

Spatial-temporal Multi-Task Learning for Within-field Cotton Yield Prediction

Long H. Nguyen¹, Jiazhen Zhu², Zhe Lin³, Hanxiang Du¹, Zhou Yang¹,
Wenxuan Guo³, and Fang Jin¹

¹ Department of Computer Science, Texas Tech University
{long.nguyen,hanxiang.du,zhou.yang, fang.jin}@ttu.edu

² Department of Computer Science, George Washington University
jiazhen.zhu@gmail.gwu.edu

³ Department of Plant and Soil Science, Texas Tech University {zhe.lin,
wenxuan.guo}@ttu.edu

Abstract. Understanding and accurately predicting within-field spatial variability of crop yield play a key role in site-specific management of crop inputs such as irrigation water and fertilizer for optimized crop production. However, such a task is challenged by the complex interaction between crop growth and environmental and managerial factors, such as climate, soil conditions, tillage, and irrigation. In this paper, we present a novel Spatial-temporal Multi-Task Learning algorithm for within-field crop yield prediction in west Texas from 2001 to 2003. This algorithm integrates multiple heterogeneous data sources to learn different features simultaneously, and to aggregate spatial-temporal features by introducing a weighted regularizer to the loss functions. Our comprehensive experimental results consistently outperform the results of other conventional methods, and suggest a promising approach, which improves the landscape of crop prediction research fields.

1 Introduction

Cotton is an important cash crop native to tropical and subtropical regions in the world. Accurate yield prediction not only provides valuable information to cotton producers for effective management of the crop for optimized production, but also is important to policymakers, as well as consumers of agricultural products. However, cotton yield prediction is challenging due to complex interactions between crop growth and weather factors, soil conditions, as well as management factors, such as irrigation, tillage, rotation, etc. Moreover, simply applying other crop yield prediction models on cotton may lead to nothing but disappointment: a prediction model that works on other crops like wheat, rice, and sugarcane, however, fails on predicting cotton yield [1].

The existing approaches estimate crop yield based either on the crop sown areas, crop-cutting experiments or market arrivals show wide variability because of their inability to capture the indeterminate nature of the crop and its response to environmental conditions [2]. One of the attempts is to apply Grey model [3] on production prediction, utilizing short-term forecasting with exponential growth.

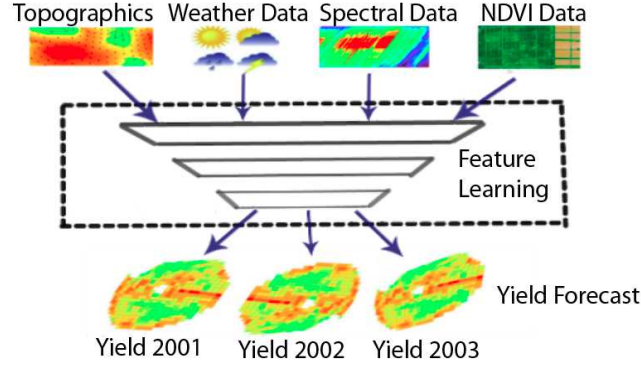


Fig. 1. Data sources in our prediction model.

The regression-based method such as time series analysis can also be applied to production prediction [4], but it suffers from great variation when the external environment is under significant variation. Another frequently used method for prediction is the differential equation model [5], which demands the system to be stable and requires extra work to solve the equation.

Conventional yield prediction treats the field uniformly despite inherent variability. Uniform assumption may result in over- or under-application of resources in specific locations within a field, which may have a negative impact on the environment and profitability [6]. However, consistent and accurate within-field yield prediction is challenging due to the high accuracy requirement under the complex interactions between yield-influencing factors, such as soil, weather, water, and spatial correlations.

With the introduction of the global positioning system (GPS), geographic information systems (GIS), and yield monitors along with other new technologies, we can quantify spatial variability in soil properties and crop yield in small areas of a field. As satellite and drone technologies develop, we are able to collect remote sensing images at fine resolutions to support within-field yield forecast. Within-field scale crop yield prediction provides valuable information for producers to site-specifically manage their crop, which can optimize crop production for maximum profitability. In the within-field prediction procedure, we use a 30-m grid to represent a continuous surface.

