Satellite Data Deluge: An Innovative Hybrid Deep Learning Model for Fusing Multi-Scale Spatio-Temporal Satellite Imagery

Zijun Yang, Chunyuan Diao
Department of Geography and Geographic Information Science, School of Earth, Society & Environment, University of Illinois at Urbana-Champaign

BACKGROUND

- The recent years have seen a great increase in the volume of data produced in almost every field.
- In remote sensing, with new data acquisition techniques and datasets emerging, more than a million images are generated every day.
- This large amount data consists of multi-sensor, multi-temporal, and multi-frequency datasets.
- Fusing multi-source remote sensing data will greatly help us gain a more comprehensive understanding of rapid changing earth.

The number of satellites launched keeps increasing in recent years.

INTRODUCTION

- Resolution trade-off in remote sensing: most sensors cannot provide images with both high spatial resolution and high temporal revisiting frequency.
- Remote sensing imagery with both high spatial and temporal resolutions is highly desired for many applications (e.g. earth surface monitoring).
- Previous methods' performances relatively degrade in the presence of rapid phenological changes.
- Deep learning provides effective ways for complex feature learning and reconstruction.

An illustration of rapid phenological change in agricultural fields from early June to early September.

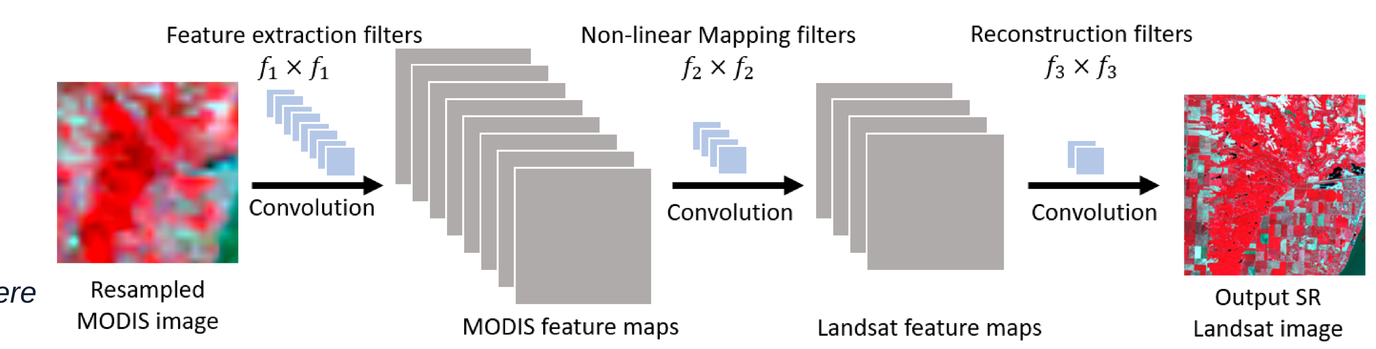
METHODS

We propose a hybrid deep learning model that hybrids convolutional neural network (CNN) and Long Short-Term Memory (LSTM).

Super-Resolution CNN (SRCNN)

