A High Spatiotemporal Resolution Framework for Urban Temperature Prediction Using IoT Data

Jingchao Yang,
Manzhu Yu, Qian Liu, Yun Li , Daniel Q. Duffy, and Chaowei Yang

Spatiotemporal Thinking, Computing and Applications
STC
https://www.stcenter.net/

AAG
AMERICAN ASSOCIATION of GEOGRAPHERS

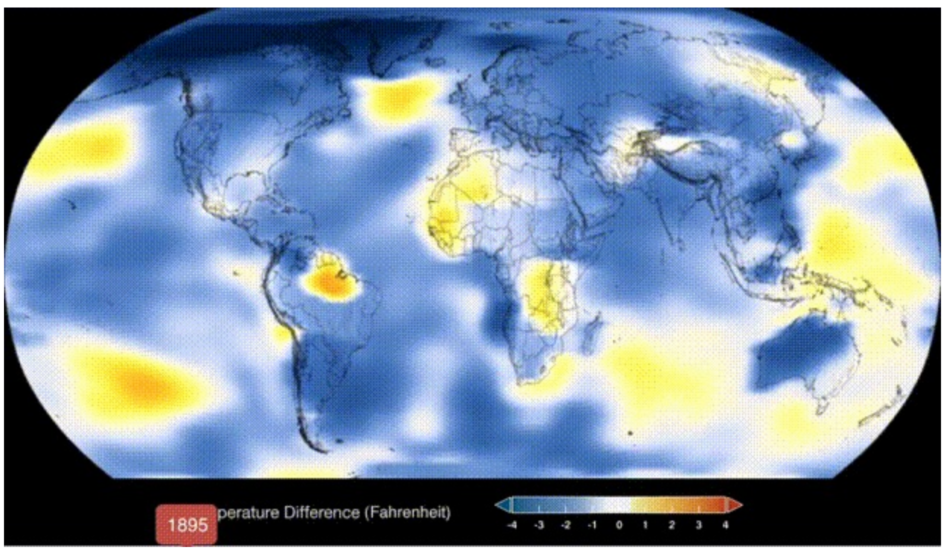
Outline

- Introduction
- Challenges
- Objectives
- Methodologies
- Results

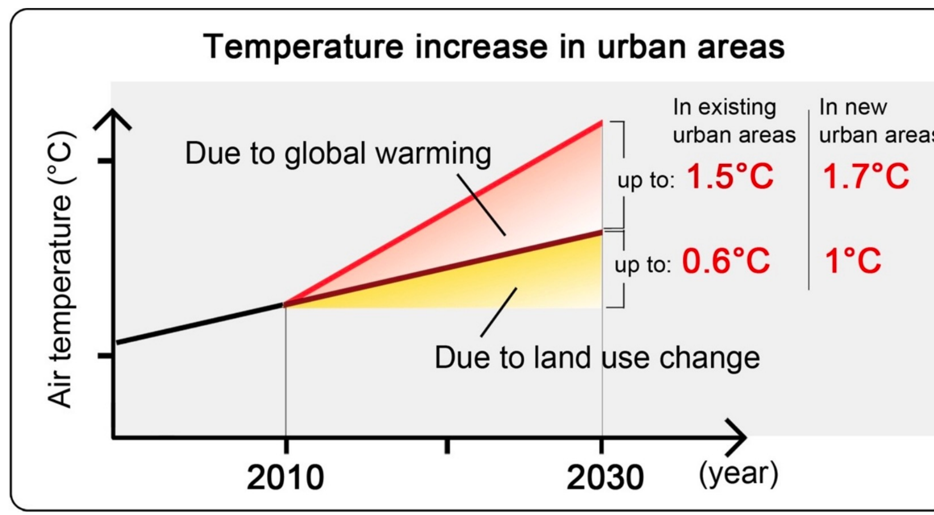
Spatiotemporal Thinking, Computing and Applications STC
https://www.stcenter.net/

Introduction: Intensified UHI

- Urban heat island (UHI) indicates an urban area is significantly warmer than its surrounding rural areas caused by human activities. The global warming trend is deteriorating the UHI by increasing the already higher temperatures in heat island areas (Luber and McGeehin, 2008; McCarthy et al., 2010; Harlan et al., 2014, Lee et al., 2017).



Intensified temperature increase
(*NASA, climate time machine*)

