# APPLYING SATELLITE REMOTE SENSING, UNMANNED SYSTEMS, AND MODELS FOR WATER QUALITY ANALYSIS

By

ABHIRAM SIVA PRASAD PAMULA

Bachelor of Technology in Civil Engineering.
Indian Institute of Technology
Bhubaneswar, India
2014

Master of Science in Environmental Engineering
Cleveland State University
Cleveland, Ohio
2018

Submitted to the Faculty of the
Graduate College of
Oklahoma State University
in partial fulfillment of
the requirements for
the Degree of
DOCTOR OF PHILOSOPHY
July, 2022

APPLYING SATELLITE REMOTE SENSING, UNMANNED SYSTEMS, AND MODELS FOR WATER QUALITY ANALYSIS

Dissertation Approved:

Dr. Mark J. Krzmarzick

Dissertation Advisor

Dr. David J. Lampert

Dr. Jorge Gonzalez-Estrella

Dr. Jamey D. Jacob

Dr. Hamed Gholizadeh

ACKNOWLEDGMENTS

I would like to thank my esteemed advisors - Dr. Mark Krzmarzick and Dr.David J. Lampert for their invaluable supervision, support and tutelage during the course of my PhD degree. Their immense knowledge in the field of water quality modeling and water quality monitoring have encouraged me in my academic research. I would like to thank all the committee members for their technical support and mentorship. Finally, I would like to express my gratitude to my parents for their support and I am grateful to have them in my life. Finally, this work is dedicated to my wife and child who understood and encouraged me in the past few years. It would be impossible to successfully complete this PhD journey without all of your support.

Acknowledgments reflect the views of the author and are not endorsed by committee members or Oklahoma State University.

*DEDICATION*

*I am dedicating this work to my mother whose love for me know no bounds. This work is also dedicated to my wife and lovely child who has been constant source of support and encouragement during the challenges of the PhD research. Finally I dedicate this to my father who was my moral compass and taught me the value of hard work.*

Dedication reflect the views of the author and are not endorsed by the committee members or Oklahoma State University.

Name: ABHIRAM SIVA PRASAD PAMULA

Date of Degree: July, 2022

Title of Study: APPLYING SATELLITE REMOTE SENSING, UNMANNED SYSTEMS, AND MODELS FOR WATER QUALITY ANALYSIS

Major Field: Civil Engineering

Abstract: Transport of soils and sediment by runoff diminish the surface water quality of lakes/reservoirs and shrink reservoir capacity. Nutrients move with transported soil and sediment due to soil erosion by floods and high wind. Nutrients from land management practices such as farming, animal feeding operations, and wastewater treatment stimulate the growth of cyanobacteria in surface water bodies.Excessive nutrient loading or eutrophication with favorable hydrodynamics result in excessive growth of cyanobacteria called Harmful Algal Blooms (HABs). These cyanobacteria sometimes release microcystins which are toxic to humans when consumed. The biological decay of HABs cause ecological dead zones due to depletion of dissolved oxygen while reducing the utility of the reservoirs. From an ecosystem service provision and human health perspective, HABs represent a great water quality threat. Therefore monitoring and modeling the formation of HABs in surface water bodies is important. In freshwater reservoirs of Oklahoma and Kansas, HAB formations and occurrences have increased due to floods, agricultural practices, and rising earth temperatures due to global warming and climate change. In Oklahoma, Grand Lake O' the Cherokees (Grand Lake) had experienced a severe HAB outbreak in 2011. In Kansas, Marion reservoir experiences HAB outbreaks almost every year. These two reservoirs provide an opportunity to study and understand HAB formation, predict them and provide warning advisories. All the in-situ data collected using unmanned systems in the freshwater reservoirs have been used to develop water quality and watershed models to understand the relationship between sediment, nutrient loadings and HAB formations.

TABLE OF CONTENTS

LIST OF TABLES

Table Page

LIST OF FIGURES

Figure Page

Figure Page

Figure Page

## ABBREVIATIONS

| | |
|---|---|
| HABs | Harmful Algal Blooms |
| USV | Unmanned Surface Vehicle |
| UAV | Unmanned Aerial Vehicle |
| ASV | Autonomous Surface Vehicle |
| HSPF | Hydrologic Simulation Program in Fortran |
| USA | United States of America |
| IR | Infrared radiation |
| EROS | Earth Resources and Observation Science |
| GCPs | Ground Control Points |
| ESA | European Space Agency |
| SVM | Support Vector Machines |
| MLPs | Multi-layer Perceptrons |
| SVR | Support Vector Regression |
| QP | Quadratic Programming |
| EC | Electrical Conductivity |
| DO | Dissolved Oxygen |
| OASIS | Oceanographic Atmosphere Sensor Integration System |
| USRI | Unmanned Systems Research Institute |
| OSU | Oklahoma State University |
| MANUEL | Mobile Autonomously Navigable USV for Evaluation of Lakes |
| ESC | Electronic Speed Controller |

| | |
|---|---|
| EPA | Environmental Protection Agency |
| USGS | United States Geological Survey |
| BASINS | Better Assessment Science Integrating Point and Non-Point Sources |
| GUI | Graphical User Interface |
| UCI | User Control Input |
| WDM | Watershed Data Management |
| BMPs | Best Management Practices |
| CHLa | Chlorophyll-a |
| RFR | Random Forest Regression |
| BGA | Blue-Green Algae/Phycocyanin concentrations |
| OLI | Operational Land Imager |
| TIRS | Thermal Infrared Sensor |
| MSS | Mineral Suspended Solids |
| GRDA | Grand River Dam Authority |
| NIR | Near Infrared |
| SWIR | Short Wave Infrared |
| RF | Random Forest |
| API | Application Program Interface |
| CSV | Comma Separated File |
| WHO | World Health Organization |
| MSE | Mean Square Error |
| RMSE | Root Mean Square Error |
| CV | Cross Validation |
| EFDC | Environmental Fluid Dynamics Code |
| OCC | Oklahoma Conservation Commission |
| NWIS | National Water Information System |

| | |
|---|---|
| NHDPlus | National Hydrography Dataset Plus |
| CDL | Cropland Data Layer |
| NCDC | National Climate Data Center |
| NSRDB | National Solar Radiation Database |
| $ET_0$ | Reference Evapotranspiration |
| PET | Potential Evapotranspiration |
| NSE | Nash and Sutcliffe Model Efficient Coefficient |
| TDS | Total Dissolved Solids |
| YSI | Yellow Springs Instrument Company Inc |
| FNU | Formazin Nephelometric Unit |
| QC | Quality Check |
| GPS | Global Positioning System |
| OWRB | Oklahoma Water Resources Board |
| DEM | Digital Elevation Model |
| LIDAR | Light Detection And Ranging |
| ABE | Acetone-Butanol-Ethanol |
| LCA | Life Cycle Assessment |
| STJ | Switchgrass to Jet Fuel |
| GHG | Greenhouse Gas |
| GWP-100 | Global Warming Potential-100 |
| CTJ | Corn to Jet Fuel |
| FTJ | Fischer-Tropsch Jet Fuel |
| HRJ | Hydro-Processed Renewable Jet Fuel |
| CPJ | Conventional Petroleum Jet Fuel |
| IPCC | Intergovernmental Panel on Climate Change |
| WTWa | Well-to-Wake |
| DDGS | Distillers Dried Grain Solubles |

| | |
|---|---|
| NG | Natural Gas |
| GREET | Greenhouse Gases, Regulated Emissions, and Energy Use in Transportation |
| CRP | Conservation Reserve Program |
| DAP | Dilute Acid Pretreatment |
| AFEX | Ammonia Fiber Expansion |

# CHAPTER I

# Introduction

## 1.1 Background and Problem Statement

The current world population is around 8 billion and is expected to reach close to 10 billion in 2050 and 11 billion in 2100 (Publications, 2019). This population growth is not uniform throughout the world, where the projected increase is expected in high-fertile lands of the globe. Excessive fertilizers usage shall maintain food security for the growing population, eventually bringing nutrients into lakes and reservoirs. Nutrient runoff from land management practices such as agriculture, animal feeding operations, water temperature, wastewater treatment plants promotes the growth of algae and cyanobacteria (Joyce, 2000). Excessive growth of algae in water causing water impairment issues are called Harmful Algal Blooms (HABs). HABs are majorly influenced by the availability of carbon, nitrogen, phosphorous, sunlight, water temperature and other trace nutrients (Tseng et al., 2014). Global warming and climate change result in extreme weather events, including a rise in the occurrence of summer storms which bring large amounts of sediment, nutrients, and other contaminants into freshwater bodies causing eutrophic conditions (excessive nutrient loadings) in lakes and reservoirs (Pielke Jr et al., 2005). In addition to eutrophication with increased temperature and sediment re-suspension, favors the growth of algae in aquatic ecosystems. Excess growth of algae concentrations can cause ecological dead zones due to depletion of dissolved oxygen and produce toxins that are harmful to aquatic species and humans when consumed (Rabalais et al., 2002; Chislock et al., 2013).

Algal toxins such as microcystins are organic molecules produced during the

growth of cyanobacteria, negatively affecting the local economy and causing water impairment issues. In addition to toxins, HABs cause taste and odor problems in drinking water systems due to release of Geosmin (Antoniou et al., 2005). It has been suggested that HABs are the greatest inland water quality threat to ecosystem well being and human health (Brooks et al., 2016a). Land management practices affect the sediment, nutrient, and heat fluxes to surface water bodies causing HABs. Since there are several interrelated factors responsible in the formation of HABs, water quality models are required to understand HAB precursors and growth. Therefore monitoring HABs in the surface water is important to ensure drinking water quality, protect public health, and provide safe ecosystem provision services.

## 1.2 Current State of HAB Monitoring

Several alternatives are available in monitoring HABs and surface water quality of inland water bodies and reservoirs. Some of these alternatives include physical grab sampling, satellite remote sensing, in-situ sensing using sensors. Grab samples for HABs should be taken carefully since they are toxic. However, processing grab samples is often slow and the sampling process does not cover the entire spatial extent of the water body. The traditional sample grab and lab analysis are useful in analyzing the water quality at the expense of time (Amejecor, 2021). Therefore studying the dynamic nature of the cyanobacteria creates the need to use remote sensing methods using satellites and unmanned systems in measuring the algae concentrations. Satellite remote sensing has proven to be an effective way to remotely sense and measure the algal pigments, including Chlorophyll-a and phycocyanin in oceans (Stumpf, 2001; Stumpf et al., 2003; Shen et al., 2012; Tang et al., 2004; Ahn and Shanmugam, 2006). Existing satellite technologies of Landsat-8 and Sentinel-2 provide image data of the earth with a spatial resolution of 30 m and (10 to 60 m) respectively. The temporal resolution of Landsat-8 and Sentinel-2 is sixteen days and ten days, respectively. In

the case of Sentinel-2, there are two satellites with revisit frequency of single satellite in ten days, and the combined constellation revisit in five days. However HAB formations are very dynamic and the concentration of algae can vary spatially and temporally. The current HAB monitoring methods have limitations in collecting data in terms of spatial and temporal scale. Therefore monitoring HABs require systems that can sample water providing large amount of water quality information for further processing in building water quality models. These water quality models will provide important information regarding HABs, enabling the policymakers and other water stakeholders to make important decisions regarding land management practices and water quality management.

## 1.3 Approach

A major step in the development of using unmanned systems for water quality monitoring and modeling is evaluation on the field. Recently unmanned systems are becoming popular in scientific applications including environmental monitoring, space exploration, search and rescue. Also using unmanned systems in surface water quality monitoring is economical and efficient. Recently development of Unmanned Surface Vehicles (USVs) has gained momentum because they are more robust and work in extreme environments which are not accessible by humans. To improve accessibility in lakes and create robust monitoring system, an Unmanned Surface Vehicle (USV) is tested in Grand Lake, Oklahoma and Marion Reservoir, Kansas for rapid water quality data collection. In addition to using unmanned systems, a remote sensing tool has been developed which can be used to develop in-lake water quality models with the in-situ data obtained from unmanned systems. As far as the remote sensing tool is concerned, empirical relationships are developed between pixel values of multi-spectral bands and in-situ water quality data. The current state of the open source remote sensing tool developed compares the utility of image data obtained

by multi-spectral imagery of Landsat-8 and Sentinel-2. The capability of this open-source remote sensing tool can be extended to the aerial image data obtained from Unmanned Aerial Vehicles (UAVs) to forecast HABs. In environmental monitoring applications, UAVs are extensively used in remote sensing applications. The remote sensing methods require high resolution spatial and temporal data sets to produce better statistical models predicting water quality parameters. The high-resolution image data sets attains better gridded estimates and forecasting scenarios to evaluate the currently developed remote sensing tool. Therefore it is very important to collect more in-situ data to provide ground truth to enhance and test the efficacy of the remote sensing methods. In addition to unmanned vehicles, unmanned systems are used for stream monitoring. The data collected from monitoring streams can be used to build watershed models to understand nutrient and sediment transport within the watershed. In this dissertation, the rich temporal and spatial data collected from unmanned systems are used in building statistical and watershed models. In Oklahoma State University, USVs developed for water quality monitoring have potential to perform fully autonomous missions in lakes and reservoirs. Monitoring of Grand Lake O' the Cherokees and Marion reservoir using USVs and other unmanned systems provides important insights in exploring rapid data collection for model building and freshwater monitoring. Both statistical water quality and watershed models will give important insights into nutrient and sediment influxes into the surface water bodies and how they influence the formation of HABs.

## 1.4 Dissertation Structure

The models built from in-situ data of reservoirs and rapid data collection of water quality using unmanned systems helps to assess the water quality conditions and causes of impairment. This dissertation is divided into seven chapters:

1. An Introduction to the environmental problems associated with HABs and why

is it important to monitor them. Also the introduction provides the importance of using unmanned systems in developing water quality models to understand the precursors to HAB formations.

2. A literature review of previous studies on remote sensing, water quality monitoring using unmanned systems, and water quality modeling.

3. the development of an open-source satellite remote sensing tool which can be integrated into providing HAB advisories.

4. the development of HSPF model to simulate a watershed of Grand Lake O' the Cherokees to understand the influx of sediment and nutrients into the watershed.

5. A demonstration of an USV to collect rapid water quality data for water quality monitoring.

6. A description of results from building a life cycle assessment of switchgrass to jet fuel using a novel co-fermentation process developed in Oklahoma State University, Stillwater, USA.

7. A summary of findings, conclusions and recommendations for future research.

# CHAPTER II

## Literature Review

HABs pose a risk to mammals and the environment because they are toxic and can negatively impact the aquatic ecosystems by depleting dissolved oxygen. In this chapter,a background on remote sensing and machine learning is presented followed by background on unmanned systems used for water quality monitoring, followed by background on watershed modeling. The remote sensing tool, water quality monitoring, and water quality modeling provides important insights to understand HAB formations. The application of unmanned systems in water quality monitoring and water quality modeling provides readers information associated with importance of rapid data that is necessary to develop models for understanding the formation of HABs.