The advancement of machine learning offers a different approach compared with the traditional ways for yield forecasting. The rapid advances in sensing technologies, the use of fully automated data recording, unmanned systems, remote sensing through satellites or aircraft, and real-time non-invasive computer vision, are additional boosts for enabling the new yield forecasting model. Due to the capability of machine learning based systems to process a large number of inputs and handle non-linear tasks, people have attempted to predict county-level soybean yield in the United States [7]. However, using deep learning for within-field cotton forecast remains as an untouched ground. In our work, the

within-field forecasting is based on each grid for one field in West Texas area across three years (2001, 2002, 2003) in order to predict the cotton yield before harvest.

On this account, we propose a Multi-Task learning model to predict within-field cotton yield. As shown in Figure 1, this model ingests many sources of data which contain features for different learning tasks, including soil topographic attributes (elevation, slope, curvature, etc.), spectral data (Blue, Green, Red, and NIR bands denoted as BAND1, BAND2, BAND3, and BAND4, respectively), normalized difference vegetation index (NDVI) during the crop seasons; and weather (temperature, rainfall, etc.) data. These multiple data sources are aggregated in the shared layer before transferring to task-specific layers. This type of design in a Multi-Task learning model makes it capable of enhancing specific learning task by utilizing all sources of information of other related tasks. In other words, this allows us to take various factors and variables into consideration to achieve a more accurate yield prediction. On the other hand, crop yield within a field is typically autocorrelated, meaning yield values close together are likely more similar than those farther apart. Hence, to incorporate the spatial relationship, we propose a spatial regularization term to minimize the yield difference between one region and the weighted average of neighboring regions. Therefore, we termed this technique as Spatial-temporal Multi-Task Learning. The main contributions of this paper are summarized as below:

- We design an innovative multi-task learning approach to predict within-field cotton yield across several years. Different from other machine learning models, to predict the cotton yield for a specific year is one of the tasks in our model; each task is enhanced through its access to all available data from prior years.
- This work provides an entirely new vision for grid-scale crop yield prediction. To the best of our knowledge, this is among the first attempts to predict fine-grain cotton yield with the Multi-Task Learning approach, as existing work focus more on county-level or country level.
- We introduce a spatial weight regularizer to overcome the effects of geographical distance on yield prediction. Each grid is trained to minimize not only the difference between the prediction and the actual value, but also the difference between its yield and its neighbors.
- We perform a comprehensive set of experiments using the real-world dataset that produced results consistently outperformed other competitive methods, which could provide guidance for achieving higher crop production.

2 Related Work

Crop Yield Prediction. Crop yield prediction is challenging. Many studies have been conducted based on NDVI derived from the new moderate resolution imaging spectroradiometer (MODIS) sensor [8], MODIS two-band Enhanced Vegetation Index [9], even future weather variables [10]. Various methodologies are employed, such as statistical models [11], fuzzy systems and Artificial

Neural Networks [12], deep long short-term memory model [7] and deep neural network [13]. You et al. [14] provided an end to end soybean yield prediction using remote sensing images as input. Ji et al. [15] investigated the effectiveness of machine learning methods and found, unfortunately, most of the academic endeavors centered on Artificial Neural Networks with one or a few data sources undermine their predict performance. Most previous studies assume that the crop yield uniformly distributed over space ignoring the spatial variations.

