Strawberry Yield Prediction Models Based on Imagery Information

Won Suk Lee, Feng Wu, Amr Abd-Elrahman, Natalia Peres, and Shinsuke Agehara

Summary

The objective of this project was to improve strawberry yield prediction models using field images and other variables. The imaging platform was upgraded with a high-performance portable computer and used to acquire images of strawberry plants. Multiple sets of images were acquired during the season, however the image analysis was interrupted due to the COVID-19 shutdown, and flower and fruit counting is currently being conducted. In February 2020, an autonomous robot was donated for this research by a company in Korea. However, it was not able to be used due to the COVID-19 shutdown. Using the dataset collected during the 2018-2019 season, plot-level yield prediction models were developed, using flower and fruit counts, weather, canopy size, days since planting, and historical yields. The results showed that flower and fruit counts and canopy size greatly improved yield prediction performance.

Methods

Field image acquisition

During the 2019-20 growing season, a strawberry field was prepared at the Plant Science Research and Education Unit (PSREU) at the University of Florida in Citra, Florida. Ten rows of strawberry plants, each 240 feet long, were used for the experiments. The strawberry cultivar was 'Florida Brilliance'.

For field image acquisition, the imaging cart was upgraded this year with a high-performance compact computer (NUC 8, Intel). Its size was 11 x 7.8 x 5 inches and weighed only 6.7 lb. It had a 3.1 GHz processer, 16 GB memory, 1 TB solid-state hard drive. Four USB ports were available so that the cameras (Grasshopper, Point Grey) could be connected simultaneously and used for image acquisition. Different artificial intelligence algorithms such as single shot multibox detector (SSD) and You Look Only Once (YOLO) v.3 were used. A total of 2,000 images were manually labeled and used for training the artificial intelligence algorithms.

A company in Korea (Unmanned Solution, Ltd., Seoul, Korea) donated an autonomous robot for this study in February 2020 as shown in Fig. 1. It is equipped with four 8-inch independent wheels, two 24 VDC geared motors, main and sub-controllers, and serial and controller area network (CAN) interfaces. It can travel at a maximum of 4 mi/hr, and its payload is up to 110 pounds.

Figure 1. A four-wheel drive autonomous robot for this study donated by a company in Korea.

As shown in Fig. 2, the robot was programmed for autonomous navigation. Once the navigation paths were programmed, it was able to navigate the strawberry field by itself. Fig. 3 shows preliminary testing in the field in mid-March 2020. Fig. 4 shows an inside view of the robot, where four cameras and lights were installed for image acquisition.

Unfortunately, due to the COVID-19 shutdown, all research activities were discontinued since mid-March including field image acquisition and testing. The images are currently being analyzed for identifying and counting the number of flowers and fruit at the time of this publication.

Figure 2. Programming navigation paths using the robot.

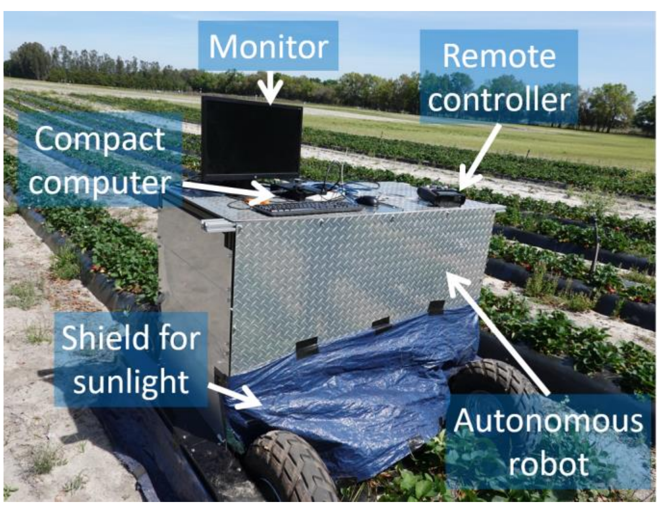

Figure 3. Preliminary testing - field image acquisition with a new high-performance compact computer and autonomous robot.

Figure 4. Inside view of the robot. Four cameras and illuminations are installed for image acquisition.

Yield prediction models

Strawberry yield prediction models were developed to aid growers in marketing, harvesting, and distribution decisions across the season. This project demonstrated the feasibility of using flower and fruit counts and canopy size for statistical yield modeling through the strawberry season.

Analysis was performed on the dataset collected during the 2018-2019 season. The dataset was collected from eighteen plots of two cultivars (Florida Radiance and Florida Beauty) at the GCREC. Strawberry flower and fruit counts, yield, and high-resolution visible and infrared image data were collected approximately two times per week (~27 acquisition session per season) from November 15, 2018 to March 4, 2019. The images were analyzed to extract canopy size variables such as canopy area, average canopy height, canopy height standard deviation, and canopy volume for each plot. Field observations were also collected from 6 plants per plot to provide actual flower and fruit counts. We acquired weather data from the Florida Automated Weather Network (FAWN).