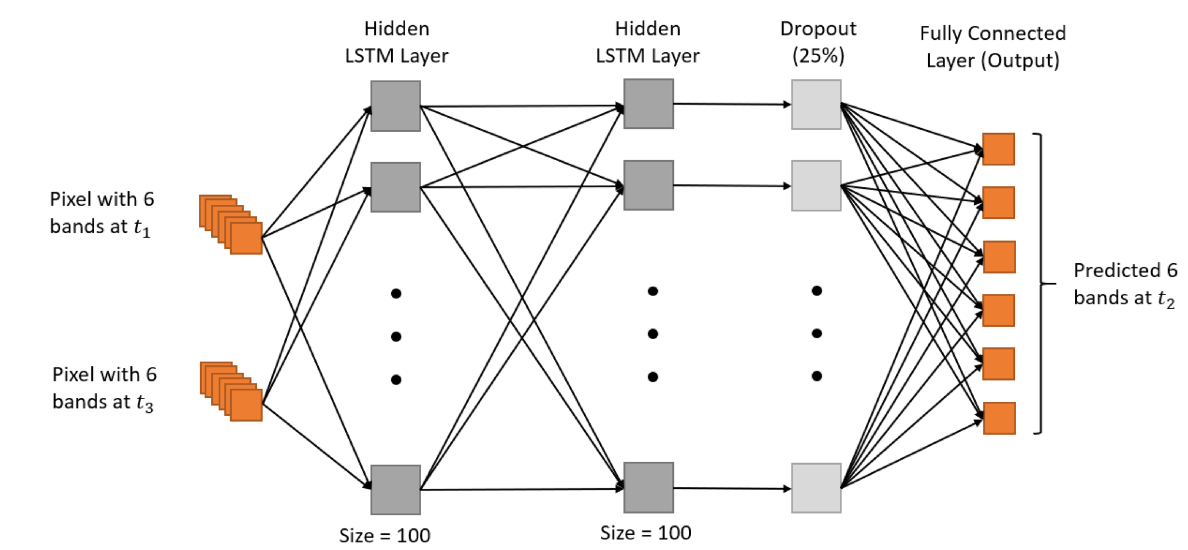
- SRCNN is for recovering high spatial resolution images from low spatial resolution images.
- We used SRCNN here to enhance the MODIS images on the prediction date with more spatial details, which is important for capturing daily changes.

Flow chart of CNN, where $f_1 = 9$, $f_2 = 5$, $f_3 = 5$.



Long Short-Term Memory (LSTM)

- LSTM is design for learning patterns in sequential data (e.g. time-series data).
- LSTM can capture long-term dependency in the sequential data.
- Pixels of SRCNN's output images are fed into the LSTM model. LSTM will make prediction of surface reflectance on t_2 based on data on t_1 and t_3.



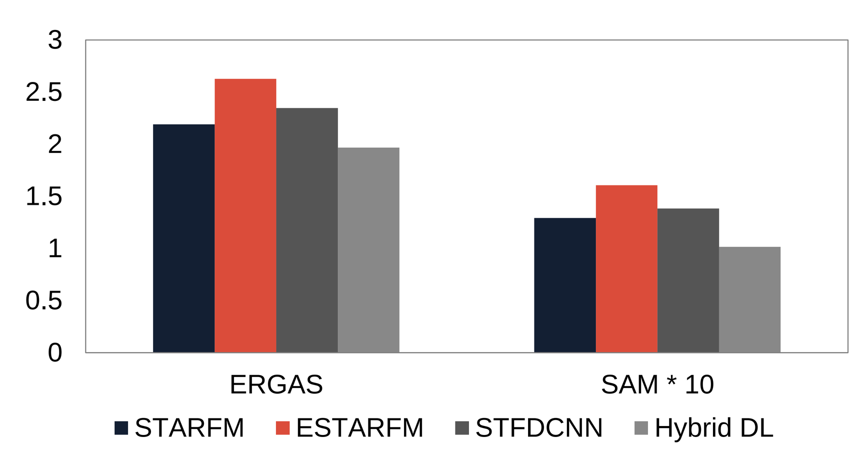
Model structure of LSTM

RESULTS

Our hybrid model is compared with spatial and temporal adaptive reflectance fusion model (STARFM), Enhanced STARFM (ESTARFM), and spatiotemporal fusion using deep convolutional neural networks (STFDCNN). STARFM and ESTARFM are two classic spatiotemporal algorithms, and STFDCNN is another method using deep learning methods (CNN).