Multi-Task Learning. Multi-Task learning (MTL) is implemented to predict spatial events due to its competence to exploit dynamic features and scalability [16]. Through learning multiple related tasks simultaneously and treating prediction at each time point as a single task, MTL captures the progression trend of Alzheimer’s Disease better [17]. MTL also has outstanding performance in event forecasting across cities [16] and fine-grain sentiment analysis [18], as well as in distance speech recognition [19]. Lu et al. [20] proposed a principled approach for designing compact MTL architecture by starting with a thin network and dynamically widening it in a greedy manner. Xu et al. [21] designed an online learning framework that can be used to solve online multi-task regression problems, although it is not a memory-efficient solution for data intensive application. To the best of our knowledge, usage of MTL in crop yield forecast in within-field practice is untouched. We propose a Multi-Task Learning model which targets at predicting each grid in the field for a crop season as an individual task. Meanwhile, we incorporate the spatial correlations as a regularization term to minimize the prediction errors.

3 Proposed Model

3.1 Overview

Figure 2 presents the framework of our prediction model. The cotton field is split into 475 grids for fine-grain prediction. We utilized the Dense and Dropout layers in the network. A shared Dense layer is used to extract latent features from all data dimensions, which are aggregated and fed into multiple sub-networks. Each sub-network represents the architecture of forecasting task in one year for all the grids. In other words, cotton yield prediction for all grids of each year are achieved in parallel via the separated sub-networks.

This aggregation is shared among all task-specific sub-networks. Therefore, it helps the task-specific sub-network to learn features from other tasks and to enhance its own prediction performance. Dense layer helps receive input from all the neurons in the previous layer with the intuition that all factors contribute each layer output neurons. Mathematically, the latent feature \hat{p}^t learned after a fully connected layer is computed as:

$$\hat{p}^t = \sigma\left(\sum_{i=1}^N x_i * w_i + b\right) \quad (1)$$

, where N is the number of neurons in a layer, x_i represents input feature, w_i

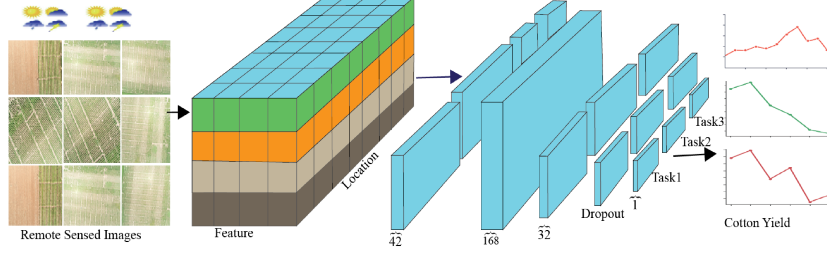


Fig. 2. The cotton yield prediction framework.

is a weight element, b is a bias and σ is the activation function. In our setting, $\sigma(x)$ is a Sigmoid function defined as $\sigma(x) = 1/(1 + e^{-x})$. Moreover, the dropout layer or dropout regularization is also used to randomly exclude some neurons (20% in ours) to avoid over-fitting.

3.2 Cotton Yield Prediction

Our feature set is enriched by concatenating the latent features and feeding the output into a shared Dense layer. Suppose $p^{t_1}, p^{t_2}, p^{t_3}$ and p^{t_4} are features from our sources, the joint feature v_{fc} is the concatenation (denote as \oplus) of those features [22]:

$$v^{fc} = p^{t_1} \oplus p^{t_2} \oplus p^{t_3} \oplus p^{t_4}. \quad (2)$$

Stacked on the top of the shared Dense layer are three separate sub-networks, and each is used for one yearly cotton yield forecasting task, as shown in Figure 2. After this layer, the latent feature is learned at time j following the equation: $h_j = \sigma(W_j * v_j^{fc} + b)$. Because cotton is usually planted by the end of May and harvest at the end of September or early October, we cut cotton's life cycle into several pieces, and each piece represents 2 weeks. Instead of taking it as time series data, we treat it as a couple of separate temporal features and utilize fully connected Dense layers and Dropout layers behind the shared Dense layer. We define the regression function for task t as:

$$\hat{y}_j^t = \sigma(W_j^t * h_j^t + b_j^t) \quad (3)$$

, where W_j^t and b_j^t are learnable parameters, h_j^t is the output of the last hidden layer, and σ is a linear activation function. The model output lies in the interval $[0, 1]$ after value normalization. We will recover them to the original values when doing performance evaluation.