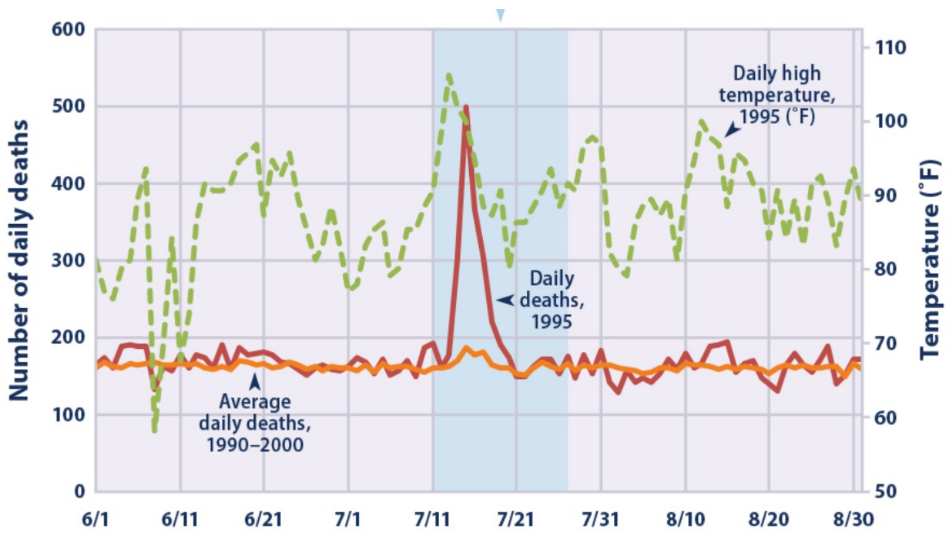


Hanoi City, Vietnam
(Lee et al., 2017)

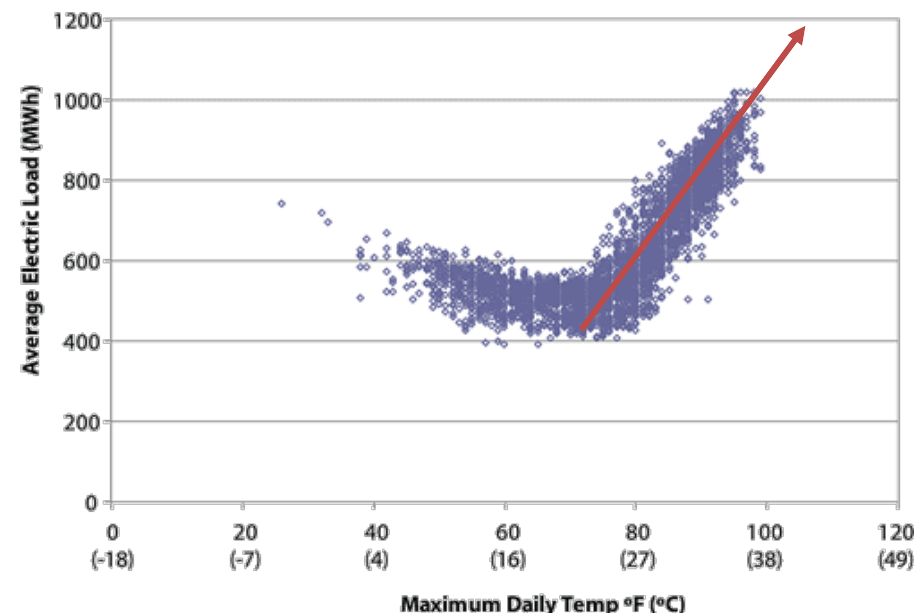
https://www.stcenter.net/

Introduction: UHI Related Issues

- Health: Heat was recorded as the underlying or contributing cause of death for nearly 8,000 Americans, between 1979 and 2010 (*EPA, Climate Change and Heat Islands*).
- Energy: Urban heat islands increase overall electricity demand, as well as peak demand (electrical load increases steadily at 68–77°F)



Heat wave in Cook County , Chicago
July 11 and July 27, 1995, 465 heat-related deaths
(*EPA, Climate Change Indicators: Heat-Related Deaths)*



New Orleans
(EPA, Heat Island Impacts)

https://www.stcenter.net/

Challenges: Microscale

- Study Scale
  - Limited data resources for micro-scale study (high resolution dataset) (Ye and Dai, 2018; Hu et al., 2016)
  - Traditionally: Numerical Weather Prediction can be computing intensive (time consuming) (Hewage et al., 2020)
- Numerical Models
  - Require very informative observations that are difficult to obtain in practice (Vautard et al., 2007; Stern et al., 2008)
  - Various types of parameters that need to be determined, resulting in limited accuracy (Pan et al., 2011; Ridder et al., 2012; T. Wang et al., 2012; Xu et al., 2017)
- Transferable learning/modeling
  - The imbalance data problem which occurs when peak or rare network target values have lesser representatives. (Ye and Dai, 2018)
  - Model across regions is challenging (different characteristics from city to city) (Hu et al., 2016)

Spatiotemporal Thinking, Computing and Applications
STC
https://www.stcenter.net/