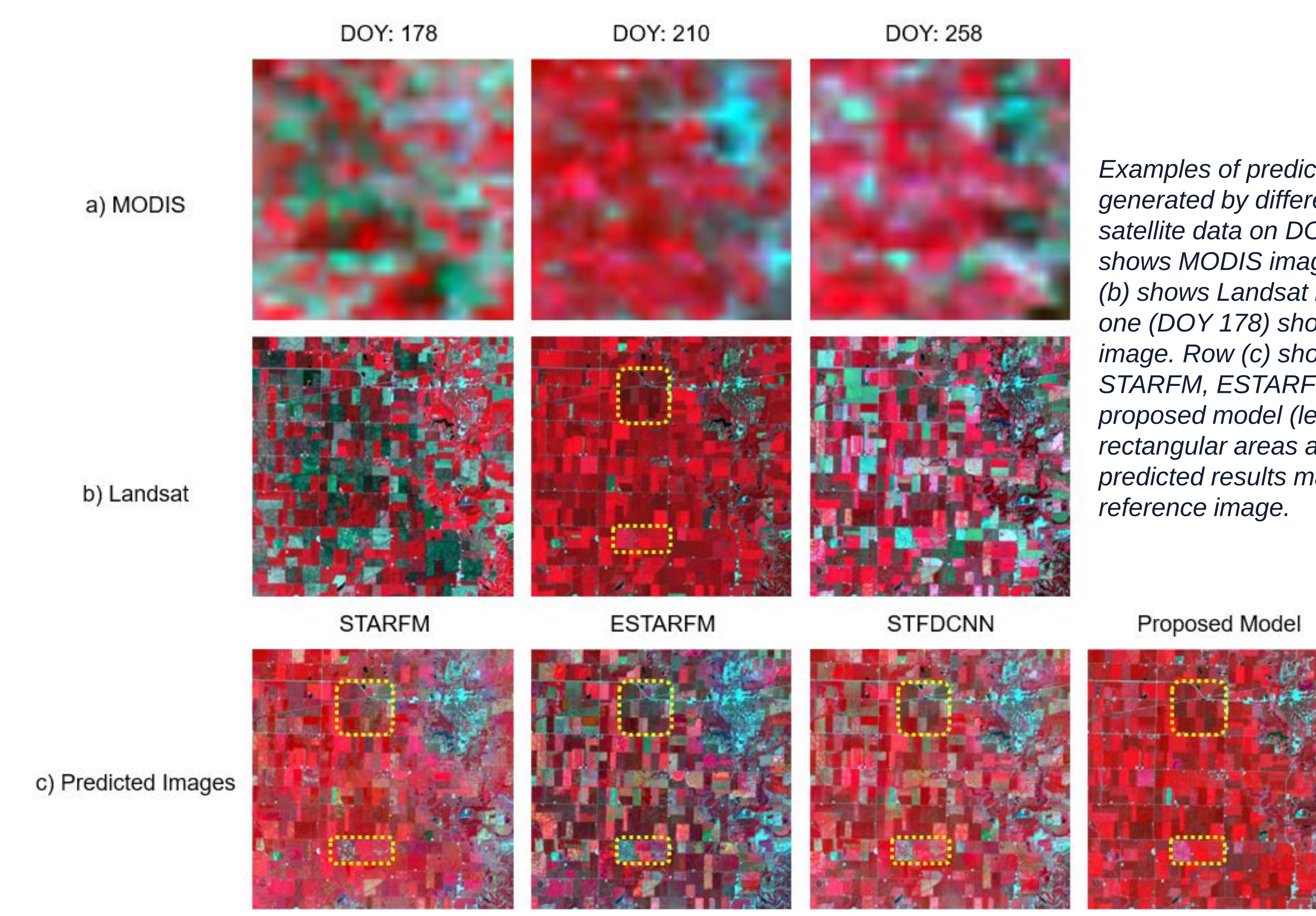
Accuracy

We use two cross-band metrics for accuracy assessment: spectral angle mapper (SAM) and erreur relative global adimensionnelle de synthese (ERGAS). SAM is used to measure the spectral distortion of the fusion results, while ERGAS assesses overall errors in reflectance values. The hybrid model shows significantly better performance compared to other models.



Hybrid DL vs.	STARFM	ESTARFM	STFDCNN
SAM	p<0.001	p<0.001	p<0.001
ERGAS	p=0.022	p<0.001	p<0.001

Results of pairwise comparisons using paired t-tests with Bonferroni-adjusted p-values. Comparisons include the proposed model vs. STARFM, ESTARFM, and STFDCNN.



