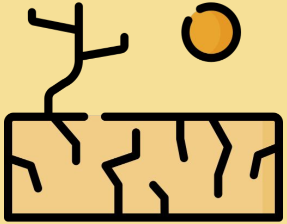
Palmer Drought Severity Index (PDSI) modelling using Deep Learning approaches

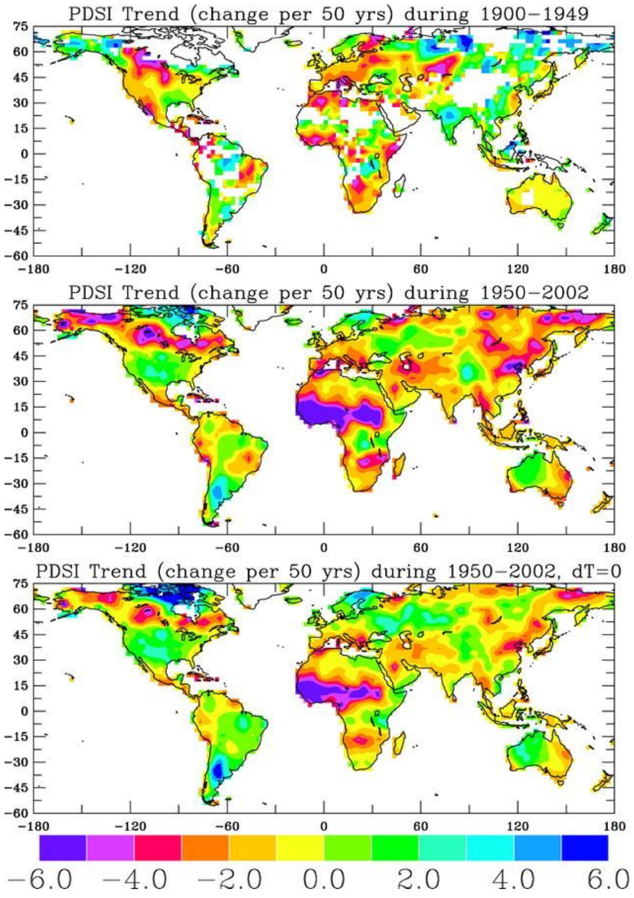
Team members:
Mikhail Kuznetsov
Victor Kozhevnikov
Ivan Gurev
Artem Gorbarenko



Motivation

Droughts represent a significant and vital challenge for regions worldwide, causing profound impacts on agriculture and ecosystems.

Palmer Drought Severity Index (PDSI) is a widely used drought index that stands out in identifying long-term drought patterns. It stakes into account readily available temperature and precipitation that generally spans from -10 (dry) to +10 (wet). However, it faces challenges when it comes to cross-regional comparability, as it may not be as easily applied across diverse geographical areas.

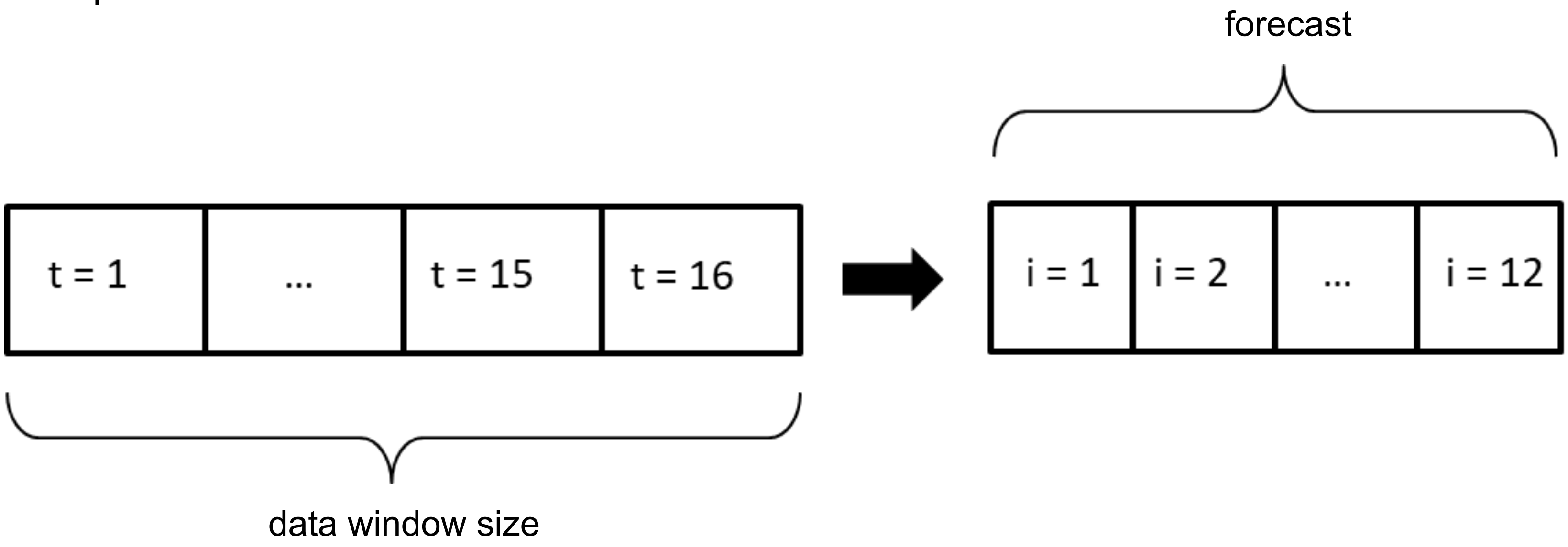


Dataset

The datasets for the selected regions, representing **monthly** PDSI values. They were obtained from the Google Earth Engine service and structured as spatiotemporal tensors -
(***T, H, W)*** where T corresponds to the temporal aspect of the data, while H and W correspond to latitude and longitude

Data preprocessing

To tackle the classification and regression tasks, data preprocessing involved creating a 16-value historical window and a 12-value future lookahead, allowing the model to consider past data and make predictions. This temporal context captures historical trends and enhances the model's ability to forecast dynamic phenomena in time series data.



