Extreme Precipitation Seasonal Forecast Using a Transformer Neural Network

Daniel Salles Civitarese
Daniela Szwarcman
Bianca Zadrozny
Campbell Watson

Climate change and extreme precipitation

– ***Climate change*** has been linked to the ***increase*** in intensity and frequency of ***extreme events***

– ***Extreme rainfall*** can cause flooding, crop damage, and ***widespread disruption*** to ecosystems

– ***Predicting*** such events ***in advance*** is critical for better ***preparedness***

– Predicting the likelihood of ***extreme precipitation*** at ***seasonal scales*** remains a ***significant challenge***

Extreme precipitation seasonal forecast

Machine learning may offer an answer:

- Recent works have shown that machine learning models offer ***encouraging performance***
- These models tend to rely on ***slowly-changing variables***, such as soil moisture and ENSO indices
- Most of these ***variables*** are ***publicly available***, but their degree of ***influence varies*** in space and time

Our proposal:

- A machine learning approach to ***forecasting*** the ***maximum precipitation*** in a week up to ***six months ahead***
- Apply the ***temporal fusion transformer*** (TFT) to improve results:
  - It combines ***multi-horizon forecasting*** with specialized components to ***select relevant inputs*** and ***suppress unnecessary features***
  - It produces ***quantiles*** as its outputs

Extreme precipitation seasonal forecast – I/O

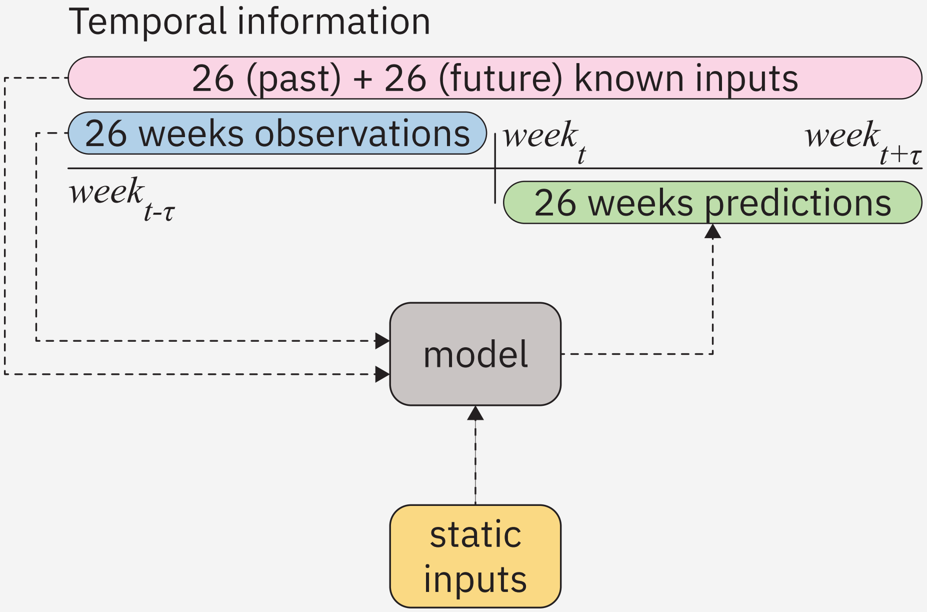
Target (green): maximum daily precipitation in each week

Input: structured into two classes:

– ***Static covariates*** (yellow) – e.g., lat/lon position

– ***Time-dependent features*** comprise:

- ***Observed inputs*** (blue) – e.g., historical rainfall
- ***Known inputs*** (pink) – e.g., day-of-week

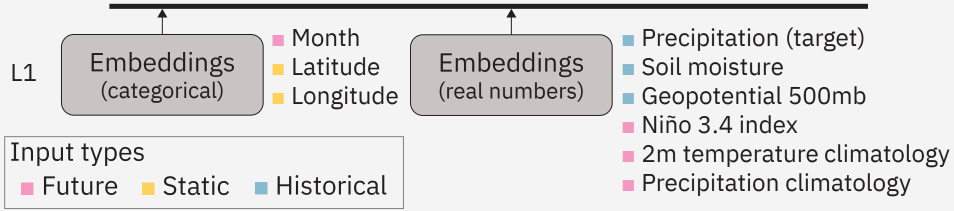


Multi-horizon forecast application

The transformer architecture

Temporal Fusion Transformer main parts:

