**Objectives:**

*1. Validate and confirm the repeatability of hyperspectral observations of soybean root diseases*. To address this objective of the project we completed field-based experiments of brown stem rot and sudden death syndrome inoculations. In short, we conducted inoculations on different varieties at a field site for the two diseases. The sites also had paired non-inoculated plants for comparison purposes. Throughout the growing season we collected handheld hyperspectral data aproximately weekly. Collecting data on the number of genotypes and throughout the growing season is imperative to determine if hyperspectral profiles are reproducibly from year to year and environment to environment. Throughout the project we extracted and analyzed our hyperspectral data. We continued to improve the accuracy and speed of our hyperspectral data extraction, which is imperative to our data analysis pipeline when working with large amounts of data and used the data along with controlled environment disease experiments to test the consistency of our detection methods.

2. *Determine the effectiveness of disease models for elite breeding material before selections for crossing.*To address this objective, we used the data from our field-based inoculation experiments. In brief, we conducted inoculations on 15 varieties at our field site. The sites also had paired non-inoculated plants for comparison purposes. We collected hyperspectral images and data approximately every week until the plots started senescing. This data is imperative to determine if hyperspectral profiles indicative of soybean plant resistance to root infection are consistent and accurate throughout a panel of soybean cultivars and breeding lines. Numerous early season measurements were taken before crossing would happen in a breeding program for data to be available to complete this objective. With this data we tested and developed early detection algorithms based on the hyperspectral data. The final output of this analysis is to attempt is to push back the detection of disease infection before crosses would need to be made in a breeder’s nursery, which we are making progress towards.

3. *Refine prediction model to reduce the amount of input information necessary to maintain accuracy.* During this project we incorporated our newly collected data into our previously developed detection models. We focused on limiting the number of wavelengths necessary for the detection of the diseases. Limiting the wavelengths is necessary to push this application into an applied setting for the data collection and analysis techniques. Currently we were able to push towards using eight wavelengths with an accuracy of detection of approximately 85%.

**Achievements:**

This report covers progress from the entire project, 5/1/2019 to 4/30/2020. This is a new project funded by the MSR&PC but builds on our knowledge and pipelines developed during our previous MSR&PC funded project the previous two years. Throughout this project we collated field and indoor growth chamber experiments data, data preservation, image data trait extraction, and analysis for completion of the project objectives. For this project, we researched two different soybean pathogens, brown stem rot (BSR) and soybean sudden death syndrome (SDS). We used 15 genotypes, 10 of which have been provided by the soybean breeding program run by Aaron Lorenz at the UMN. The addition of these lines helped us determine the ability of our methods in current breeding germplasm. The selected cultivars were based on maturity rating and being elite breeding material. We choose lines that would mature around the same time in Rosemount, MN and scored these lines for pathogen response to the pathogens (from resistant to highly susceptible) as it was largely unknown in this breeding material. The SDS trial was grown at the Rosemount Research and Outreach Center.  This was due to limitations of applying different diseases at the field sites, with the St. Paul fields being more restrictive. Starting around 30 days after planting hyperspectral measurements were made approximately once a week. During the project we worked to collate the collected data from the summer field growing season and incorporated the data into our data analysis pipelines for early disease detection. At the SDS trial in Rosemount the genotypes were planted in two different treatments (non-inoculated and inoculated) and replicated five times for 150 total plots. Within each of these plots we phenotyped 6 different leaves per plant and are using plot averages in our downstream data analyses. Growth chamber and greenhouse experiments were also run in parallel throughout the project with the same set of lines and data collection and analysis techniques. This allowed us to test the reproducibility in different environments for robustness.

For the second objective the collected handheld hyperspectral data of the breeding material in our Rosemount SDS trial was used for analysis. We focused our efforts to be able to detect/predict infection at or before when selections would be conducted in a conventional soybean breeding program. By focusing our analysis with this in mind this project is of interest to the breeding community and growers as well.

Throughout the project we ensured all the data was maintained in our management system to preserve for potential future use and public sharing. The data generated in the project was quality controlled and analysis techniques were used to determine unique spectral signatures for each soybean pathogen tested. We generated a large amount of phenotyping data for this project and we worked out important data storage and handling techniques that are necessary for successful completion of the project objectives and long-term data preservation.

We tested and implemented numerous data pipelines for this project. In addition, we have developed tools to automate data handling, manipulation, and analyses. We were able to process raw data to normalized data in less than a day, which is very important for broader applications of the techniques. By looking at data during late vegetative and early reproductive stages of growth, we detected unique spectral profiles for these two soybean diseases. We also used the data to improve, test, and validate our detection models based on different machine learning techniques. We are especially interested in reducing the number of wavelengths needed to detect the two disease, as this will most likely increase the adoption of the detection techniques. We are excited with the high reproducibility of spectral profiles we obtained for these diseases and are hoping to test them against other diseases and stresses after the projection completion. We are also exploring other data analytical techniques that might afford a more easily adaptable way for this research to be disseminated in a production field setting.

**Challenges:**

Our biggest challenge in this project is ensuring adequate and consistent inoculation of plants with the two diseases we are studying, brown stem rot and sudden death syndrome. This challenge is both prevalent in controlled and field conditions. We are taking all the necessary steps to mitigate this challenge as best we can, but we can’t control everything within experiments or between them either. We are continuing to perfect our inoculation methods and growing conditions to limit inoculation variation as much as possible. Also the challenges at the end of the project that Covid-19 placed on limits of personnel interaction and lab access we difficult and new territory for the project completion.

**Publication(s)/Symposium:**

This project work has led to the near completion of a manuscript of the data and analysis conducted in this proposal and our previous MSR&PC projects. This manuscript is planning on being submitted and accepted for publication in the coming months. Data and analysis techniques will be shared publicly upon publication.

**Tech Transfer:**

We were fortunate to have different avenues to show this work off to the broader community throughout the project. Parts of this project and previously related funded MSR&PC projects to our group were presented at as an oral presentation by Cory Hirsch at The Plant and Animal Genome Conference to about 50 people, mainly academic with some industry and government members. Parts of this project and previously related funded MSR&PC projects to our group were also presented at a Corteva sponsored seminar series in Brazil at the University of Vicosa to about 100 people, mainly academic with some industry members. Also, this project was presented as a poster at the Minnesota AgExpo in Mankato.