Evaluation metrics

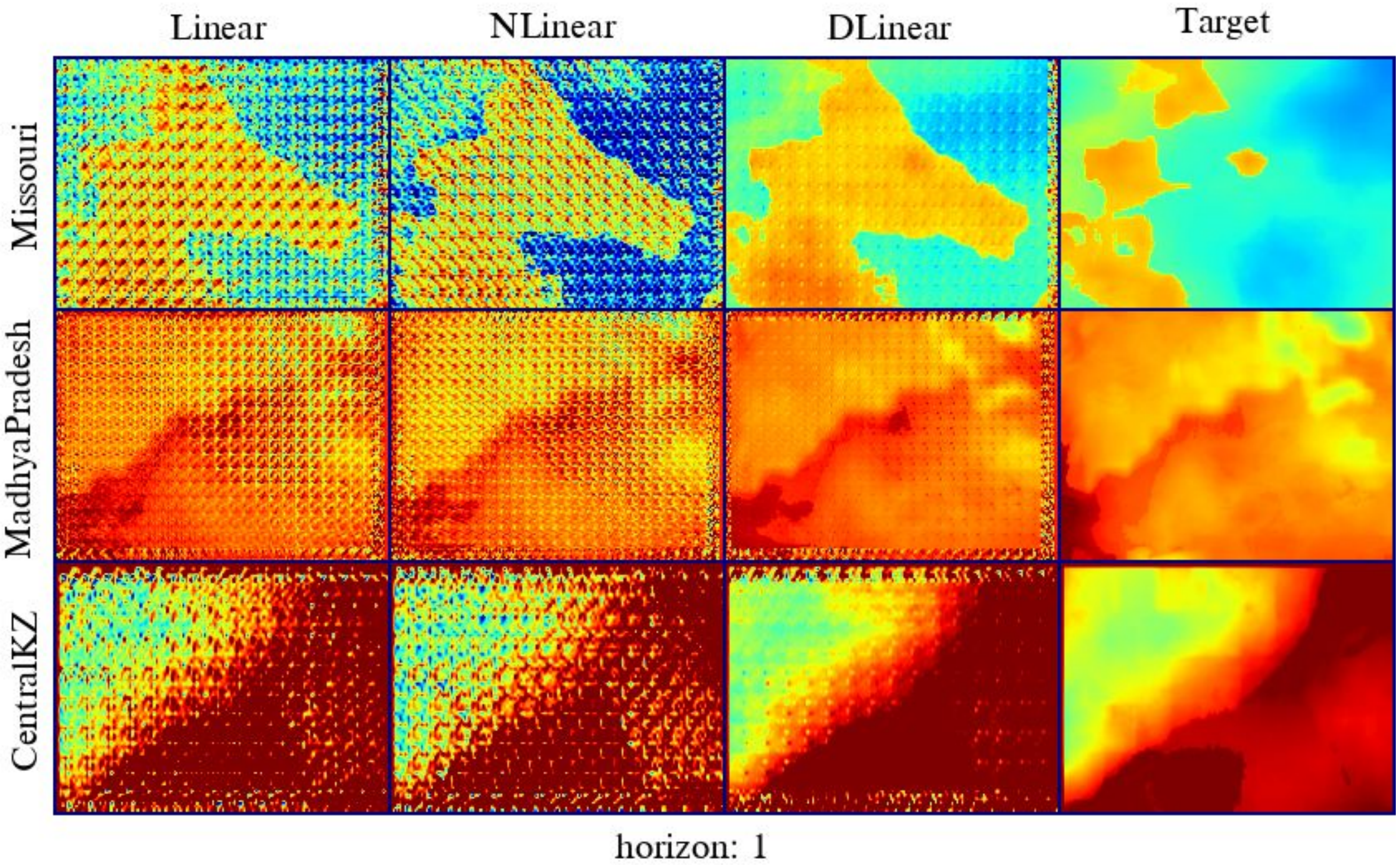
In our experimental setup, we have employed three distinct tasks:

- regression (direct prediction of PDSI)
  - metrics: *R^2, MAE, MSE, RMSE*
  - loss: *MSE*

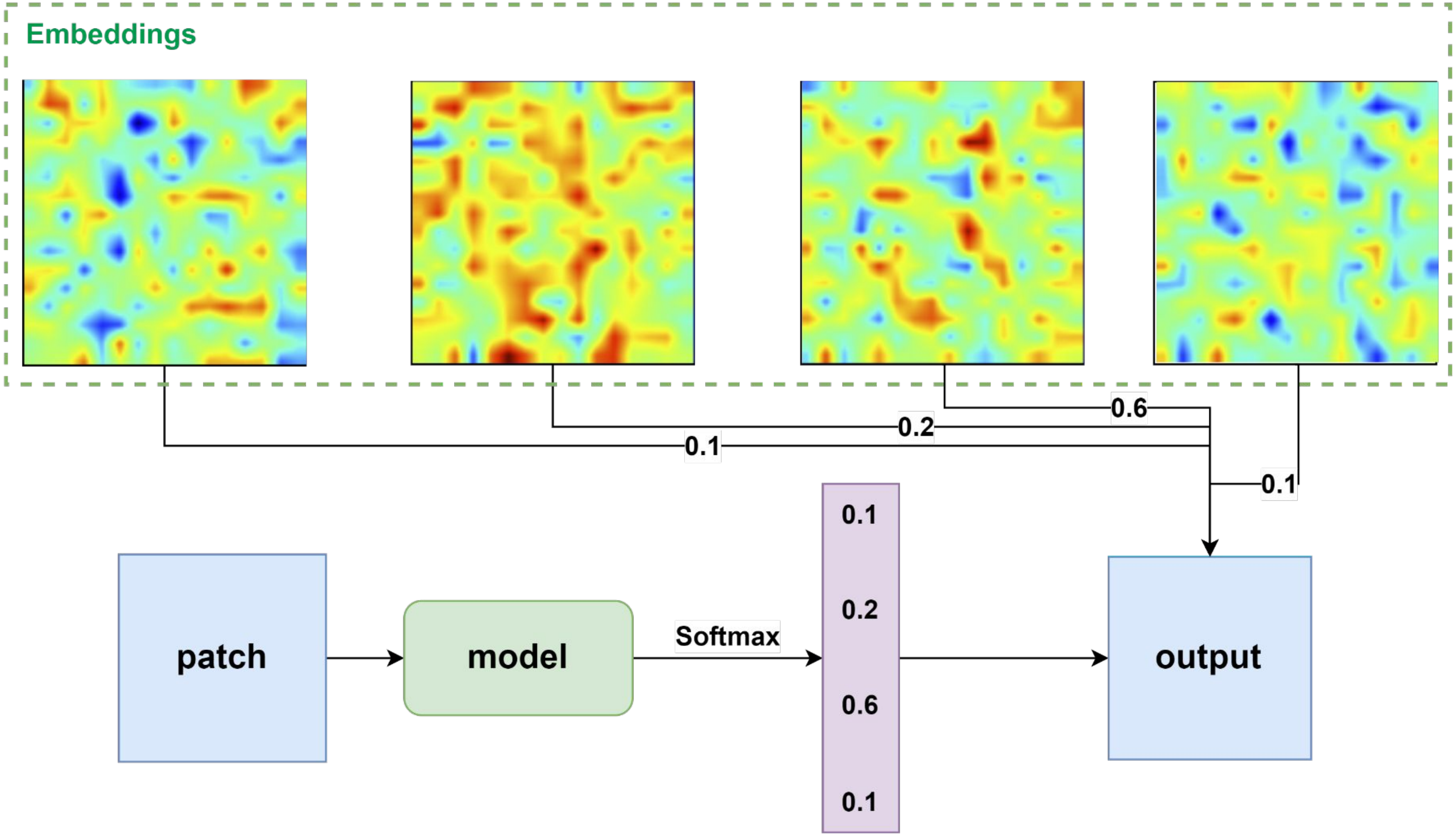
- binary classification (PDSI < -2 = ”drought”, otherwise ”non-drought”)
  - metrics: *ROC-AUC, PR-AUC, F1 Score*
  - loss: *BCEWithLogitsLoss*

