IMPROVING THE CROP TYPE CLASSIFICATION CAPABILITY FOR LARGE GEOGRAPHIC AREAS

Yaping Cai[1] , Shaowen Wang[1] , and Kaiyu Guan[2]

[1] Department of Geography & GIScience
University of Illinois at Urbana-Champaign
[2] Department of Natural Resources and Environmental Sciences
University of Illinois at Urbana-Champaign

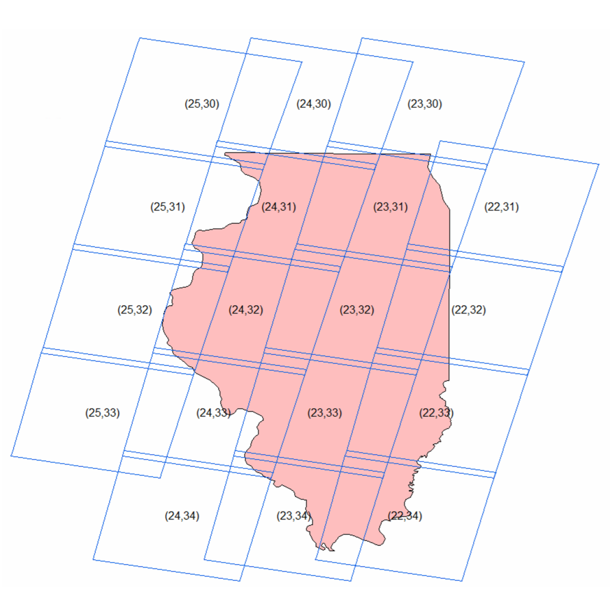
INTRODUCTION

Accurate crop type classification based on remote sensing data is important for both scientific and practical purposes. The state-of-the-art research on crop-type classification has been shifted from relying on only spectral features of single static images to combining both spectral and time series information. Facing the data-intensive and computationally-intensive challenges raised by the advanced crop type classification method based on both spectral and time series information through machine learning approaches (especially for the deep neural network) for large geographic areas, we proposed a high-performance workflow to preprocess the data and speed up the model training and testing. In this case study, we collected Landsat Surface Reflectance Data covering IL State in 2016 as model input information and collected Cropland Data Layer in 2016 as ground truth. We first compared model performance at county, agricultural statistics districts (ASD), and state-level as reference. Then, we explored the model performance of simple spatial partitions and climate partitions using K-Means. Since the model input is the crop phenology information that will be influenced by different geo-locations and climate, we combined both geo-locations (i.e. latitude and longitude) and climate factors (i.e. temperature and precipitation) as features (assign different weights) to conduct spatial partitions using K-Means. We built models for different spatial partitions to determine the optimized spatial scale for crop type classification. We found the state-level model can already achieve a high performance, whereas the spatial partition with more weight on climate factor can have a better performance (overall accuracy ~ 0.897).

RESEARCH QUESTION

(1). How can we optimize the computational workflow to handle state-level satellite data?

(2). What is the optimized spatial scale for crop-type classification if it exists?

