

CS CAPSTONE MID TERM PROGRESS REPORT WINTER 2018

FEBRUARY 16, 2018

Linking Seasonal Weather Data to AgBizClimateTM

PREPARED FOR

OREGON STATE UNIVERSITY

CLARK SEAVERT

PREPARED BY

GROUP26

THOMAS NOELCKE

SHANE BARRANTES

SHENGPEI YUAN

Abstract

Contents

1	Introduction	3
1.1	Purpose	3
1.2	Overview	3
1.3	Scope	3
1.4	Definitions, Acronyms and Abbreviations	3
1.5	Product Function Overview	4
1.6	References	4
2	Current Project State	4
2.1	Resolved Items	4
2.2	In progress	5
2.3	ToDo's	5
2.4	Blockers	6
2.4.1	netcdf problems	6
3	Weekly summary of progress	6
3.1	Week 1	6
3.1.1	Plans	6
3.1.2	Progress	6
3.1.3	Problems	6
3.1.4	Summary	6
3.2	Week 2	7
3.2.1	Plans	7
3.2.2	Progress	7
3.2.3	Problems	7
3.2.4	Summary	7
3.3	Week 3	7
3.3.1	Plans	7
3.3.2	Progress	7
3.3.3	Problems	8
3.3.4	Summary	8
3.4	week 4	8
3.4.1	Plans	8
3.4.2	Progress	8
3.4.3	Problems	8
3.4.4	Summary	8
3.5	week 5	9
3.5.1	Plans	9
3.5.2	Progress	9
3.5.3	Problems	9
3.5.4	Summary	9
3.6	week 6	9
3.6.1	Plans	9
3.6.2	Progress	10
3.6.3	Problems	10
3.6.4	Summary	10
4	Individual Section	10
4.1	Thomas Noelcke	10
4.1.1	My Contributions	10
4.1.2	Peer Review	11
4.2	Shane Barrantes	12
4.3	Shengpei Yuan	12

List of Figures

1 Introduction

1.1 Purpose

The purpose of this document is to describe the progress we have made so far on the *AgBizClimate* project. In this document we will give a brief introduction to the *AgBizClimate* project. In this section we will discuss the purpose of the project. Additionally, we will also discuss the scope of the project and an overview of the project functions. This document is designed for the project owners. This document is also designed for the development team so we can evaluate our progress so far on this project. This project is also designed to fulfill the minimum requirements for the CS461 class for the OSU computer science program.

1.2 Overview

Seasonal climate is one of the essential factors that affects agricultural production. As a module of *AgBiz Logic*, *AgBizClimate* delivers essential information about climate change to farmers, and help professionals to develop management pathways that best fit their operations under a changing climate. This project aims to link the crucial seasonal climate data from the Northwest Climate Toolbox database to *AgBiz Logic* so that it can provide changes in net returns of crop and livestock enterprises through powerful graphics and tables.

1.3 Scope

This project is a part of a much larger *AgBiz Logic*TM program. However, the purpose of this project is to add a short term climate tool to the *AgBizClimate* module. This limits the scope of the project to the *AgBizClimate* Module. Additionally, we will only be adding the short term climate data tool as the long term climate data tool already exists.

Currently *AgBizClimate* has a long-term climate tool but no such tool exists for short term climate data. We will implement a tool to extract short-term climate data from the Northwest Climate Toolbox database, display it to the user and allow the user to adjust crop and livestock yields or quality of products sold and, production inputs. Moreover, a landing tool will be developed to allow users to switch between short-term seasonal tool and long-term climate data tool.

1.4 Definitions, Acronyms and Abbreviations

REST - Representational State Transfer, This is a type of architecture that manages the state of the program. This is especially popular in web development.

API- Application Programming Interface. This is a piece of software that allows a connection to another piece of software providing some sort of service.

NWCTB - Northwest Climate Toolbox. This is the database we will be connecting to that will provide the short term climate data we plan to use.

Thredds Data Server - This is a web server that provides meta-data and data access for scientific data sets using OPeNDAP along with some other remote data access protocols.

OPeNDAP - Open-source Project for a Network Data Access Protocol. This is the protocol we will be using to retrieve the data sets from the Thredds data server.

