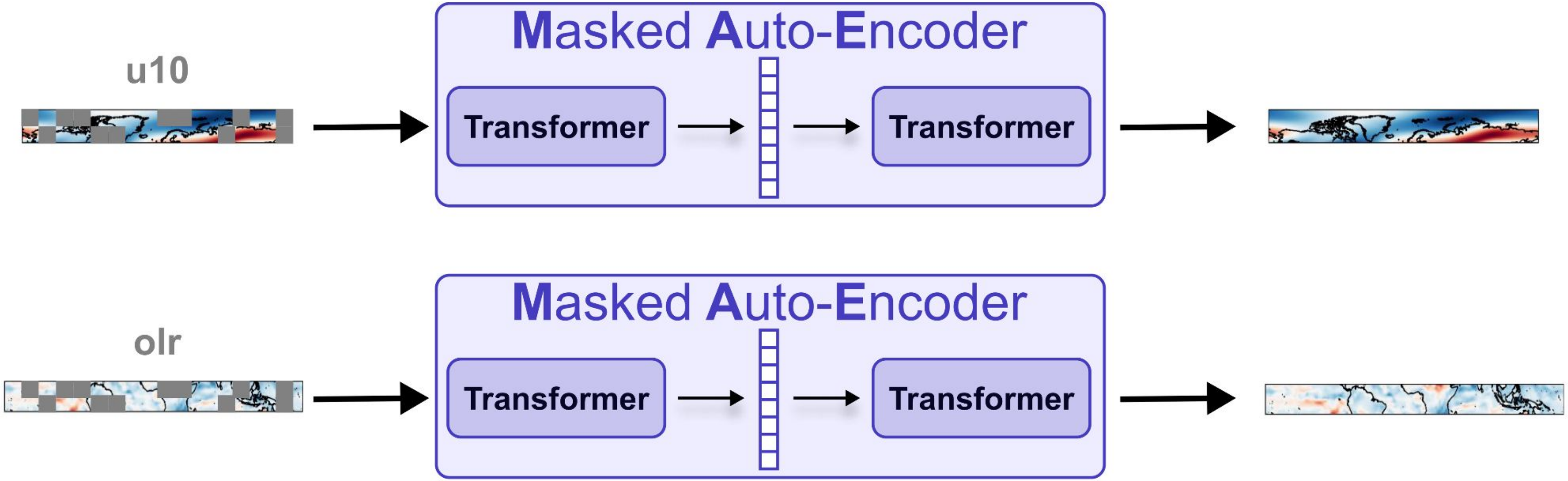
Physics-based Representations?