- multiclass classification (PDSI thresholds = [-2, 2])
  - *Accuracy*
  - loss: *CrossEntropyLoss*

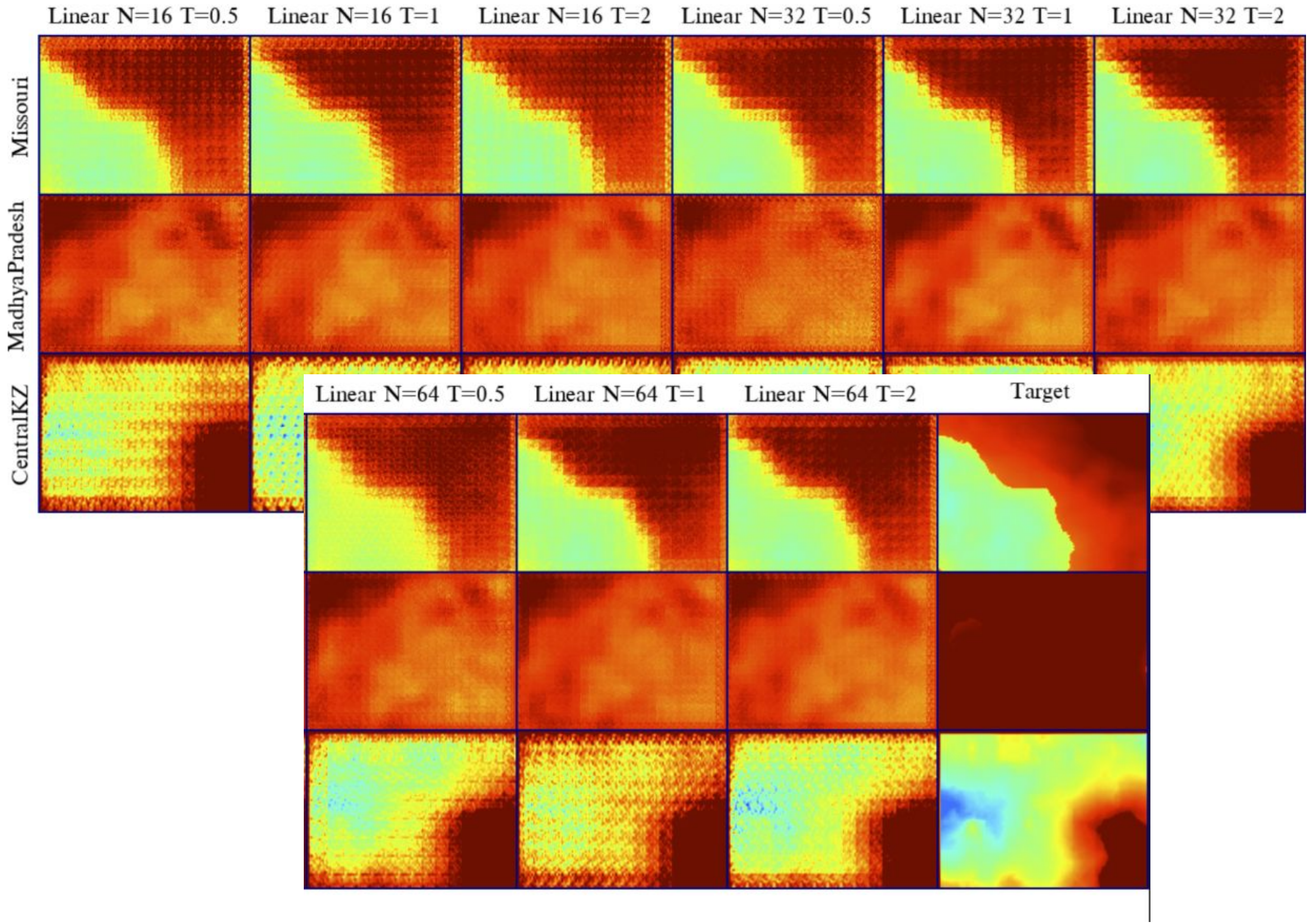
Linear / NLinear / DLinear



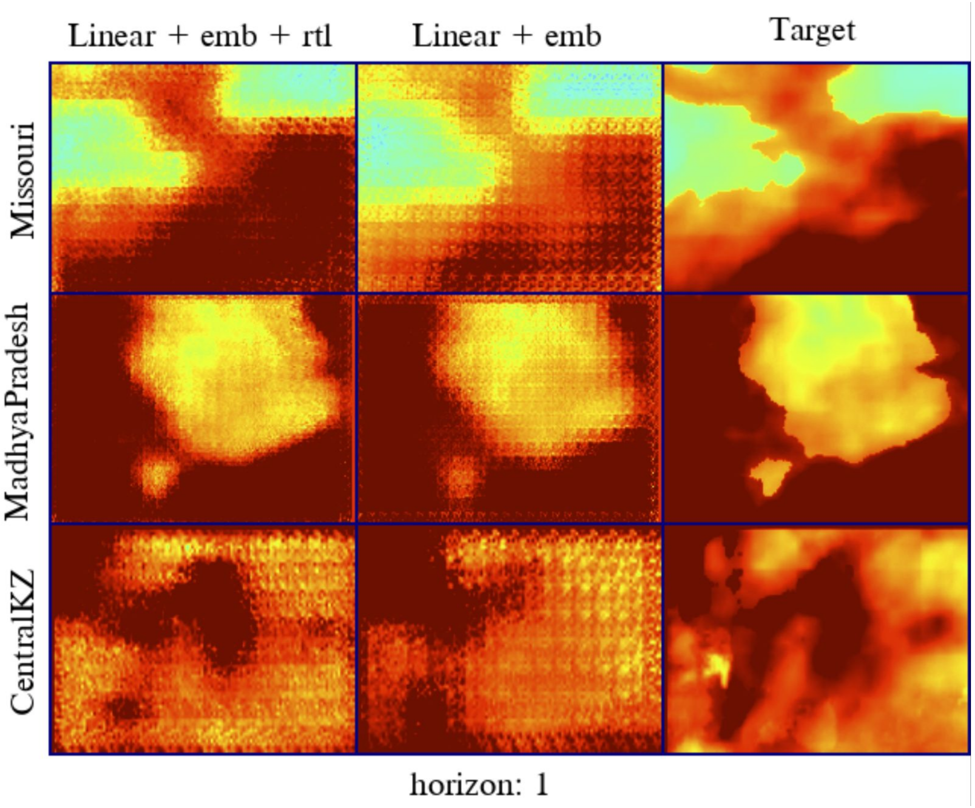
Linear + Emb

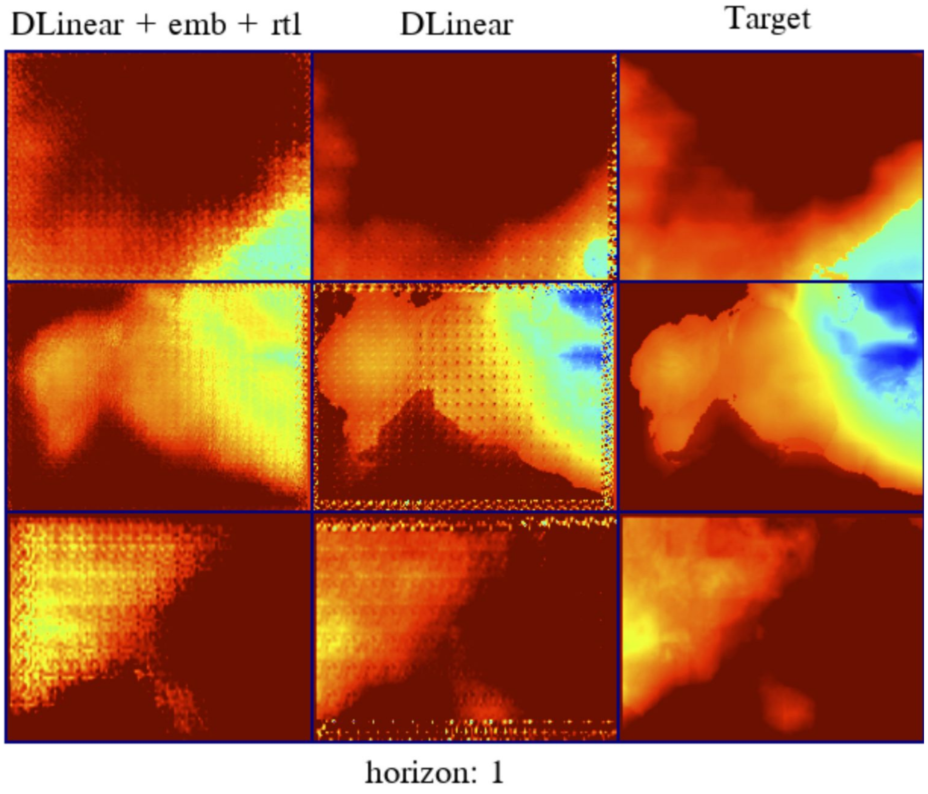


Linear + Emb + Temperature

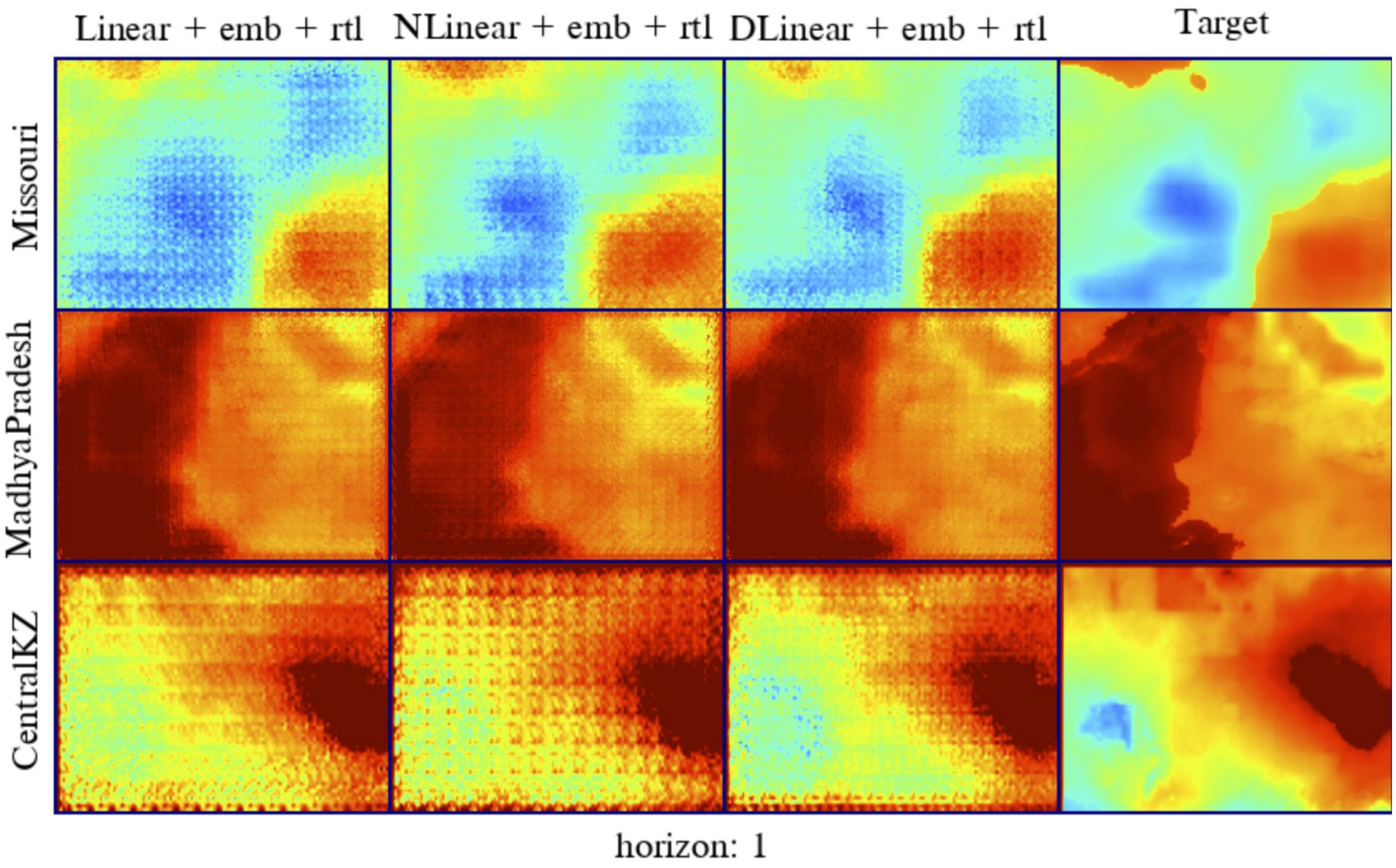


Add RTL like feature





After all enhancements



Some metrics

reg r2_1_test

model	Linear	NLinear	DLinear
data			
CentralKZ	0.904825	0.905052	0.918569
