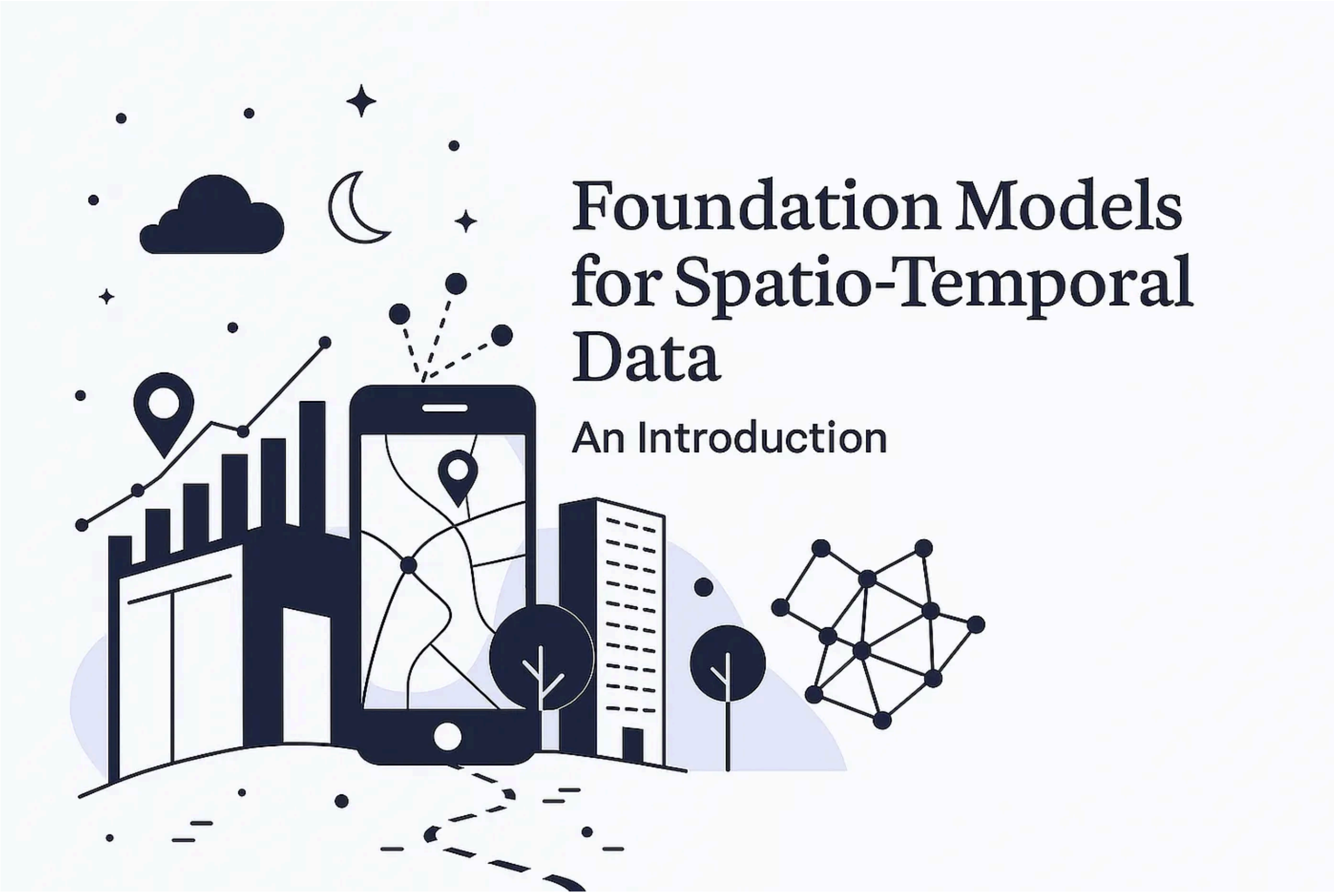

How AI Learns Space and Time: An Introduction to Spatio-Temporal Models

By Dev Mulchandani
MSSE San José State University, Fall 2025

Introduction

Everyday, vast amounts of data are coming out about our world, such as the movement of human populations, traffic flow patterns, and changes in the weather.

Because this type of data contains location and timeline characteristics, it is known as **Spatio-Temporal (ST) data.**

Spatio-temporal data is very powerful yet very challenging for machine learning models. Many changes happen every minute in cities, and human traffic behavior changes every minute as well. Most machine learning models are designed for single tasks and single sets of data only, and are unable to generalize or reason beyond what they were trained on.

That is when a new type of AI enters the scene:

Spatio-Temporal Foundation Models (STFMs)

These types of models attempt to unravel patterns in both space and time and apply the same principles of Foundation Models that made large language models (LLMs) revolutionary.

In this article, I will provide an overview of the paper **"Foundation Models for Spatio-Temporal Data Science: A Tutorial and Survey" (2025)** and explain why STFMs are such an enormous improvement compared to current data mining and machine learning methods.

Spatio-Temporal Data (ST Data)

Spatio-temporal data refers to any data type that can vary in both **space** and **time**.