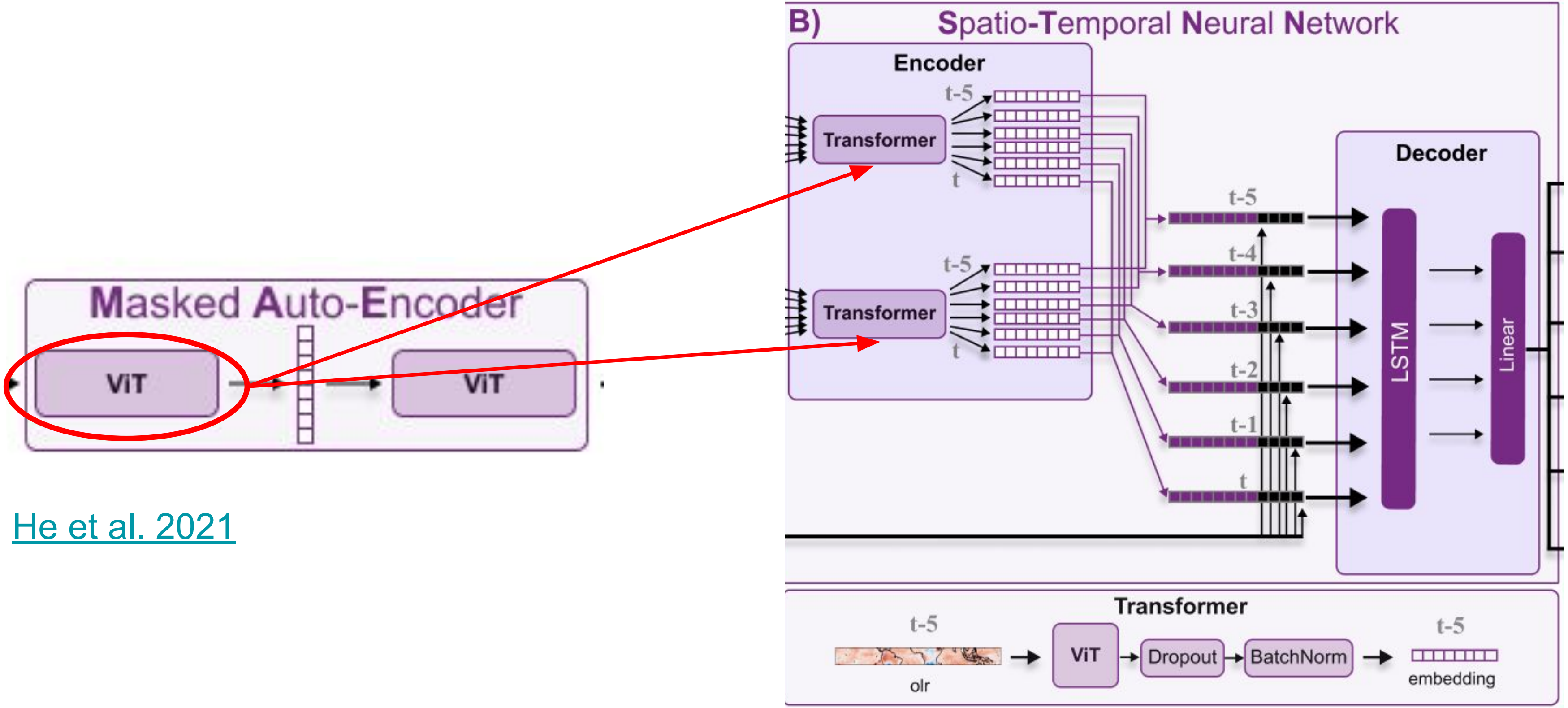


Fel et al. 2023

Pretraining of Transformer