MadhyaPradesh	0.883341	0.817849	0.882852
Missouri	0.934189	0.935407	0.942752

multiclass acc_1_test

model	Linear	NLinear	DLinear
data			
CentralKZ	0.227791	0.341454	0.221902
MadhyaPradesh	0.331068	0.330183	0.334281
Missouri	0.783762	0.451154	0.790940

binary auc_1_test

model	Linear	NLinear	DLinear
data			
CentralKZ	0.991082	0.989548	0.985282
MadhyaPradesh	0.993866	0.994716	0.994463
Missouri	0.996567	0.997907	0.997801

ConvLSTM

$$i_t = \sigma(W_{xi} * X_t + W_{hi} * H_{t-1} + W_{ci} \odot \mathcal{C}_{t-1} + b_i)$$

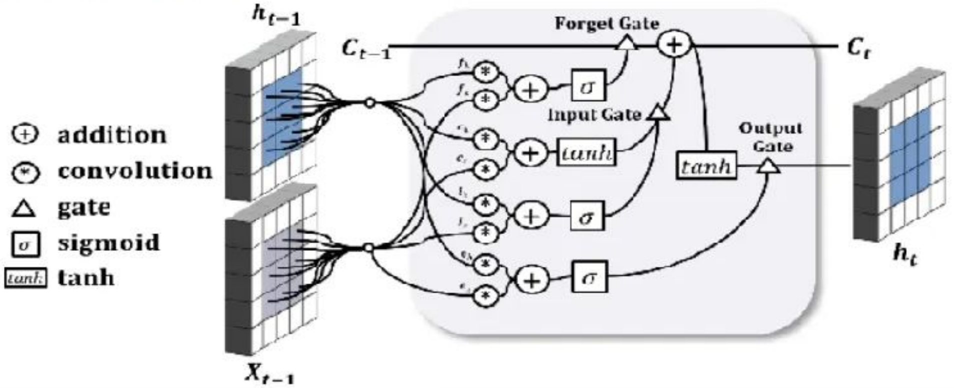
$$f_t = \sigma(W_{xf} * X_t + W_{hf} * H_{t-1} + W_{cf} \odot \mathcal{C}_{t-1} + b_f)$$

$$\mathcal{C}_t = f_t \odot \mathcal{C}_{t-1} + i_t \odot \tanh(W_{xc} * X_t + W_{hc} * \mathcal{H}_{t-1} + b_c)$$

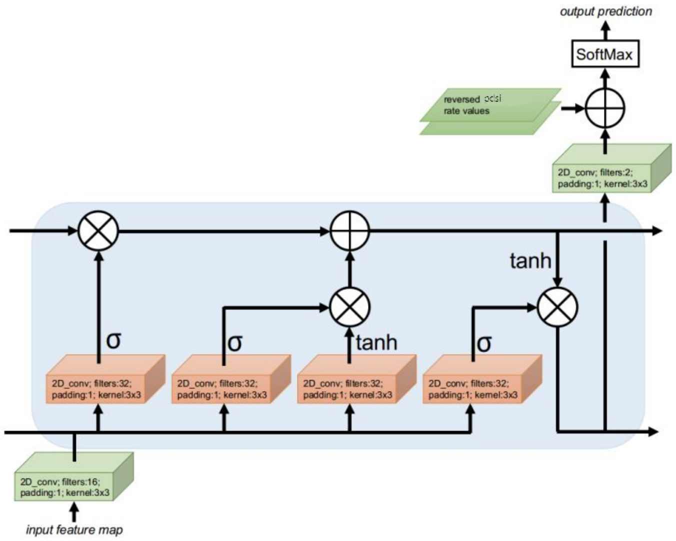
$$o_t = \sigma(W_{xo} * X_t + W_{ho} * \mathcal{H}_{t-1} + W_{co} \odot \mathcal{C}_t + b_o)$$

$$\mathcal{H}_t = o_t \odot \tanh(C_t)$$

The key equations of ConvLSTM where * denotes the convolution operator and ⊙ the Hadamard product.