Objectives

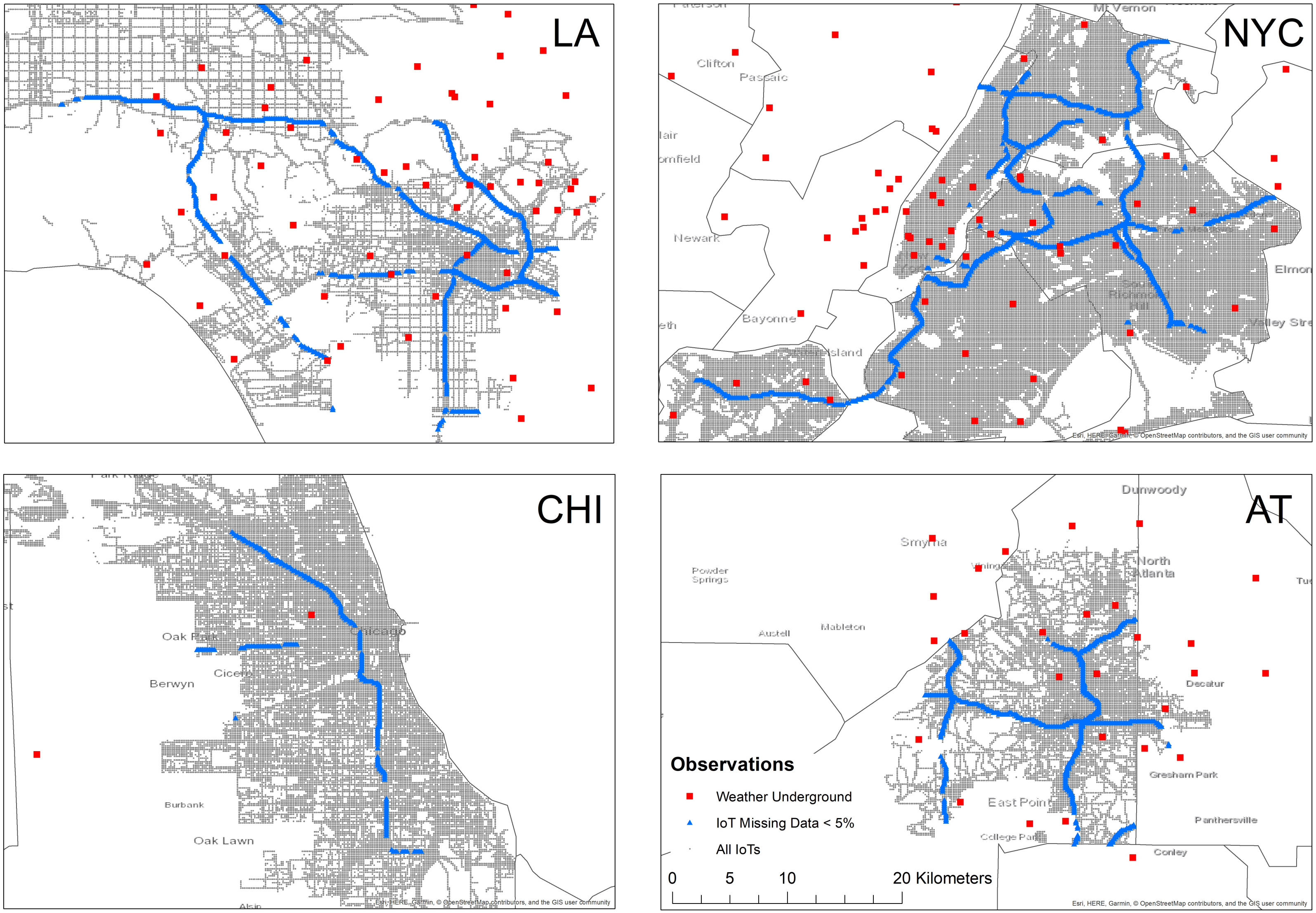
- An integrated **high-resolution dataset** for model input
  - Data fusion
- High-resolution UHI **Predictions**
  - High resolution predictions based on data fusion and DL
  - Analyzing transfer learning across different cities and building prediction models to fit different regions by tuning with target region weather parameters
  - Result interpolation, producing full coverage grid-based (200m*200m) hourly product for target region