NMME - North American Multi-Model Ensemble. This is a data set that brings together a variety of different weather models into one data set.

Climate Scenario - This is a theoretical calculation of yields, inputs and of the overall budget for one situation based on the climate data.

NETCDF - This is a file storage format for large scientific data sets especially good for any data that is referenced on a grid and related to is geo-location.

1.5 Product Function Overview

AgBizClimate is a web based decision tool that will allow users to gain specific insight on how localized climate data for the next seven months will affect their crop and livestock yields or quality of products sold and production inputs. The *AgBizClimate* tool will allow users to input their location (state, county) and a budget for the specific crop or livestock enterprise. *AgBizClimate* will select climate data for the next seven months for that location and provide graphical data showing temperature and precipitation. Users will then be able to change yields or quality of product sold by a percentage they think these factors will affect and modify production inputs. Finally the tool will calculate the net returns.

1.6 References

- [1] C. F. Seavert, “Negotiating new lease arrangements with the transition to direct seed intensive cropping systems,” 2017.
- [2] S. Y. Thomas Noelcke, Shane Barrantes, “Problem statement,” 2017.
- [3] S. Y. Shane Barrantes, Thomas Noelcke, “Requirements document,” 2017.
- [4] —, “Tech review,” 2017.
- [5] —, “Design document,” 2017.

2 Current Project State

In this section of this document we will review what items we have resolved, List and explain what items we still have to do, discuss major blockers and give a week by week summary of what we have done so far. This section is intended to give a good overall picture of the status of the *AgBizClimate* project along with what we have accomplished so far.

2.1 Resolved Items

Updated Requirements Document: The requirements document has been updated to reflect changes to the climate data API. We were expecting to have an endpoint we could call with a latitude and a longitude however accessing the climate data is going to require more work that we had originally expected. So we’ve updated the requirements document to reflect this change.

Create Concept script for Thredds database: We’ve created a concept script to use as an example in our climate API that connects to the Thredds database and read back the climate data for the Latitude and Longitude we’ve request. However, we are getting errors on reads latter in the data set that we can’t explain.

Install Script for NETCDF4 and Dependencies: Another problem we encountered while creating the concept script was setting up and installing the python NETCDF4 library. Luckily we were able to write a scrip that installs this library and its dependencies.

Added option to select Short term Climate Scenario: We added a drop down to the region selection page that allows users to select short term climate scenarios and long term climate Scenarios. This will allow people to use the tool we are creating.

Created The Charts Page: We've also created the page where the users will view the climate data. This page will also allow the users to enter a change in yield for their crops and will then redirect them to the Budget review page. Though this page has a few issues most of the work is completed.

2.2 In progress

In this section we will discuss work items that are in active development. We will describe each item giving a brief description. We will also discuss how much progress has been made on each item.

Climate Data API: We have started work on the back end Climate data API however currently it is just an end point that distributes mocked data as we have run into issues with the netcdf4 library we are using to connect to the database with all the data in it. Once we have a solution for connecting to and reading all the values from the database it won't be hard to get real data from the API.

Document Updates: We have had some serious changes to the expected design as a result we will need to make some document updates. We have started some of these updates. However, I don't anticipate being able to finish all the document updates this term.

2.3 ToDo's

In this section we will discuss items we will need still need to do over the next two terms. We will give a brief description of each item. We will also state an approximate time for when we expect each item to be completed.

Backend Testing: Given that our Climate data has not yet been well defined we haven't started writing test to test the back end code. Once we get the API written in something close to its final form we will want to start writing some test to ensure that the API meets the requirements in the requirements document.

Front End testing: Similarly as the front end work is still in active development we haven't written any tests to cover this code. However we will be done with front end development soon and should be able to get to work doing front end testing.

Finding a solution to NETCDF4 Problems: We will need to find a work around to get around errors we are having while reading from the database using the netcdf4 library. Currently, we think we can get around this by writing the bit that communicates with the server in c++. However we are trying installing dependencies for the netcdf library in a different way first.

QA: Finally, before we are done with our project will want to provide some QA on the work that we will have done over the term. This will include running some manual integration tests to ensure that we have completed the work we agreed to complete.