### 2.1 Literature on Remote Sensing

The application of remote sensing is based on the interaction (absorption and reflection) of the earth's land, natural resources, and objects with solar radiation. It is observed each earthly object has its own spectral signature (Bell, 1995). In remote sensing applications each earthly object that can be detected by finding a relation between the prediction variable and the multi-spectral input features based on unique spectral signatures. Therefore electromagnetic spectrum and understating the spectral characteristics of remotely sensed object is important in spatio-temporal studies.

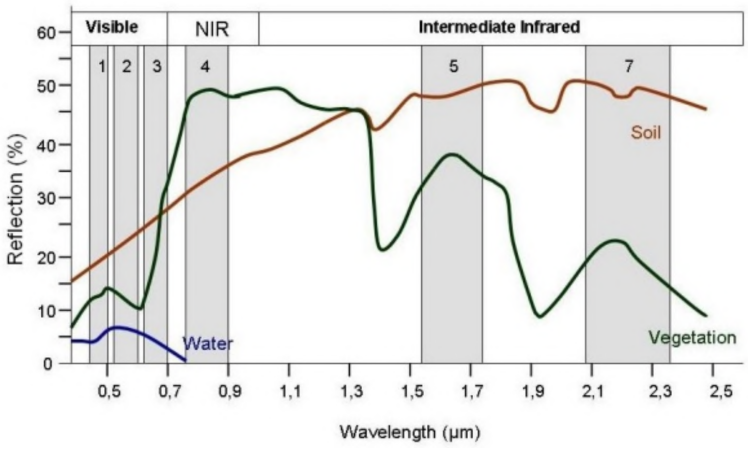


Figure 2.1: Spectral reflectance of soil, vegetation, and water

### 2.1.1 Electromagnetic Spectrum

The electromagnetic spectrum ranges from shorter to larger wavelengths i.e. from gamma rays to radio waves. In remote sensing applications several regions of electromagnetic spectrum is useful for earth observation. Especially ultra-violet region (1 nm to 400 nm) of the electromagnetic spectrum is useful in remote sensing of earth surface materials such as rocks, and minerals (Aggarwal, 2004). The light which our eyes detect is in the visible spectrum and there is a lot of radiation in and outside the visible spectrum that is absorbed or reflected around different objects of the earth's surface which can be detected using remote sensing instruments. The visible spectrum ranges from 400 nm to 700 nm. This portion of the electromagnetic spectrum can be associated with the detection of several colors by human eyes. Next spectrum of interest after visible spectrum is the infrared region where the radiation (IR) can be divided into two categories including reflected IR and emitted/thermal IR region respectively. The reflected IR and thermal IR wavelengths are in the range of 700 nm to 3,000 nm and 3,000 nm to 100,000 nm respectively. Finally the micro wave wavelengths ranging from 1 nm to 1 m is a region of interest in the remote sensing applications. Some of the important spectral signatures of the earth is as shown in Figure 2.1.

Algae are aquatic species that have the capability to perform photosynthesis. These photosynthetic algae has light absorbing molecules called pigments that can absorb the electromagnetic radiation. There is a certain range of electromagnetic spectrum the pigment can absorb which is called as the absorption spectrum and the rest of the radiation is reflected. In case of algae, chlorophyll-a is a pigment that absorbs red, green wavelengths and reflects the green and near infrared radiation (NIR) wavelengths (Datt, 1998; Murphy et al., 2005). Therefore the visible and NIR spectrum's can be used to remotely sense the chlorophyll-a pigment.

### 2.1.2 Remote Sensing of Water Quality/Clarity

Water quality in surface water bodies is dependent on natural and anthropogenic activities in a watershed, hydrodynamics and climate change. Accessing and frequency of monitoring water bodies frequently is challenging sometimes and this negatively impacts monitoring water quality. Satellite remote sensing is a solution to this challenge which has been identified as a potential alternative to traditional water quality monitoring methods. There was a case study of Ozark/Ouchita-Appalachian ecoregion in eastern Oklahoma where satellite remote sensing sensing has been performed and developed predictive statistical models to predict chlorophyll-a and turbidity (Barrett and Frazier, 2016). However the resultant tool seems to perform well during winter months of a year rather than in summer. In addition, water quality models in inland water bodies can be developed using satellite remote sensing. These articles document the use of satellite imagery in efficiently monitoring water quality and limitations that researchers need to overcome (Ritchie et al., 2003; Flores-Anderson et al., 2020). However Landsat-8 bands are efficient in detecting chlorophyll-a, turbidity concentrations in lakes and reservoirs (Barrett and Frazier, 2016; Ma et al., 2021). Similarly multi-spectral instruments on Sentinel-2 provide reflectance bands which help in detecting phycocyanin pigment of HABs (Sòria-Perpinyà et al., 2020).

### 2.1.3 Challenges in Remote Sensing

Advantages of remote sensing have improved the understanding of HABs and help in identifying the HAB dynamics in surface water bodies. However certain limitations and challenges in remote sensing of HABs include unsystematic understanding of HABs, insufficient utilization of remote sensing. HABs are usually complex dynamic mechanisms that involves nutrient inputs, favorable climatic conditions, and hydrodynamics. Previously Chen et al. (2003) reported that HAB formations are based on physical, chemical, and biological processes with the associated contributors. Moreover, linking both influx of nutrients and HAB formation is difficult unless significant in-situ data is available (Paerl et al., 2016). Therefore comprehensive understanding of relationship between nutrient inputs and HAB formations is important. However, there have been few studies which focused on comprehensive understanding of HAB formations in the area of remote sensing. In this thesis, a comprehensive understanding of HAB formations will be explored based on the amount of sediment and nutrient transport into the freshwater bodies. In addition to understanding the complex HAB formation dynamics, most satellites have medium spatial and temporal resolution images when compared to imagery collected by Unmanned Aerial Vehicles (UAVs) equipped with multi-spectral cameras. Also UAVs are flexible for remote sensing applications because the users can choose their own flight path, revisit times, which are critical in monitoring algal blooms which are very dynamically formed.

### 2.1.4 Image Errors in Remote Sensing

Water quality monitoring using remote sensing should include methods that incorporate methodological and analytical protocols to reduce image error significantly. Sources of error in remote sensing includes geometric and radiometric distortions from earth's system and remote sensing sensors. Errors with sensors are systematic and correction protocols help in improving quality of the image. In case of Landsat,

the sensors have either whisk broom or push broom scanners. In case of Sentinel satellites, the sensors have push broom scanners. The whisk broom scanner used in the earth observing satellite scans the image across the satellite path which collects data one pixel at a time. As the scanner moves to capture the pixel information the back and forth motion at different angles of incidence and reflections can cause image distortions which needs to be corrected. Before the launch of Landsat-8, all Landsat satellites were launched with sensors equipped with only whisk broom scanners. While whisk broom scanners move constantly to capture pixel information, push broom scanners move few times comparatively and collect pixel information perpendicular to the flight path. Therefore push broom sensors have fewer image corrections as far as edges are concerned. The Earth Resources and Observation Science (EROS) located in South Dakota performs correction of all the images before making them available for public access.

The sources of error while getting image data by satellite scanners is either shot noise or line stripping. Shot noise indicates bad pixel and line stripping means dropping lines while scanning an image object. Bad pixels occur when scanner skips certain points in the image and records wrong pixel readings. Traditionally interpolation methods are used to correct bad pixel readings. However if the land cover is heterogeneous and dense then using interpolation is not a proper approach. Moreover change in seasons can also impact the spectral characteristics of land cover providing incorrect information while image correction. Apart from satellite sensors, the position of the sun can also result in error in images. To avoid these errors, selecting spectral bands that are not majorly affected by solar radiation is preferable in building the remote sensing algorithms or tools. In addition to solar radiation and satellite sensors, distortion in synchronization between satellite and earth orbits can result in image errors. Earth is not a perfect sphere which cannot make the angle of incidence $90^o$ at the nadir resulting in spatial error. Moreover, circular error of probability is

another possibility to be aware and corrected if the error is too large. The circular error can be estimated as the average distance between the missed and intended target pixels of the remote sensing object (Claverie et al., 2018). Correction of these displaced pixels is performed using indirect Geo-referencing such as using Ground Control Points (GCPs) and other mathematical models. Again the EROS center in South Dakota and European Space Agency (ESA) perform all these corrections for Landsat and Sentinel before making the image data accessible to public. Moreover heterogeneous land cover in a small region can also result in pixel errors. Furthermore, low spatial resolution of satellite imagery makes it difficult to remotely sense objects or concentrations based on a spectral signature. To address these issues of assigning a spectral signature, researchers use methods for testing to find the correct pixel value (Nagol et al., 2018; Vaesen et al., 2001; Witharana and Civco, 2014).

### 2.1.5 Remote Sensing Algorithms and Tool Development

The main objective of developing remote sensing tools in water quality monitoring involves developing a relationship between target variable and independent variables using statistical modeling. In statistical model development, selection of models that perform better and serve the purpose of monitoring is of critical importance. As we know there are several statistical machine learning models including supervised, unsupervised, and object based methods. Out of all the statistical methods, regression which is a sub-field of supervised machine learning is useful in building remote sensing algorithms and tools. Regression methods are useful in estimating a target variable by using several independent variables. However, dropping certain independent variables needs to be clearly investigated if there is any multi-collinearity between them (Daoud, 2017). Since the environmental processes like formation of HABs are complex and dynamic, it is very important to check the sensitive aspects of statistical modeling including data distribution, multi-collinearity, and chances of over-fitting.

Previous literature including (González-Márquez et al., 2018; Xu et al., 2021) reported that statistical models are useful in remote sensing application of measuring the concentrations of surface water quality parameters. There are three types of approaches in remote sensing that can be used including: Empirical approach, Semi-empirical approach, and analytical approach.

1. **Empirical Approach**: Typical empirical methods include linear and multiple regression, single band method, multiple band-combination methods, artificial neural networks, principal component analysis, and other machine learning methods. The data used in the empirical methods include in-situ data (dependent or target variable), and satellite data (independent variables). The approach is based on creating empirical relationship between input and output variables.

2. **Semi-empirical Approach**: Semi-empirical methods require both measured spectral analysis and statistical analysis. Most of these methods have inherent assumptions such that they help in providing information to improve the accuracy of the model.

3. **Analytical Approach**: In this approach, parameters (independent variables) are used to find the inherent optical properties. These optical properties can help in quantifying the pixel information to measure the concentrations of different water quality parameters.

### 2.1.6 Regression Models

Regression models are most widely used empirical methods to associate algal biomass with spectral signatures of the satellite sensors or UAV imagery. Regression models attempt to minimize the deviation between measured concentrations and predicted values of the water quality. The difference between the predicted and observed values

is called error. Minimizing the error and improving the prediction performance is the main objective of the statistical regression models. Remote sensing methods utilizes the coefficient of determination, mean square error, and Nash-Sutcliffe efficiency to evaluate the performance of the regression model. with values ranging from 0 to 1 such that the model close to 1 has less error with good fit.

1. **Multiple regression**: Multiple regression is a statistical machine learning technique that can be used to analyze the relationship between several independent variables and one dependent variable. The main objective of the multiple regression is to use the values of independent variables to predict the value of dependent variable. The weights denote the relative contribution to the whole prediction such that the multiple regression is mathematically represented as follows:

$$Y = a + b_1 X_1 + b_2 X_2 + ..... + b_n X_n \tag{2.1}$$

   The variables in the Equation 2.1 here includes Y as dependent variable and $X_1$, $X_2$,.....,and $X_n$ are independent variables. Although multiple regression is a preferable method of remote sensing in measuring the concentration of water quality parameters, the error is significantly large. Therefore there is a need to explore other machine learning algorithms to understand the non-linear dynamics of Algal blooms.

2. **Support vector regression**: Algal blooms are dynamic and therefore it is very important to explore machine learning algorithms that can handle non-linear data. Although Support Vector Machines (SVMs) are developed to solve classification problems, the domain has been extended to solve regression problems. With the training data $(x_i, y_i)_{(i=1)}^{N}$ such that $x_i \in R^n$, and $y_i \in R$, the SVM maps x into a higher dimensional space using non-linear mapping. If $z = \varphi(x)$

such that a non linear model $y = f(x, w)$ can be constructed in this feature space. The cost function $y = wz + b$ can be written as follows as shown in Equation 2.2:

$$f(x, w) =< w.\varphi(x) > +b \tag{2.2}$$

where w is the weight which decides how much influence the input will have on the output parameter. $\varphi(x)$ is the vector data in the feature space used for prediction or training. b is the bias which is a constant. $< w.\varphi(x) >$ denotes the dot product between the weight and the feature vector. The main objective of the regression equation is to minimize the loss function using the training data. However, one of the main criteria of the SVM is that it attempts to minimize the generalization error instead of the training error. This behavior of the support vectors prevents the phenomenon of the over-fitting and results in high generalization performance. This generalization error is the combination of training error and regularization term which controls the complexity of the statistical model. The $\varepsilon$-insensitive loss function given by Vapnik et al. (1995) as $L = |y - f(x, w)|_\varepsilon$ such that:

$$|y - f(x, w)|_\varepsilon = \begin{cases} 0, \text{ if } |y - f(x, w)| \leq \varepsilon \\ |y - f(x, w)| - \varepsilon, \quad \text{otherwise} \end{cases} \tag{2.3}$$

In the Equation 2.3, $\varepsilon$ is the tolerance to the error such that if the predicted value is less than the tolerance then the loss is zero. If the predicted value is greater than the tolerance then the loss is the magnitude of the difference between the predicted value and the radius of the $\varepsilon$. Therefore Support Vector Regression (SVR) is robust to small errors in the training dataset to the least square errors used for Multi-Layer Perceptrons (MLPs). The $\varepsilon$ insensitive loss

function as described by the Vapnik et al. (1995) introduces non-negative slack variables such as $\xi_i, \xi_i^* i$ = 1,....,N, which measures the deviation of training data outside the tolerance zone. The SVR is ultimately posed as a convex optimization problem as shown below:

$$\begin{aligned} &\min \frac{1}{2}||w||^2 + C\sum_{i=1}^{N}(\xi_i, \xi_i^* i) \\ &\text{subject to} \begin{cases} y_i - < w.\varphi(x) > -b \leq \varepsilon + \xi_i^* i \\ < w.\varphi(x) > -b - y_i \leq \varepsilon + \xi_i^* i \\ \xi_i, \xi_i^* i \geq 0, i = 1, ..., n \end{cases} \end{aligned} \tag{2.4}$$

In Equation 2.4, C is a fixed regularization constant which determines the trade-off between training error and model complexity. The convex optimization problem can be solved using Lagrange multipliers, and the regression function is as shown in Equation 2.5:

$$f(x) = \sum_{i=1}^{N}(\alpha_i - \alpha_i^*)K(x_i, x) + b \tag{2.5}$$

where $\alpha_i - \alpha_i^*$ are the Lagrange multipliers which are obtained by solving the Quadratic Programming (QP) problem. $K(x_i, x) =< \varphi(x_i), \varphi(x) >$ is a kernel function that can satisfy the Mercer's condition. The QP problem can be solved as illustrated in Equation 2.6:

$$\begin{aligned} &\min_{\alpha,\alpha^*} -\frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}(\alpha_i^* - \alpha_i)K(x_i, x_j)(\alpha_j^* - \alpha_j) + \varepsilon\sum_{i=1}^{N}(\alpha_i + \alpha_i^*) + \sum_{i=1}^{N} y_i(\alpha_i - \alpha_i^*) \\ &\text{subject to} \sum_{i=1}^{N}(\alpha_i^* - \alpha_i) = 0, \alpha_i, \alpha_i^* \in [0, C] \end{aligned} \tag{2.6}$$