Dataset

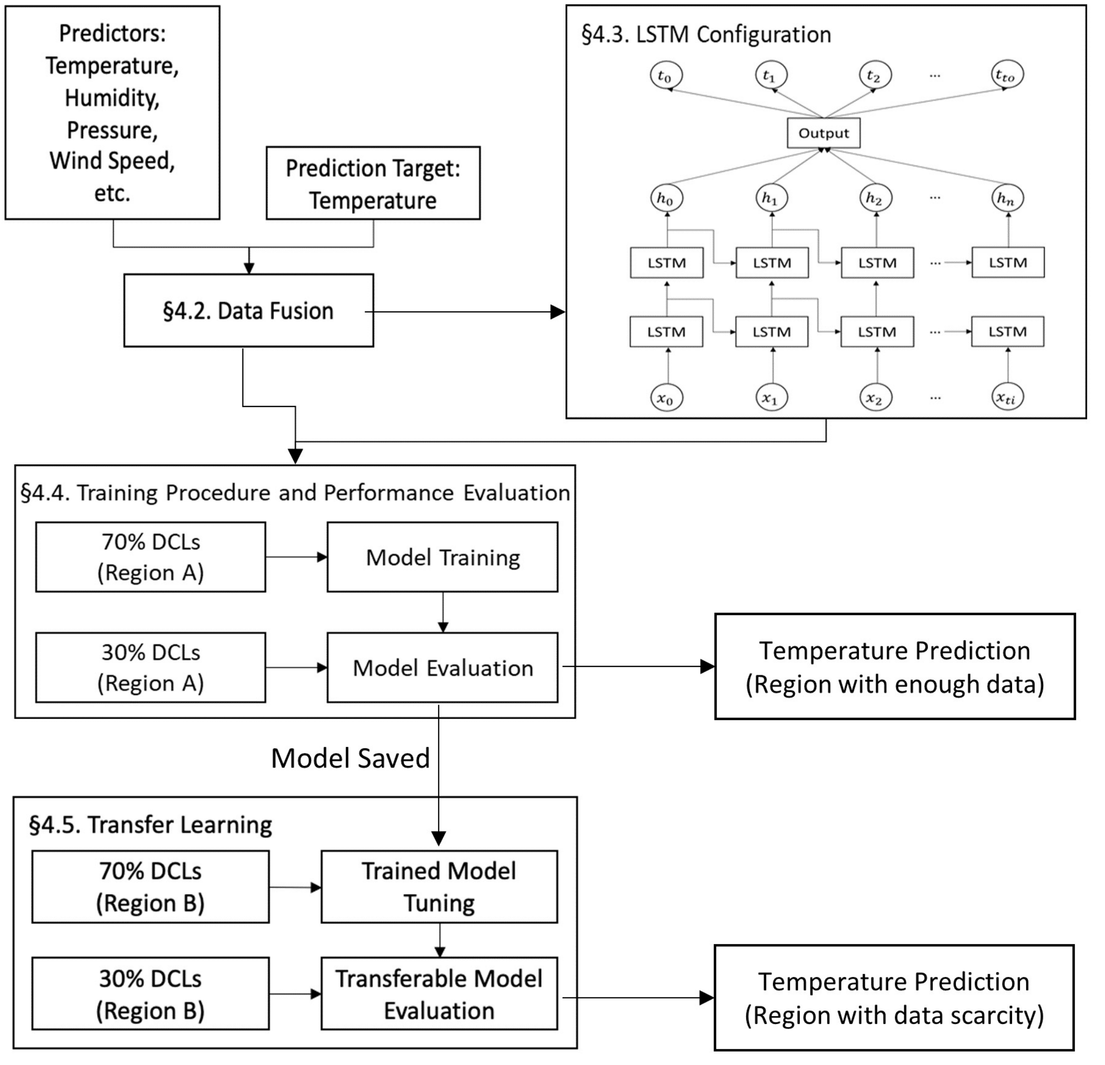
Dataset	Org. & Data Source	Spatial Resolution	Temporal Resolution	Time Coverage	Variable	Role
GeoTab IoT	GeoTab, Inc. https://data.geotab.com/weather/temperature	927 DCLs (LA) 1045 DCLs (NYC) 508 DCLs (AT) 269 DCLs (CHI)	hourly	2019-2020	Temperature	Data fusion. Both predictors and prediction target.
Weather Underground (WU)	IBM Weather Underground https://www.wunderground.com/about/data	100 stations (LA) 136 stations (NYC) 27 stations (AT) 12 stations (CHI)	hourly	2019-2020	Temperature, humidity, pressure, wind speed, UV index, etc.	Data fusion. Predictors.

https://w

Study Area



Framework


Predictors:
Temperature,
Humidity,
Pressure,
Wind Speed,
etc.
Prediction Target:
Temperature
§4.2. Data Fusion
§4.3. LSTM Configuration
Output
LSTM
§4.4. Training Procedure and Performance Evaluation
70% DCLs
(Region A)
Model Training
30% DCLs
(Region A)
Model Evaluation
Temperature Prediction
(Region with enough data)
Model Saved
§4.5. Transfer Learning
70% DCLs
(Region B)
Trained Model
Tuning
30% DCLs
(Region B)
Transferable Model
Evaluation
Temperature Prediction
(Region with data scarcity)

https://www.stcenter.net/

Spatial Data Fusion

$$P = WU[x_0, \dots, x_k, y_0, \dots, y_k]^2 + IoT[x_0, \dots, x_i, y_0, \dots, y_i]^2 \tag{1}$$

$$= \begin{bmatrix} p_{[0,0]} = WU[x_0]^2 + IoT[x_0]^2 + WU[y_0]^2 + IoT[y_0]^2 & \cdots & p_{[k,0]} \\ \vdots & \ddots & \vdots \\ p_{[0,i]} & \cdots & p_{[k,i]} \end{bmatrix}$$

$$M = \begin{bmatrix} IoT[x_0] & IoT[y_0] \\ \vdots & \vdots \\ IoT[x_i] & IoT[y_i] \end{bmatrix} \begin{bmatrix} WU[x_0] & \cdots & WU[x_k] \\ WU[y_0] & \cdots & WU[y_k] \end{bmatrix} \tag{2}$$

$$= \begin{bmatrix} m_{[0,0]} = WU[x_0] + IoT[x_0] + WU[y_0] + IoT[y_0] & \cdots & m_{[k,0]} \\ \vdots & \ddots & \vdots \\ m_{[0,i]} & \cdots & m_{[k,i]} \end{bmatrix}$$

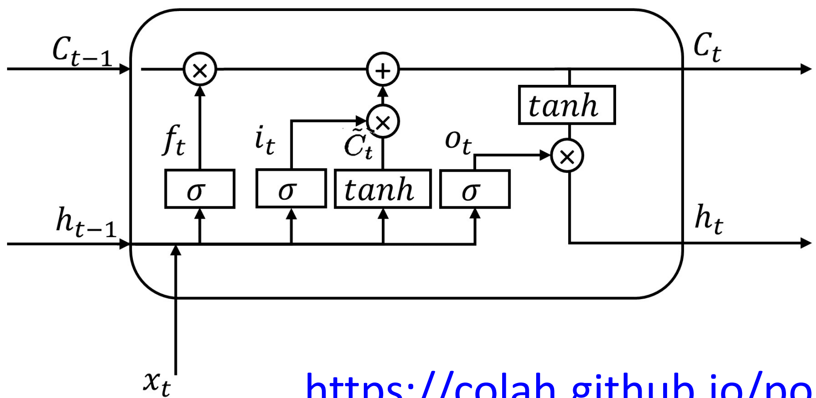
$$D(S_i, S_k) = \sqrt{(P - 2M)} \tag{3}$$