3.3 Spatial Feature in the Loss Function

A loss function is defined as the mean square error between the observation and the prediction:

$$\mathcal{L}(\theta) = \sum_{k=1}^N (y_k - \hat{y}_k)^2 \quad (4)$$

, where θ means all learnable parameters, N is the number of regions in the field, y_k represents actual yield value, and \hat{y}_k represents the predicted yield value. To train θ by minimizing the loss function may introduce overfitting. Therefore, for the grid-scale crop forecasting within a field, the spatial correlations depend heavily on the factor of distance. This drives us to define the spatial influence via a regularization term that the yield difference between the predicting region and the weighted average yield of the neighboring regions should be minimized. In particular, suppose $G(k)$ is the set of neighbors of region k (determined by all regions whose distance to region k less than a threshold), and $w(k, j)$ is the inverse distance weight between region k and region j , the loss function now becomes:

$$\mathcal{L}(\theta) = \sum_{k=1}^N [(y_k - \hat{y}_k)^2 + \lambda \sum_{j \in G(k)} \frac{w(k, j)}{|G(k)|} * |\hat{y}_k - y_j|^2] \quad (5)$$

, where λ is the hyper parameter, $d(k, j)$ is the Euclidean distance between these regions $k \neq j$. The spatial weight $w(k, j)$ is computed as:

$$w(k, j) = \frac{1}{d(k, j)^p} \quad (6)$$

, where p is the power parameter (which equals to 2 in our experiment).

4 Experiments

4.1 Dataset and Feature Extraction

Our dataset includes weather data, soil properties, spectral data, and NDVI. Spectral data and NDVI are extracted from Lantsat 5 and Landsat 7 remote sensing images. The multi-spectral images were collected from 2001 to 2003 of a cotton field in west Texas. The total area is approximately 48 ha. The sensed images spatial resolution is 30 m. Hence, there are 475 grid cells under investigation. Figure 3 shows the distribution of some features over the field.

Weather data. Weather data includes the daily temperature and rainfall level. For simplicity, we use the average of every two weeks' weather data as features to match the sensed images.

Soil properties. Topographic variation is a common characteristic of large agricultural fields that has effects on spatial variability of soil water and ultimately on crop yield [23]. Besides, soil electrical conductivity (ECa) is also a reliable measurement of field variability. The relationship between ECa and crop yield depends on climate, crop type and other specific field conditions [24]. The variables that are considered in this paper include elevation, slope, curvature, the average electrical conductivity of soil, etc.

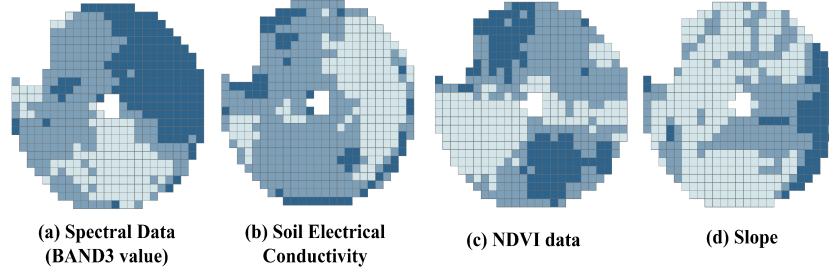


Fig. 3. (a) to (d) Within-field feature value distributions. Darker colors indicate higher values. Each feature value is normalized into $[0,1]$.

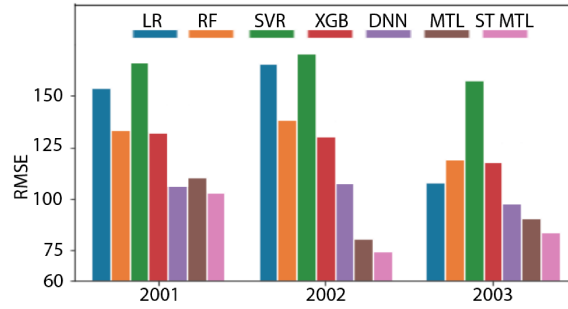


Fig. 4. Performance comparison in each year measured in RMSE, using whole dataset until September.