Plot-level yield prediction models were developed using linear regression models. In addition to weather variables and canopy size variables, we also included time (days since planting) and historical yields as predictor variables. The models were estimated by the least square method. We used the adjusted R^2, or goodness of fit, to select the prediction model. The model with the largest value of adjusted R^2 was selected as the best model. Given the large number of available predictor variables, the stepwise regression method was used to fit regression models and choose predictor variables that provided the highest predictive power.

Results & Observations

We first started a prediction model with easily accessible, basic variables, such as time and previous yields, and then added weather variables, until all imagery metrics were included in the model. We implemented a time series analysis to predict yield 3-4 days ahead of harvest. Table 1 presents the goodness of fit associated with different combinations of predictor variables.

The goodness of fit reached 72% using only previous yields and time as predictors for 3-4 days ahead of harvest yield (Model 1). With actual weather data used in the model, the goodness of fit increased to 89% (Model 2). Our results show that flower and fruit counts and imagery data, such as canopy size variables, are particularly instrumental in improving the prediction performance. The goodness of fit reached 92% when lagged harvest yields, time, weather, lagged flower and fruit counts, and canopy size variables were used (Model 3). Fig. 5 shows that the actual yield and predicted yield generated from Model 3 matched closely. The performance of yield prediction at 1-week ahead of harvest was also encouraging (Table 1 and Fig. 6). Lastly, three-week-ahead yield prediction models presented an even stronger fit when exploiting imagery data (Fig. 7), and the goodness of fit was as high as 99% (Table 1).

Conclusions

The results provided strong evidence that image-based calculation of flower and fruit counts and canopy sizes could be a valuable tool for strawberry yield prediction. These variables could substantially improve predictive power.

In this phase of the project, the field imaging system was upgraded with a compact portable high-performance computer and an autonomous robot. The developed system is expected to be more robust and reliable, allow easy field setup and maintenance and take us one step closer to commercial implementation.

Acknowledgments

The authors would like to thank Mr. Doug Thomas at Crown Nursery, LLC for providing strawberry transplants and Mr. Hee Chang Moon at Unmanned Solution, Ltd. In Korea for donating an autonomous robot for this study.

Contact

Dr. Won Suk "Daniel" Lee
Dept. of Agricultural and Biological Engineering
University of Florida
1741 Museum Road
Gainesville, FL 32611-0570
P: 352-294-6721
E: wslee@ufl.edu

Dr. Feng Wu
Gulf Coast Research and Education Center
University of Florida
14625 CR 672, Wimauma, FL 33598
P: 813-419-6591
E: fengwu@ufl.edu

Table 1. Goodness of fit of different yield prediction models.

Predictor	Time	Historical yield	Weather	Flower and fruit counts	Canopy size	Goodness of fit
Yield prediction 3-4 days ahead of harvest						
Model 1	X	X				72%
Model 2	X	X	X			89%
Model 3	X	X	X	X	X	92%
Yield prediction 1 week ahead of harvest						
Model 1	X	X				92%
Model 2	X	X	X			95%
Model 3	X	X	X	X	X	97%
Yield prediction 3 weeks ahead of harvest						
Model 1	X	X				96%
Model 2	X	X	X			98%
Model 3	X	X	X	X	X	99%