Matrix manipulation for Euclidean Distance based nearest neighbor calculation

https://www.stcenter.net/

LSTM

x_t: input
f_t: forget gate
i_t: input gate
$\tilde{C}_t$: cell update
C_t: cell state
o_t: output gate
h_t: output



$f_t = \sigma(w_f[h_{t-1}, x_t] + b_f)$	(4)
$i_t = \sigma(w_i[h_{t-1}, x_t] + b_i)$	(5)
$o_t = \sigma(w_o[h_{t-1}, x_t] + b_o)$	(6)

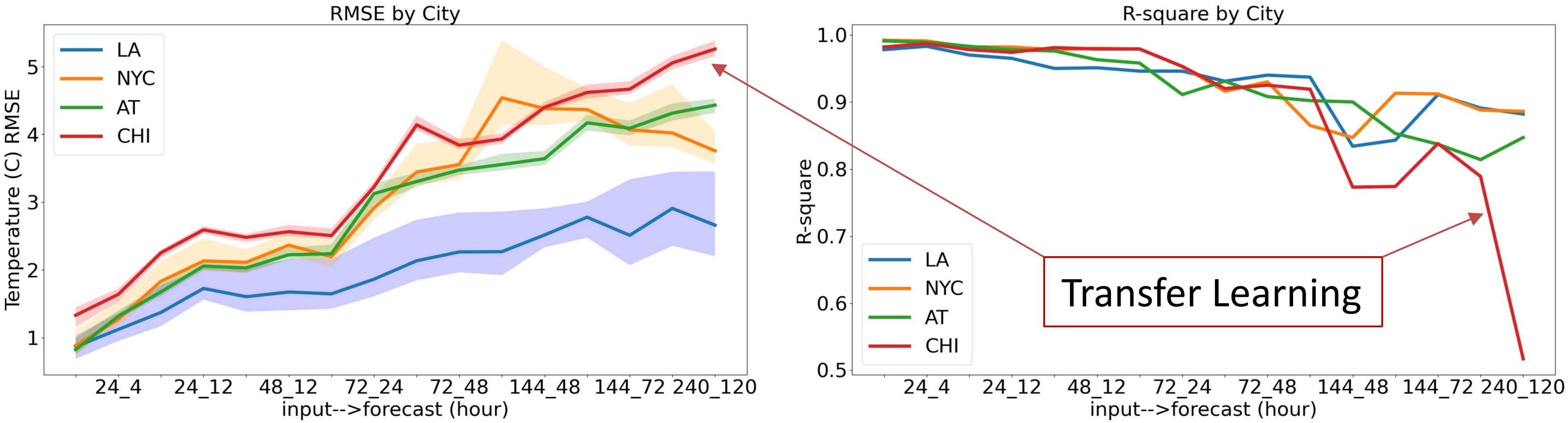
https://colah.github.io/posts/2015-08-Understanding-LSTMs/

ML Model Comparison for 12-step Prediction with 24-hour Input

ARIMA	XGBoost	LSTM
$R^2 = 0.94$	$R^2 = 0.94$	$R^2 = 0.97$
Max RMSE = 5.51	Max RMSE = 2.13	Max RMSE = 1.94
Min RMSE = 4.25	Min RMSE = 1.83	Min RMSE = 1.43

Spatiotemporal Thinking, Computing and Applications STC https://www.stcenter.net/