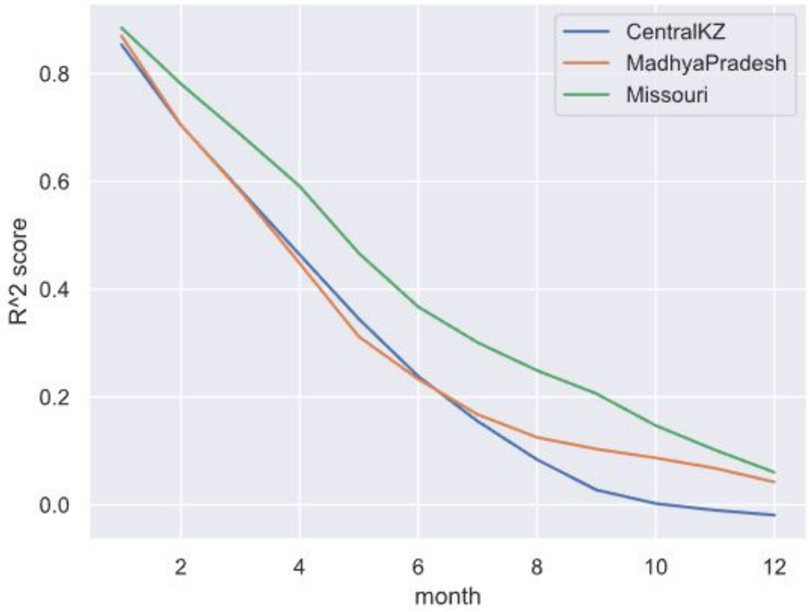


A ConvLSTM cell



ConvLSTM

Regression



MSE				
Month	1	6	12	mean
CentralKZ	0.97599	5.14295	7.11523	4.89017
MadhyaPradesh	1.57936	9.17617	11.3918	8.21442
Missouri	1.15433	6.23825	8.78758	5.77095

RMSE				
Month	1	6	12	mean
CentralKZ	0.98792	2.26781	2.66744	2.21137
MadhyaPradesh	1.25673	3.02922	3.37518	2.86608
Missouri	1.0744	2.49765	2.96439	2.40228

MAE				
Month	1	6	12	mean
CentralKZ	0.74574	1.79321	2.29918	1.74961
MadhyaPradesh	0.86791	2.35592	2.71859	2.17318
Missouri	0.85249	2.06514	2.46412	1.90556

ConvLSTM

Binary classification

PRAUC				
Month	1	6	12	all
CentralKZ	0.770206	0.427325	0.401143	0.564756
MadhyaPradesh	0.888006	0.503397	0.524974	0.628264
Missouri	0.530576	0.357761	0.443726	0.454797

F1 score				
Month	1	6	12	all
CentralKZ	0.783703	0.02236	0	0.449896
MadhyaPradesh	0.835772	0.532438	0	0.466221
Missouri	0.638919	0.193939	0	0.294882

ROC-AUC				
Month	1	6	12	all
CentralKZ	0.895444	0.747296	0.616885	0.753303
MadhyaPradesh	0.921624	0.595824	0.669981	0.731231
Missouri	0.772521	0.66588	0.655321	0.73539

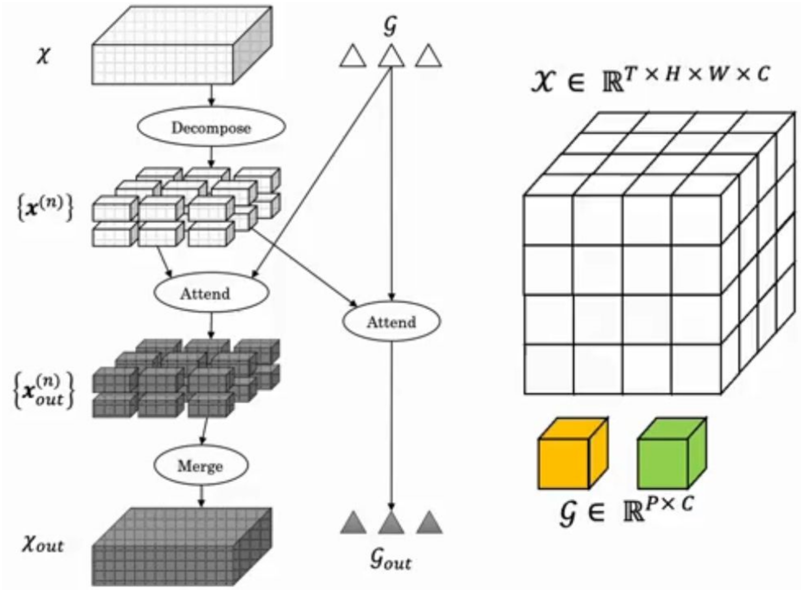
Multiclass classification

Accuracy				
	1	6	12	all
CentralKZ	0.82219	0.52503	0.42196	0.55029
MadhyaPradesh	0.83079	0.52262	0.38374	0.53915
Missouri	0.86291	0.60602	0.43279	0.6018

Earthformer

- transformer utilizing cuboid attention (spatio-temporal dependency) and global vectors $\mathcal{G} \in \mathbb{R}^{P \times C}$ initially designed for earthquakes prediction.

Params:
enc_depth = dec_depth = 1
num_heads = 4
cuboid_size = (4,4,4)