Moreover, the value of bias b can be obtained by the Karun-Kush-Tucker conditions as discussed in previous literature (Vapnik, 1999).

$$\begin{aligned} b &= y_i - < w, \varphi(x_i) > -\varepsilon, \text{for } \alpha_i \in (0, C) \\ b &= y_i - < w, \varphi(x_i) > +\varepsilon, \text{for } \alpha_i \in (0, C) \end{aligned} \tag{2.7}$$

We know that K is a positive definite and Equation 2.7 represents a convex problem that has a unique global minimum avoiding several local minima. Equation 2.5 the difference in Lagrange multipliers differs from zero and the support vectors. These data points typically are with in the tolerance of $\varepsilon$. If the data points are outside the $\varepsilon$-tube border or outside then they are not useful and do not contribute to the regression function. Therefore larger the tolerance fewer are the support vectors with sparser representation of the solution. However, larger tolerance can decrease the accuracy of the SVM. Therefore the tradeoff between closeness of data and sparseness is important and needed to be maintained. As far as the kernel functions are concerned for generalization there are generally three types including:

$$K(x_i, x) = \exp\left(-\frac{\|x - x_i\|^2}{2\sigma^2}\right) \tag{2.8}$$

$$K(x_i, x) = (x^T x_i + 1)^p \tag{2.9}$$

$$K(x_i, x) = tanh(\beta_0 x^T x_i + \beta_1) \tag{2.10}$$

The kernel functions shown in Equation 2.8 are the Radial Basis Function (RBF) where $\sigma$ is the width parameter. Equation 2.9 shows the polynomial regression of degree $p$ that will convert to linear function when p=1. Finally Equation 2.10 is the two layer sigmoid perceptron kernel function. Both the polynomial and RBF kernel functions satisfy mercer's condition. In case of two-layer perceptron

mercer's condition is satisfied only for some values of $\beta_0$ and $\beta_1$. Although the concept of SVMs is very similar to that of the multi-layer perceptrons, the number of hidden nodes in SVMs are based on number of SVs and their values. In case of MLPs the number of hidden layers are based on trial and error.

3. **Random forest**: Random forest is an ensemble machine learning algorithm which is a collection of tree predictors. These algorithms generate large number of decision trees which acts as regression functions on their own. The collection of predictors or trees are in the form of function $h(x;\theta_k)$ such that k = 1,2,3,.....,K. The x in the hypothesis function represents the observed input vector which is of the length p. The associated random vector X and $\theta_k$ are independent and identically distributed. In the regression scenario, the focus is on the numerical outcome as a scalar $S_n$ such that the observed in-situ or training data is assumed to be independently drawn from the joint distribution of (X,y). The training set contains n number of observations which can be expressed as shown in equation:

$$S_n = (X_1, Y_1), (X_2, Y_2), ...., (X_n, Y_n), X \in \mathbb{R}^n, Y \in \mathbb{R} \tag{2.11}$$

Random forests can be built by random sampling which is called as a bootstrap sample. The prominence of using bagging is to avoid the correlation between the decision trees which is a problem especially in multi-spectral satellite remote sensing. When it comes to random forests there is flexibility of using the training data with bagging where some data is used and some not at all. This flexibility makes the random forest to prevent overfitting the training dataset. Random forest can perform both classification and regression. In case of regression, the random forest is the unweighted average of the collection of decision trees such that:

$$h(x) = \frac{1}{K}\sum_{k=1}^{K} h(x;\theta_k) \tag{2.12}$$

As K increases and tends to infinity then the law of large numbers ensures such that:

$$E_{X,Y}(Y - h(X))^2 \rightarrow E_{X,Y}(Y - E_\theta h(X;\theta))^2 \tag{2.13}$$

The quantity on the left of Equation 2.13 is the prediction error such that it tends to be the generalization error for the random forest. The convergence in Equation 2.13 shows that random forests do not tend to over-fit the training data. The generalization/prediction error for random forest is designated as $PE_f$. The prediction error for individual tree is important and random forests have only two parameters to tune including number of trees and number of random features for each split. The average prediction error for an individual tree can be mathematically represented as shown in Equation 2.14:

$$PE_t = E_\theta E_{X,Y}(Y - h(X;\theta))^2 \tag{2.14}$$

For all values of $\theta$, if the tree is unbiased which means that EY $= E_X h(X;\theta)$ then the prediction error of random forest is as follows:

$$PE_f \leq \bar{\rho} PE_t \tag{2.15}$$

For independent values of $\theta$ and $\theta'$ the weighted correlation ($\bar{\rho}$) between residuals is $Y - h(X;\theta)$ and $Y - h(X;\theta')$. Equation 2.15 shows the requirement for accurate random forest regression. There is a need for low correlation between different individual tree members of the random forest to avoid over-fitting. Also, the prediction error is supposed to be low for the individual decision trees. Furthermore, the random forest will decrease the tree error by the ($\bar{\rho}$)

factor as shown in 2.15. Some of the strategies involved to achieve better performance of random forests include: growing the trees to maximum depth to avoid individual error, and keeping the residual correlation error low. The correlation of the residuals can be kept low by growing each tree on bootstrap samples and specifying the number of covariates and splitting the node based on the covariates.

In case of remote sensing, the diversity of input data is important to get better regression outputs without sacrificing model accuracy. Previously (Yajima and Derot, 2018) studied to forecast algal blooms using random forest models . Chlorophyll-a in-situ data with monthly time resolution was used to calibrate the random forest model. The result of the random forest model has shown that the chlorophyll-a concentrations could be predicted. However, Ahn et al. (2007) reported that Phycocyanin concentrations are important for developing HAB advisories. Therefore it is important to include both Chlorophyll-a and phycocyanin concentration to develop HAB alert systems.

## 2.2 Literature on Water Quality Monitoring

Periodic sampling of water is necessary to identify the changes in water quantity and quality over time. Monitoring water quality helps in developing land use and watershed management plans to ensure safe drinking water, control pollution, and protect human health. Important water quality indicators of the health of surface water includes temperature, Electrical Conductivity (EC), turbidity, pH, algal pigments including chlorophyll-a concentration, and phycocyanin concentration, Dissolved Oxygen (DO). Among these parameters pH influences the solubility of chemical constituents such as nutrients (nitrogen, phosphorous, and carbon) and heavy metals (lead,arsenic, and chromium) in surface water. Also, aquatic organisms require pH in the ranges between 6.5 and 9.0 for survival. Temperature governs the types of aquatic

life and influences the chemical and biological reactions in water. EC is a measure of water's ability to flow electrical current and is directly related into presence of ions in water. These ions indicate the presence of dissolved salts and inorganic materials including chlorides, sulfides, and carbonates. A sudden change in Electrical conductivity indicates that there is a sediment input into the water bodies. Turbidity is the measure of cloudiness of water and it indicates the presence of suspended solids in surface water bodies. Algal pigments including chlorophyll-a and phycocyanin indicate the presence of algae in water. DO is an indicator of non-compound oxygen in water and helps in assessing the water quality. Decrease in DO in surface water bodies results in decomposition of organic material by microbial species negatively affecting the aquatic life and water quality.

Based on previous literature, lots of polluted water and untreated municipal wastewater ends up in surface water bodies negatively affecting the health of aquatic ecosystems. In addition, excessive use of fertilizers in agriculture can result in leaching of nutrients causing eutrophic conditions in surface water bodies and negatively impacting surface water quality and human health. Therefore it is very important to perform water quality monitoring. Traditional water quality monitoring involving manned field sampling poses limitations in terms of spatial and temporal variability of water quality data. Monitoring water quality close to mining zones, water bodies surrounded by steep terrain are difficult to access for getting samples. Therefore there is a critical need to assess alternative methods including remote sensing and unmanned systems to monitor surface water bodies. Unmanned systems are becoming popular in the field of water quality monitoring due to their ability to collect high temporal and spatial resolution datasets. Unmanned Surface Vehicles (USVs) and UAVs have been used in monitoring water quality. USVs with on-board water quality sensors can collect high quality spatial and temporal data in in-accessible areas and operate in toxic aquatic environments.

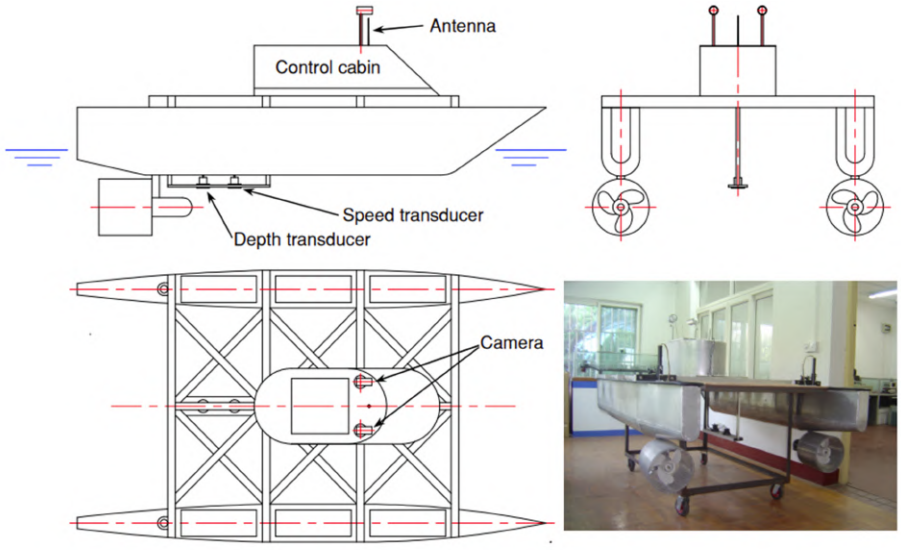


Figure 2.2: Design of ASV by Wang et al. (2009a)

Previously Wang et al. (2009a) designed an Autonomous Surface Vehicle (ASV) using the body of a high speed boat for monitoring oceans. The ASV was designed to use power efficiently and have a long endurance. The maximum wireless communication distance of the ASV designed by Wang et al. (2009a) was 10 nautical miles. The design of the USV developed by Wang et al. (2009a) is shown in Figure 2.2.

Recently ASVs have the capability to sample ocean water for long duration (days to weeks) by solar power integration (Higinbotham et al., 2008). The ASV developed

Figure 2.3: OASIS (ASV3) Oceanographic research platform Higinbotham et al. (2008)

by (Higinbotham et al., 2008) provides a low-cost, reusable, long duration ocean

observing unmanned system for understanding carbon-dioxide air-sea flux and phytoplankton productivity. Apart from water quality monitoring an ASV was developed by (Dunbabin and Grinham, 2010; Dunbabin et al., 2009) for greenhouse gas emission monitoring. The ASV is powered with solar panels which has unique capabilities to profile water vertically during motion. The ASV was tested on Lake Wivenhoe located west of Brisbane, Queensland, Australia is shown in Figure 2.4.

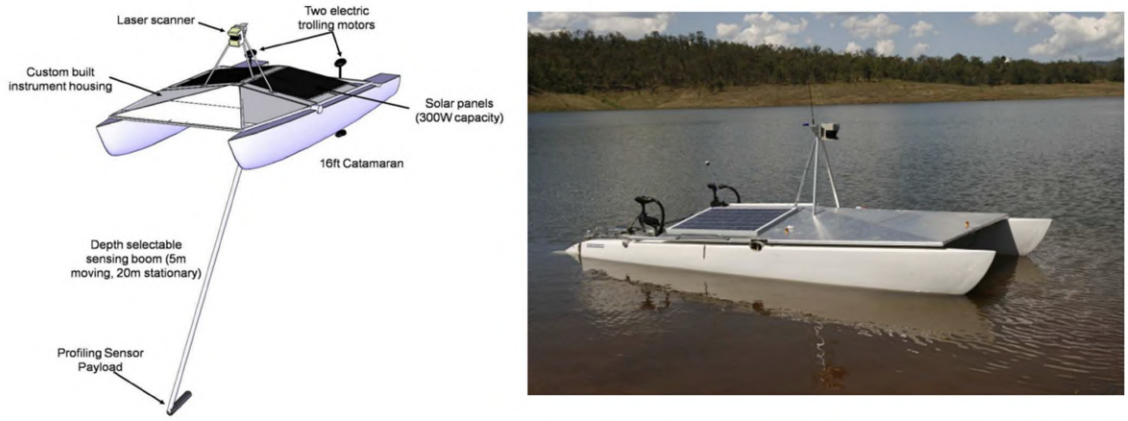


Figure 2.4: ASV system design by Dunbabin et al. (2009)

In the field of aerosol modeling a USV was developed by Powers et al. (2018) to understand precipitation and other bio-geological processes. The sensors on board the USV are shown in Figure 2.5.

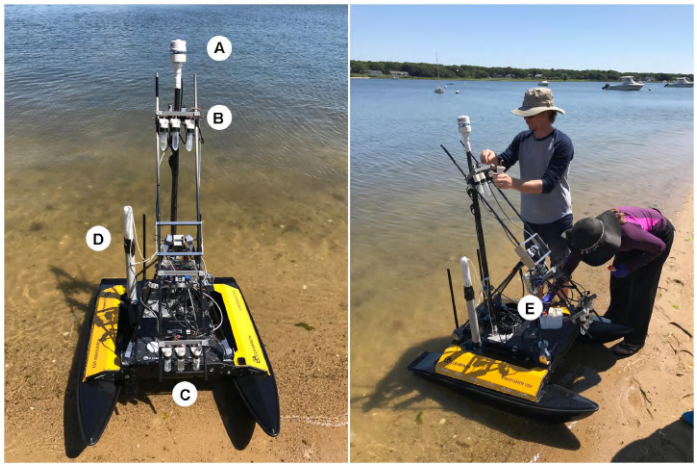


Figure 2.5: USV system design by Powers et al. (2018) integrated with A) Airmax 200 WX sensor to capture meteorological data, B) Impingers and particle counter at 1.1 m and C) 0.1 m , D) Turbidity sensor, and E) extender for vertical profiling

The motivation to develop the USV was to understand the distribution of biological aerosols in aquatic environments and how they influence in formation of clouds and precipitation processes. The geographic locations where water meets land such as wetlands are essential for aquatic life. To sample and analyze shallow water bodies an ASV was developed and tested by Odetti et al. (2020) researchers from Europe. The ASV shown in Figure 2.6 is completely autonomous with on-board WiFi for communication.

Figure 2.6: USV system design by Odetti et al. (2020)

Also (Banerjee et al., 2020; Koparan et al., 2018) developed a UAV system shown in Figure 2.6 that can collect physical water samples from contaminated aquatic environment. The water samples collected by the UAV system reduces the risk of humans coming in contact with toxic water while sampling. Figure 2.6 shows the UAV equipped with water sampling design.

