STCast: Adaptive Boundary Alignment for Global and Regional Weather Forecasting

Authors: Hao Chen, Tao Han, Jie Zhang, Song Guo, Lei Bai

Affiliations: Hong Kong University of Science and Technology (HKUST) | Shanghai AI Laboratory

Conference: CVPR 2026 **(Highlight Paper)**

Introduction & Motivation

The Challenge: Kilometre-Scale Regional Forecasting

🎯 **Goal:** Achieve accurate, high-resolution (km-scale) regional weather forecasting.
🌐 **Significance:** Far-reaching socio-economic impact on disaster preparedness, agriculture, and energy.

⚠️ **Current Limitations:**

- **Dedicated Regional Models:** Neglect critical cross-regional dependencies.
- **Global Forecasts Downsampling:** Loss of fine-grained local details; training a 1km global model is computationally infeasible.
- **Hybrid Frameworks:** Use **static, imprecise regional boundaries** (contradicts the **Atmosphere–Ocean–Land–Biosphere Coupling Theory**).

Our Approach: STCast

An innovative AI-driven framework designed to overcome these limitations by enabling:

✅ **Adaptive regional boundary optimization**
✅ **Dynamic monthly forecast allocation**

"The true boundary for a region is not its neighbors, but the entire Earth."

Related Work & Our Position

Global-Regional Weather Coupling

NWP Models: Solve PDEs with fixed boundary conditions (e.g., sponge layers). Computationally expensive.

AI Approaches: Bypass PDEs by concatenating a **fixed neighborhood** from the global area to the region, leading to **static and local coupling**.

Recent Work (OneForecast): Concatenates neighboring low-res global forecasts with high-res regional variables.

Our Difference: Replace static concatenation with a **transformer-based framework** that adaptively models global-regional correlations, guided by a learned prior. This enables **dynamic boundary refinement** and **long-range dependency modeling**.

Data-Driven Weather Forecasting

Evolution: From NWP (PDEs) to data-driven methods (FNO, Transformers).

Current Trends: Neural operators (KNO, SFNO) and specialized neural networks (FengWu, Graphcast, FuXi).

MoE in Weather Forecasting: VAMoE, EWMoE extend MoE to forecasting.

Our Difference: Propose **TMoE (Temporal Mixture-of-Experts)** that explicitly partitions inputs by **month** and dynamically routes them to specialized temporal experts, capturing inter-month variability and intra-month correlation.

Proposed Method: STCast Framework

Overall Architecture

A unified deep-learning framework integrating four key tasks: **Global (Φ_g), Regional (Φ_r), Typhoon Track (Φ_{tc}),** and **Ensemble (Φ_{ens})** forecasting.