2.4 Blockers

In this section we will discuss hurdles we are facing that are impeding progress on this project. We will discuss why they will keep us from moving forward on this project and we will also discuss what we intend to do to move past these problems.

2.4.1 netcdf problems

Initially, we had issues getting netcdf4, a python library for dealing with the data files that contain the

3 Weekly summary of progress

In this section we will give a weekly summary of our progress on this project. For each we will list out our plans, problems we have encountered during the week and will show a summary of what we have accomplished during the week.

3.1 Week 1

3.1.1 Plans

- Meet with group to set up iteration one of project development.
- Meet with Sean to set up git branch and discuss git workflow.
- set tasks for iteration 1.

3.1.2 Progress

- Forked github repo from AgBiz-Logic
- Set up a meeting with Sean to discuss project development.
- Started setting up tasks for iteration one on the git repo.
- Started working on the Wiki page with common help items for the project.

3.1.3 Problems

- Still haven't heard any thing from the NWCTB team regarding API access for the climate data.

3.1.4 Summary

This week we tried to set up a meeting with Sean to do some project planning and set up for iteration one of our project. However, Sean was unavailable this week so we set up a meeting for next week. We also got our git repo, forked from *AgBiz-Logic* set up. We started planning the first iteration of development on the project by adding issues to the github repository. We also started compiling some help pages on the Wiki of our repo.

3.2 Week 2

3.2.1 Plans

- Meet with Sean.
- Start Iteration One.
- Get UI elements implemented along with most of the front end functionality.
- Plan iterations 2 and 3.

3.2.2 Progress

- Set up meeting with Sean Hammond for Friday at 1 pm.
- Finished setting up iteration one tasks.
- Finished adding content to the help wiki on the github repository.
- Finally defined Climate Data API Access.
- Set a Weekly status meeting time to meet with the group. We plan to meet every week at one pm.

3.2.3 Problems

- API Access is less than ideal and will require more work than we were planning on but is still better than having to write our own service from scratch.
- Finding time to meet up as a group has been more challenging than I had anticipated.

3.2.4 Summary

This week we didn't get much development work done on our project like we had planned on. However, we did do some set up work. we finished setting up the github repository and finished laying out tasks on our story board. We also started defining what tasks we'd like to have in future iterations of our project. Additionally, we finally know what our API access to the climate data looks like. this will allow us to get the data we will need to plot. However, this will also require much more work than we had planned on and may set us back a bit in terms of our project schedule. That being said we worked in some flex time in to our schedule so we should be able to make it work.

3.3 Week 3

3.3.1 Plans

- Create proof of concept script for connecting to the database and getting data.
- start working on front end changes.
- Update design document and requirement document.
- Meet with Sean for status update at 1pm on Friday.

3.3.2 Progress

- Started working on concept script.
- Managed to get dev environment set up instructions completed.
- Installed netcdf.

- Created example script for getting climate data from the thredds database. However we get some errors on certain reads.
- Updated requirements document.

3.3.3 Problems

- We had a hard time getting NETCDF4 to install. We ended up using anaconda however we are guessing Sean doesn't want to use Anaconda and will want us to produce an install script.

3.3.4 Summary

This week was primarily a week of setup. we spent most of our time trying to get the dev environment set up along with installing NETCDF4 and its dependencies. This week we did find a way to install NETCDF4 using anaconda. However, we anticipate that we will be required to find a better way to install it. In the mean time this will allow us to develop a concept script. We also managed to the development environment for AgBiz-Logic set up. This took us more time than we had anticipated but wasn't as difficult as we thought it might be. This week we also made some updates to the requirements document to reflect the changes to the climate data API.

3.4 week 4

3.4.1 Plans

- Start working on the front end of the application.
- Refine the proof of concept to be more dynamic.
- Write script to install netcdf and dependencies.
- Start working on backend changes.
- Update documents.

3.4.2 Progress

- Started working on refining proof of concept script to search for points if the point we asked for doesn't have data also added more advanced bounds checking.
- Determined that NETCDF4 is having issues reading in blocks for chunk three.

3.4.3 Problems

- requests past index 435 on latitude cause a runtime error.