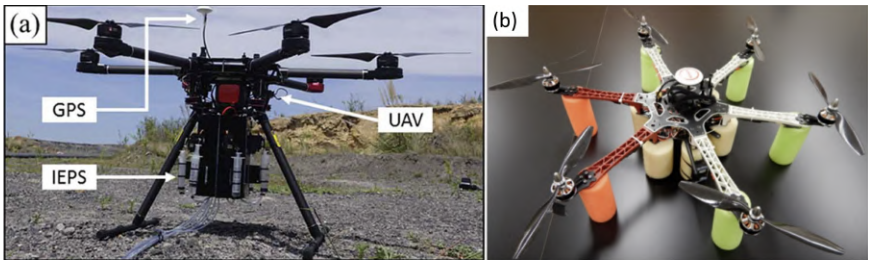


Figure 2.7: UAV system design by (a) Banerjee et al. (2020) and (b) Koparan et al. (2018) for water sampling

In addition to UAVs, a robotic boat was designed to work as a USV with an

unmanned water sampler with 12 bottles and storage capacity of 1 liter each bottle to collect water samples as shown in Figure 2.8.

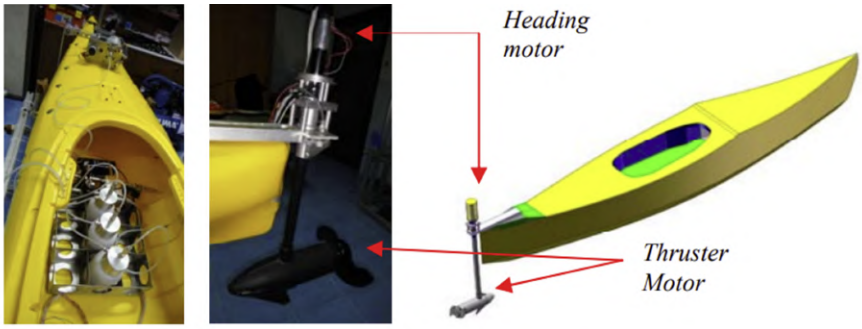


Figure 2.8: ASV system design by Prempraneerach and Kulvanit (2010) for water sampling

Unmanned systems are explored in water sampling of aquatic environments by researchers and academic scholars. The Unmanned Systems Research Institute (USRI) at Oklahoma State University (OSU) developed two autonomous USVs for monitoring lakes in Oklahoma. The two USVs include Aquarius-1 and Mobile Autonomously Navigable USV for Evaluation of Lakes (MANUEL). Aquarius-1 shown in Figure 2.9 is the first prototype used by the research team at OSU to monitor lakes in Oklahoma and Kansas.

Figure 2.9: USV prototype system design by USRI

Aquarius-1 is capable in performing autonomous missions and has on-board sensors that can collect water quality parameters such as temperature and depth. How-

ever there are certain challenges that are associated with range, mobility, and maneuverability using Aquarius-1 which are successfully dealt with by building MANUEL. MANUEL shown in Figure 2.10 is built by using an existing kayak designed to carry a maximum weight of 130 pounds. The kayak is 6 feet long and 2 feet wide with differential thrust for better steering and maneuvering.

Figure 2.10: MANUEL sampling water in Grand Lake

The avionics on-board MANUEL shown in Figure 2.11 include Orange Cube Pixhawk, GPS, Electronic Speed Controller (ESC), Receiver, Transmitter, and Telemetry. The navigation and path planning of MANUEL is similar to that of an Unmanned Aerial Vehicle (UAV) using Mission Planner software. Mission Planner provides the GPS way points for the MANUEL to operate in AUTO and smart RTL modes. All the specification of MANUEL is provided in Table 2.1.

The sensors on-board MANUEL include a ping sonar to measure depth and YSI EXO-3 sonde. The ping sonar is a single beam echo-sounder that measures depth up to 100 feet and the theory of operation is shown in Figure 2.12. The sonar produces a beam width of 30 degrees to measure depth and operates in water temperature

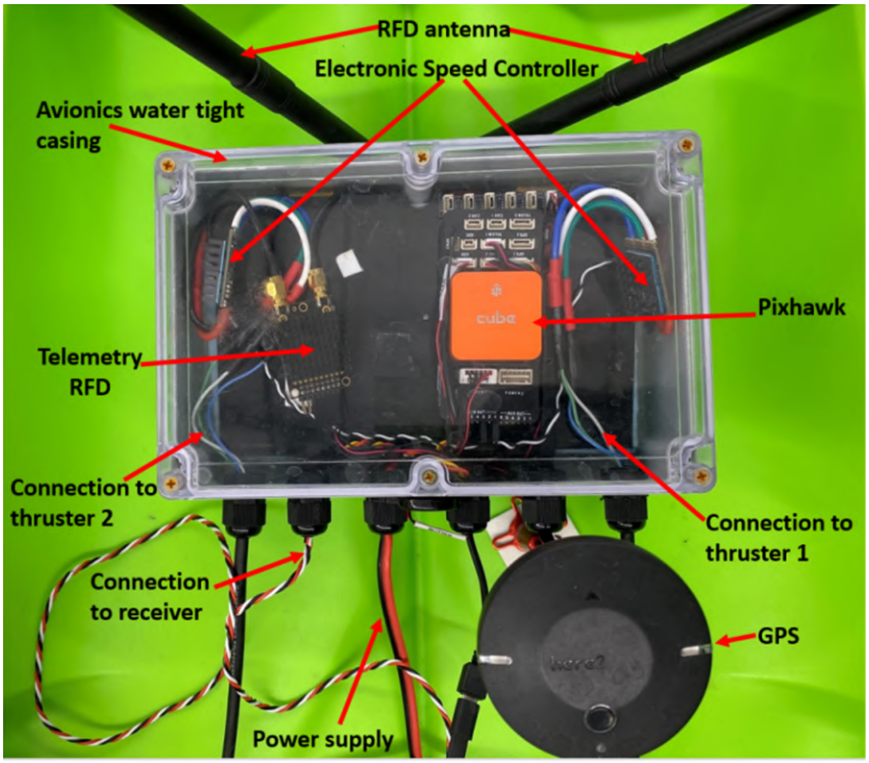


Figure 2.11: Avionics on-board MANUEL

Table 2.1: Specifications of MANUEL

| | |
|---|---|
| Vehicle structure | Single Hull kayak, 6 feet by 2 feet |
| Steering type | Thrust differential |
| Propulsion System | Dual T200 Thrusters |
| Powered | 4 6S LiPo 14,000 mAh |
| Endurance | 3.5 hours at cruising speed |
| Cruise speed | 4.5 ft/s |
| Maximum Speed | 9 ft/s |
| Maximum weight | 130 lbs |
| Maximum thrust | 24 lbs |
| Autopilot | Orange Cube Pixhawk 2.0 |
| Modes of operation | Auto, Manual and SmartRTL |
| Maximum range of operation | <1 mile |

ranging from 0 to 30 ($^oC$). The ping sonar is connected to a EnviroDIY mayfly micro-controller which is programmed as an Arduino board to log depth data.

The EXO-3 sonde shown in Figure 2.13 is a bundled sensors with water quality probes. It can measure parameters including pH, turbidity, temperature, chlorophyll-a concentration, phycocyanin concentration, conductivity, and dissolved oxygen. The

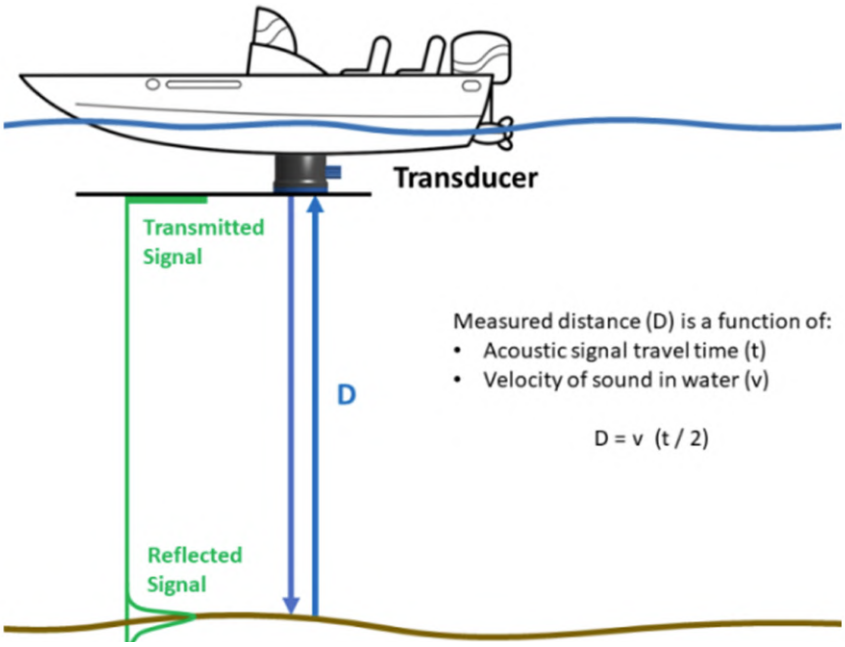


Figure 2.12: Ping sonar theory of operation

EXO-3 sonde is capable of Bluetooth connection and SDI-12 communication with titanium sensors. All the sensors on-board MANUEL are calibrated before performing an auto-mission to monitor lakes.

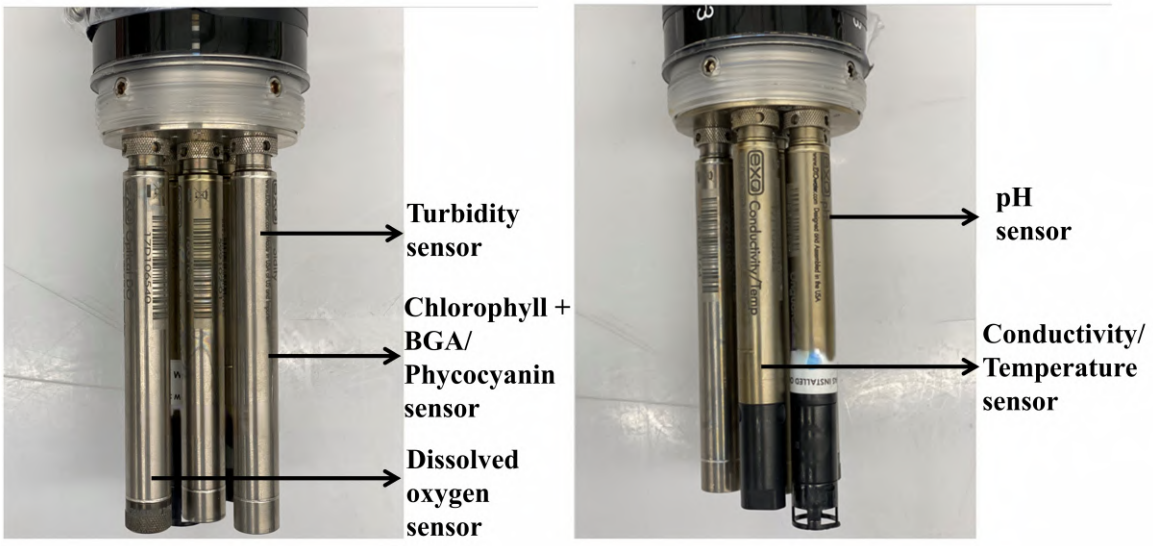


Figure 2.13: Bundled sensors of EXO-3 sonde

Previously the utilization of USVs and UAVs have been explored to understand the climatic conditions, bio-aerosol, and Green house gas emissions. However, integrating

water quality sensors on a USV can provide rapid data collection of surface water quality data possible. This high resolution datasets can provide ground truth while integrated with UAV imagery using remote sensing tools can help in developing in-lake water quality models. These models helps to assess the strategies for in-situ HAB treatment approaches by removing precursors or the algal biomass.

## 2.3 Literature on Watershed Modeling

Controlling water pollution in a watershed is important because nutrients like nitrogen and phosphorous are precursors in formation of HABs in lakes and reservoirs. Understanding pollutant sources from point and non-point sources is important to understand how the land management practices are impacting watershed health. Watershed modeling has been increasingly used to assist in managing the pollution from non-point sources (Srinivas et al., 2020; Benaman et al., 2001; Moges et al., 2018).Understanding the importance of non-point sources that contribute nutrients into surface water bodies, U.S EPA developed the Hydrologic Simulation Program in FORTRAN (HSPF) (Bicknell et al., 1997).

### 2.3.1 HSPF

Development of a HSPF model requires data related to land-use, catchment areas, stream networks, climate time series, and flow time series for calibration purposes. Using HSPF is flexible and helps in applying the model to different climatic conditions around the world. HSPF has modules that can simulate hydrology, temperature, sediment transport, nutrient transport, and snow. Due to its universal applicability and reliability, the HSPF model has remained relevant even today to simulate watersheds and understand sediment, and nutrient transport (Singh and Frevert, 2010).

HSPF has been maintained by Environmental Protection Agency (EPA) and United States Geological Survey (USGS) jointly with the Better Assessment Science

Integrating Point and Non-Point Sources (BASINS) version 4.5. BASINS is useful in providing integrated Geographic information system (GIS) data for data extraction, pre-processing, calibration, and visualization. However there are certain limitation of using BASINS including software crashes and incompatibility with Windows Visual.Net and open source Unix operating systems. Also, the GUI based tool reduces the flexibility of using some modules of HSPF. Therefore alternative methods are needed for pre-processing, simulating, calibrating, and post-processing HSPF. PyHSPF using Python provides the flexibility of using HSPF modules and is supported by both windows and UNIX operating systems.

The Python programming contains useful tools for data analysis and is relevant for hydrology and water quality modeling. The PyHSPF uses both the hydrology and water quality capabilities of HSPF which integrates with the data analysis tools in Python creating a flexible framework in performing modeling tasks to develop watershed models. A HSPF simulation requires minimum of two input files including a User Control Input (UCI) file and Watershed Data Management (WDM) file. The UCI file is a plain text file with instructions for a simulation. The WDM file is an unformatted file with input timeseries data. The HPSF version 12.2 can handle multiple WDM files with large amounts of timeseries input data.

### 2.3.2 PyHSPF Software

Building a HSPF model requires input hydrography data on catchments and stream reaches, land use category, climate data, and hydrology process parameters to model land segments. All these data can be translated as input data to HSPF by creating UCI and WDM input files. A series of Python classes in PyHSPF can be used to build input and output files, run simulations, calibrate hydrology process parameters, and postprocess simulation results. PyHSPF uses several third party software packages to gather data from world wide web for data processing and postprocess HSPF

simulations. In this dissertation, a new class has been added to the extract climate data from Oklahoma Mesonet which are used to simulate Horse Creek Watershed of Grand Lake.

The simulated Horse Creek Watershed provides important information to understand the behavior of non-point sources of nutrients and sediment into the Grand Lake. The information can improve the decisions related to best management practices (BMPs) of land in the Horse Creek watershed. These BMPs reduce the non-point source runoff and provides important insights about the nutrient precursors that initiates HAB formation in the Grand Lake. The remote sensing tool integrated with the watershed model using HSPF provide forecasting models for HABs and provide early warnings of swim advisories due to HABs in freshwater bodies.