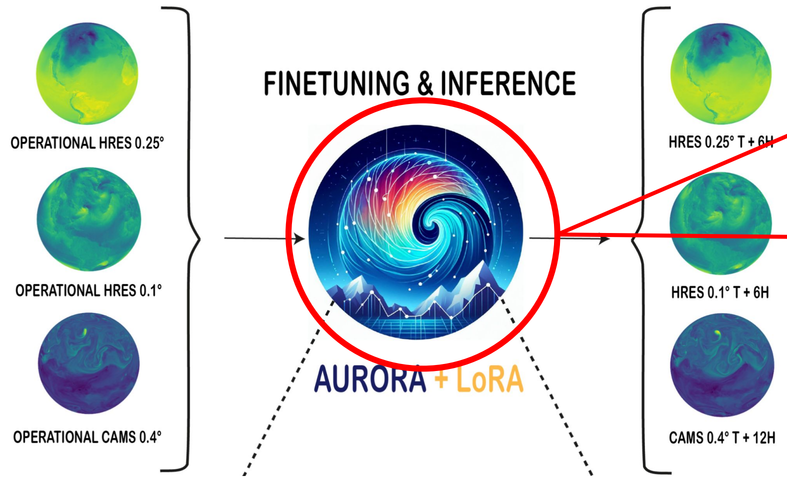


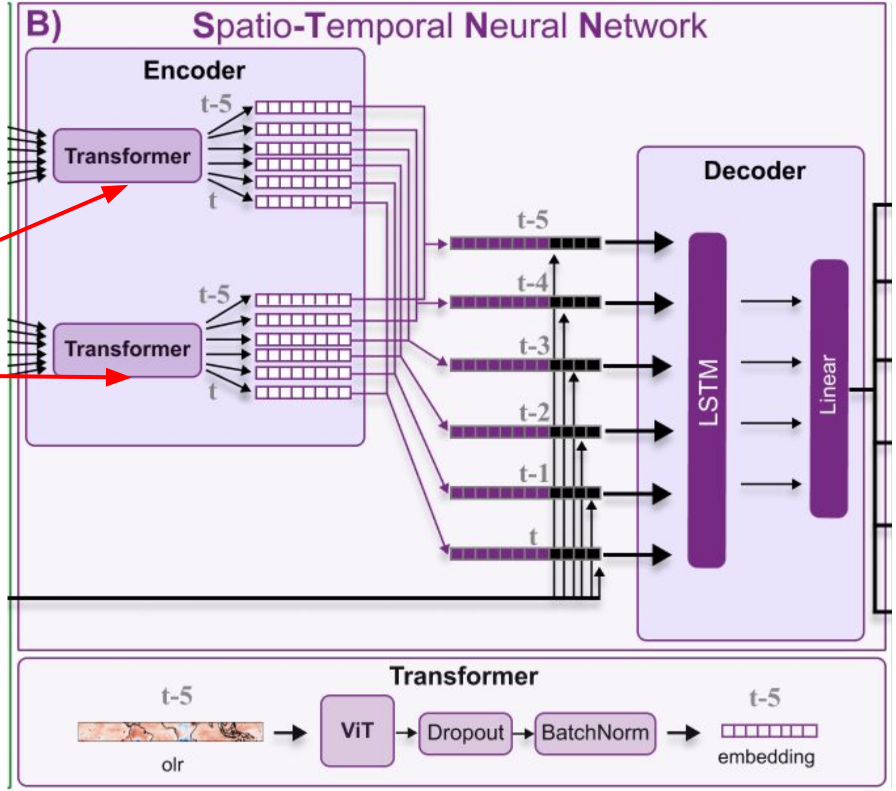
Architecture



[He et al. 2021]