3.4.4 Summary

This week was mostly focused on working on the proof of concept scrip that will be used later to access the data from the thredds server. This week we discovered that the NETCDF4 library throws errors on any lat index greater than 435. We also discovered through the database administrator that this is the boundary between chunk two and chunk three of the file we are trying to read. We think that the NETCDF4 library may have a bug in it. Regardless we are going to need to find a work around moving forward. Shane and Thomas also got together on Saturday and started working on front end changes.

3.5 week 5

3.5.1 Plans

- Work on front end changes.
- Follow up with NETCDF4 developers about potential bug.
- Continue to work on concept script to see if we can tease out the runtime error.
- Finish NETCDF5 install script.
- Research other potential options other than python or netcdf4 for reading in data from the server.

3.5.2 Progress

- Followed up with netcdf4 people.
- Shane finished the netcdf4 install script.
- Researched alternatives to netcdf4 We can write a c program that will do the same thing. There are a few other libraries for reading data via opendap.
- made progress on frontend changes.

3.5.3 Problems

- Issues with netcdf4 library.
- netcdf4 developers will not fix unless I can produce a self contained example of the read failing.
- We think that netcdf4 dependencies may not be installed correctly.

3.5.4 Summary

This week we made progress on the front end development and installing the dependencies for netcdf. However, we've run into some issues with netcdf. We think we maybe able to fix it by installing the dependencies for netcdf from source with certain flags enabled but we aren't totally sure on that. We also started work on setting up the end point where the API will live. The plan for now is to have it serve mock data as to enable us to continue with the rest of the development work without getting behind.

3.6 week 6

3.6.1 Plans

- Finish development of the charts page.
- Set up API to mock data.
- Figure out work around for netCDF problems.
- write the midterm progress report.
- make the midterm progress presentation.
- finish the poster rough draft.

3.6.2 Progress

- Finished Development on the charts page.
- Finished the midterm progress report.
- Made the midterm progress report presentation.
- Finished the Poster rough draft.
- Found a work around for NETCDF4 issues.

3.6.3 Problems

- Created some bugs by introducing short term climate scenarios.

There are many ways we can fix this problem we will need to discuss with Sean how he wants this solved.

3.6.4 Summary

This week was a busy week. This week we accomplished most of the front end development required by this project in the maps page. However, we also introduced some bugs. Mostly we created an issue where if you have multiple budgets its not possible to tell what page you need to redirect to once you save your budget. There are many ways we can solve this so I want to ask Sean how he thinks the best way to go about this is. This week we also made our expo poster along with creating the midterm progress report document and presentation. Additionally, Shengpei found a work around to the NETCDF problems we've been having.

4 Individual Section

In this section we will each write about our individual parts of the project and the progress we have made. We will also use this section to provide a peer review for our group mates rating their performance and contribution to the project.

4.1 Thomas Noelcke

In this section I will discuss the various contributions I've made to the project along with some of the problems I've come across including any solutions to problems that I've found. I will also discuss the contribution of my team mates to this project giving each of my group mates a peer review.

4.1.1 My Contributions

Through out this project I've had the opportunity to work on many aspects of the project. So far I've worked on the concept script, documentation, and worked on the front end. However, these are not my sections below in the section titled My Sections I will explain why what I've worked on so far does not relate to my sections in the technical review.

Concept API Script

One of the first things I worked on this term was creating a concept script that connected with the Thredds server and read in NETCDF climate data from the server. The goal of this script was to prove that it was possible to use the NETCDF4 library and python to retrieve the data we needed for our project. In the end the script I worked on proved to be a dead end. Though through the process of writing this script I learned a lot about how scientific data

storage works. In the end we ran in to a dead end in a bug that the creator of the NETCDF4 library stated they would not fix. Ultimately Shengpei ended up finding another solution to this problem with out using the NETCDF library so I passed the development of the concept script off to him.

Documentation

Another item I worked on this term was updating the requirements document as a result of changes to the climate API. I made modifications to the glossary and to the functional requirements to reflect the changes made to the project. The big change made to the project was a change in the API Access. This necessitated the updating of the requirements of the project.

Front-end Work

The other work I've been spending a lot of my time on this term is front end development. So far I've done most of the development on the Charts page. This page is the meat and potatoes of the UI in that its the page where we are showing the user the climate data and allowing them to make adjustments to their yields. I also added a drop down menu to the region selection page the allows the user to select a short term climate scenario. If you are interested in viewing the result of this work please see our midterm presentation as inserting code snip-its is impractical given the nature of Angular code.