# CHAPTER III

## A Remote Sensing Tool for Real-Time Monitoring of Harmful Algal Blooms and Turbidity in Reservoirs

### 3.1 Abstract

Harmful Algal Blooms (HABs) diminish the utility of reservoirs for drinking water supply, irrigation, recreation, and ecosystem service provision. HABs decrease water quality and are a significant health concern in surface water bodies. Real-time monitoring of HABs in reservoirs and small water bodies is essential to understand the dynamics of sedimentation and HAB formation. This study uses satellite imagery to remotely sense Chlorophyll-a concentrations (CHLa), Phycocyanin concentrations (BGA), and turbidity in two reservoirs, the Grand Lake O' the Cherokees and Hudson Reservoir, OK, USA, to develop a tool for real-time monitoring of HABs. Landsat 8 and Sentinel-2 imagery from 2013 to 2017 and from 2015 to 2020 were used to train and test three different machine learning models that include multiple regression, Support Vector (SVR) Regression, and Random Forest Regression (RFR). Performance was assessed by comparing the three machine learning models to estimate CHLa, BGA, and turbidity. The results showed that RFR achieved the best performance, with $R^2$ values of 0.75, 0.82, and 0.79 for CHLa, turbidity, and BGA, while multiple regression had $R^2$ values of 0.29, 0.51, and 0.46, and SVR had $R^2$ values of 0.58, 0.62, and 0.61 on the testing datasets, respectively. The current study enhances the scope to utilize publicly available satellite imagery to estimate the real-time water quality of reservoirs in Oklahoma using the developed open-source remote sensing tool.

Keywords: Harmful algal blooms; remote sensing; machine learning; water quality monitoring; lakes

## 3.2 Introduction

Nutrient runoff from agriculture, animal feeding operations, and wastewater treatment plants promote the growth of algae and cyanobacteria. Extreme rainfall events can transport large amounts of sediment into surface water bodies that bring nutrients and other pollutants into reservoirs, which adversely affects the ecosystem health of surface water bodies and causes Harmful Algal Blooms (HABs). Sedimentation, sediment resuspension, and eutrophication diminish the utility of reservoirs and negatively impact human health (Zhang et al., 2020).Increasing agricultural activity and excessive use of fertilizers exacerbates HAB events, resulting in water impairment issues. HABs cause taste and odor problems in drinking water due to the release of Geosmin (Deng et al., 2017; Wurtsbaugh et al., 2019; Antoniou et al., 2005). The decay of HABs creates ecological dead zones due to the depletion of dissolved oxygen in surface water creating unfavorable conditions for the sustenance of flora and fauna in water bodies (Bennett, 2017; Chislock et al., 2013; Zohdi and Abbaspour, 2019). Besides, HABs produce toxins in water which pose health risks to mammals when consumed (Anderson, 2009; Park et al., 2017). High toxic risk upon exposure and increasing occurrences of HAB events creates a critical need for regular water quality monitoring. Successful monitoring of water bodies requires observations of spatial and temporal trends to understand sedimentation and HAB formation.

Existing satellite technologies provide spatial and temporal observations that can be used to study water quality in inland water bodies. For instance, Landsat-8 and Landsat-9 orbits the earth in sun-synchronous near-polar orbit with a 16-day repeat cycle, where it collects multi-spectral imagery of the earth's surface (Carpenter and Carpenter, 1983; Blackwell, 1979; Masek et al., 2020). Landsat 8 has two instru-

ments, including the Operational Land Imager (OLI) and Thermal Infrared Sensor (TIRS). Landsat-8 imagery has been widely used to monitor water quality, including Chlorophyll-a (CHLa), turbidity, Mineral Suspended Solids (MSS), and surface temperature (Watanabe et al., 2015; Silva et al., 2016; Olmanson et al., 2016; Trinh et al., 2017; Harrington Jr et al., 1992; Fraser, 1998; Potes et al., 2012). Similarly, the Sentinel-2 mission uses two satellites, S2A and S2B, in sun-synchronous orbit which are phased at $180^{\text{o}}$ to each other to collect multispectral data of the earth's surface. Sentinel-2 satellites have thirteen spectral bands with variable spatial resolutions from 10 to 60 m.

Multispectral data from these platforms provide the reflectance of water surface at spectral regions beyond the visible range enhancing our ability to monitor water quality trend in space and time. For instance, algal pigments including CHLa and Phycocyanin (BGA) which are fluorescent to light reflect electromagnetic radiation that can be estimated using remote sensing. Previously, remote sensing methods have been used to estimate the concentrations of CHLa, turbidity, and BGA using satellite data in the oceans and reservoirs (Chen et al., 2013; Porcar-Castell et al., 2014; Babin et al., 1996; Crosta and MOORE, 1989; Brezonik et al., 2005; Chipman et al., 2004; Kutser et al., 2006). However, most of the previous work in remote sensing of water quality has been focused on oceans and large lakes. In addition to large water bodies, monitoring reservoirs and inland water bodies are important because they provide water to several households (Afroz et al., 2014). Besides acting as potential water source, inland water also bodies provide recreation and ecosystem provision services which helps in developing local economy (Ward et al., 1996; Cowx et al., 2010).

In the U.S., there are around 90,000 dams that have inland water bodies that can be monitored remotely using satellite sensors for water quality (Song et al., 2021). The Grand Lake O' the Cherokees (Grand Lake) and Hudson Reservoirs are located in northeastern Oklahoma, where they are monitored and operated by the Grand

River Dam Authority (GRDA). Recently, Grand Lake has experienced an increase in HAB outbreaks, including a shutdown on the fourth of July weekend in 2011, that significantly affected the local economy (Mason and Triplett, 2016). These increased occurrences of HAB events in Oklahoma and consistent monitoring from GRDA creates an opportunity for developing an operational near real-time HAB detection tool based on remote sensing.

The main objective of this study is to develop a remote sensing tool to monitor HABs for inland water bodies. To achieve this objective, empirical relationships were built using the surface reflectance values from both the Landsat-8 and Sentinel-2 satellites and in-situ data collected at Grand Lake and Hudson Reservoir. For active HAB monitoring, we developed an open-source software tool in Python that acquires satellite images from the web, extracts values of pixels within each reservoir, and estimates in-situ water quality data, including CHLa, turbidity, and BGA concentrations. The tool is then used to visualize a HAB in Oklahoma reservoirs and compare the utility of the Landsat-8 and Sentinel-2 satellites for remote sensing. This tool can help guide HAB advisories and provide observations for research into HAB dynamics.

## 3.3 Materials and Methods

### 3.3.1 Description of the Study Area

The study area includes Grand Lake and Hudson Reservoir located in northeastern Oklahoma. Detailed information of Grand Lake and Hudson Reservoir involving storage capacity and area are provided in Table 3.1. Grand Lake has a total area of 188.2 square kilometers, a mean depth of 11.0 m, a maximum depth of 40.5 m (Hunter et al., 2020). Three significant rivers, including Neosho, Spring, and Elk, drain into Grand Lake. The lake spans across parts of Craig, Delaware, Mayes, and Ottawa counties in Oklahoma, making it the third largest lake in terms of both surface area and water holding capacity in the state (Hunter et al., 2020). The watershed of Grand

Table 3.1: Characteristics of Lakes under study

| Lake | Surface area ($km^2$) | Shoreline length (km) | Mean Depth (m) | water volume ($m^3$) | References |
|---|---|---|---|---|---|
| Grand Lake | 188.2 | 2,092 | 11.0 | 2,062,378,560 | (Authority, 2009) |
| Hudson Reservoir | 49 | 322 | 5.5 | 247,066,044 | (Board, 2008) |

Lake contains cultivated agriculture, livestock, and wastewater treatment plants that act as sources of HAB nutrient precursors. Nutrient loadings into Grand Lake vary due to differences in land management practices across its many sub-watersheds. In Grand Lake, Horse Creek Cove is a point of interest in testing remote sensing methods, because most of the HAB events start in this area during the summer months when nutrient-rich runoff from agricultural sources and other land use in the Horse Creek Watershed flows into the lake. Overall land use in the watershed consists of 57% grassland (including planted pasture and natural grassland), 20% agricultural land, 6% developed land, and 3% surface water and wetlands. The Hudson Reservoir is located in Mayes County in the Arkansas River Basin on Neosho River. It is the fourteenth largest reservoir in Oklahoma and supplies water to public and private stakeholders authority (Authority, 2009). Hudson Reservoir receives water mainly from releases of Grand Lake, and it is fed by other tributaries originating in Oklahoma and Arkansas. The study areas of both Grand Lake and Hudson Reservoirs are shown in Figure 3.1

### 3.3.2 Water Quality

Grand Lake and Hudson Reservoir were sampled for water quality parameters by GRDA using EXO-1 and EXO-2 multiparameter sondes regularly since 2011, in-situ water quality throughout the duration of the Landsat-8 and Sentinel-2 missions. The in-situ data collected by GRDA were used to develop remote sensing algorithms for

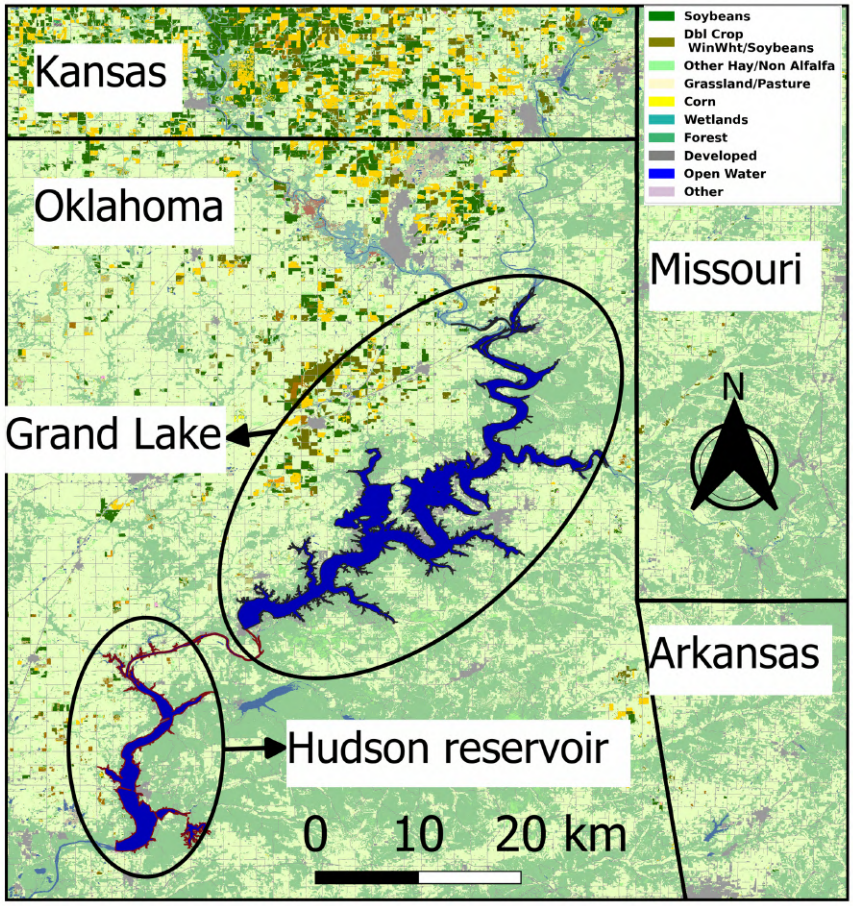


Figure 3.1: Study area of Grand Lake and Hudson Reservoir

estimating CHLa, turbidity, and BGA at Grand Lake and Hudson Reservoir using the spectral imagery from Landsat-8 since 2013 and Sentinel-2 since 2015 as described in Section 3.3.3. Observations of in-situ water quality from Grand Lake and Hudson Reservoir were collected by the GRDA water quality monitoring program. (YSI) multi-parameter water quality probes were used to measure CHLa, turbidity, and BGA. An EXO-1 and EXO-2 sondes with water quality probes were assembled, prepared, and calibrated according to procedures detailed in the instrument's operation manual (YSI, 2017). All the sampling sites used to collect in-situ water quality data for both Grand Lake and Hudson reservoirs are shown in Figure 3.2. Parameters used to build the remote sensing tool include the CHLa concentration in µg/L, turbidity in NTU, and phycocyanin concentration in cells/mL. For model building, the water

quality data selected corresponds to within ± 3 days of sampling for CHLa, turbidity, and ± 5 days of sampling for BGA. The in-situ data represent the water quality conditions close to the water surface where all the sensor data were collected within 3 feet of the surface.