– ***Embeddings*** for ***categorical*** and ***continuous*** variables

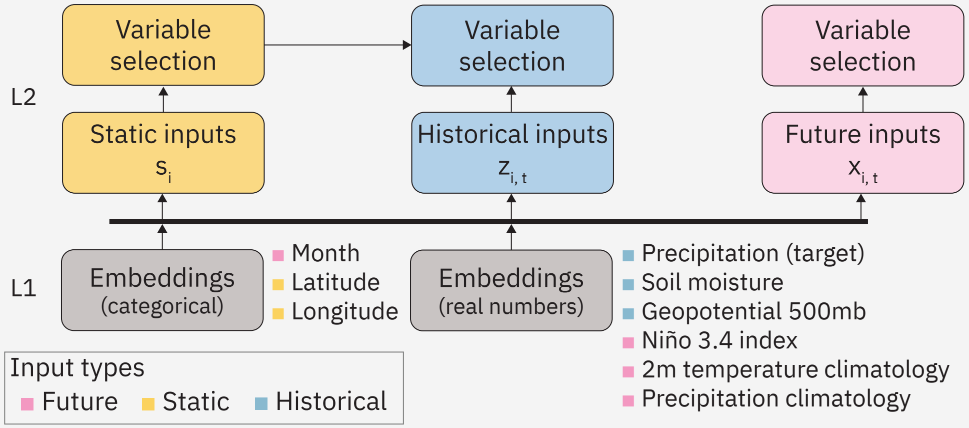


The transformer architecture

Temporal Fusion Transformer main parts:

– ***Embeddings*** for ***categorical*** and ***continuous*** variables

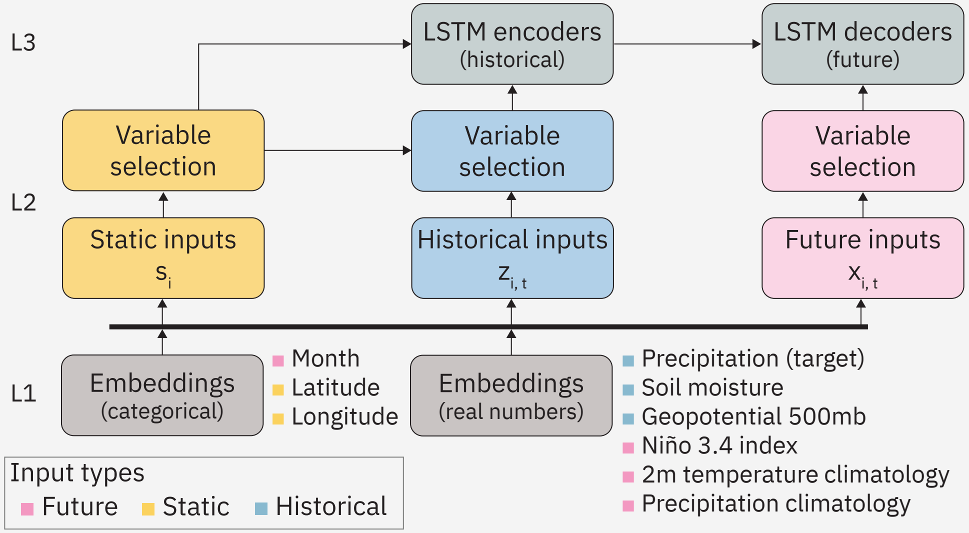
– ***Gating mechanisms*** select the most relevant parts of the data



The transformer architecture

Temporal Fusion Transformer main parts:

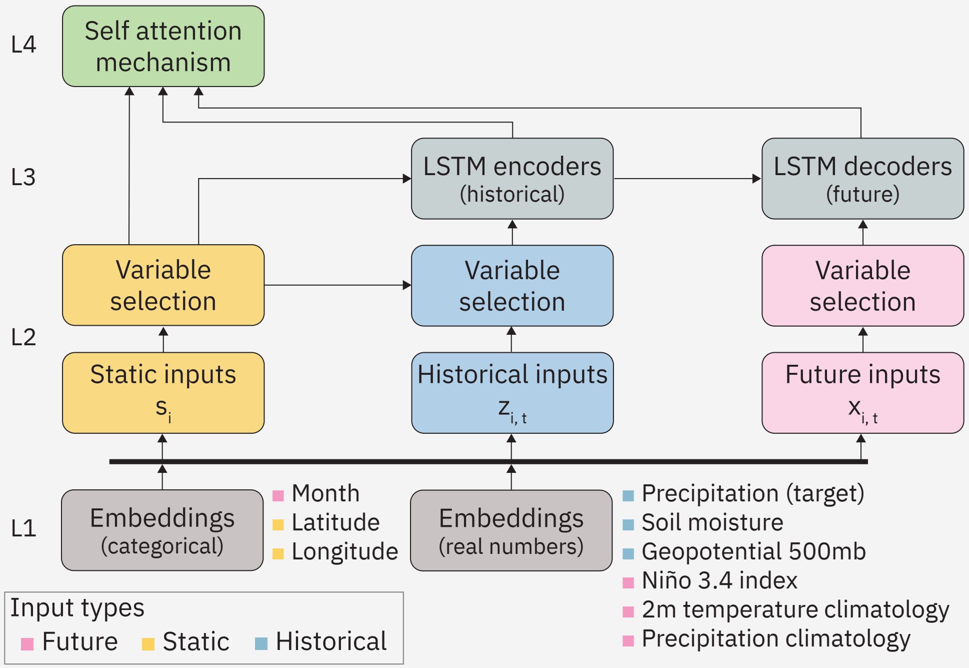
- ***Embeddings*** for ***categorical*** and ***continuous*** variables
- ***Gating mechanisms*** select the most relevant parts of the data
- ***LSTM*** nodes ***capture temporal correlations***



The transformer architecture

Temporal Fusion Transformer main parts:

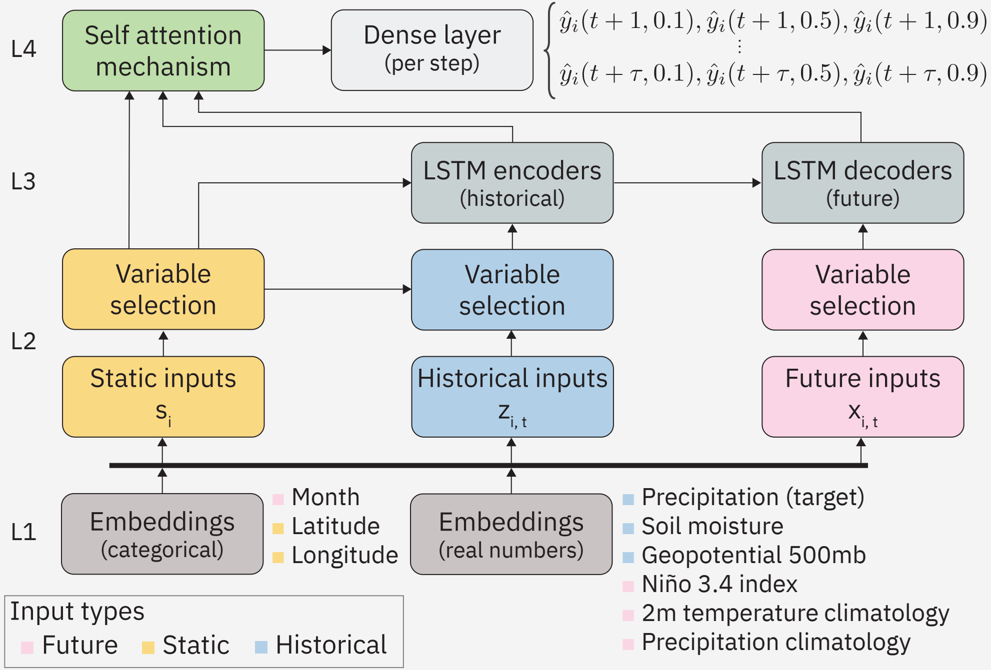
- ***Embeddings*** for ***categorical*** and ***continuous*** variables
- ***Gating mechanisms*** select the most relevant parts of the data
- ***LSTM*** nodes ***capture temporal correlations***
- ***Self-attention*** mechanism to ***learn long-term*** relationships across different time steps