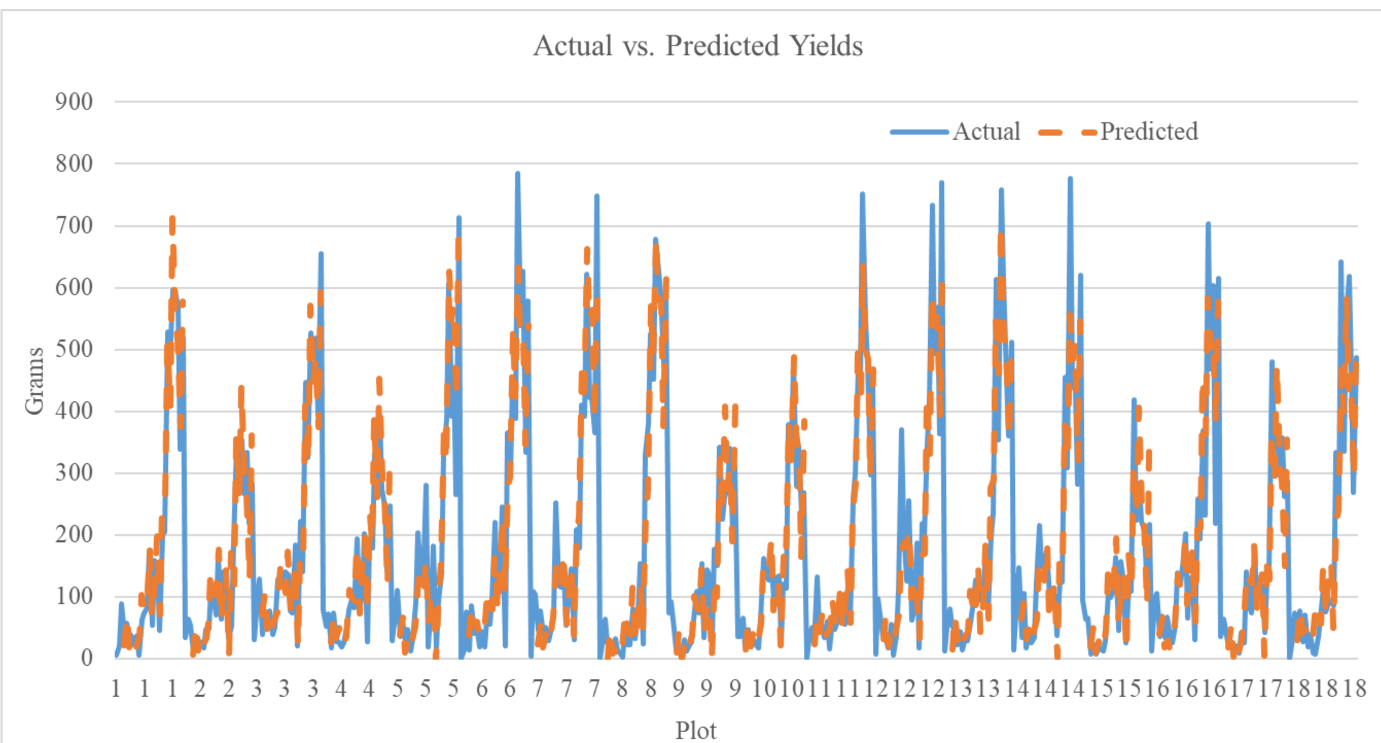

Figure 5. Actual yield at 3-4 days ahead of harvest across 18 plots and predicted yield generated from Model 3.