Field spectral data before planting. Field spectrum before planting may influence the entire crop yield. This data has four spectral bands extracted from the sensed images with 30-metre spatial resolution. Band1, Band2, Band3, and Band4 represent blue, green, red and near infrared value, respectively.

NDVI data. NDVI represents Normalized Difference Vegetation Index. It is typically related to amount or density of vegetation, which is calculated as the difference between the reflectance in near-infrared (which vegetation strongly reflects) and red wavelengths divided by the sum of these two. NDVI is computed as: $NDVI = \frac{NIR - RED}{NIR + RED}$, where NIR represents the spectral reflectance in near-infrared wavelength and RED is the spectral reflectance in the red wavelength.

4.2 Competing Approaches and Comparison Metrics

Competing Approaches In our experiment, a list of classical forecasting models are used for comparison and analysis:

Linear Regression This is a traditional linear regression model. Its standard formula is: $y = \sum_{i=1}^n \alpha_i x_i + \epsilon$ where y is the response variable, x_i is the feature and ϵ is the deviation.

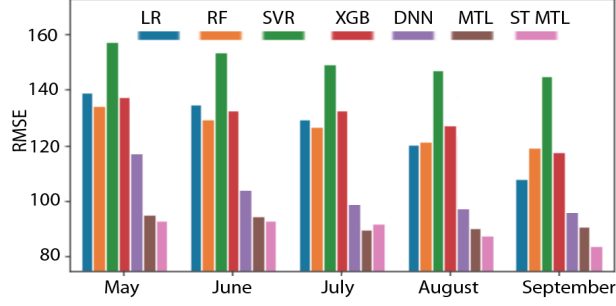


Fig. 5. Model performance in each month of 2003 measured in RMSE.

Random Forest This model tries to fit a number of regression trees on various sub-samples of the dataset and uses averaging to improve the predictive accuracy and control overfitting. Each leaf of the tree contains a distribution for the continuous output variable.

Support Vector Regression (SVR) SVR is a nonparametric technique that aims to find a function $f(x)$ that produces output deviated from observed response values y_n by a value no greater than ϵ for each training point x , and meanwhile, as flat as possible.

XGBoost This is an extension of gradient boosting machine (GBM) algorithm that tries to divide the optimization problem into two parts by first determining the direction of the step and then optimizing the step length.

Deep Neural Network (DNN) A fully connected deep neural network composed of three Dense layers, connecting to a Dropout and followed by another Dense layer is developed and used as another baseline for comparison.

Evaluation Metrics Let N be the number of grids under forecast. We denote A_i as the actual crop yield and F_i is the forecast yield for grid i . A set of classical metrics such as Mean square error (MSE), Root mean square error (RMSE), Mean absolute error (MAE), Mean absolute percentage error (MAPE) and Max error (ME) are used to elaborate the performance. These measures are computed as: $MSE = \frac{1}{N} \sum_{i=1}^N (A_i - F_i)^2$, $RMSE = \sqrt{MSE}$, $MAE = \frac{1}{N} \sum_{i=1}^N |A_i - F_i|$, $MAPE = \frac{100\%}{N} \sum_{i=1}^N \frac{|A_i - F_i|}{A_i}$ and $ME = \max(|A_i - F_i|)$ where $i = 1, \dots, N$.

4.3 Experimental Results

Average Performance Figure 4 shows the performance of our proposed model compared with other baselines in term of RMSE metric. The Multi-Task learning and our proposed Spatial-Temporal Multi-Task learning model have the least error in all three years. Figure 6 and Table 1 take yield prediction performance of 2003 as an example. The Multi-Task Learning and our Spatial-Temporal Multi-Task learning methods show significant superiority than all the other approaches.