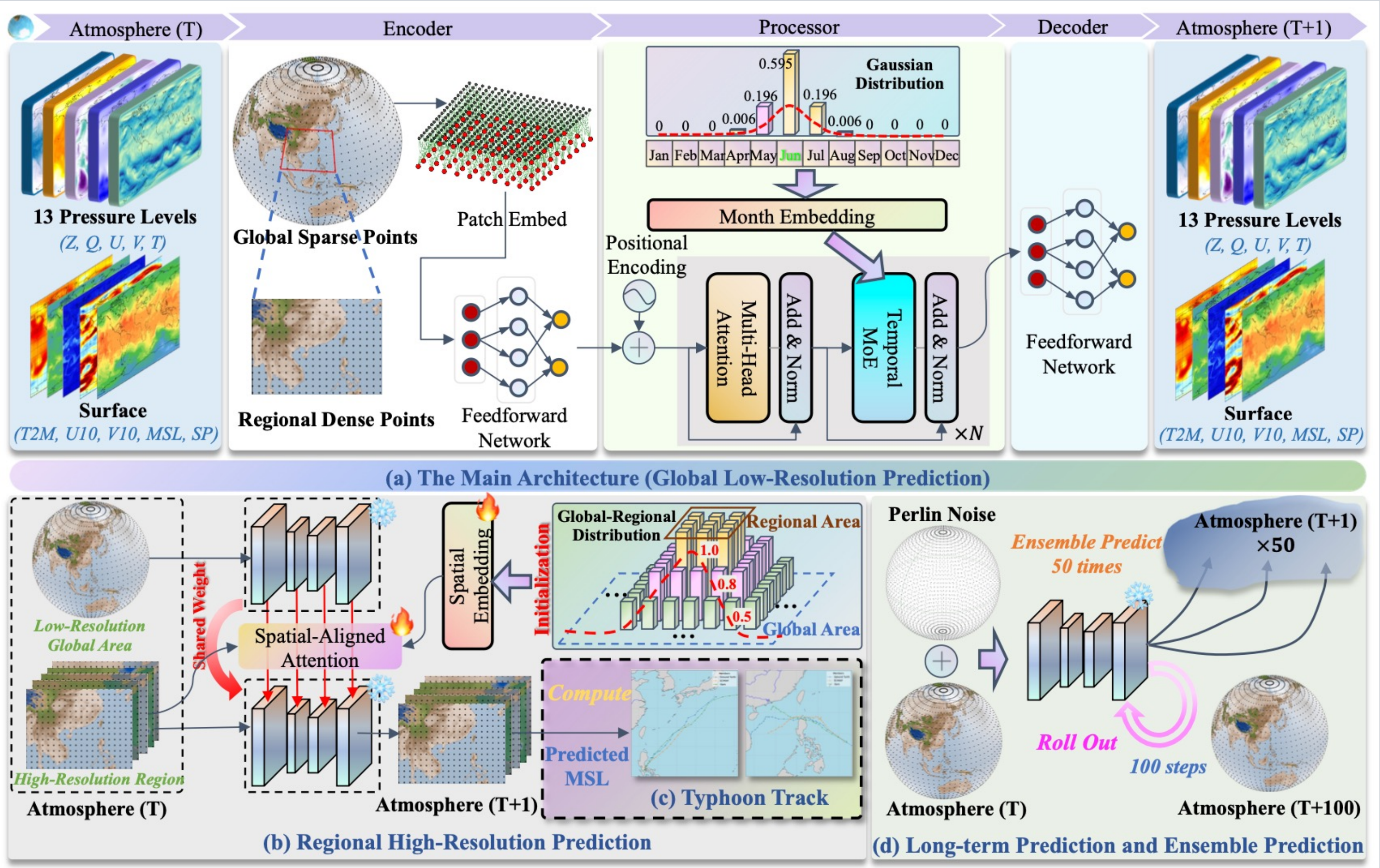
Core Components

- **Encoder:** Processes multi-level input atmospheric variables (13 pressure levels + surface).
- **Processor:** Combines window-based and self-attention to model local & global dependencies.
- **Decoder:** Generates future atmospheric states and forecasts.

Key Innovations

- **SAA (Spatial-Aligned Attention):** Enables dynamic coupling between global & regional domains.
- **TMoE (Temporal Mixture-of-Experts):** Effectively captures complex temporal variability.

Fig.1 Illustration of our method. *(a) The overall structure of low-resolution global weather forecasting; (b) High-resolution regional weather forecasting structure; (c) Typhoon track prediction; and (d) The long-term weather forecasting and ensemble weather forecasting.*



Key Innovation 1: Spatial-Aligned Attention (SAA)

Purpose

Dynamically couple global and regional features to enable accurate, high-resolution regional weather forecasting.

Limitation of Static Boundaries

Traditional methods fail to capture the true, dynamic global influence on a local region due to fixed spatial partitions.

01. Feature Fusion

Inputs: Global context as (Query, Key) and Regional data as (Value). Uses linear cross-attention to aggregate global information for regional prediction.

02. Global-Regional Dist.

Learns a dynamic distribution that quantifies how global features spatially relate to and influence the target regional area.

03. Physical Prior Init.

Initializes weights with**Great Circle Distance**and exponential decay to align with real-world spatial physics.

04. Adaptive Refinement

The model further refines the initialized distribution during training to learn the true, non-linear spatial dependencies.