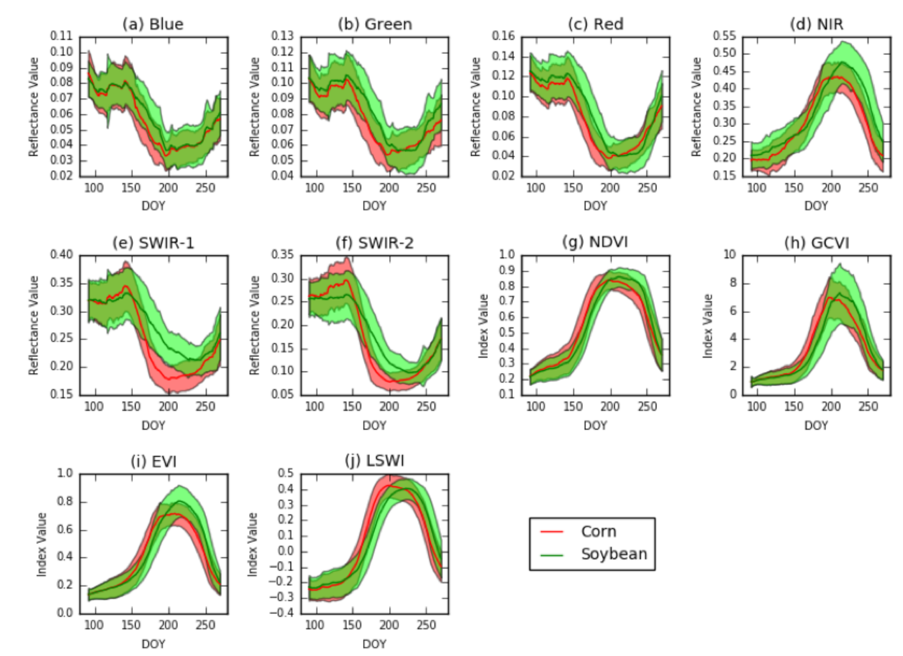


102 counties

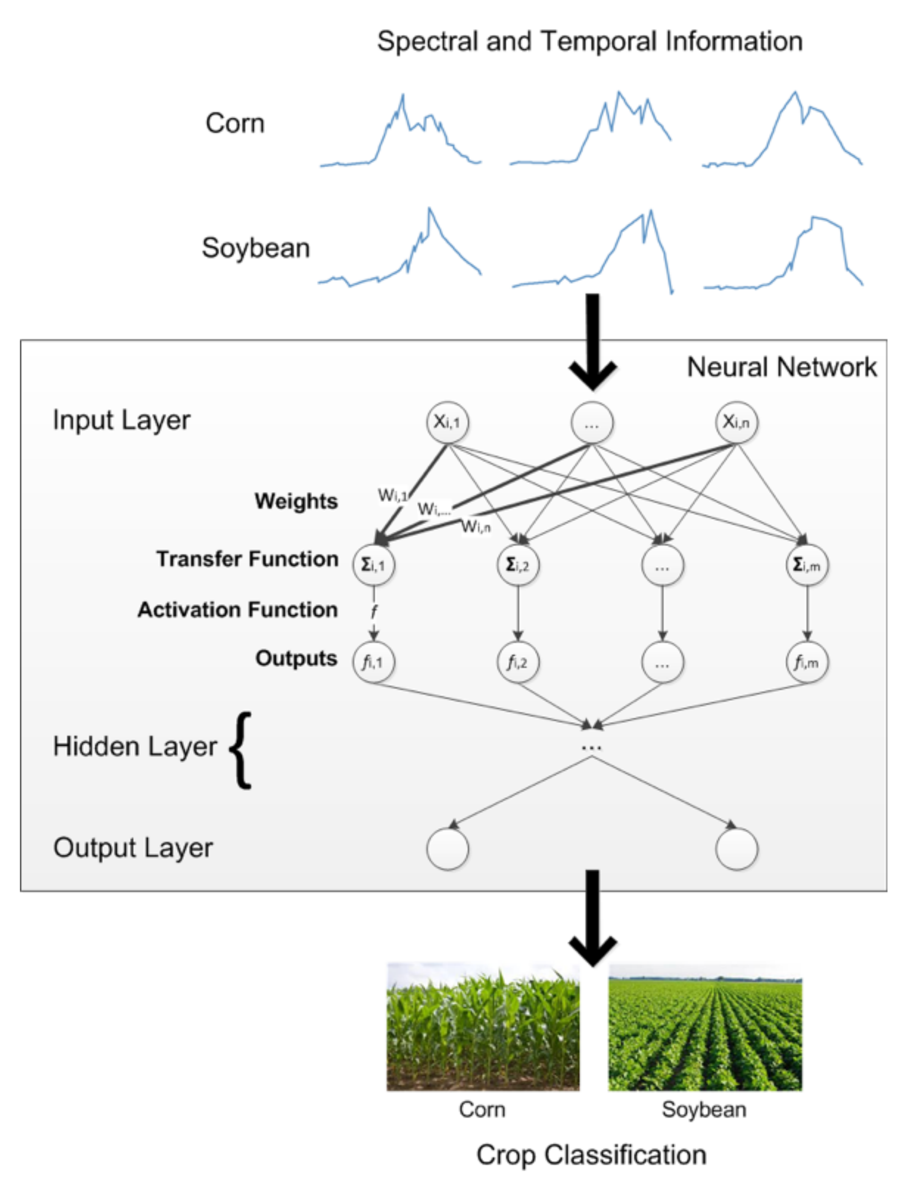
28 scenes covering IL

817 scenes in total in 2016

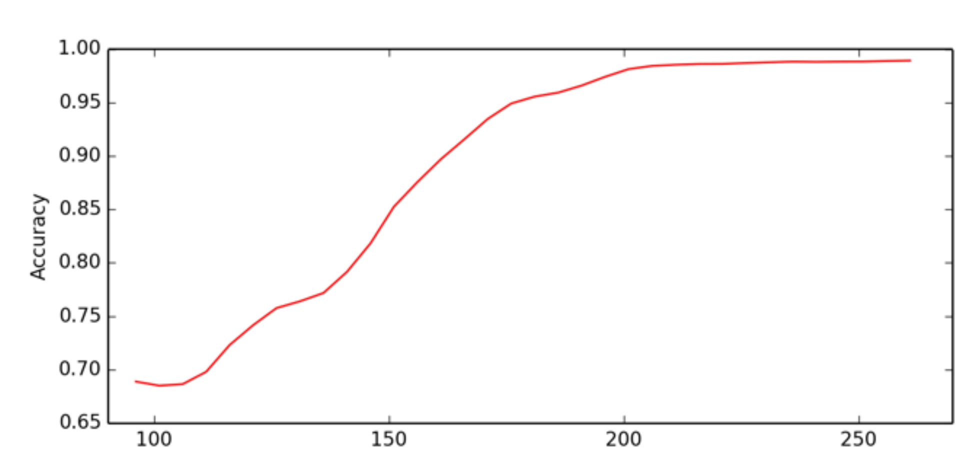
COUNTY-LEVEL FINDINGS