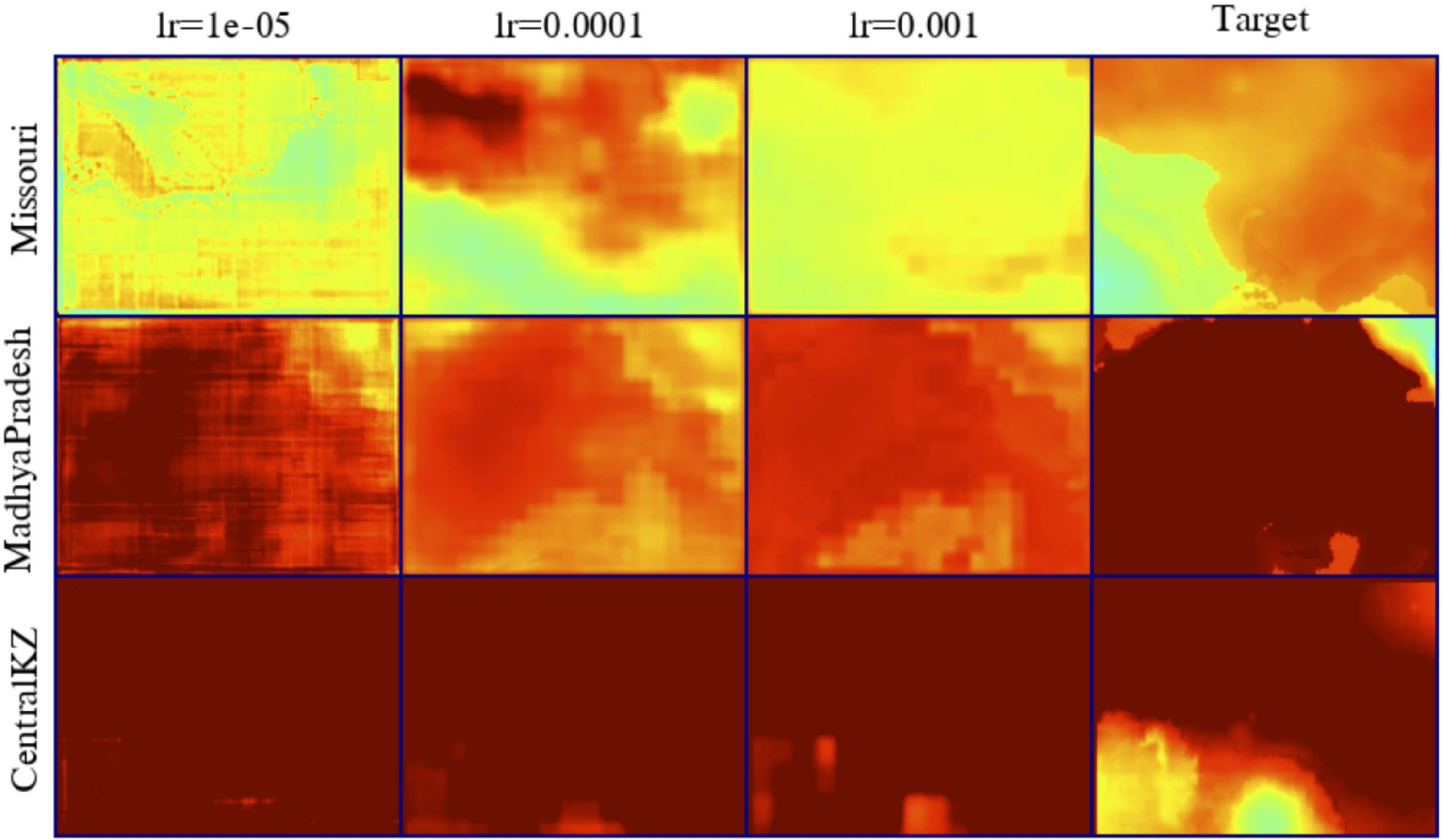


$$\mathcal{X}_{\text{out}}, \mathcal{G}_{\text{out}} = \text{CubAttn}_{\Theta}(\mathcal{X}, \mathcal{G}, \text{cuboid\_size}, \text{strategy}, \text{shift})$$

$$\Downarrow$$

$$\begin{aligned}
\{x^{(n)}\} &= \text{Decompose}(\mathcal{X}, \text{cuboid_size}, \text{strategy}, \text{shift}) \\
x^{(n)}_{\text{out}} &= \text{Attention}_{\Theta}\left(x^{(n)}, \text{Cat}(x^{(n)}, \mathcal{G}), \text{Cat}(x^{(n)}, \mathcal{G})\right) \\
\mathcal{G}_{\text{out}} &= \text{Attention}_{\Phi}\left(\mathcal{G}, \text{Cat}(\mathcal{G}, x^{(n)}), \text{Cat}(\mathcal{G}, x^{(n)})\right) \\
\mathcal{X}_{\text{out}} &= \text{Merge}\left(\left\{x^{(n)}_{\text{out}}\right\}_n, \text{cuboid_size}, \text{strategy}, \text{shift}\right)
\end{aligned}$$

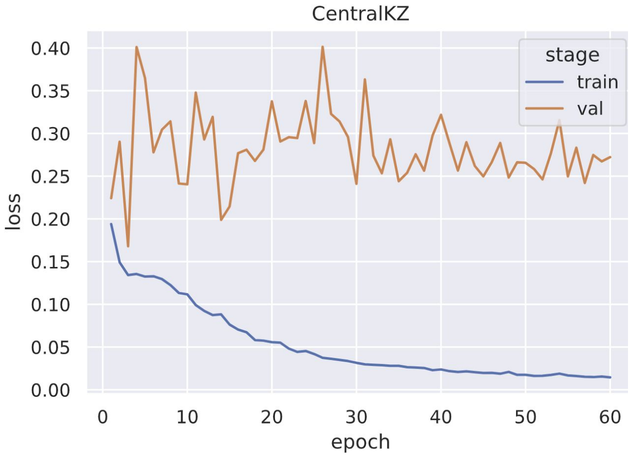
Training

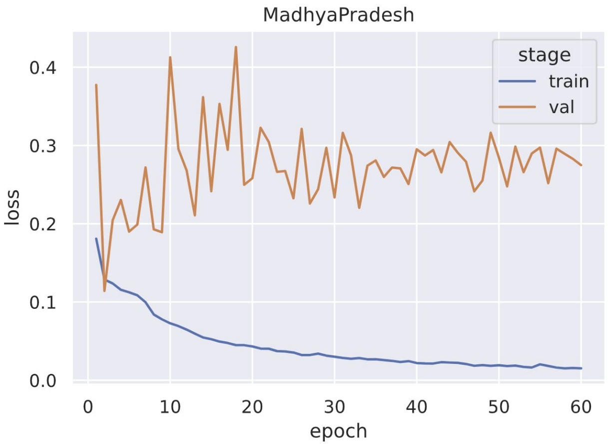

lr=1e-05
lr=0.0001
lr=0.001
Target
Missouri
MadhyaPradesh
CentralKZ

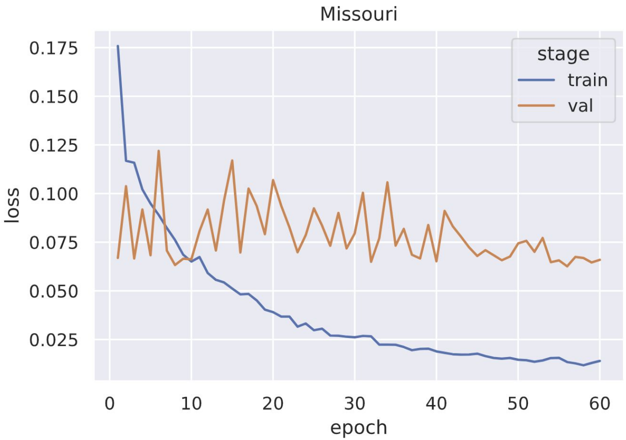
Why is the performance so poor ?