Model Local Testing Evaluation



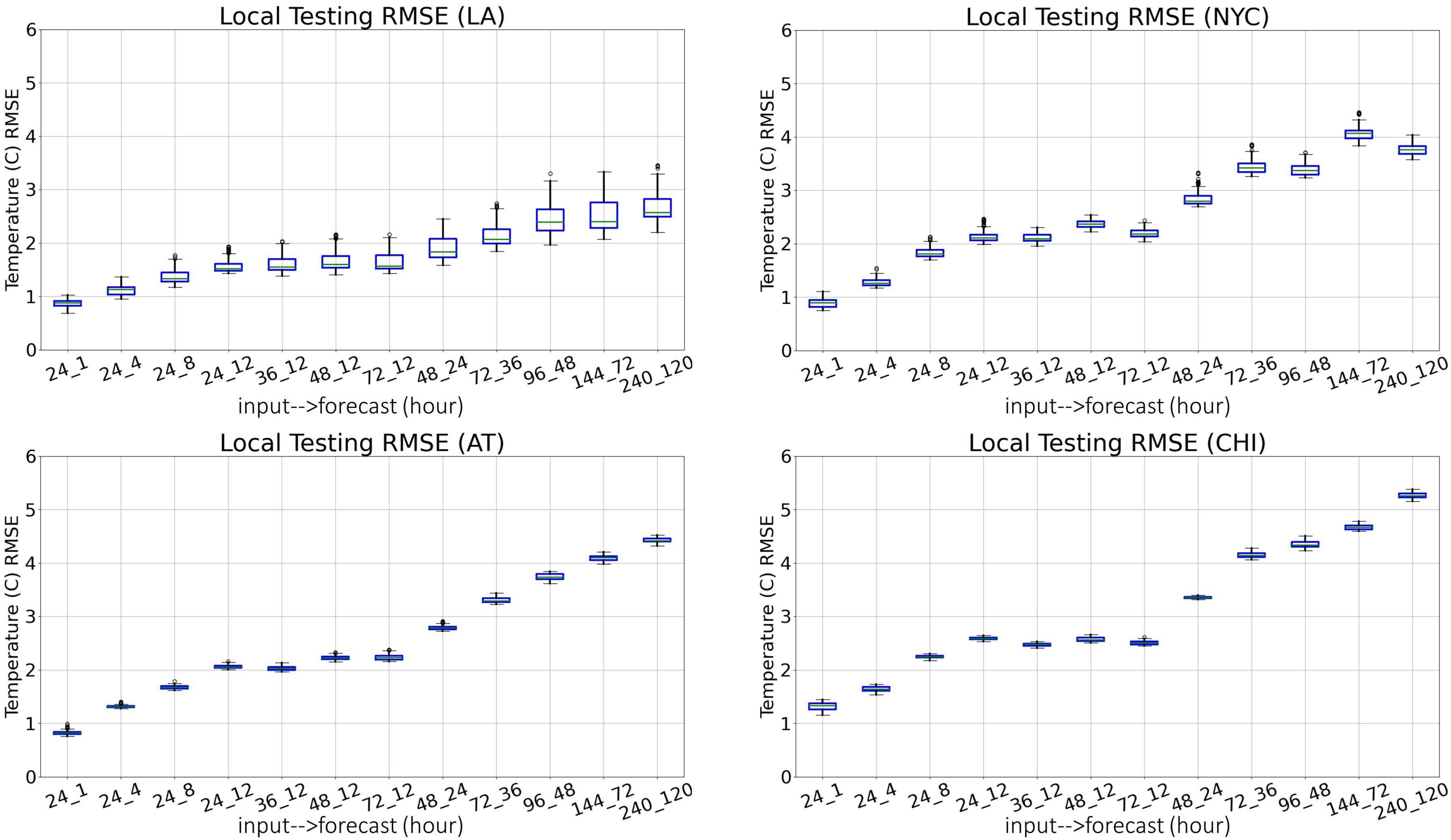
(a) (b)

Multi-step prediction from the proposed framework in comparison to the best results reviewed by Cifuentes et al. (2020)

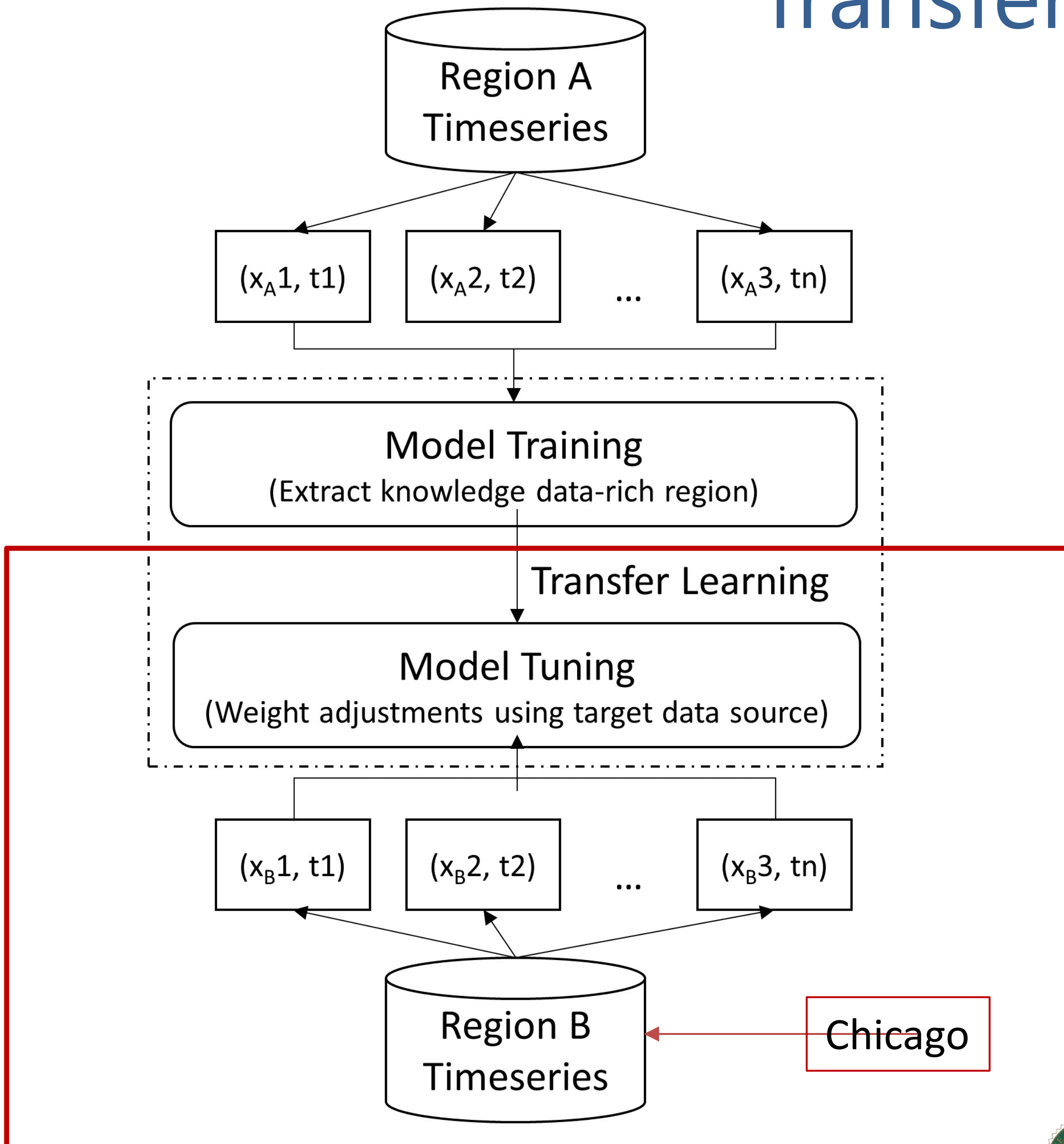
MAE (°C)	Previous Studies	Proposed Framework (study region)	Max Accuracy Change
4-step(hour)	1.20 SVM; Chevalier et al., 2011	0.91 (LA) 0.88 (NYC) 0.96 (AT)	-26.7%
8-step	1.62 Ward MLPNN; Smith et al., 2009	0.98 (LA) 1.37 (NYC) 1.22 (AT)	-39.5%
12-step	1.87 Ward MLPNN; Smith et al., 2009	1.13 (LA) 1.45 (NYC) 1.58 (AT)	-39.6%