Time-series spectral band information and vegetation indices are aggregated for all the corn and soybean fields and years for the Champaign County, IL.

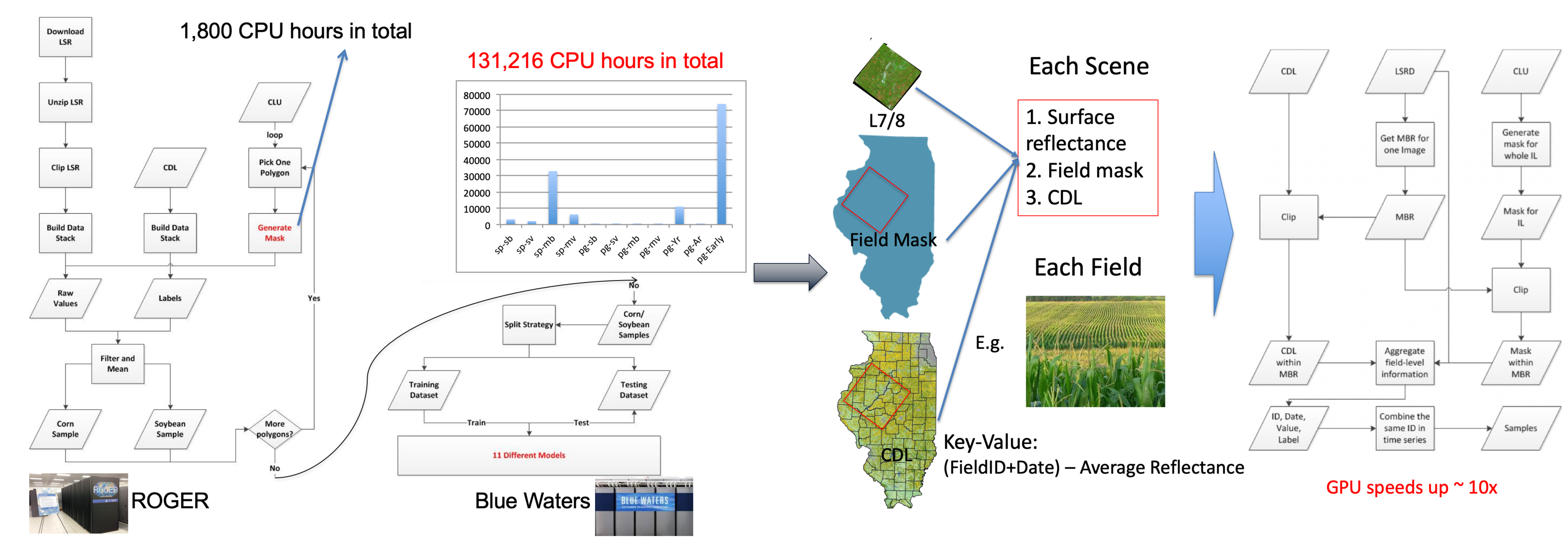


Framework of the ANN based crop type classification.