Architecture

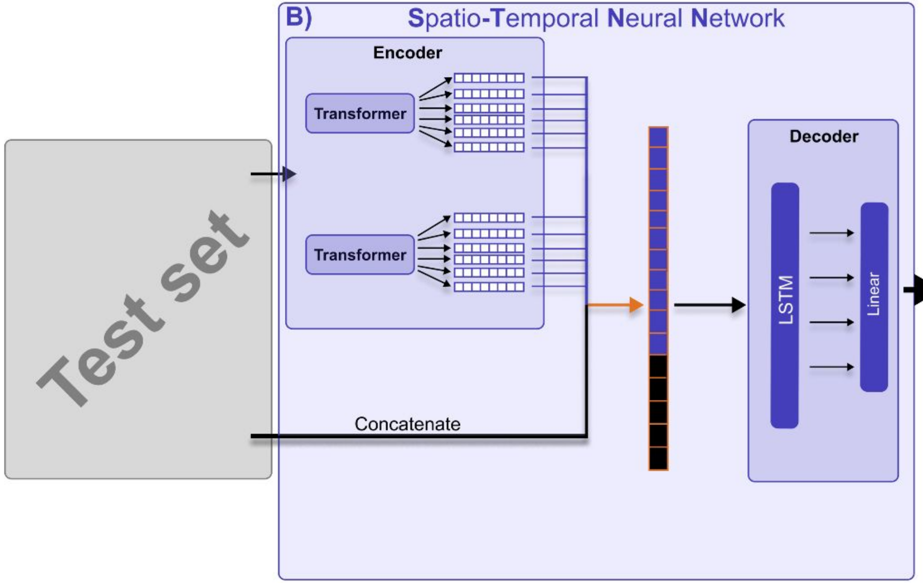


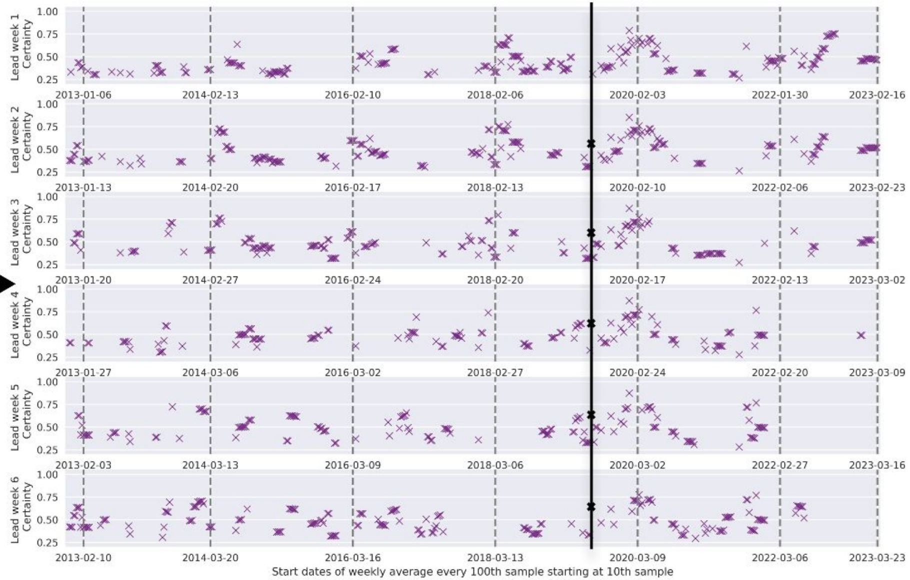


[Bodnar et al. 2024]()

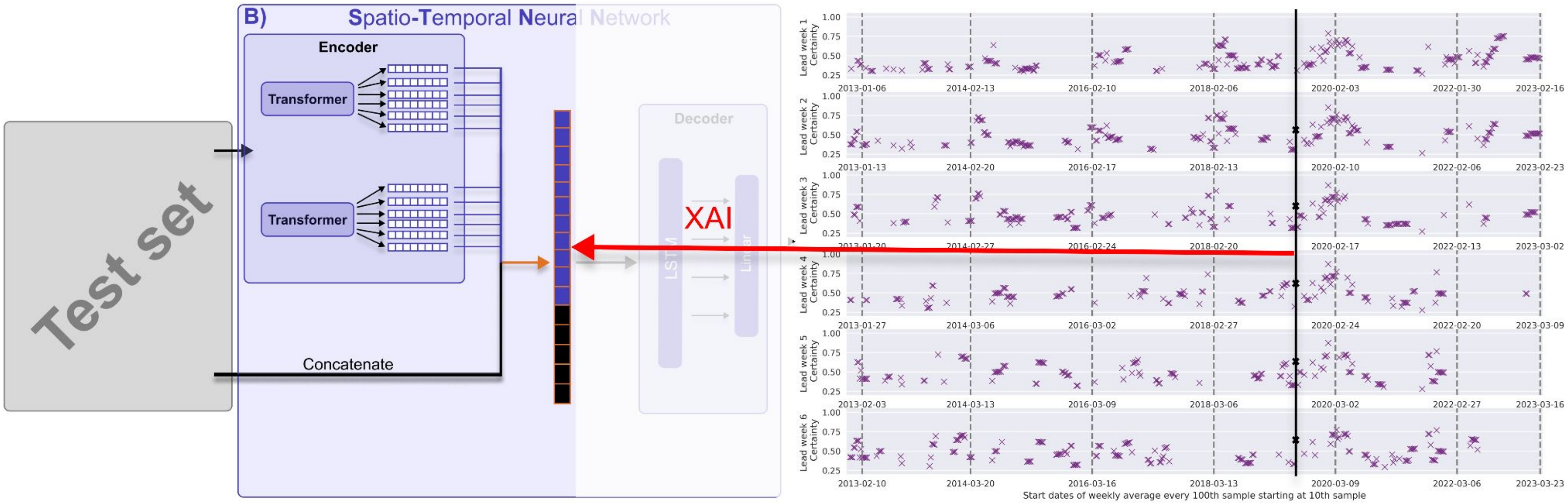
Explainable AI (XAI)

XAI for teleconnections





XAI for teleconnections



XAI for teleconnections