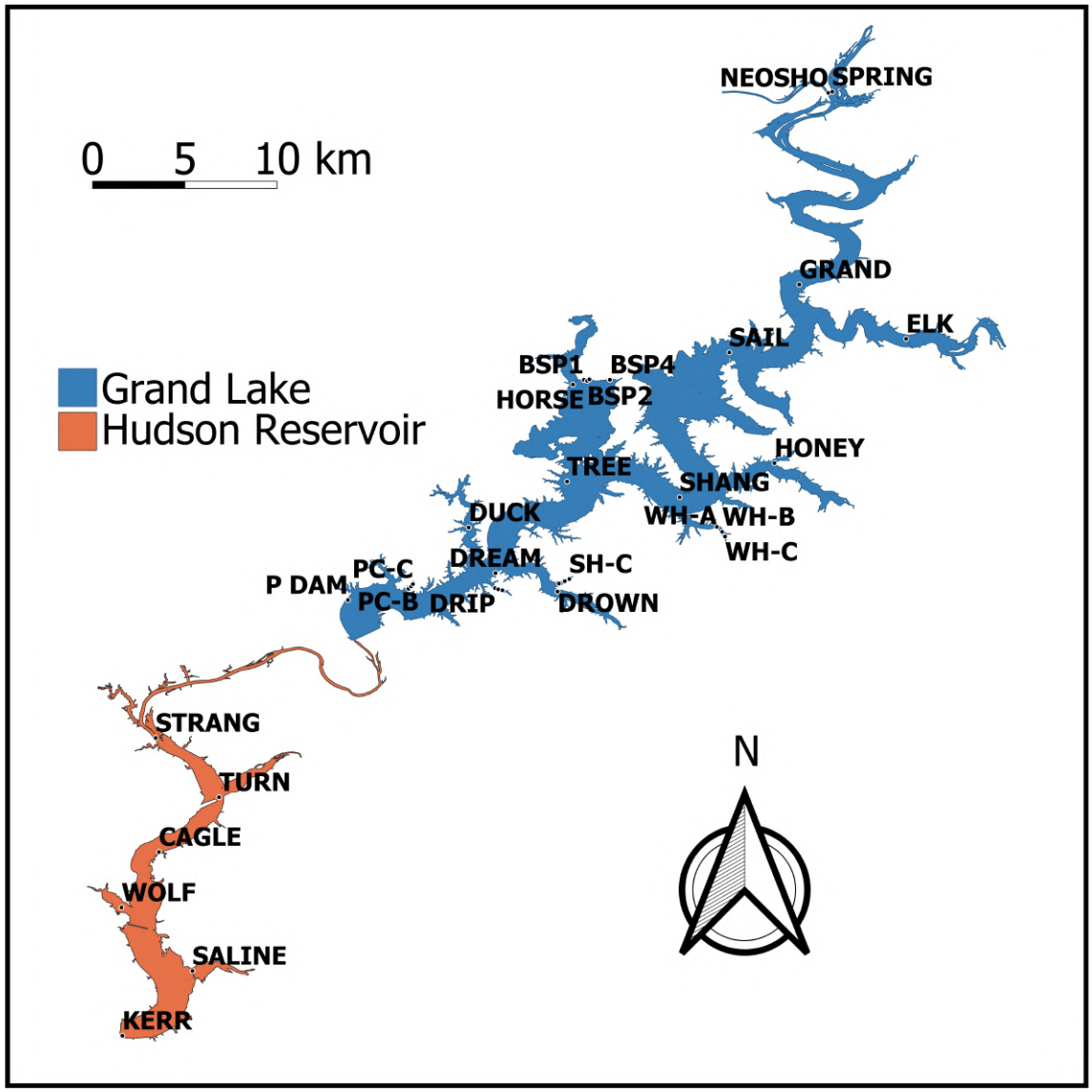


Figure 3.2: Sampling sites of reservoirs under study

### 3.3.3 Satellite Data

Satellite imagery was obtained from both Landsat-8 and Sentinel-2 and co-located with in-situ observations for development of remote sensing algorithms for CHLa, turbidity, and BGA. The OLI is suitable to assess water quality in surface water bodies such as lakes, rivers, and coastal marine zones. The OLI measures the reflectivity of earth's surface across nine spectral bands including Panchromatic band at 15 meters, Ultra-Blue Coastal Aerosol, Blue, Green, Red. Near Infrared (NIR), Short Wave Infrared (SWIR) 1, SWIR2, and Cirrus bands at 30 meters resolutions on the ground. The TIRS has two thermal spectral bands including Thermal Infrared 1, and Thermal Infrared 2 with spatial resolution of 100 meters. Sentinel-2 is an earth observation mission from the Copernicus program to acquire satellite imagery with spatial resolution ranging from 10 meters to 60 meters on the ground. Currently, the two Sentinel satellites S2A and S2B collects imagery with twelve spectral bands including Blue, Green, Red, NIR, at 10 meters, Vegetation Red edge, Narrow NIR, SWIR bands at 20 meters, Water Vapor, and Cirrus bands at 60 meters resolutions. The detailed information about the Landsat-8 and Sentinel-2 bands used in this study are shown in Table 2. Also, the information related to the bands available for both Landsat-8 and Sentinel-2 are provided in the supplemental document in Table A.1, Table A.2 and Table A.3.

## 3.4 Software Tool

### 3.4.1 Empirical Basis of the Tool

Statistical models including multiple regression, SVR and RFR were used to develop algorithms for estimating CHLa concentrations, turbidity, and BGA concentrations from the co-located observations of reflectance and in-situ water quality. The CHLa, turbidity, and BGA data were not normally distributed; therefore, we log-transformed

Table 3.2: Spectral bands with band wavelength used to remotely sense CHLa, turbidity, and BGA

| Remote sensing parameter | Spectral band notations and band ratios used | Wavelength (nm) |
|---|---|---|
| CHLa with Landsat-8 sensors | SR_B2 (blue) | 432-512 |
| | SR_B3 (Green) | 533-590 |
| | SR_B4 (Red) | 636-673 |
| | SR_B5 (Near Infrared) | 851-879 |
| | SR_B6 (Shortwave infrared 1) | 1566-1651 |
| | ST_B10 (Surface temperature from level-1 Band 10) | 10600-11190 |
| Turbidity with Landsat-8 sensors | SR_B2 (blue) | 432-512 |
| | SR_B3 (Green) | 533-590 |
| | SR_B4 (Red) | 636-673 |
| | SR_B5 (Near Infrared) | 851-879 |
| | SR_B7 (Shortwave infrared 2) | 2107-2294 |
| BGA with Sentinel-2 sensors | B2 (Blue) | 496.6 (S2A) / 492.1 (S2B) |
| | B3 (Green) | 560 (S2A) / 559 (S2B) |
| | B4 (Red) | 664.5 (S2A) / 665 (S2B) |
| | B5 (Red Edge 1) | 703.9 (S2A) / 703.8 (S2B) |
| | B6 (Red Edge 2) | 740.2 (S2A) / 739.1 (S2B) |
| | B12 (Short Wave Infrared 2) | 2202.4 (S2A) / 2185.7 (S2B) |
| | B8A (Red Edge 4) | 864.8 (S2A) / 864(S2B) |

the data (Figure 3.5) to make the distribution closer to normal. Regression models are widely used to associate algal pigments and other water quality parameters with spectral signatures of satellites. Regression minimizes the sum of the squared errors using a cost function given by Equation 3.1

$$C = \frac{1}{l}\sum_{i=1}^{n}((y_i - f(x_i, v))^2 \tag{3.1}$$

Given a training data set containing dependent variables $y_i$ and independent variables $x_i$ denoted $x_i, y_i \in \mathbb{R}^k$, $f(x_i, v)$ is the function used in the regression for prediction, n is the sample size and v is the vector that defines the function f with same dimension i.e., the number of parameters in both vector v that defines the function f are same. If the noise generated in the predicted data does not follow a Gaussian

distribution, then a new loss function is needed to quantify and minimize the error between predicted and expected concentrations of CHLa, turbidity, and BGA. In case of multiple regression, the dependent variables include CHLa and turbidity while the independent variables include the spectral bands from Landsat 8 Operational Land Imager (OLI) and Thermal Infrared Sensor (TIRS). A model was developed using multiple regression to find a relationship between each dependent variable and the independent variables. Algal dynamics in surface water are non-linear, and therefore machine learning models including SVR, and RFR were also analyzed to develop remote-sensing algorithms for water quality parameters. SVR is an extension concept of Support Vector Machines, which are applied to regression statistical analysis. SVR has been successfully applied to develop machine learning algorithms for non-linear, chaotic systems including alternating current cycles, urban traffic flows, medical applications, and stock markets (Xu et al., 2014; Alipour et al., 2022; Hong et al., 2011; Weng et al., 2013; Liang et al., 2010; Khan and Zubek, 2008; Agarwal et al., 2011; Das and Padhy, 2012; Yu et al., 2008). In SVR, a linear $\in$ loss function given by Vapnik (Drucker et al., 1996; Vapnik et al., 1995) and the function is shown in equation 3.2:

$$L = \mid y - f(x, v) \mid_{\in} \tag{3.2}$$

Where L is the loss function in SVR, y is the expected value and $f(x, v)$ is the predicted regression function. The loss function takes a value 0 when L $\leq \in$, otherwise the L can be $|y - f(x, v)|$ - $\in$ i.e. when L $> \in$. In this study, as the predicted CHLa,turbidity, and BGA were tested with non-linear machine learning models, a kernel function is needed for SVR. Typically, a kernel function used to transform features from original to higher dimensions to improve group separation of expected data around hyper-plane. An optimal prediction was found by iterating through the available kernel functions including linear, polynomial, Gaussian, Sigmoid, spectral angle, and Radial Basis Function. In this study, a Radial Basis Function (RBF)

kernel was selected after testing all the available kernel functions to give an optimum prediction for both CHLa and turbidity. The RBF kernel was used to predict water quality and has been widely used in regression problems (Chitralekha and Shah, 2010; Nadadoor et al., 2012).

A Random Forest (RF) is an ensemble machine learning algorithm that contains a large number of decision trees such that the algorithm obtains a prediction based on a consensus of responses in all the trees. The sample data were partitioned to random subsets for predicting both CHLa, turbidity, and BGA. RF can handle large datasets and can predict co-related water quality data without decreasing prediction accuracy. In multispectral imagery there is strong correlation between bands. The RF algorithm uses numerous de-correlated decision trees randomly and assigns weighting based on the performance of each tree, using Gini impurity. The Gini importance can be computed as the sum of the node impurity decrease over the number of splits (across all the decision trees). Although RF seems to be a promising model it is important to check if the model is overfitting the training dataset. To check whether a statistical model is not overfitting the trained dataset, both k-fold cross-validation and hyper parameter tuning was performed on different training and validation datasets to check how the model is performing.

### 3.4.2 Structure and Workflow of Python Tool

The software developed in this study is written in Python 3.7. Python is free, open source, and extensively utilized for scientific applications because of its object-oriented, interpretive, and interactive programming features (Lutz, 2013; Millman and Aivazis, 2011). Python possesses a wide range of extension libraries that were utilized to ensure smooth operation of the procedure to extract imagery and analyze the data. The Google Earth Engine Python Application Program Interface (API) was used to acquire level-2 Landsat-8 imagery of the study area. Similarly, the United

States Geological Survey (USGS) Application Programming Interface (API) was used to obtain Sentinel-2 level-1 images from Google Cloud Services. Landsat-8 level-2 images were obtained using spatial bounds for the lakes in scenarios where the cloud cover was less than 10% of the image. Geographical coordinates (latitude and longitude) are provided as a Comma Separated Value (CSV) file such that the python software can find the satellite images within the bound spatial coordinates. In case of Sentinel-2, after acquiring the level-1 images, Sen2Cor was used to process the single date Level-1 top of atmosphere images to level-2 bottom of atmosphere or surface reflection products (Main-Knorn et al., 2017). Based on the geographic coordinates and in-situ sampling dates, the Python tool selects raster images from the available data for each mission, then prints the appropriate candidates that are within the spatial bounds of the study area while having a cloud cover of less than 10%. The level-2 data obtained from Earth engine API has atmospherically corrected level-2 images containing five visible bands, two short-wave bands, and one thermal infrared band. The bands were orthorectified to surface reflectance and surface temperatures. In total there are 19 bands available from the level-2 images which are shown in Table A.1 and A.2. Sentinel-2 has two satellites which are operated by the European Space Agency (ESA). Both the satellites S2A, and S2B carries multispectral instruments that images the earth both in visible and infrared band spectrum. Sen2Cor was used to resample the level-1 images to level-2, which contain bands with spatial resolutions of 10, 20, 60 meters. The 20-meter spatial resolution is considered in building the remote sensing tool to measure the concentration of BGA because the band 8A which is obtained by level-2 processing has more precise spatial spectral information. The detailed description of the bands after processing to level-2 with surface reflectance using sentinel-2 are shown in Table A.3.

After acquiring and processing the satellite images for both the Landsat-8 and Sentinel-2 satellites, the pixel values were extracted at the in-situ sampling location

that most closely matched the sampling time (within a week). All the pixel values of the reservoirs were parsed based on the availability of in-situ water quality data. The pixel data obtained was split to train the machine learning algorithms, then test the performance of the models using cost functions. The statistical training process involves improving the model performance based on using the error metrics, including Root Mean Square Error (RMSE) and coefficient of determination ($R^2$). In addition to error metrics, feature selection, and hyper-parameter tuning were used to improve the performance of the remote sensing tool. The detailed information regarding the feature selection and hyper-parameter tuning are discussed in sections 3.5.2. Based on the model performance, the RF model was selected for subsequent testing using k-fold cross-validation to estimate model accuracy. After gaining confidence that the RFR did not overfit the training dataset, it was selected as the final model to be used in the remote sensing tool. As discussed in section 3.4, a remote sensing algorithm was developed between the multispectral band data, CHLa concentrations, and BGA concentrations. Similarly, another algorithm was developed between the multispectral data and turbidity. The steps involved in developing the remote sensing software are shown in Figure 3.3.

Table 3.3: Recreational action levels for CHLa, BGA, and Microcystin levels

| Probability of acute health effects | CHLa ($\mu$g/L) | Cyanobacteria (cells/mL) | Estimated corresponding Microcystin concentrations |
|---|---|---|---|
| Low | $\leq$10 | $\leq$20,000 | $<$ 10 |
| Moderate | 10 to 50 | $>$ 20,000 to 100,000 | 2 to 20 |
| High | $\geq$50 | $>$ 100,0000 | $>$ 20 |

The information required to use the remote sensing tool requires the user to register on Google cloud and Earth Resources Observation and Sciences (EROS). In addition to login information to the publicly available satellite imagery, users must provide the desired cloud cover (in percentage), geographical coordinates, and tempo-

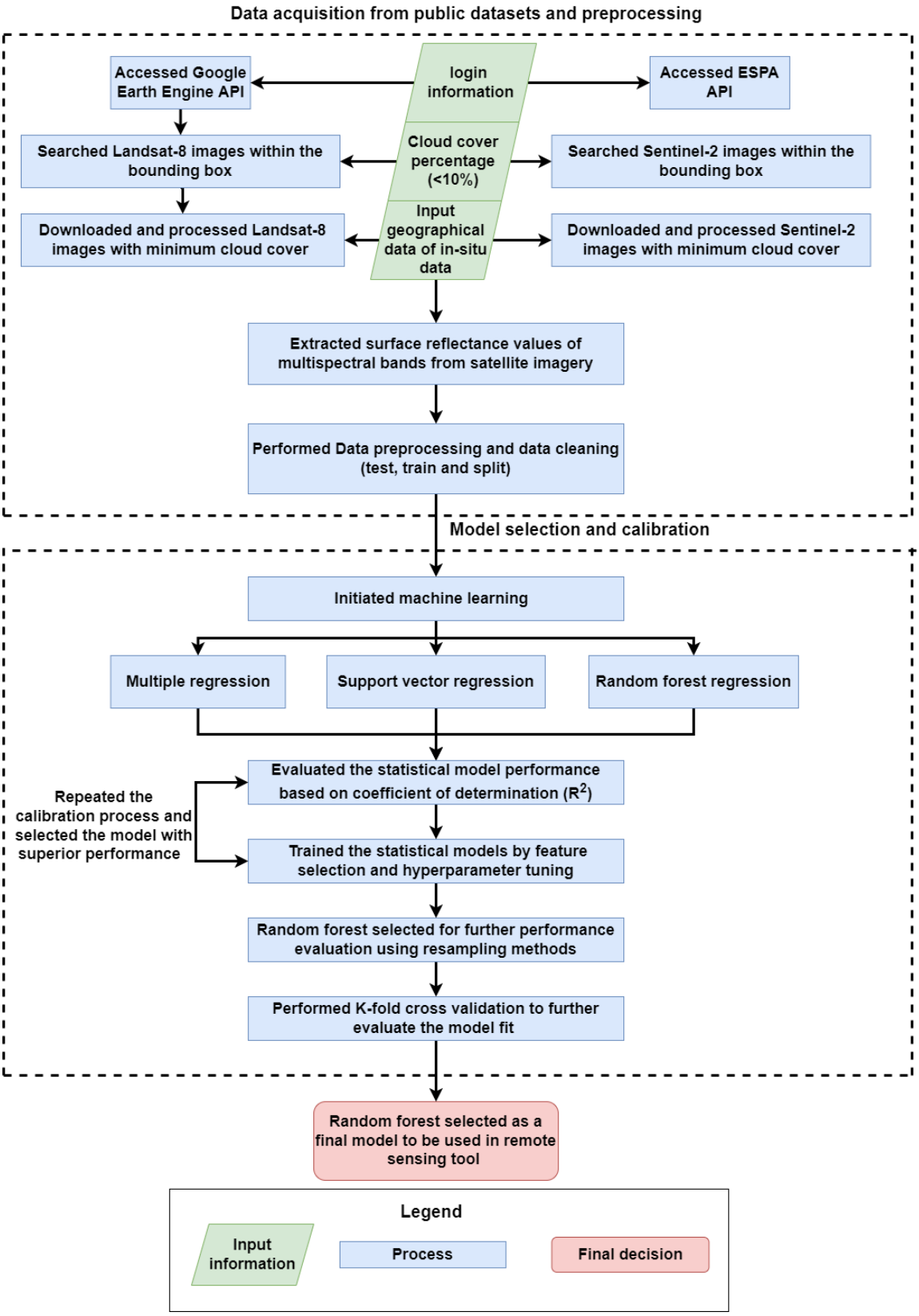


Figure 3.3: Flow diagram showing the operational structure of the python tool

ral information (start date and end date) of the imagery to select the desired satellite images within the region of interest. In addition, a boundary shapefile of the water-body must be provided by the user to clip the region of interest and remove unwanted regions in the raster. Using all the input information provided by the user, the tool searches publicly available datasets from Landsat-8 and Sentinel-2 imagery in Google cloud and downloads them to the local computer. The spectral reflectance values in the region of interest are then extracted by the remote sensing tool. The RF algorithm is applied to the clipped region of interest to produce gridded estimates, and the unwanted pixels of land are masked by the tool. Using the surface reflectance values of water, the tool estimates the concentrations of CHLa, turbidity, and BGA to produce gridded images of water quality. The gridded estimates on the satellite raster can be used to provide HAB advisories for further analysis based on the concentrations of CHLa and BGA as shown in Figure 3.4.