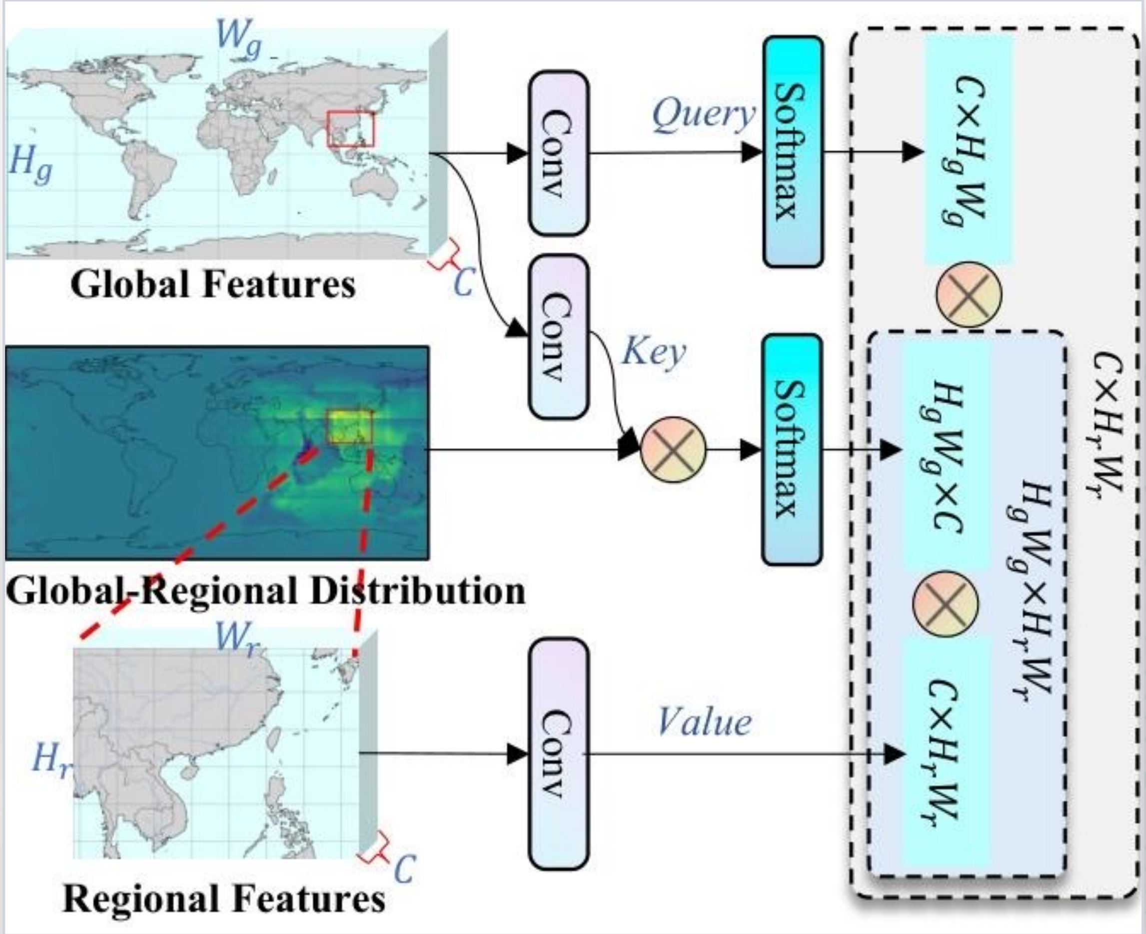


Figure: Schematic overview of the Spatial-Aligned Attention (SAA) module.

Key Innovation 2: Temporal Mixture-of-Experts (TMoE)

🎯 Purpose

Capture temporal variability by assigning monthly forecasting tasks to specialized experts.

🔍 **Observation:** Atmospheric dynamics vary significantly across different months.

01. Discrete Gaussian Dist.

Learns a Gaussian for each month, with peaks rotating to match input month.

02. Guided Expert Routing

Modulates weights to decay with increasing temporal distance from an expert.

03. Multi-Expert Activation

Activates multiple experts to enhance diversity and capture cross-month relationships.

04. Explicit Guidance

Uses month embedding to guide selection, preventing homogenization.