https://www.stcenter.net/

Multi-step LSTM With Multivariate Prediction Result Evaluation

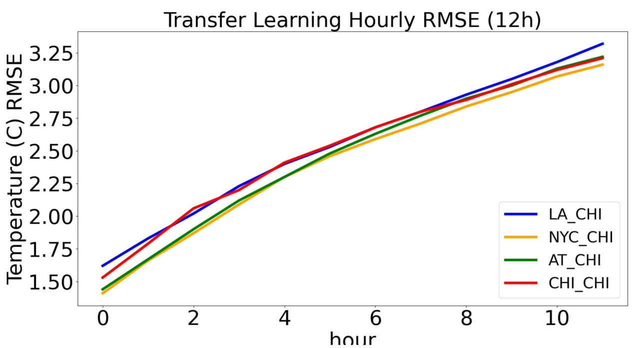


Transfer Learning

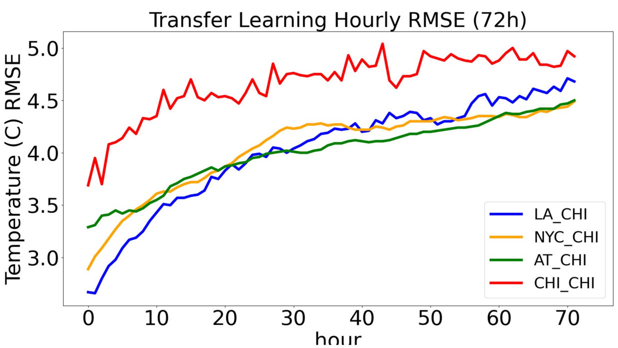


Transferable Prediction

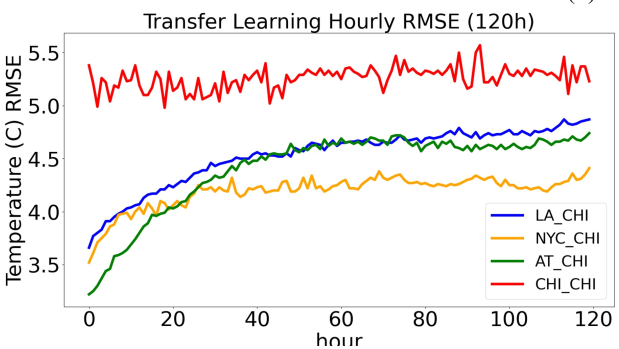
Accuracy Change	CHI (baseline)	LA predicting CHI	NYC predicting CHI	AT predicting CHI
24 → 12h	R^2: 0.96 - -	R^2: 0.97 -0.7% RMSE -2.9% MAE	R^2: 0.98 -5.8% RMSE -7.8% MAE	R^2: 0.98 -7.7% RMSE -9.3% MAE
144 → 72h	R^2: 0.70 - -	R^2: 0.92 -24.9% RMSE -24.2% MAE	R^2: 0.91 -24.3% RMSE -24.4% MAE	R^2: 0.87 -25.2% RMSE -24.2% MAE
240 → 120h	R^2: 0.51 - -	R^2: 0.83 -13.9% RMSE -16.1% MAE	R^2: 0.87 -21.0% RMSE -25.7% MAE	R^2: 0.88 -16.5% RMSE -17.9% MAE