The criteria to establish a HAB alert are based on World Health Organization (WHO) alert levels of CHLa and BGA concentrations more than 50 µg/L and 100,000 cells/mL, respectively (Izydorczyk et al., 2009; USEPA, 2019; Ahn et al., 2007). The detailed information on alert concentration of algal pigments are shown in Table 3.

## 3.5 Results and Discussions

The distribution of CHLa, turbidity, and BGA from the in-situ observations in the study area are shown in Figure 3.5. CHLa concentrations used in this study ranged from 0 to 60 µg/L, turbidity ranged from 0 to 150 NTU, and BGA ranged from 0 to 20000 cells/mL. The resolution of BGA sensor indicates that 0.1 RFU corresponds to estimated cyanobacteria concentration of 1 cell/mL (Lorenzen, 1966). Based on the WHO guidelines for recreational water, CHLa concentrations more than 50 µg/L and BGA of 200,000 cells/mL indicate a HAB event and requires further testing of water samples in the laboratory for microcystins. Both CHLa and turbidity can be

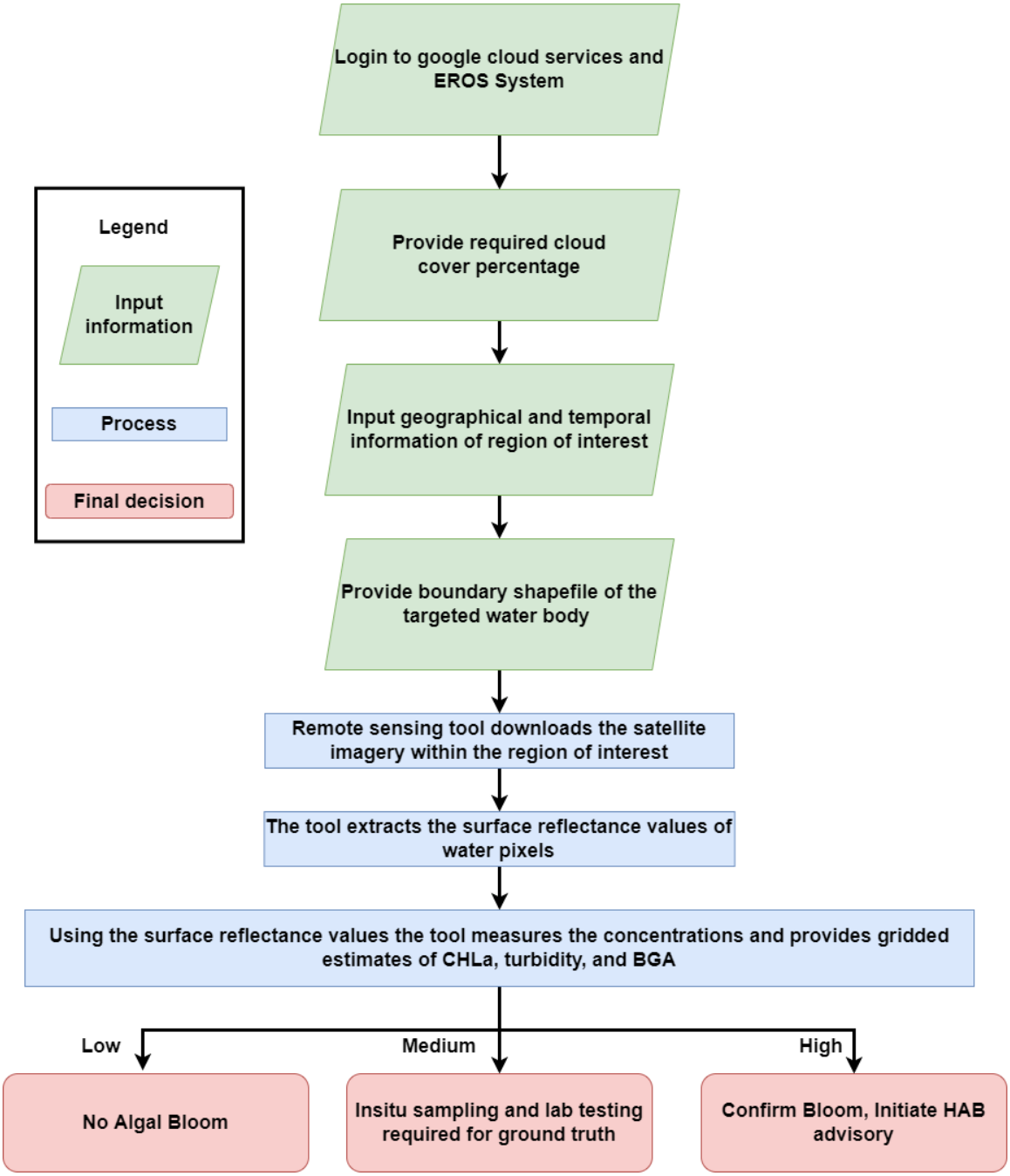


Figure 3.4: The process flow diagram showing the steps to use the remote sensing tool

remotely-sensed using Landsat-8 satellite imagery. Landsat-8 satellite data since the mission launched in 2013 were compiled to match 1709 in-situ observations for CHLa and turbidity in the study area. Sentinel-2 data were compiled to remotely sense BGA that were matched to 1095 data points from 2015 to 2020. The in-situ water quality data were positively skewed for CHLa, turbidity, and BGA as shown in Figure 3.5. Each of the three water quality parameters was log-transformed such that the fit is close to a Gaussian or normal distribution. However, non-parametric machine learning models like SVR and RFR do not need data to fit a normal distribution.

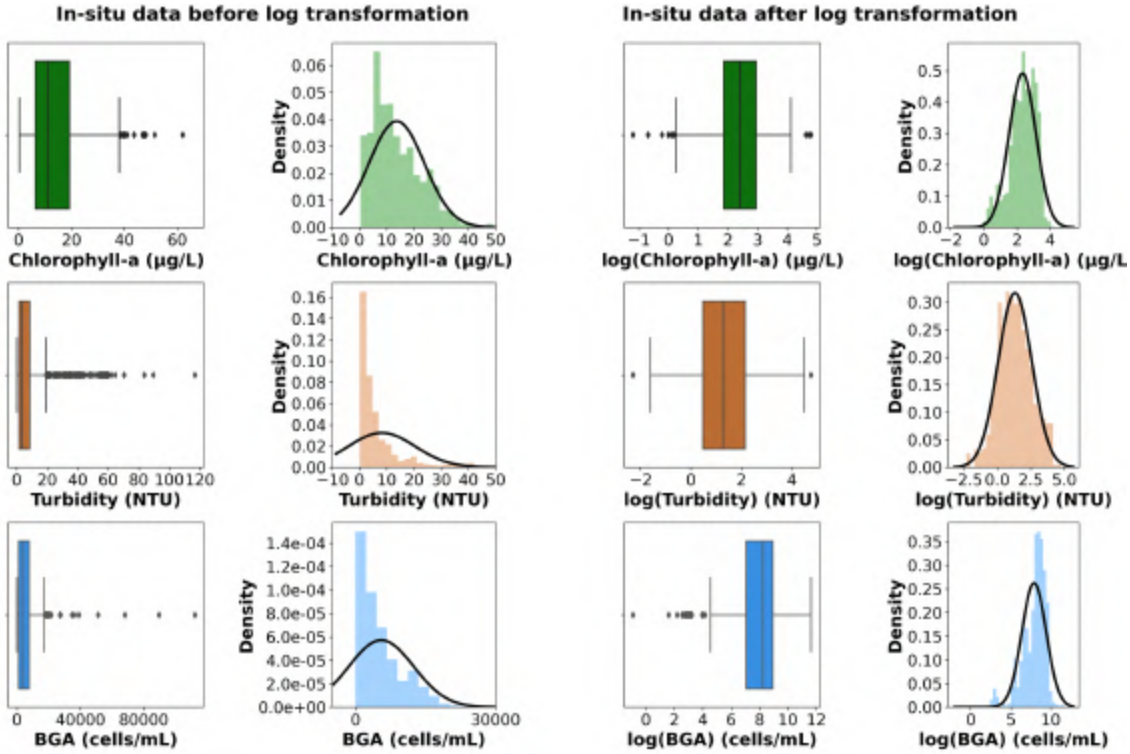


Figure 3.5: Whisker and distribution plots of in-situ water quality parameters

After fitting the in-situ data to a Gaussian distribution, algorithms that translate observations from single bands and band combinations into water quality parameters were generated using co-located pixels based on coordinates. The bands used in building and testing regression algorithms are shown in Table 3.2 for each water quality parameter and satellite mission. For remote sensing of both CHLa, turbidity, and BGA, features were extracted using individual bands from Landsat-8 and Sentinel-2

respectively.

### 3.5.1 Multicollinearity of Multi-spectral Bands

Some of the multispectral bands of both Landsat-8 and Sentinel-2 images are strongly correlated, which causes data redundancy. Therefore, we dropped highly-correlated bands with coefficient of determination ($R^2$) of greater than 0.9. Figure 3.6 shows the correlation matrix between all the observations from the Landsat-8 and Sentinel-2 imagery. The bands that were dropped included Ultra-Blue, Coastal Aerosol (SR_B1) and Shortwave Infrared 2 (SR_B7) while estimating CHLa concentrations using Landsat-8. Similarly, surface reflectance bands including Ultra-Blue, Coastal aerosol (SR_B1), and Shortwave Infrared 1 (SR_B6) bands were dropped while estimating turbidity concentrations using Landsat-8. In case of BGA estimation, surface reflectance bands including Red Edge 3 (B7) and Short-Wave Infrared 1 (B11) were dropped while using Sentinel-2 imagery for remote sensing. In addition to surface reflectance bands, few surface temperature bands were retained because surface temperature is very important in detecting CHLa and turbidity concentrations which are discussed in section 3.5.2 (Harod et al., 2021). The bands used in machine learning algorithms for model calibration were selected based on feature importance as discussed in section 3.5.2.

### 3.5.2 Feature Selection for Regression

Feature selection for regression was done based on the feature importance scores determined by a RF (Esteban et al., 2019; Lee et al., 2020; Belgiu and Drăguţ, 2016). All the three regression algorithms are further investigated (multiple regression, SVR, and RFR) and RF seems to estimate the water quality parameters better which is explained in section 3.5.3. Feature importance data provides important insights into the significance of the various spectral bands to remotely sense the water quality

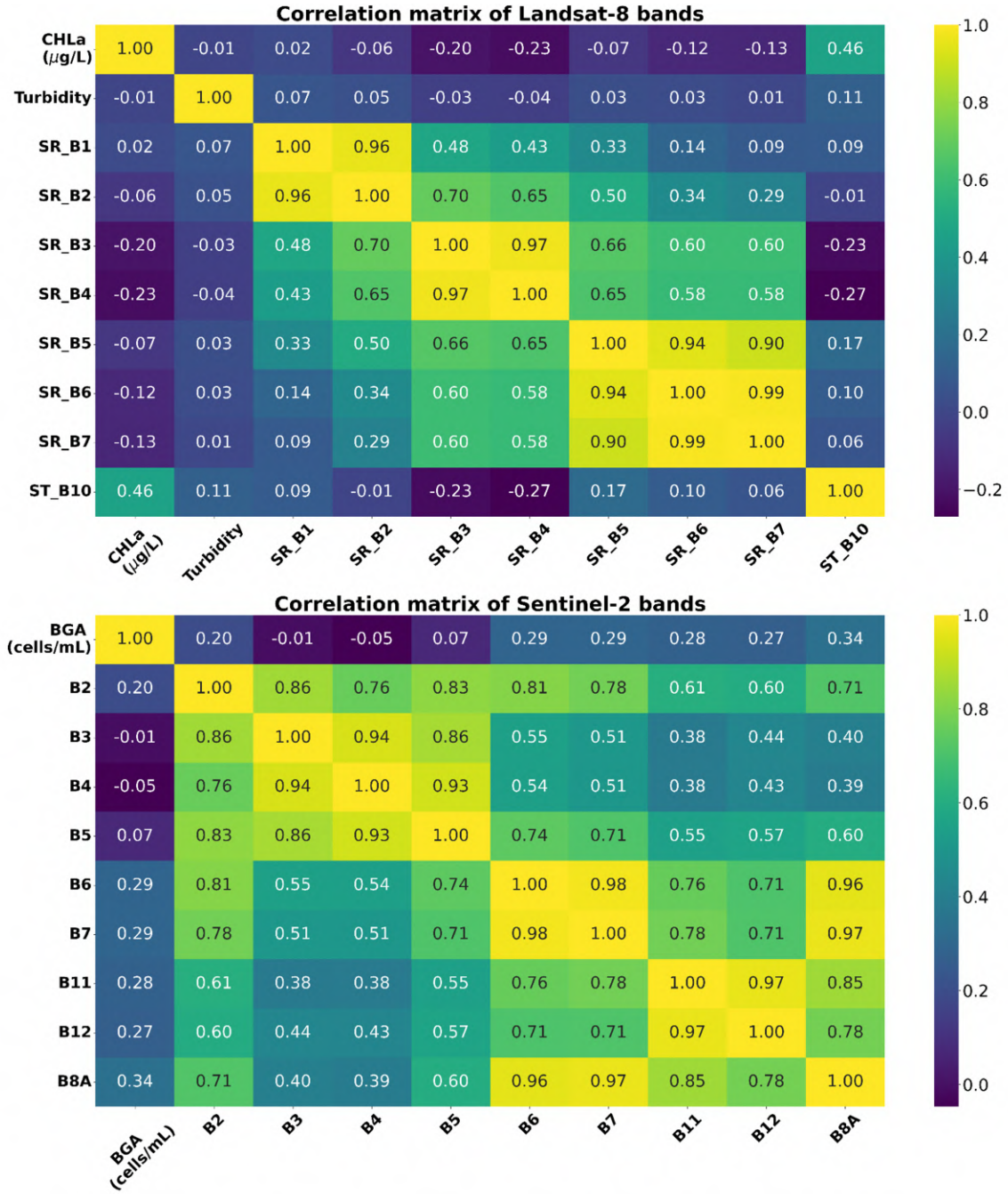


Figure 3.6: Correlation matrices between water quality parameters, Landsat-8 and sentinel-2 bands

parameters. The RF algorithm has built-in feature importance method that can be computed using Gini importance. Importance of computed bands including surface temperature bands of Landsat-8 play an important role in predicting CHLa and turbidity, as shown in Figure 3.7.