Examples of predicted images (DOY 210) generated by different fusion models using satellite data on DOY 178 and 258. Row (a) shows MODIS images on the three dates; row (b) shows Landsat images. Note the middle one (DOY 178) should be the reference image. Row (c) shows fusion results by STARFM, ESTARFM, STFDCNN, and the proposed model (left to right). Yellow dashed rectangular areas are examples of how predicted results may deviate from the reference image.

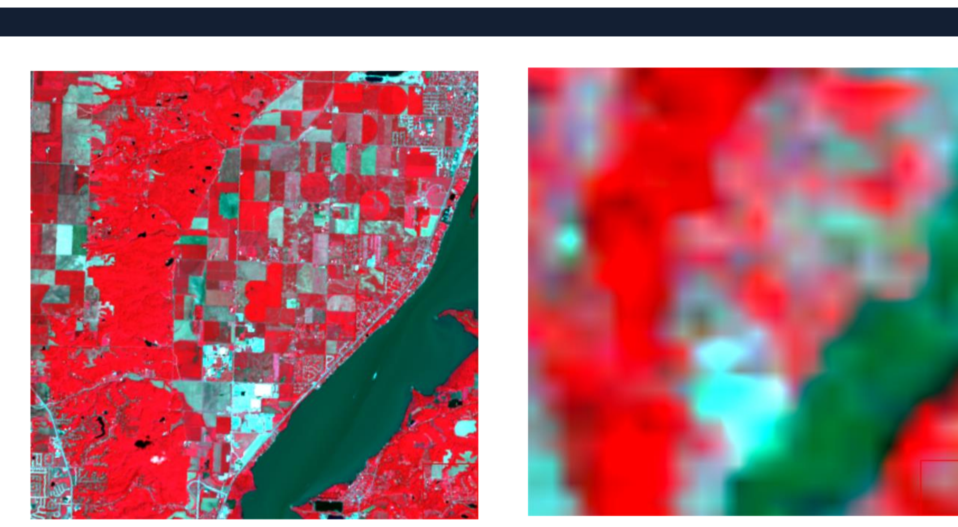
OBJECTIVES

Develop a hybrid deep learning model to fuse satellite imagery of various spatial and temporal scales for earth system dynamic monitoring. Specifically, this hybrid deep learning model aims to:

- 1) generate accurate and reliable daily reflectance images with both high spatial and temporal resolutions.
- 2) better capture rapid earth system changing dynamics, particularly agricultural phenological dynamics.

DATA & STUDY AREA

- MODIS: 500-meter spatial resolution with daily observation
- Landsat: 30-meter spatial resolution with a 16-day revisiting frequency
- Our study area is in Champaign County, with a composition of agricultural fields (corn and soybean), forest, urban areas, and water bodies.
- In this study, we use Landsat-MODIS images pairs on six dates (day of year: 162, 178, 210, 258, 290, 354). Data were in growing season when the phenological change was considered fast.



A subset of our study site in Landsat-8 OLI 30-meter imagery (left) and MODIS 500-meter imagery (right); both images acquired on June 27, 2017.

CONCLUSIONS

- Our hybrid deep learning model gives more satisfactory results presence of rapid phenological changes compared to benchmark methods.
- The hybrid model leverages CNN's ability to process spatial features and LSTM's ability in learning sequential information, demonstrating its potential of providing a high-quality time-series dataset that has both high spatial and temporal resolutions.

FUTURE WORK

- Generate datasets that has high spatiotemporal resolution using the hybrid deep learning model.
- Explore the possibility of fusing multiple datasets (e.g. Landsat, MODIS, Sentinel-2, Planet, etc.).

ACKNOWLEDGEMENTS

- The deep CNN training process used Blue Waters which is supported by the National Science Foundation through awards ACI-0725070 and ACI-1238993.

I ILLINOIS