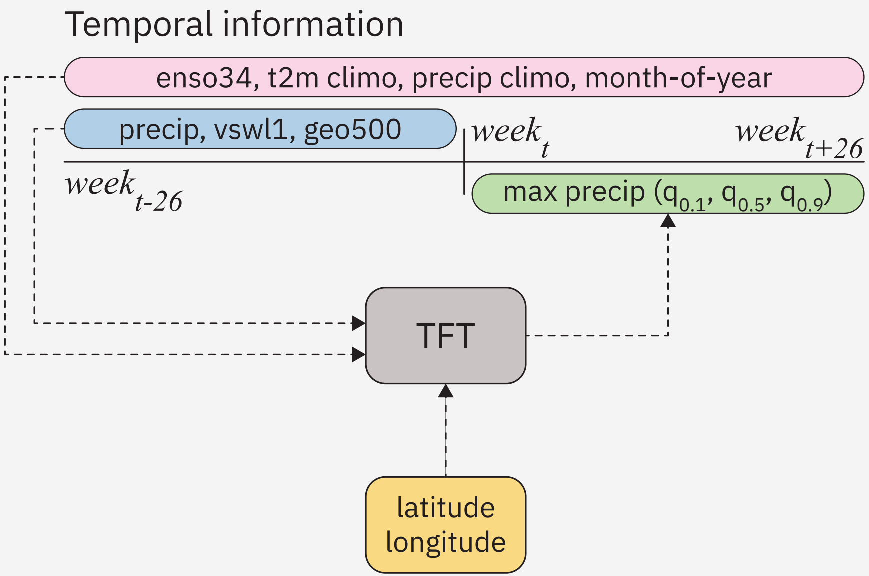
The transformer architecture

Temporal Fusion Transformer main parts:

- ***Embeddings*** for ***categorical*** and ***continuous*** variables
- ***Gating mechanisms*** select the most relevant parts of the data
- ***LSTM*** nodes ***capture temporal correlations***
- ***Self-attention*** mechanism to ***learn long-term*** relationships across different time steps
- Three ***quantile outputs***, 0.1, 0.5, and 0.9