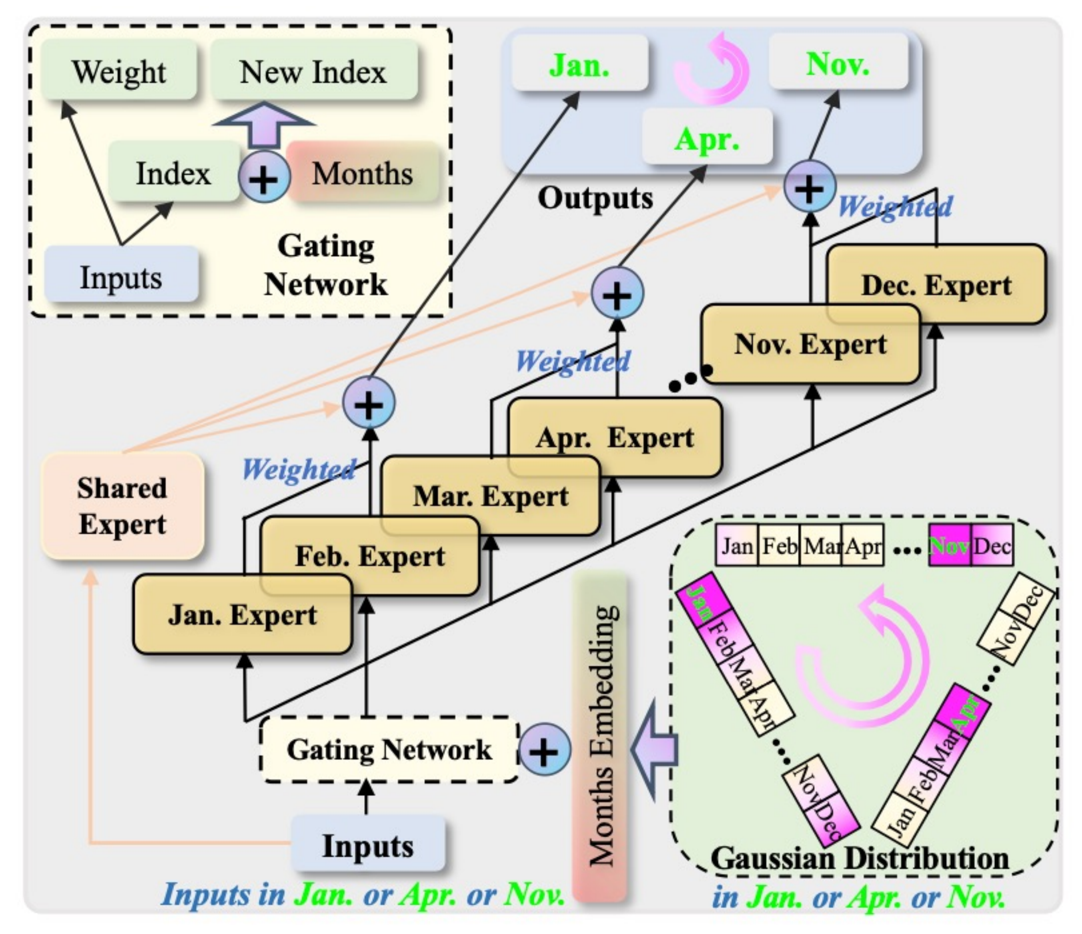


Figure: TMoE Module Architecture

Experimental Setup

Dataset

ERA5 reanalysis dataset (ECMWF) from 1979 to 2019.

Global Model: 1.4° resolution (70 variables).

Regional Model: 0.25° resolution over Eastern Asia (5 surface variables).

Implementation

Optimizer: AdamW with a learning rate of 0.0002.

Training: 100 epochs, batch size 16.

Hardware: 16 NVIDIA Tesla A100 GPUs.

Evaluation Metrics

RMSE (Root Mean Square Error): Lower is better.

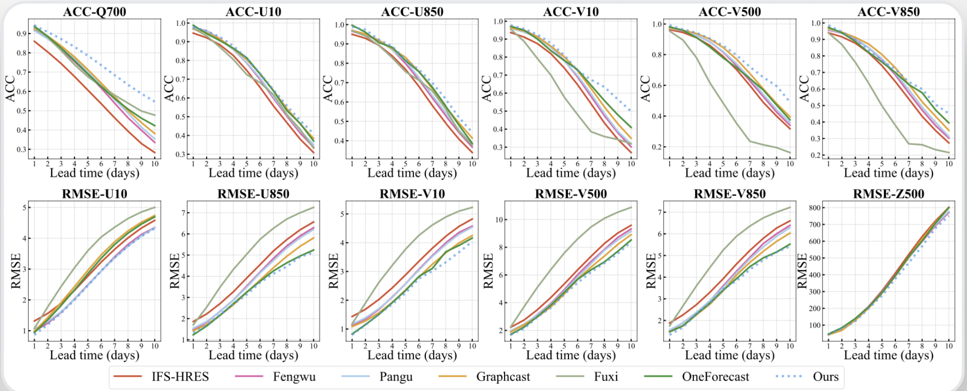
ACC (Anomaly Correlation Coefficient): Higher is better.

Main Results: Global & Regional Forecasting

Low-resolution Global Forecasts (1.4°)

•**STCast achieves state-of-the-art performance** across all lead times (6h–10d) vs. Pangu-Weather, Graphcast, Fuxi, and OneForecast.

