DI

TN

BW

FS

BI

# LARGE-SCALE REMOTE MONITORING OF INVASIVE SPECIES DYNAMICS THROUGH A PETASCALE HIGH-PERFORMANCE COMPUTING SYSTEM

**Allocation:** Innovation and Exploration/110 Knh
**PI:** Chunyuan Diao[1]

[1]University of Illinois at Urbana–Champaign

## EXECUTIVE SUMMARY

The rapid expansion of exotic saltcedar along riparian corridors has drastically altered landscape structures and ecosystem functions throughout the United States. Conducting the large-scale monitoring of spatio-temporal dynamics of this invasive species over the past 40 years is essentially critical to understanding its invasion mechanism. Previous studies indicated that the leaf senescence stage is the optimal time window to remotely monitor saltcedar distributions. However, the computational complexity in predicting the leaf senescence timing, along with the massive volume of satellite data in both spatial and temporal dimensions, makes large-scale invasive species monitoring prohibitive. Instead, the research team developed a parallel computational network model on Blue Waters that can accommodate the spatio-temporal variation in plant phenology to facilitate the large-scale monitoring of invasive species dynamics. Blue Waters provides unprecedented opportunities to achieve regionwide monitoring of saltcedar distribution to revolutionize scientific understanding of saltcedar invasion processes.

## RESEARCH CHALLENGE

Remote monitoring of invasive saltcedar dynamics is essential for conservation agencies to develop cost-effective control strategies [1,2]. Acquiring satellite imagery during the leaf senescence stage of saltcedar is necessarily crucial to facilitate the large-scale repetitive mapping of this invasive species [3–5]. However, owing to climate variability and anthropogenic forcing, the timing of saltcedar leaf senescence varies over space and time [6]. Given that the leaf senescence stage of saltcedar lasts only for three to four weeks, it is challenging to pinpoint the appropriate satellite imagery that can accommodate this spatial and interannual variation over wide geographical regions [7]. The dearth of high-performance computational systems that can tackle this phenological issue makes the large-scale saltcedar mapping prohibitive and hinders the understanding of its invasion mechanism.

## METHODS & CODES

To accommodate the spatio-temporal variation in plant phenology, the research team developed a complex network-based phenological model using satellite time series. The network-based phenological model constructs an undirected network for each pixel based on its spectral reflectances along the phenological trajectory. Specifically, the spectral reflectance of the pixel obtained on each day is represented as a node. Those nodes that share similar spectral similarity are connected by edges. The network model groups the spectral reflectances along the trajectory into three clusters: a "pretransition cluster," a "transition cluster," and a "posttransition cluster" (Fig. 1). Further, the team developed several network measures to accommodate the spatio-temporal variation in invasive species phenology.

To overcome the computational limits, the research group proposed an innovative pixel-based remote sensing parallel computational system for large-scale saltcedar monitoring. The parallel system adopts hybrid computation models, including a core-level computation model and a node-level data distribution model. In the core-level computation model, a computing node stores the information of a number of pixels and processes them using OpenMP. The node-level data distribution model includes a two-level data decomposition strategy. The first level uses a space decomposition. All computing nodes are assigned to different node groups such that each group is responsible for all imagery tiles covering the same spatial extent. At the second-level decomposition, both space and time decomposition are applied within each group to improve the I/O performance. Using this strategy, the huge number of data and I/O operations are evenly distributed among all computing nodes. This greatly reduces the memory requirement and time for reading data files. Furthermore, this data distribution model is highly scalable and can be used efficiently on Blue Waters.

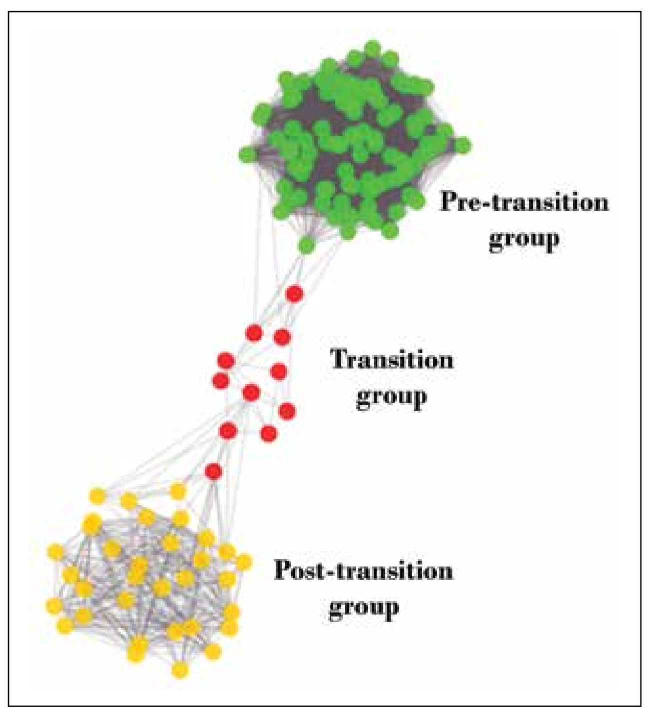


Figure 1: Complex network of phenological progress of invasive species with three clusters.