1. **Hyperparameter Tuning:** needed careful parameter search.
2. **Data Scarcity:** big model tend to have more training data to be on the level with baselines (model do not generalize well)
3. **Resource Constraints**: to train with minimal config we need about 18 GB GPU memory + slow training

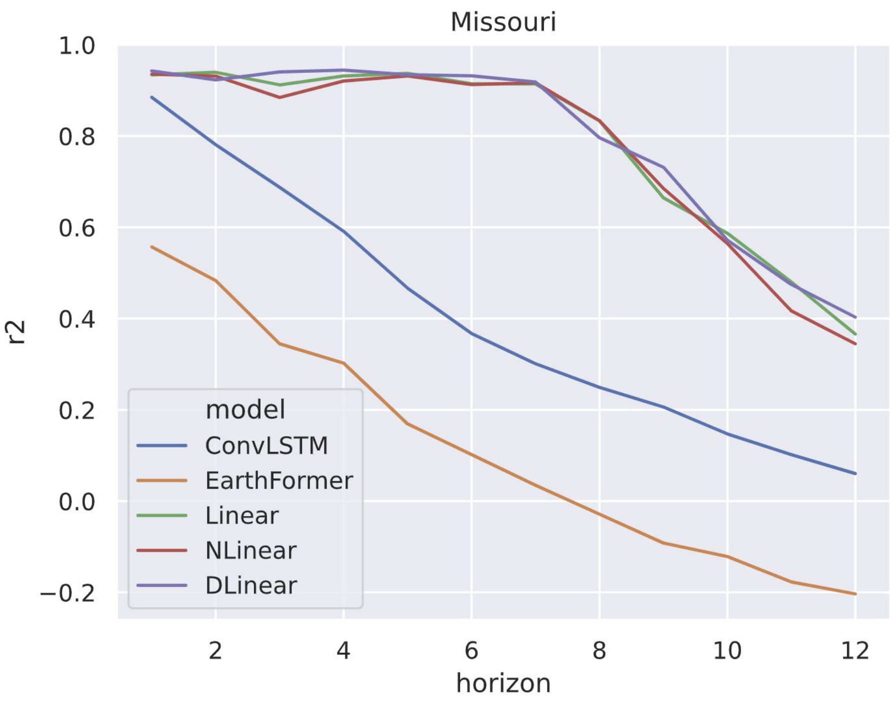
	time (min)
	median
model	
Linear	1.107213
NLinear	1.140015
DLinear	1.823801
ConvLSTM	13.326720
EarthFormer	30.614867







Results



multiclass acc_1_test

model	Linear	NLinear	DLinear	ConvLSTM	EarthFormer
data					
CentralKZ	0.227791	0.341454	0.221902	0.822194	nan
MadhyaPradesh	0.331068	0.330183	0.334281	0.830793	nan
Missouri	0.783762	0.451154	0.790940	0.862910	0.600142

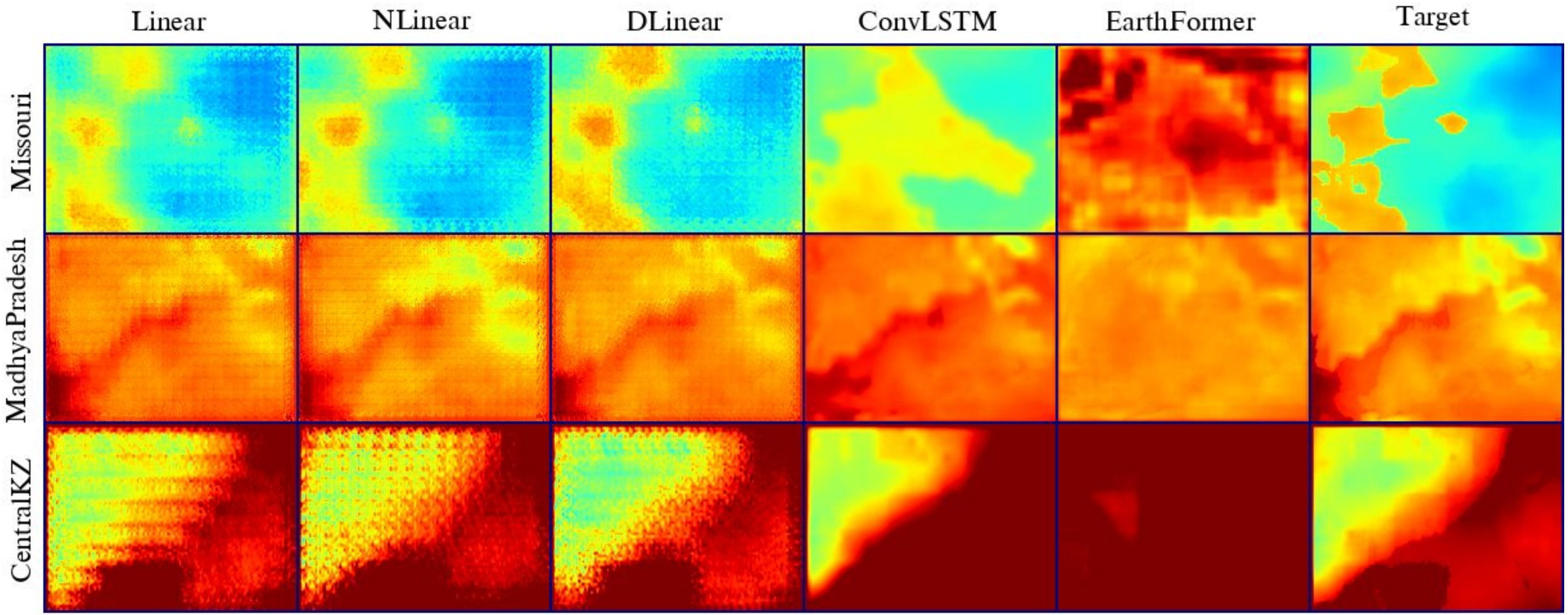
binary auc_1_test

model	Linear	NLinear	DLinear	ConvLSTM	EarthFormer
data					
CentralKZ	0.991082	0.989548	0.985282	0.895444	0.703174
MadhyaPradesh	0.993866	0.994716	0.994463	0.921624	0.900445
Missouri	0.996567	0.997907	0.997801	0.772521	0.671940

reg r2_1_test

model	Linear	NLinear	DLinear	ConvLSTM	EarthFormer
data					
CentralKZ	0.904825	0.905052	0.918569	0.853802	0.183712
MadhyaPradesh	0.883341	0.817849	0.882852	0.869848	-0.220142
Missouri	0.934189	0.935407	0.942752	0.884976	0.556904

Appendix A



HPO models, horizon: 1

Roles

1. Mikhail Kuznetsov - Linear model (improvements: emb + rtl) + pipelines
2. Victor Kozhevnikov - Transformer
3. Ivan Gurev - Linear model (expanding to H, W, C)
4. Artem Gorbarenko - ConvLSTM