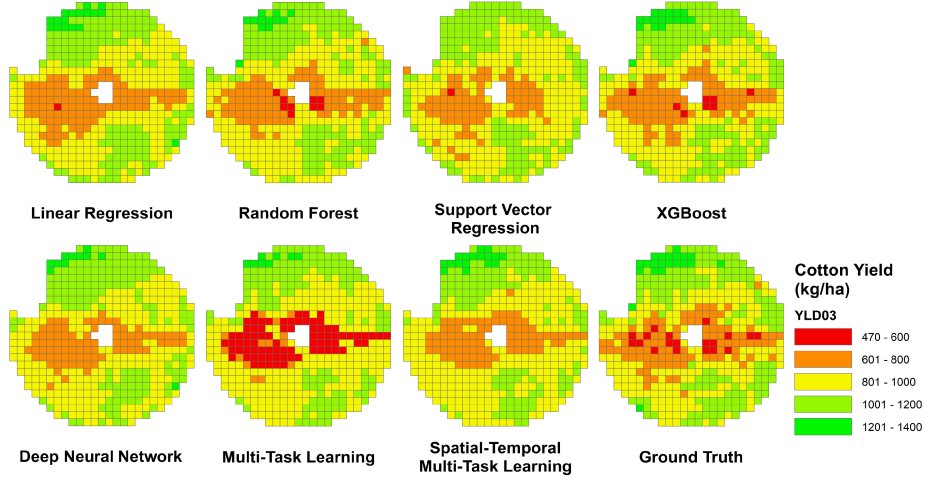


Fig. 6. Within-field cotton yield prediction on different algorithms versus ground truth yield monitor data in 2003.

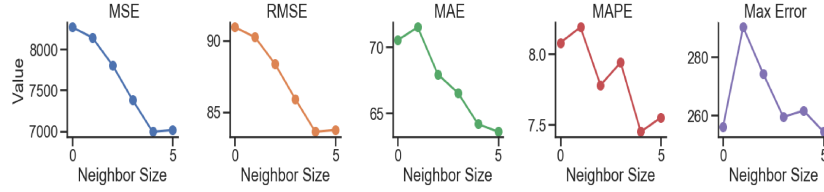


Fig. 7. Impact of neighborhood size on model performance

With the whole data package, our model achieves the smallest error metrics (MSE, RMSE, MAE, MAPE, ME) which are 7,013.5, 83.7, 63.6, 7.55 and 254.4, respectively. The overall performance of Multi-Task Learning is second and close to our proposed approach, while Support Vector Regression shows the worst performance.

Real-time Prediction Throughout the Year Considering the life cycle of cotton in the U.S., we train the model using partially available input features and predict the cotton yield in each month in an online manner. Figure 5 shows the performance when we try to make a prediction in May, June, July, August and September, using only the data available up to that point. As more information is available, most of the models improve. The improvement during the first three months is less than that of later two months. All models perform better in August and reach the best in September.

Spatial Correlations: We also vary the neighborhood distance of each region from 1 to 5 to verify the impact of spatial correlation among regions under prediction. As shown in Figure 7, MSE, RMSE and MAE gradually decrease

Table 1. Cotton yield prediction performance comparison of year 2003 while combining all data sources. Bold values represent the best results.

	MSE	RMSE	MAE	MAPE	Max Error
Linear Regression	17,875.6	133.7	109.0	10.96	276.7
Random Forest	19,295.3	138.9	112.0	10.68	321.7
Support Vector Regression	33,588.5	183.3	142.2	13.35	475.4
XGBoost	19,498.4	139.6	111.4	10.05	353.0
Deep Neural Network	9,504.8	97.5	77.8	9.96	269.4
Multi-Task Learning	8,267.5	90.9	70.5	8.08	256.2
Spatial-Temporal M.T.L.	7,013.5	83.7	63.6	7.55	254.4

Table 2. Performance of the spatial-temporal Multi-task Learning model using individual source of data.