(a) 24h → 12h



(b) 144h → 72h



(c) 240h → 120h

https://www.stcenter.net/

Highlight

- The LSTM as a DL algorithm is proven to offer advanced prediction capability for multi-step temperature prediction with multivariate on our fusion dataset compared to ARIMA and XGBoost. The generalized prediction model performs well on the majority of local testing DCLs;
- Even for the most extreme case tested (120-step prediction), the proposed framework outperforms the best model reviewed for 12-step prediction. In a future study, the influence of climatic pattern will be considered for better long-term predictions;
- Despite the encouraging results, a greater prediction length leads to a larger error, and the optimized input-output combination for different multi-step predictions needs to be explored;
- Transfer learning minimize the prediction error for regions with data scarcity problem and improves the predicting MAE up to 25.7%; and
- The effectiveness of models trained in different regions depends on the similarity shared by source and target locations. The ideal density of DCLs for transferable model training is a future research initiative

https://www.stcenter.net/

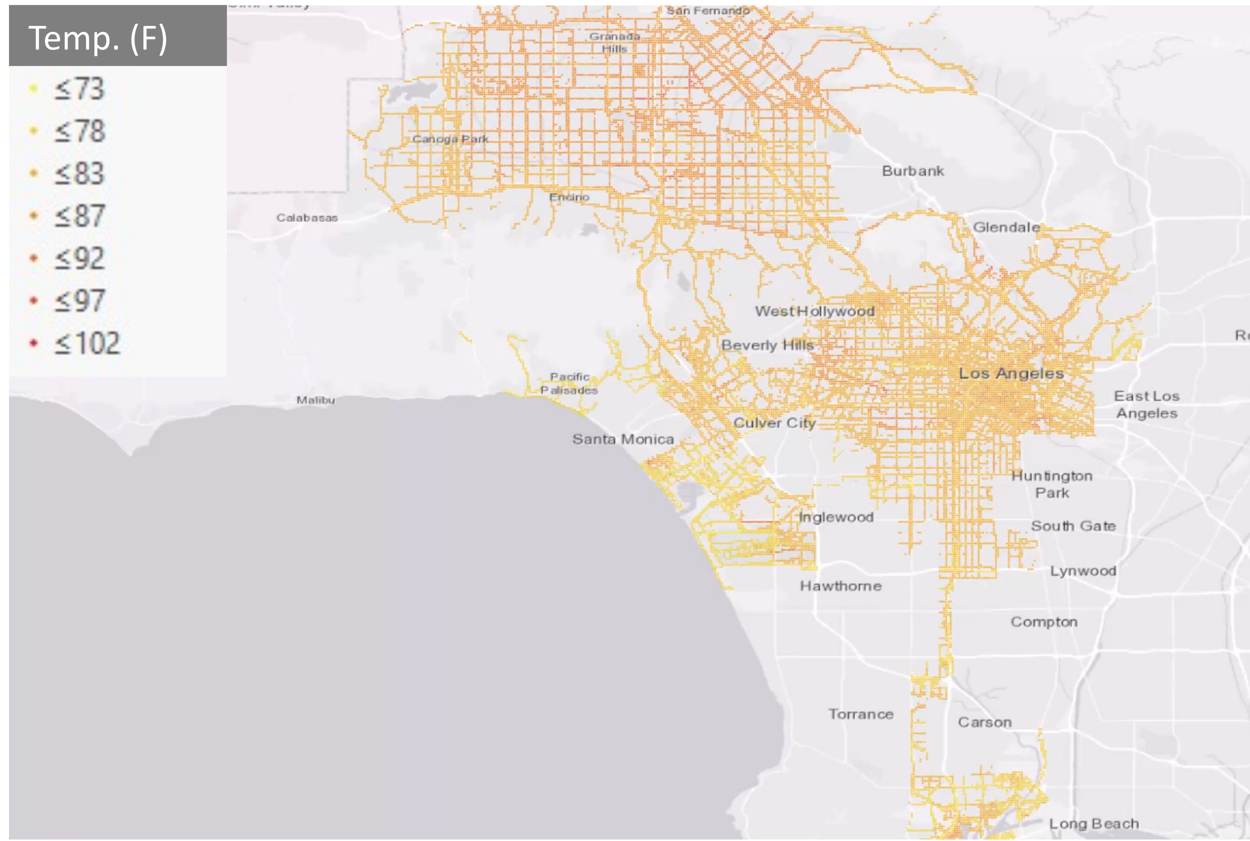
What's next

- Understanding when a transferable model would give optimal prediction
  - NYC model outperforms LA model when predicting Chicago temperature
- Providing full coverage for all study area by integrating forecasting dataset (e.g., HRRR)

IoT Data Interpolation

Kriging

$$\hat{Z}(s_0) = \sum_{i=1}^{N} \lambda_i Z(s_i)$$



Jul 25-28, 2019 LA, CA

Thank You!

Jingchao Yang

- jyang43@gmu.edu

Code availability

https://github.com/stccenter/IoT-based-Temperature-Prediction

Spatiotemporal Thinking, Computing and Applications

STC

https://www.stcenter.net/

AAG
AMERICAN ASSOCIATION of GEOGRAPHERS