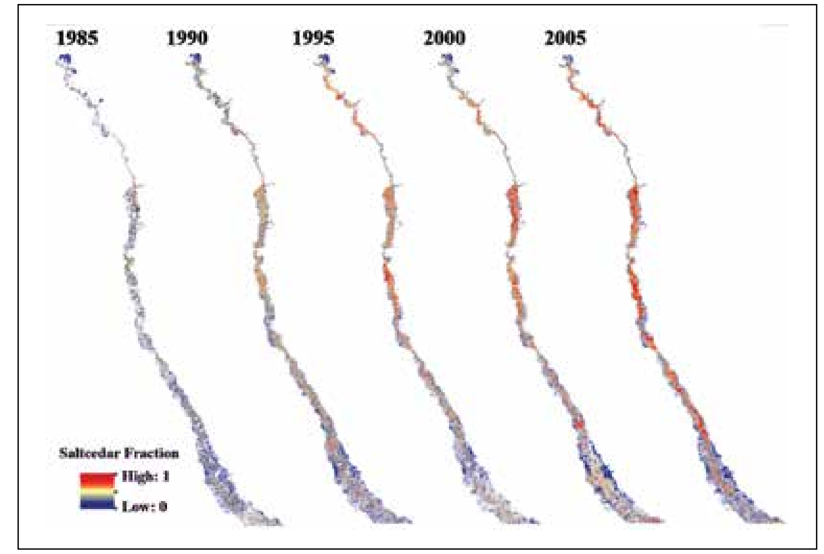


Figure 2: The spatial distributions of invasive saltcedar from 1985 to 2005 along the Forgotten River Reach.

## RESULTS & IMPACT

The network-based phenological model presents a new representation of the complex phenological process of vegetation. This representation characterizes the vegetation's phenological status through tracking the continuing changes of spectral signatures of vegetation along the temporal trajectory. The network results revealed a gradual phenological shift from north to south along the latitudinal gradient, following the Hopkins bioclimatic law. It suggested that temperature decreased with increasing latitude and would likely be an important force driving the phenological variations in vegetation over wide geographic regions. Along the latitudinal gradient, the interannual variation in the estimated average timing of leaf senescence was generally less than 10 days.

The team has successfully implemented this parallel computational system on Blue Waters and analyzed the massive amount of time series of remotely sensed data to monitor the phenological dynamics, as well as mapping the large-scale invasive species distributions. The team found that invasive saltcedar has expanded dramatically over the riparian corridors of the Southwestern United States (Fig. 2). It has amalgamated many smaller patches into fewer, larger patches; formed a greater number of high-density patches; and outcompeted native vegetation, especially along the river. The landscapes along the riparian corridors, as a consequence, have been transformed enormously. The monitoring results show great promise in helping conservation agencies to conduct systemic restoration of the affected riparian ecosystems. The parallel system is transferable to many other types of competing vegetation species aside from saltcedar, enhancing the system's utility and value across a wide range of agencies concerned with invasions.

## WHY BLUE WATERS

The computational challenges of the project lie in two aspects: (1) the high computational costs of processing a single pixel, and (2) the large number of pixels to process at the large-scale. The network algorithm devised for detecting saltcedar leaf senescence outperforms most existing algorithms but has a high computational cost. A huge number of pixels and associated time series data are required for the large-scale monitoring. This number adds to the total amount of computation time. Further, it requires a large amount of storage space. Blue Waters is the most ideal resource to tackle these computational challenges. In addition, Blue Waters' file system provides high-speed data access. This could dramatically reduce the time of I/O operations of the massive imagery involved in this project, and hence reduce the total computation time.

## PUBLICATIONS & DATA SETS

C. Diao, "Complex network-based time series remote sensing model in monitoring the fall foliage transition date for peak coloration," *Remote Sensing Environ.*, vol. 229, pp. 179–192, 2019.

C. Diao and L. Wang, "Landsat time series-based multiyear spectral angle clustering (MSAC) model to monitor the inter-annual leaf senescence of exotic saltcedar," *Remote Sensing Environ.*, vol. 209, pp. 581–593, 2018.

C. Diao, "Innovative pheno-network model in estimating crop phenological stages with satellite time series," *ISPRS J. Photogrammetry Remote Sensing*, vol. 153, pp. 96–109, 2019.