Experiments – variables

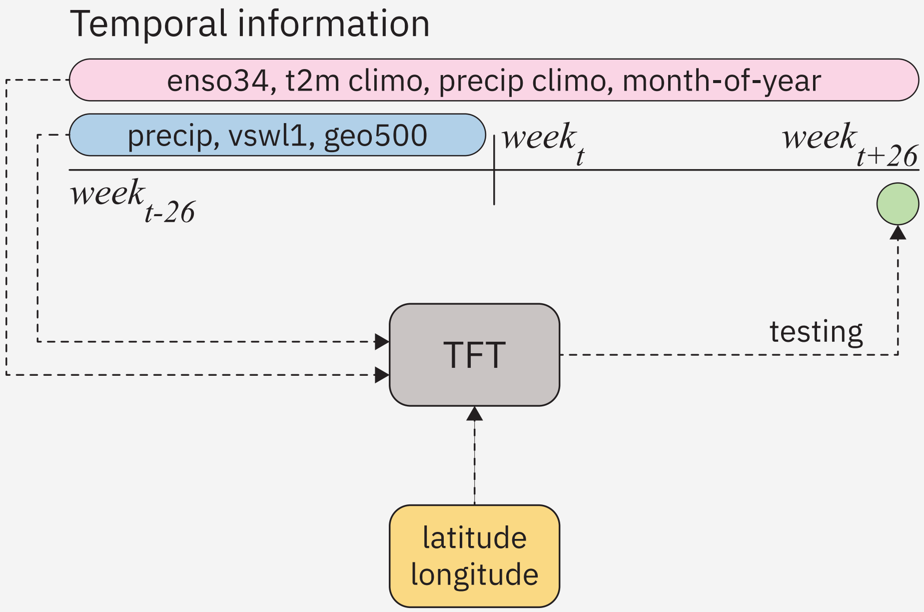


Datasets

- ***Historical data***
  - ***CHIRPS*** v2 (USGS/UC Santa Barbara)
    - Precipitation
  - ***ERA5*** reanalysis data (C3S)
    - Volumetric soil water layer 1 (single-level)
    - Geopotential 500 mb (pressure level)
- ***Future data***
  - ***Niño 3.4*** index (JAMSTEC)
  - ***Climatology***
    - 2-meter temperature (ERA5)
    - Precipitation (CHIRPS)
- ***Known data***
  - Month of the year
- ***Static data***
  - Latitude
  - Longitude