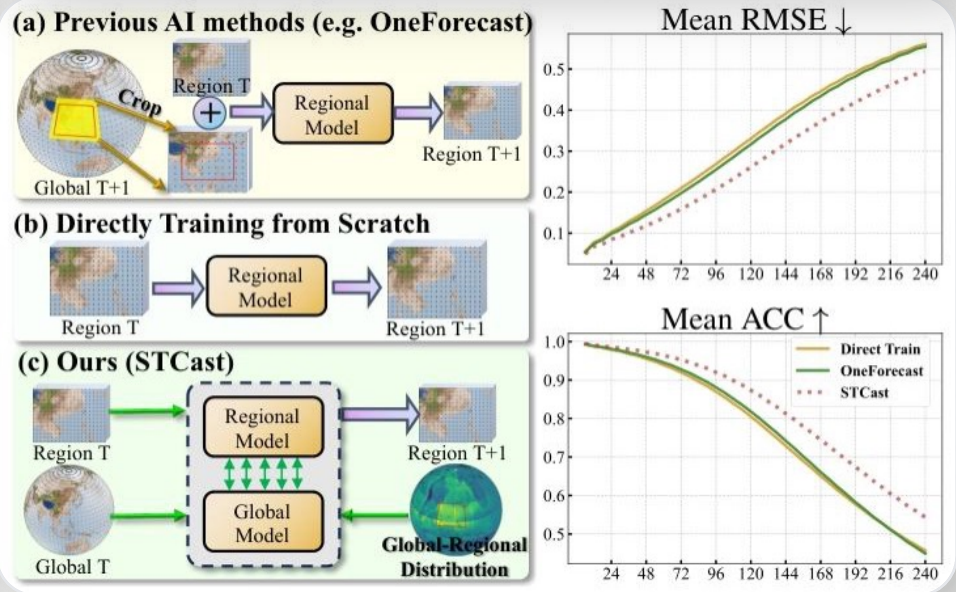
•**Key Finding:** Significant long-term prediction gains (*e.g.*, 10-day RMSE: STCast 0.5763 vs. OneForecast 0.4457), attributed to TMoE's month-specific training strategy.



High-resolution Regional Forecasts (0.25°)

•**STCast outperforms** both Direct Training (STCast w/o SAA) and OneForecast baselines.

•**Key Finding:** Dynamic boundary (SAA) yields **RMSE ↓0.05** & **ACC ↑0.1**, confirming the critical role of adaptive boundary modeling.



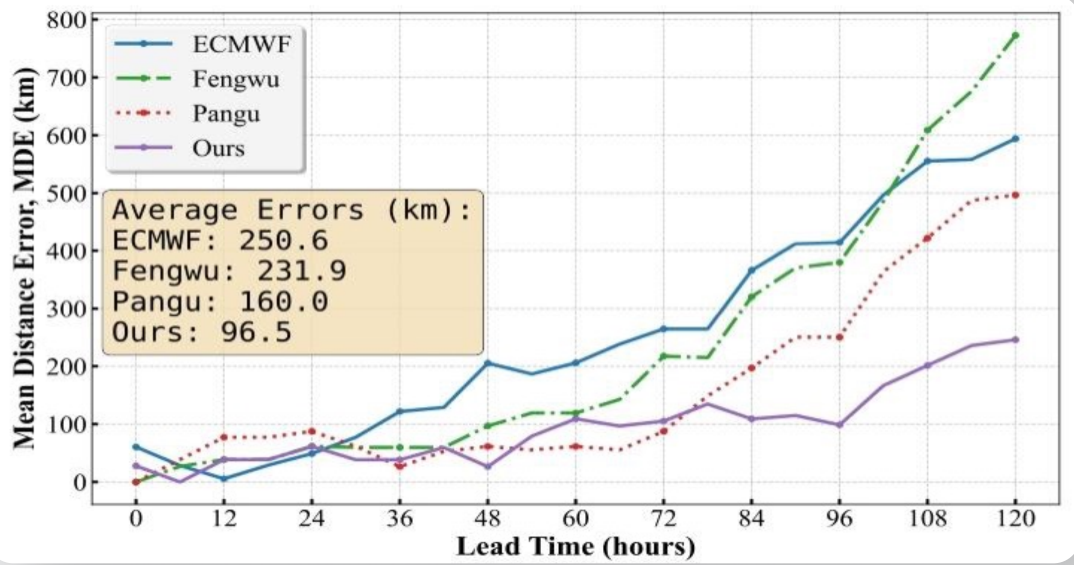
Main Results: Extreme Events & Long-term Forecasts