Input Source	MSE	RMSE	MAE	MAPE	Max Error
Soil Properties	15,791.9	125.7	98.2	12.04	372.7
Spectral Data	21,292.5	145.9	114.4	13.62	450.1
NDVI	15,134.5	123.0	94.7	11.80	470.4

when the distance increases. This trend stops when neighborhood size equals to 4. The performance becomes more stable afterwards. Even though we see a random value in MAPE and ME metrics with respect to the neighborhood distance, there is also a decreasing trend on MAPE and ME when neighborhood size increases. Therefore, in our experiment, we set the neighbor distance as 5.

Understanding the Importance of the Features: Since weather data is shared in all regions under prediction, we do not evaluate its impact. Instead, we explore the impacts of soil properties, spectral conditions before planting and NDVI on the cotton yield prediction. We split the data by dimensions and conduct two experiments: in Table 2, we use one data source at a time to compare the importance of this single source, in Table 3 we remove one data source and use the rest input each time to compare the performance.

Table 2 shows that NDVI data contributes most significantly to better precision. It gets MSE and RMSE values at 15,134.5 and 123.0 while the spectral data produces the worst results, whose MSE and RMSE are 21,292.5 and 145.9, respectively. Table 3 indicates that if we ignore the spectral feature, the model achieves the best results compared with ignoring the soil properties or NDVI features. These results demonstrate that the NDVI impacts the prediction the most, then soil properties, while spectral data before planting has the minimal impact.

Table 3. Discover the importance of sources of data by removing one source at a time.

Removed Source	MSE	RMSE	MAE	MAPE	Max Error
Soil Properties	13,383.0	115.7	88.8	10.79	481.6
Spectral Data	13,189.6	114.8	93.0	11.54	276.5
NDVI	15,642.0	125.1	100.3	12.39	330.0

5 Conclusion

This paper proposes a novel Multi-task Learning framework for within-field scale cotton yield prediction, which ingests multiple heterogeneous data sources, such as soil type, weather, topographic, and remote sensing, and is capable of predicting within-field cotton yield throughout the growing season. By aggregating these multiple data sources in the shared layer before transferring to task-specific layers, this creative strategy is able to enhance specific learning task by utilizing sources from other related tasks. To minimize the spatial errors in prediction, this work introduces a spatial regularization to measure the correlations between a certain grid and its neighboring grids. The experimental results show the proposed approach consistently outperforms other competing approaches, and has a promising future in the crop yield prediction research field.

Acknowledgement. This work was supported by the U.S. National Science Foundation under the Grant CNS-1737634.

References

1. W. G. Bastiaanssen and S. Ali, “A new crop yield forecasting model based on satellite measurements applied across the indus basin, pakistan,” *Agriculture, ecosystems & environment*, vol. 94, no. 3, pp. 321–340, 2003.
2. K. Hebbar, M. Venugopalan, M. Seshasai, K. Rao, B. Patil, A. Prakash, V. Kumar, K. Hebbar, P. Jeyakumar, K. Bandhopadhyay *et al.*, “Predicting cotton production using infocrop-cotton simulation model, remote sensing and spatial agro-climatic data,” *Current Science*, pp. 1570–1579, 2008.
3. D. Akay and M. Atak, “Grey prediction with rolling mechanism for electricity demand forecasting of turkey,” *Energy*, vol. 32, no. 9, pp. 1670–1675, 2007.
4. G. Sugihara and R. M. May, “Nonlinear forecasting as a way of distinguishing chaos from measurement error in time series,” *Nature*, vol. 344, no. 6268, p. 734, 1990.
5. O. Garcia, “A stochastic differential equation model for the height growth of forest stands,” *Biometrics*, pp. 1059–1072, 1983.
6. W. Guo, S. Cui, J. Torrión, and N. Rajan, “Data-driven precision agriculture: Opportunities and challenges,” in *Soil-Specific Farming*. CRC Press, 2015, pp. 366–385.