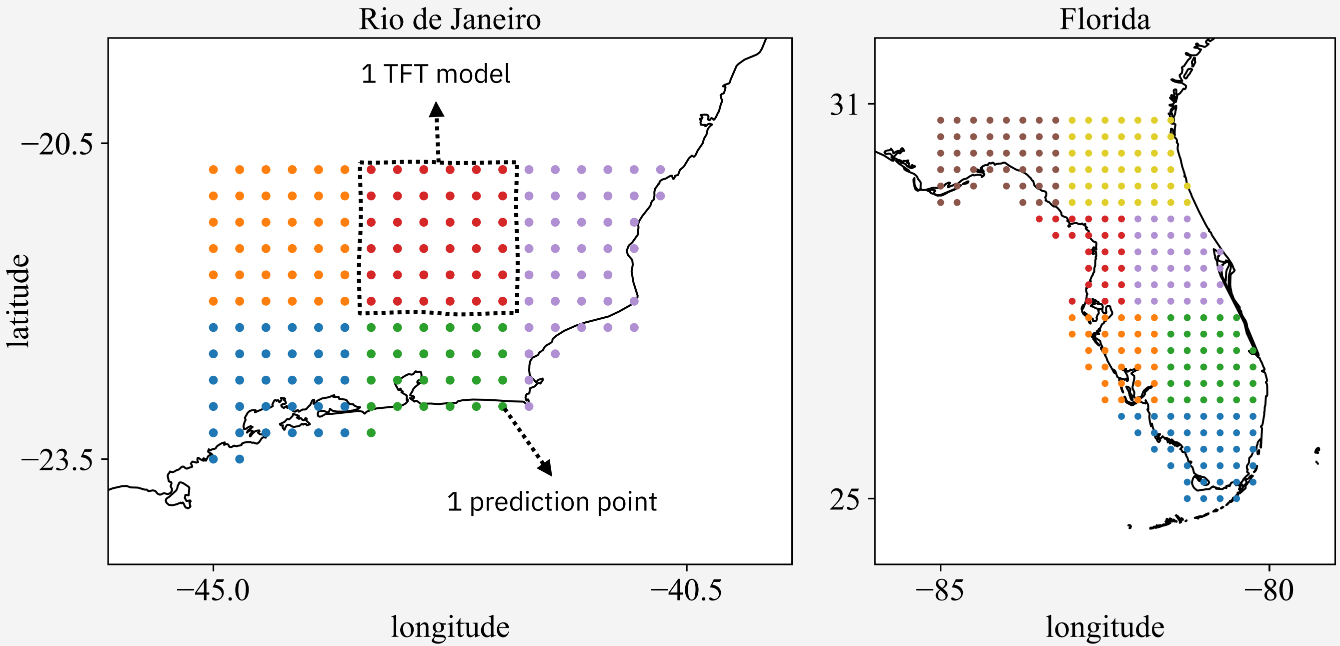
Experiments – pre-processing

- ***Spatial resolution***: 0.25
  - CHIRPS: spatial max pooling to go from 0.05 to 0.25
- ***Temporal resolution***: week
  - Weekly maximum for precipitation
  - Weekly mean for soil moisture and geopotential
- ***Dataset split***
  - 1981-2010: train
  - 2011-2014: valid
  - 2015-2019: test



Experiments – regions

- ***Rio de Janeiro***, Brazil
- ***Florida***, USA
- The ***areas*** were ***divided*** into smaller ***subregions***
- ***Each model*** is responsible for ***one subregion***



Experiments – q-risk results

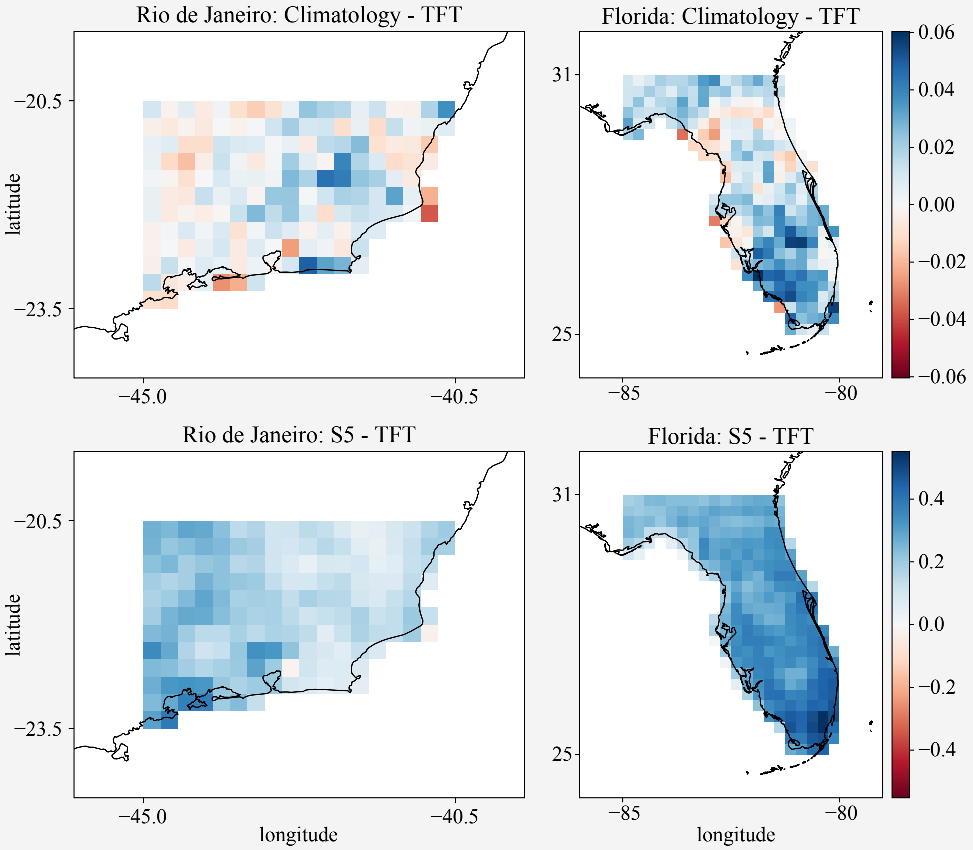
Q-risk (w_{26})

- Metric ***based*** on the ***quantile loss***
  - It divides the quantile loss by the sum of absolute values of the targets