My Sections

In the tech review my sections were the database, the testing frameworks and the HTTP frame work. I have only spend minimal time working on these sections because there just isn't that much work as relates to this project in these sections. I have not written many unit test yet because we have not developed that much testable code yet. I also havent spent any time working on the database because we've not made any data changes requiring migrations just yet. I have however spend some time working on with our HTTP framework while making request to the back end from the front end while doing front end development.

4.1.2 Peer Review

In this section I will provide a peer review for each of my group mates. For each group mate I will describe contributions made by each group member pointing out things each group member does well

Shane Barrantes

Throughout this project this term Shane has been a helpful team mate and has contributed to the project both by writing feature code and by preforming complex environment setup. Shane has provided much input into the execution and planing of this project throughout the beginning of the term. He has been an above average team mate in showing up to and contributing to meetings and work sessions.

I will be first to admit that package management and environment setup are areas of software development that I am weak in. As a result early in this project I struggled getting the environment set up and running. Shane was a huge help in setting up the environment and getting the project running. Not only did Shane help me fix issues with getting the project to build, he also taught me about package managers and project set up in general. Shane also helped the team out by writing install scripts for complicated installs. For instance early in the project we ran into many issues installing the Python netcdf4 library. Shane figured out how to install it and wrote script that installed the netcdf4 library along with its dependencies. Shane has also been very helpful in regards to developing the frond end and back end code. Shane started developing a complicated charts page on the front. He helped get the page set up so that the front end development could be carried out. However, he never had the chance to get that page completed because I needed him to create an elaborate install script that installed the HDF5 library from source enabling certain options on that package.

The only criticism that I may provide on Shane's performance so far this term is that outside of group meetings he hasn't been able to put in as much time as I had hoped. Generally, I think this has been difficult for Shane due a few things beyond his control such as deadlines at work and a heavy course load. As this project progresses this is one area where I hope to see improvement.

Shane has also shown his value as a team mate in the way that he communicates and contributes to meetings. Shane is not able to make it to every meeting, however Shane is very good about communicating when he is unable to make it to a meeting and does a good job of warning me of potential scheduling conflicts. When Shane is in meetings he is present, listening and contributing to the conversation. Shane generally focuses well on his work when at work sessions and group meetings.

Overall I would give Shane 810 as a group mate. I've taken two points off because Shane has had a hard time contributing to the project outside of group meetings and work sessions. I anticipate that this will improve as Shane has expressed that he is close to finishing a big project at work. Shane as also expressed that he would also like to see this aspect of his performance Improve.

Shengpei Yuan

Shengpei was largely absent from this project and the beginning of the term. How this was due to some extenuating circumstances. In the last few weeks Shengpei has been a much larger part of this project and has been contributing.

Early this term Shengpei expressed that he would be busy this term as he is taking a large credit load and also studding for his GMAT exam. As a result Shengpei was largely absent from the first four weeks of this project. This is not a personal attack on Shengpei but rather a statement of fact. However, in the last few weeks Shengpei has been more involed in this project.

In the last few weeks Shengpei has been showing up to more meetings and has been able to contribute more to the project. This is a direct result of Shengpei completing his GMAT exam. In his last few weeks Shengpie has being working on a work around for the NETCDF problems we've been having and has found a potential solution.

Overall I would give Shengpei a 510 as a group mate. This ratting has little to do with Shengpei's performance in meetings and when working on the project. But rather is a result of Shengpei being unable to contribute to the project through the first 4 weeks of the project. I anticipate that this will improve as Shengpei has finished his GMAT exam.

4.2 Shane Barrantes

4.3 Shengpei Yuan

Briefly recaps the project purposes and goals

The project is dedicated to call essential climate data from a professional climate toolbox database and deliver them in various ways to agriculture producers and land managers in order to help them develop better crop plans and achieve better harvest. Generally, the project contains two key tasks to accomplish the final goal. First, appropriate climate forecast database has to be chosen so that the system could call desired climate data in various ways according to the application contexts. Secondly, the immediately collected climate data should be appropriately stored and presented to users with different demands. In particular, the system would offer comprehensive data APIs for AgBizClimate web programs to extract and demonstrate interested climate data including the changing trends.