7. Z. Jiang, C. Liu, N. P. Hendricks, B. Ganapathysubramanian, D. J. Hayes, and S. Sarkar, "Predicting county level corn yields using deep long short term memory models," *arXiv preprint arXiv:1805.12044*, 2018.
8. M. Mkhabela, P. Bullock, S. Raj, S. Wang, and Y. Yang, "Crop yield forecasting on the canadian prairies using modis ndvi data," *Agricultural and Forest Meteorology*, vol. 151, no. 3, pp. 385–393, 2011.
9. D. K. Bolton and M. A. Friedl, "Forecasting crop yield using remotely sensed vegetation indices and crop phenology metrics," *Agricultural and Forest Meteorology*, vol. 173, pp. 74–84, 2013.
10. W. Schlenker and M. J. Roberts, "Nonlinear temperature effects indicate severe damages to us crop yields under climate change," *Proceedings of the National Academy of sciences*, vol. 106, no. 37, pp. 15 594–15 598, 2009.
11. D. B. Lobell and M. B. Burke, "On the use of statistical models to predict crop yield responses to climate change," *Agricultural and Forest Meteorology*, vol. 150, no. 11, pp. 1443–1452, 2010.
12. S. S. Dahikar and S. V. Rode, "Agricultural crop yield prediction using artificial neural network approach," *IJIREEICE'14*, vol. 2, no. 1, pp. 683–686, 2014.
13. I. Oliveira, R. L. Cunha, B. Silva, and M. A. Netto, "A scalable machine learning system for pre-season agriculture yield forecast," *arXiv preprint arXiv:1806.09244*, 2018.
14. J. You, X. Li, M. Low, D. Lobell, and S. Ermon, "Deep gaussian process for crop yield prediction based on remote sensing data." in *AAAI*, 2017, pp. 4559–4566.
15. B. Ji, Y. Sun, S. Yang, and J. Wan, "Artificial neural networks for rice yield prediction in mountainous regions," *The Journal of Agricultural Science*, vol. 145, no. 3, pp. 249–261, 2007.
16. L. Zhao, Q. Sun, J. Ye, F. Chen, C.-T. Lu, and N. Ramakrishnan, "Multi-task learning for spatio-temporal event forecasting," in *Proc. KDD'15*. ACM, 2015, pp. 1503–1512.
17. J. Zhou, L. Yuan, J. Liu, and J. Ye, "A multi-task learning formulation for predicting disease progression," in *Proc. KDD'11*. ACM, 2011, pp. 814–822.
18. G. Balikas, S. Moura, and M.-R. Amini, "Multitask learning for fine-grained twitter sentiment analysis," in *Proc. SIGIR'17*. ACM, 2017, pp. 1005–1008.
19. Y. Zhang, P. Zhang, and Y. Yan, "Attention-based lstm with multi-task learning for distant speech recognition," *Interspeech 2017*, pp. 3857–3861, 2017.
20. Y. Lu, A. Kumar, S. Zhai, Y. Cheng, T. Javidi, and R. S. Feris, "Fully-adaptive feature sharing in multi-task networks with applications in person attribute classification." in *CVPR*, vol. 1, no. 2, 2017, p. 6.
21. J. Xu, P.-N. Tan, J. Zhou, and L. Luo, "Online multi-task learning framework for ensemble forecasting," *IEEE Transactions on Knowledge and Data Engineering*, vol. 29, no. 6, pp. 1268–1280, 2017.
22. H. Yao, F. Wu, J. Ke, X. Tang, Y. Jia, S. Lu, P. Gong, and J. Ye, "Deep multi-view spatial-temporal network for taxi demand prediction," *arXiv preprint arXiv:1802.08714*, 2018.
23. A. Hanna, P. Harlan, and D. Lewis, "Soil available water as influenced by landscape position and aspect 1," *Agronomy Journal*, vol. 74, no. 6, pp. 999–1004, 1982.
24. N. Kitchen, K. Sudduth, and S. Drummond, "Soil electrical conductivity as a crop productivity measure for claypan soils," *Journal of Production Agriculture*, vol. 12, no. 4, pp. 607–617, 1999.