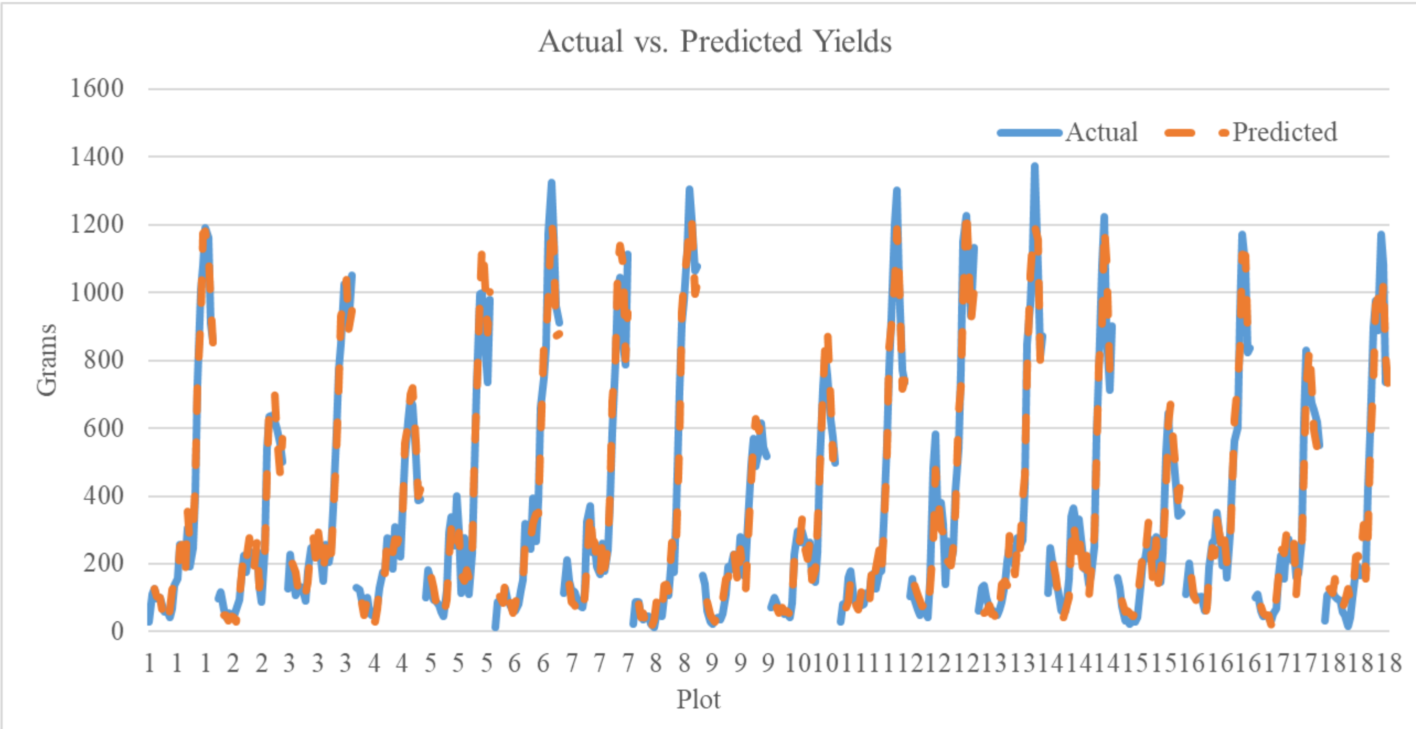

Figure 6. Actual yield at 1-week ahead of harvest across 18 plots and predicted yield generated from Model 3.

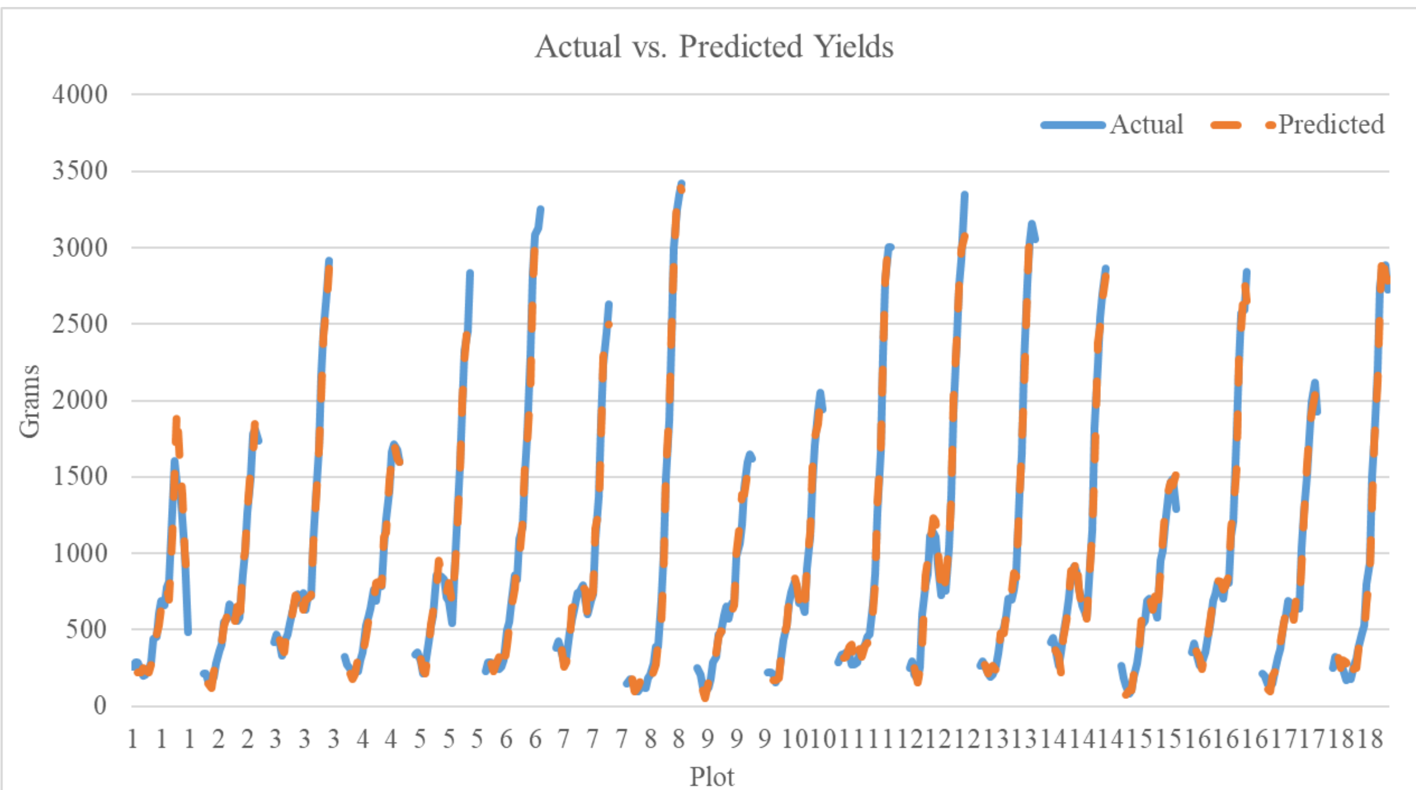

Figure 7. Actual yield at 3 weeks ahead of harvest across 18 plots and predicted yield generated from Model 3.