Comparison	***Region***	***0.1***	***0.5***	***0.9***
$\frac{(climo - TFT)}{climo}$	Rio	2.45%	0.90%	***1.08%***
	Florida	-2.16%	-0.41%	***3.70%***
$\frac{(S5 - TFT)}{S5}$	Rio	3.71%	11.18%	***29.54%***
	Florida	5.70%	16.15%	***41.87%***

Experiments – q-risk maps

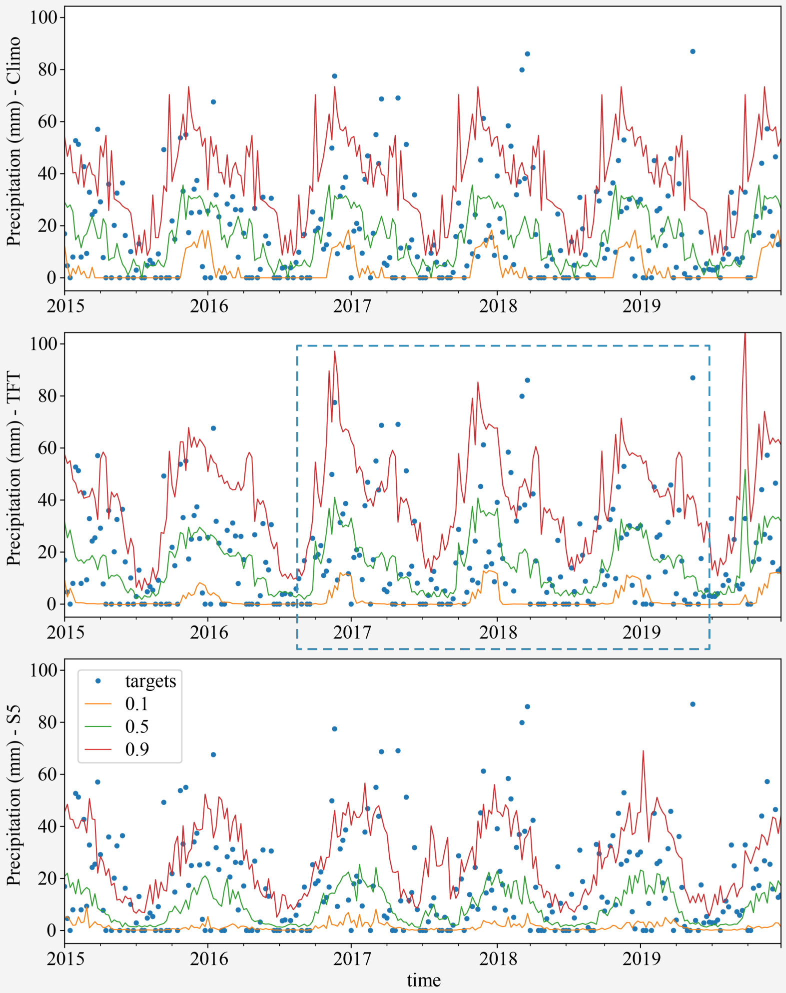
- ***Q-risk difference for quantile 0.9*** in each prediction point of the regions of interest
- Reference – TFT (w_{26})
  - ***Blue***: TFT is better
  - ***Red***: reference is better



Experiments – time-series predictions

- ***Predictions*** and ***targets*** in a location in Rio:
  - ***latitude*** -21.5, ***longitude*** -41.75

- Q-risk for quantile 0.9 (w_{26})
- (Climatology - TFT) → ***1.9%*** improvement
- (S5 - TFT) → ***15.8%*** improvement



Conclusions and future work

– Conclusions

- ***TFT*** generated significantly ***improved q-risks compared to the S5***
- Comparing the 0.9 quantile prediction in one location in Rio, we showed that TFT could accurately raise the quantile level and respond to changes that climatology cannot

– Future work

- Incorporate other input variables, such as dynamical model predictions
- Modify the model's input to ***support 2D*** spatial information
- Apply additional pre-processing, such as POD to capture teleconnections
- Use the interpretable multi-head attention block to identify connections between the input variables and extreme rainfall

© 2021 IBM Corporation