It is clear that we would mainly employ the Python program language and the powerful Django web framework to finish the project. And we adopt the popular MVC (Model-View-Controller) design pattern for the overall system architecture. During the past one term, I have finished some substantial work for the project. First, I have reviewed three candidate web design techniques for the project including containers, back end models and frameworks, and styling frameworks. Although I have little practical experiences on some of the web development tools or models, I

did roughly learn their general application scenarios and advantages and disadvantages over each other, which may be quite useful for my future studies or work in the area of web design. Second, I have drafted and revised the controller design that is an important part of the overall design document. As an interface between Model and View, the controller plays an important rule in the in the Django MVC idea of the system architecture.

The parts currently working on and the left work to do

As a member of the energetic team for the project, it is currently my responsibility to define and implement the climate data API. More specifically, the system could call the API to get climate forecast data back from some authoritative and professional climate database in a flexible form. In addition, it would be expected that the API could be set up like a service on the back end and called directly by any programs at any time when it is necessary. Also it is one of my partners work to transform the workable climate data API scripts into the Django app.

With regards to the future tasks, I would also focus on the backend climate data API. During the entire term of development of the project, the detailed requirements of API may change in some ways. Also the parameters such as model and climate variables need to be considered thoroughly when design the API interfaces. Last but not least important, it is necessary to write both unit and integration tests to ensure that each single function code and the entire system works exactly the same as we have expected in most scenarios.

Action items for the project currently include the Python program scripts for proof of conceptual design of API that call climate data from remote climate databases and updated documents of the project.

The troublesome problems and relevant solutions

There are several troublesome problems we may have or need to solve to make the climate forecast data API work correctly.

First, how to call or retrieve climate data of one point that are concerned from the website URL providing downscaled climate forecast data without downloading the entire dataset which may be quite large. In other words, in most application scenarios we may only need climate data for some small points on the map instead of the overall data of one grid or the USA. This is crucial for the system at least in sense that the entire climate data for one grid or USA is too large to load immediately from the remote database.

Second, how to connect with the climate database and get desired climate data including certain variables within acceptable time like a few seconds. This is necessary since we would like to call for climate data of some relevant variables for specific forecast targets.

Thirdly, for this project we are going to use an OPeNDAP dataset in NETCDF format served from a Thredds database that is provided by <https://climate.northwestknowledge.net/>. However, the NETCDF format is not as convenience to transmission or handle by Python programming languages as other formats like JSON. We thus may need to dig deep to convert the NETCDF format to JSON or other formats.

```
url = self.prepare_url()
r = requests.get(url)
r.close()
text = r.content
lines = text.replace('<br />', '\n').replace('<br />', '\n').split('\n')
for ln in lines:
    ln = ln.strip()
    if ln == '':
        continue
    if ln[0] == '#': # skip comment lines
        self.comments += ln + '\n'
    else: # climate forecast data
        self.climate_data.append(ln.split(','))
```

Although not complete and accurate, we have explored and found some general solutions to above three problems. For instance, we have tracked down the template URL to connect to the climate database and call desired climate data of certain point within a grid by some parameters. Through a formatted HTTP GET request, we could call specific climate data of a point on the map. The code I have showed above shows the sample python code snippets

to download climate forecast data of one point. The code first requests climate data from a formatted URL address and then transform and parse the returned text to get specific figures of climate data. It should be noted that the process is fast as we could specify the exact latitude and longitude and other control parameters to make the climate data to be downloaded each time very small. And we have determined that the ensemble mean forecast data from several available models would be introduced to make the results more convincing.

Moreover, to guarantee that the design and implementation of climate data API being finished both correctly and efficiently, we have laid out a set of five iterations with milestone tasks. First, we would make proof of the concept that the climate forecast data set could be called by Web interfaces like HTTP. Second, a preliminary climate data API would be built to get climate data of one model for any location of the US. Thirdly, the API would work correctly for multiple models available from the remote database. Fourthly, unit and integration tests are written to cover most functionalities of the system to find as many bugs as possible. Fifthly, we would fix bugs and revise the API if necessary to meet the requirements of the system.