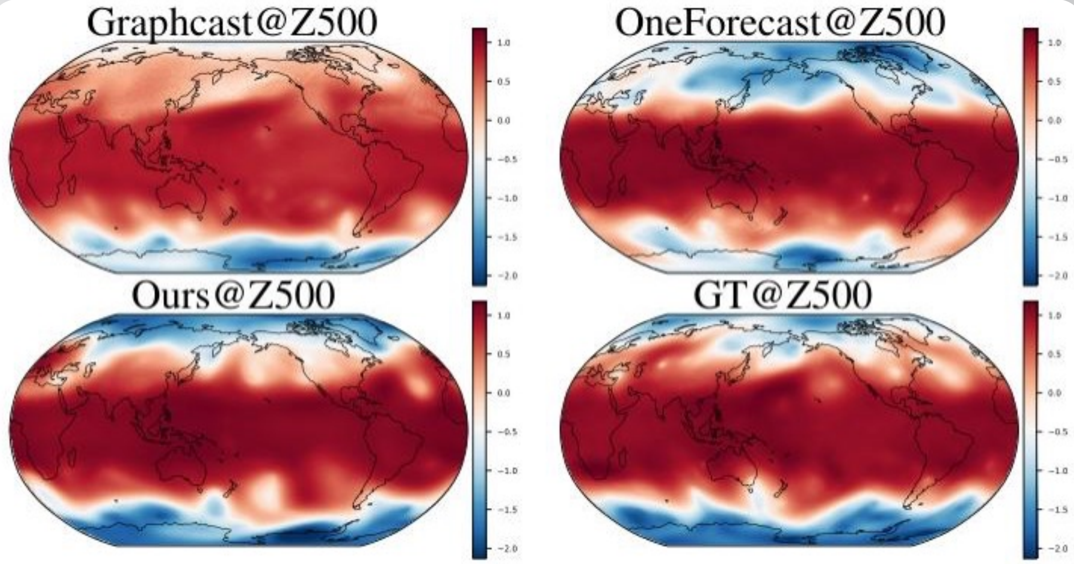
Extreme Events: Typhoon Track Prediction

STCast's 72-hour forecasts for Typhoon Ewiniar & Yinxing showed superior alignment with observed paths. For a 120h prediction of Yinxing, STCast achieved a **mean error of only 96.5 km**, outperforming Pangu-Weather (160 km) and other leading models in long-term accuracy.

Long-term & Ensemble Forecasting

•**100-day Z500:** Consistent alignment across all latitudes (vs. degraded performance at high latitudes for Graphcast).
•**Ensemble:** The ensemble variant achieved the **lowest error distribution** across initialization times.





Ablation Studies

Validating the Effectiveness of SAA and TMoE

Ablation on Regional Forecasts (w/o SAA)

Removing SAA degrades performance significantly (e.g., 10-day RMSE increases by 0.22). Removing the global-regional distribution initialization also causes a large drop.

Ablation on Global Forecasts (w/o TMoE)

Replacing TMoE with an MLP degrades performance. Removing the month embedding from TMoE causes the most substantial drop (e.g., 10-day RMSE increases by 0.13).

Conclusion: Every component (SAA, global-regional distribution, TMoE, month embedding) plays a critical role in enhancing the overall effectiveness of STCast.

Further Ablations

TMoE Expert Number

Increasing experts improves performance (12 experts > 6 > 4), but requires a balance with computational cost.

TMoE Embedding

Gaussian embeddings outperform one-hot embeddings, suggesting a more nuanced representation of temporal information.

SAA Distance Metric

Great Circle distance yields the lowest RMSE compared to Manhattan and Euclidean, aligning with Earth's spherical geometry.

Conclusion & Future Work

Key Conclusions

01. Proposed STCast: A novel framework that unifies global, regional, extreme event, and ensemble forecasting tasks into a single model.

02. Spatial-Aligned Attention (SAA): Successfully models adaptive global-regional boundaries, leading to a significant improvement in regional forecast accuracy.

03. Temporal Mixture-of-Experts (TMoE): Effectively captures complex monthly temporal patterns, enhancing the performance of both global and long-term forecasts.

04. State-of-the-Art Performance: STCast consistently outperformed all competing baselines across all evaluated weather tasks.

Future Directions

Broader Earth System Applications

Expand STCast beyond weather forecasting to address other critical Earth science challenges, including climate modeling, ocean current prediction, and air quality monitoring.

Model Optimization & Efficiency

Investigate further optimizations to the SAA and TMoE modules to improve inference speed and reduce computational costs, making it more practical for real-world operational use.

Thank You!

Code: https://github.com/chenhao-zju/STCast

Q & A