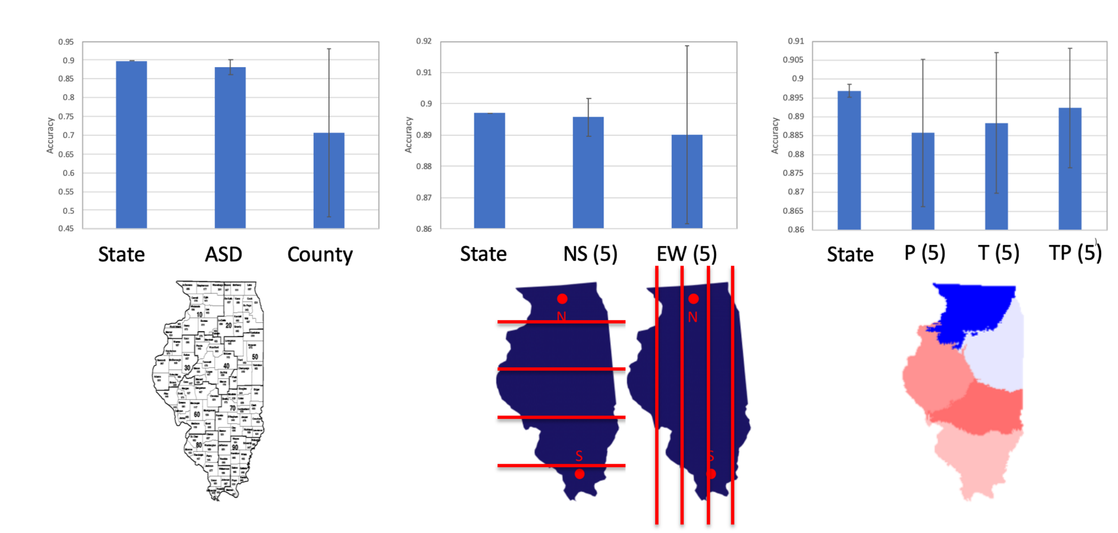


Classification performance as a function of time

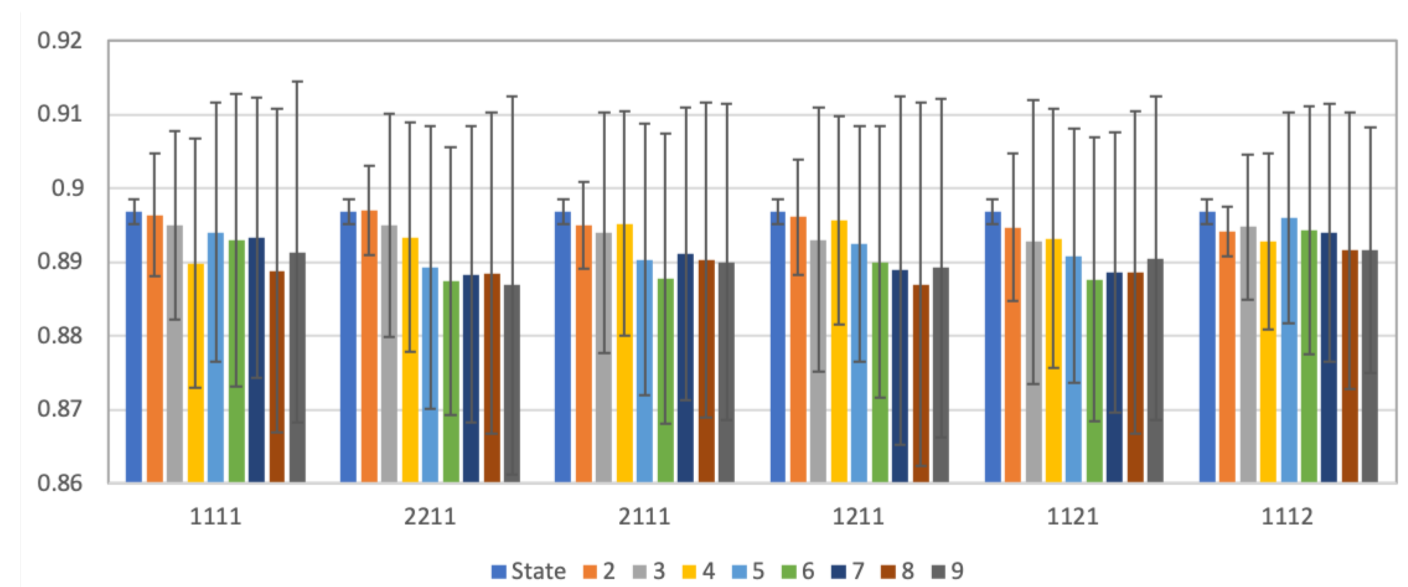
SCALE UP TO STATE-LEVEL



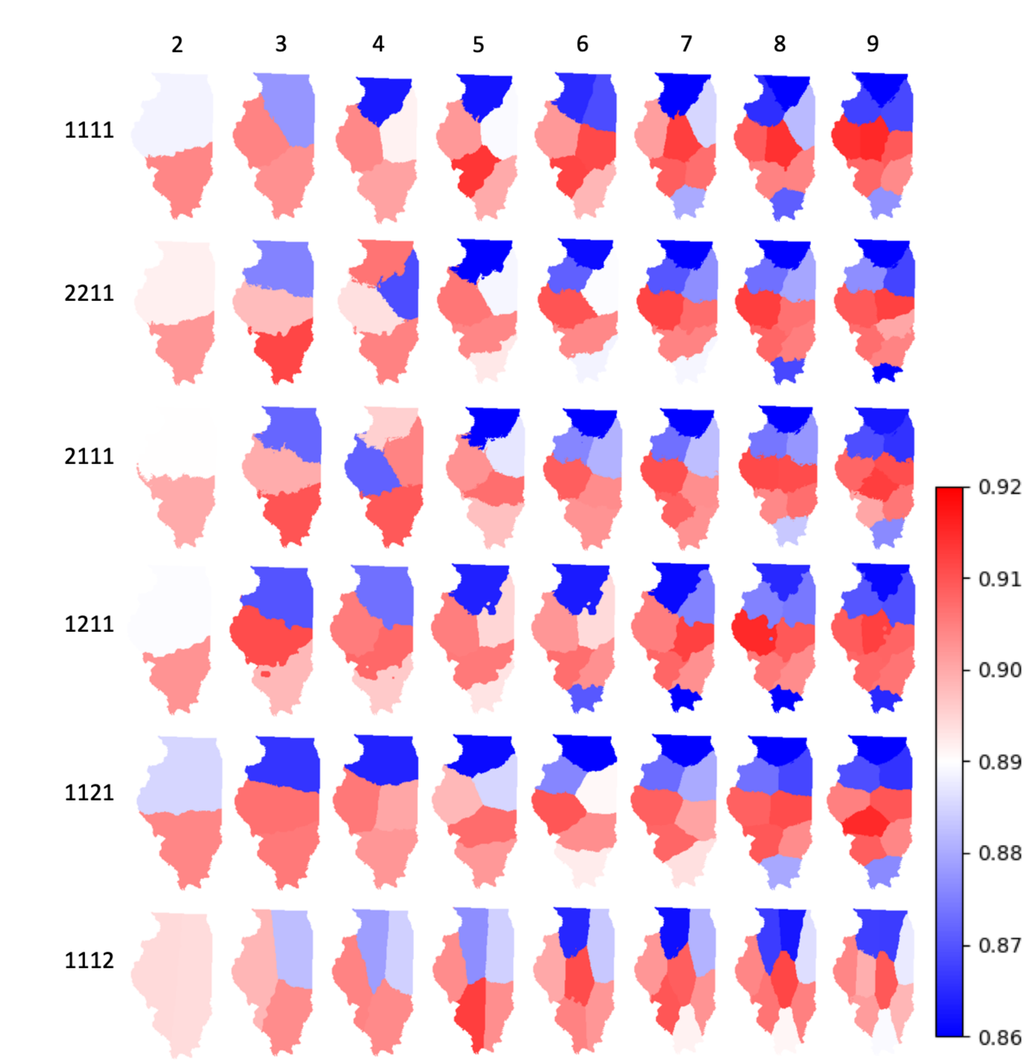
Workflow Optimization



Explore the model performance at different spatial scales



The performance of different partitions based on the combined geo-locations and climate with different weights.



The spatial patterns of the performance for different weights and numbers of partitions.

CONCLUSIONS

We paralleled the previous sequential workflow of county level for preprocessing the data and used GPU to speed up the model training/testing.

We conducted four groups of experiments to explore the optimized spatial scale and found that the two partitions based on the combined geo-locations and climate factors with weights "2211" achieved the best performance.

ACKNOWLEDGEMENTS

Illinois Distinguished Graduate Fellowship

ROGER is supported by the National Science Foundation under grant number 1429699

Blue Waters is supported by the National Science Foundation under award numbers 0725070 and 1238993