Examples include:

- Maps on temperature changes in different regions
- Traffic sensors monitoring traffic flow and congestion during the day
- GPS coordinates and human movement routes
- Changes in satellite images over the seasons
- Natural events such as earthquakes, storms, or road closures

The paper groups ST data into:

- **Locations** (specific points such as lat/long)

- **Trajectories** (movement paths)
- **Events** (something happening at a place/time)
- **ST Rasters** (grid data like satellite images, weather maps)
- **ST Graphs** (road networks, sensor networks)

Such data is ubiquitous — but understanding it requires models that can handle relationships in both space and time.

Foundations of Foundation Models

Foundation Models (FMs), including GPT, BERT, and CLIP, have made immense progress in AI.
They are trained on very large datasets and can be adapted for many types of tasks.

Strengths of Foundation Models:

- Learn generalized representations
- Require little fine-tuning
- Process multimodal data
- Quickly adapt to new areas

Currently, the FM philosophy is also being extended into spatio-temporal data.

Spatio-Temporal Foundation Models (STFMs)

Conventional approaches are specific to tasks: for example, separate models for traffic, weather, or mobility.
STFMs aim to go beyond this.

According to the paper:

- **STFMs include space, time, and context**

- **They learn how patterns emerge over space and time**
- **STFMs can perform multiple tasks**, such as forecasting, imputing, event detection, anomaly detection, decision making, and many others
- **STFMs integrate the entire ST process pipeline** — from **Sensing → Managing → Mining → Real-world Applications**
- **They combine perception, optimization, and reasoning**, which is far better than classical deep learning

In simple terms, STFMs are like engines that understand movement and change in the real world.

STFM Workflow

The paper explains a workflow in four steps, and STFMs support all of them:

1. ST Data Sensing

Gathering real-world data:

- Sensors
- GPS
- Satellites
- Social media

Also includes **synthetic data generation**, where models like Trajectory-LLM and LLMob create trajectories or routes.

2. ST Data Management

Cleaning and organizing large ST datasets:

- Filling missing data
- Correcting noisy signals
- Retrieval systems (PLMTrajRec, UrbanLLM)
- Knowledge graph construction (UrbanKG)

LLMs are also used for natural language queries like:

> “Which areas had increased pollution last month around schools?”

3. ST Data Mining

This is the “ML/AI” portion.

* Perception

Learning patterns (examples: STEP, Pangu, UniST)

* Optimization

AI agents helping with tasks such as traffic control (AgentMove, TrafficGPT)

* Reasoning

Common sense, numeric, and causal reasoning (UrbanGPT, NuwaDynamics)

STFMs are much more than predictive models — they can **explain**, **optimize**, and **reason.**

4. Practical Applications

STFMs are currently used for:

- Weather forecasting (Pangu, GraphCast)
- Traffic forecasting and control
- Air quality forecasting
- Disaster response planning
- Self-driving cars
- Mobility analysis
- Simulation scenarios

These examples show the huge potential of STFMs in improving city systems and climate systems.

Representative Models

Some of the top STFM families are:

* LLM-based Models

UrbanGPT, TrafficGPT, GeoGPT

* Raster Models

Pangu, ClimaX

* Graph Models

Graph-based models like OpenCity

* Trajectory Models

TrajFM

* City-Scale Modeling

UrbanDiT

Each one focuses on different data types and problem areas.

⚠️ Challenges and Future Perspectives

The paper also points out several limitations:

1. Interpretability

STFMs can act like “black boxes” — hard to understand why they make certain predictions.

2. Multimodal Integration

Combining maps, text, images, and sensor data is still challenging.

3. Scalability

Earth-scale systems are extremely expensive to train.

4. Universal STFM Vision

Having one model for every spatio-temporal task is still a long-term challenge.

My Perspective

In my perspective, STFMs are the next level of progress in data mining and machine learning.They bring together tasks that were previously handled by multiple separate models, and they combine perception, reasoning, and decision-making in one unified approach.

It struck me as very interesting that STFMs are able not just to predict but also to simulate and optimize complex systems in the real world.However, interpretability poses difficult questions — especially now that these methods are being used in real-world decision-making like traffic control, disaster response, and travel planning.

Overall, STFMs look like the next generation of intelligent systems that can understand the physical world. This is what made this survey different and more interesting for me.

Conclusion

STFMs represent a powerful shift in how we as humans analyze spatio-temporal data.They unify sensing, management, mining, and applications — while learning general patterns across tasks.

This research paper survey cleared my confusion of how AI can move from simple and easy prediction to true reasoning about our physical world.

...

You can find my full project repository here :- GitHub Repo

...

Reference

Liang, Y., Wen, H., Xia, Y., Jin, M., Yang, B., Salim, F., Wen, Q., Pan, S., & Cong, G. (2025).
Foundation Models for Spatio-Temporal Data Science: A Tutorial and Survey.
arXiv:2503.13502

Machine Learning · Data Mining · Artificial Intelligence · Deep Learning · Data Science

Written by Dev Mulchandani