Besides high correlation, the surface reflectance band Ultra-Blue Coastal Aerosol (SR_B1) was dropped because of low feature importance with Gini impurity close to zero. The Surface Temperature (ST_B10) was retained for algorithm development because it has been previously been used to remotely sense the concentrations of CHLa in water (Tanaka et al., 2004; Allan et al., 2011). For remotely-sensing BGA, the Aerosol Optical Thickness (AOT) band of Sentinel-2 was dropped because aerosols in air do not aid in detecting phycocyanin pigments. The Red Edge 3 (B7) band was dropped because of low feature importance and high correlations between Red Edge 2 (B6) and Red Edge (B7). Also, between bands Short Wave Infrared 1 (B11) and Short Wave Infrared 2 (B12), Short Wave Infrared 1 (B11) was dropped because of low feature importance and high correlation. However retaining one of the short wave infrared bands (B12) is important because they help in detecting cyanobacterial blooms (Yan et al., 2018). Most of the remote sensing methods predicted BGA in the spectral range from 600 to 754 nm (Ogashawara et al., 2013; O'Shea et al., 2021; Yacobi et al., 2015). However, SWIR bands have been less susceptible to turbidity in detecting algal pigments and distinguish surface cyanobacterial blooms and macrophytes (Mansaray et al., 2021; Oyama et al., 2015). Therefore, the B12 band was retained in building the remote sensing tool to estimate the BGA pigment.

### 3.5.3 Machine Learning Algorithms for Remote Sensing of HABs and Turbidity

Using data from the two reservoirs (Grand Lake and Hudson Reservoir), three machine learning algorithms were developed and tested. Since the in-situ data of CHLa,

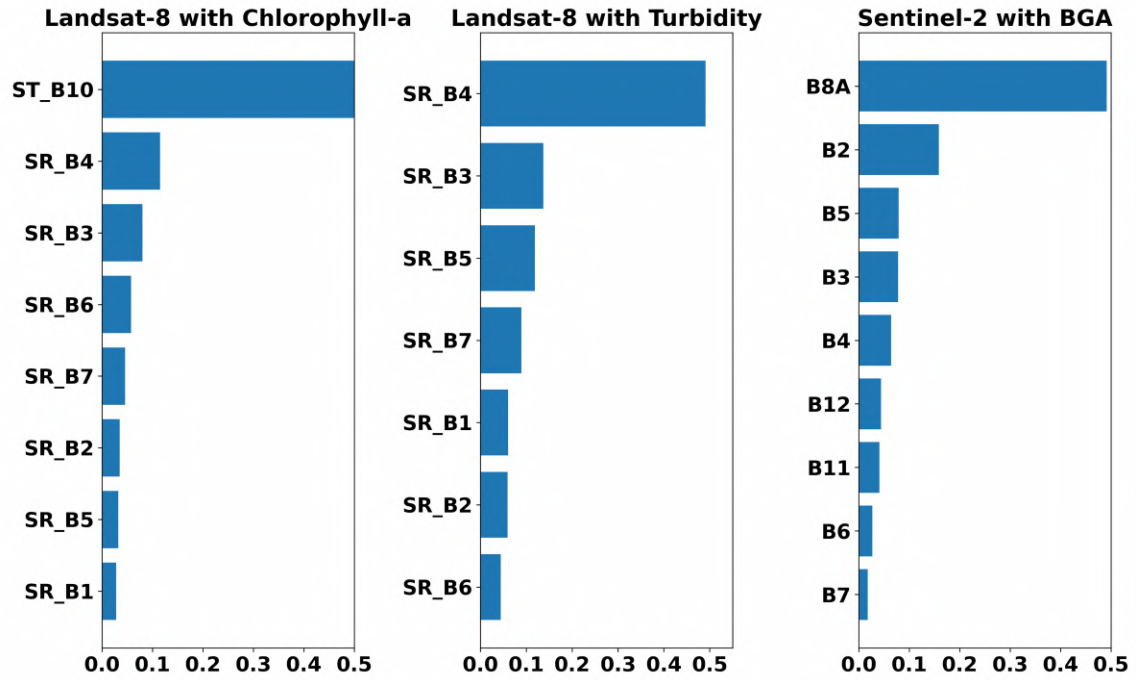


Figure 3.7: Feature importance using random forest regression, Landsat-8 and sentinel-2 bands

turbidity, and BGA were close to a normal distribution, standardization of all the features was performed using a standard scaler (Ferreira et al., 2019; Garreta and Moncecchi, 2013). After scaling, the data were split by 70% to train and by 30% to test the machine learning models. As different models were tested using different band combinations, the most efficient regression models were selected for CHLa, turbidity, and BGA based on $R^2$ values and kernel density plots as shown in Figure 3.8. The kernel density plots in Figure 3.8 show that the multiple regression model explains roughly half of the relative variability in the concentrations of BGA, CHLa, and turbidity.

Also, the SVR model overestimates the water quality parameters and does not have definitive peaks at high concentrations of CHLa. Out of the three machine learning algorithms tested, RFR estimates the water quality parameters close to the measured testing dataset. The RFR algorithm is further tested on testing datasets to observe the model performance based on RMSE and $R^2$. Based on previous literature

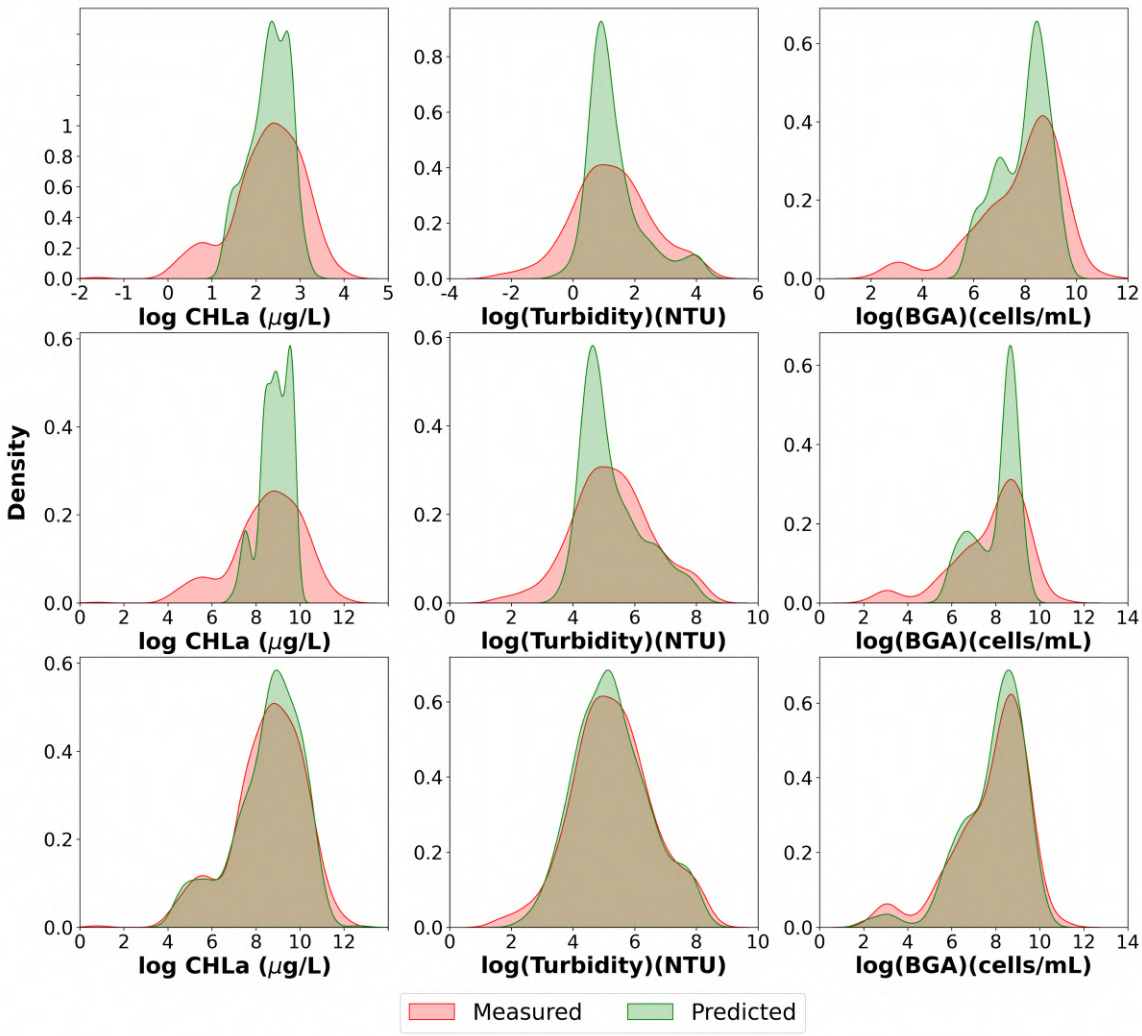


Figure 3.8: Kernel density plots for measured and predicted water quality data using multiple regression (top), SVR (middle), and RFR (bottom)

Xie et al. (2012); Xu et al. (2015) and trials, the RBF kernel was used in SVR for remote sensing of water quality parameters. The performance of the statistical models was assessed using RMSE and $R^2$. In both the testing and training cases, the $R^2$ values for CHLa, turbidity, and BGA using the SVR and RFR algorithms improved model performance compared to multiple regression, as shown in Table 3.4. The RFR model explained observations of CHLa, and BGA better for training than testing. A k-fold cross validation using random search CV was used to validate the model for different potential hyper-parameters, including maximum depth, maximum features, samples split, and sample leaf (Krstajic et al., 2014; Bates et al., 2021). The results of the k-

fold cross validation for the RFR model after validation are reported in Table 3.4. To visualize the relationship between measured and modeled water quality parameters, a scatter plot is shown in Figure 3.9.

Table 3.4: Statistics metrics for best fit regression models

| Regression model used | Water quality parameter | RMSE (training dataset) | RMSE (test dataset) | $R^2$ (training dataset) | $R^2$ (test dataset) |
|---|---|---|---|---|---|
| Multiple regression | CHLa | 0.71 | 0.68 | 0.29 | 0.29 |
| | Turbidity | 0.90 | 0.89 | 0.47 | 0.51 |
| | BGA | 1.11 | 1.21 | 0.41 | 0.46 |
| Support vector regression | CHLa | 0.54 | 0.52 | 0.59 | 0.58 |
| | Turbidity | 0.79 | 0.78 | 0.59 | 0.62 |
| | BGA | 0.88 | 1.03 | 0.63 | 0.61 |
| Random forest regression | CHLa | 0.36 | 0.40 | 0.81 | 0.75 |
| | Turbidity | 0.44 | 0.54 | 0.87 | 0.82 |
| | BGA | 0.47 | 0.75 | 0.89 | 0.79 |

The best regression fit for all the water quality parameters provided visualization of outliers which were carefully examined and removed if necessary. The $R^2$ values are displayed at the top of each subplot to illustrate the ability of the model to explain the data. The k-fold cross validation, scatter plots, and statistical metrics demonstrate that RFR is the best machine learning method of the three approaches to remotely sense HABs using imagery from Landsat-8 and Sentinel-2.

Overfitting is always a problem while using non-parametric models like RF while building remote sensing tools. To avoid overfitting of the RF regressors to the training datasets, hyperparameter tuning was performed manually using different tree depths, maximum features, and minimum samples split using manual search and Random Search CV (Schratz et al., 2019). The number of estimators used in building the RFR algorithms were 300. The estimators used in building the RFR algorithms have a maximum tree depth of 15, 16, and 18 for remote sensing of CHLa, turbidity, and BGA respectively. The maximum tree depth used in the estimator trees is based on hyperparameter tuning, because excessive tree depth can result in overfitting of the

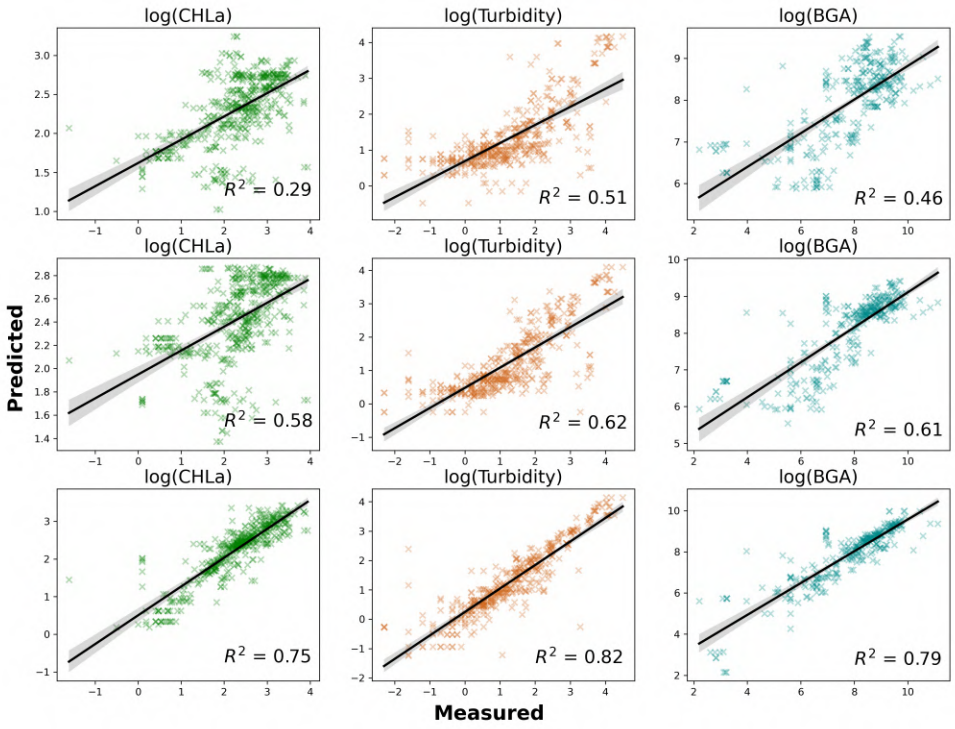


Figure 3.9: Scatter plots showing the relationship between remotely-sensed in-situ observations of water quality parameters using multiple regression (top), SVR (middle), and RFR (bottom)

trained dataset. The maximum number of features used are 10, 7, and 7 in training the RFR algorithm. The decision to include the features in building the RFR was based on multi-collinearity of different bands, feature selection, and hyperparameter tuning. Also, the minimum number of data points inside a node before the split was two, and the minimum sample split was determined using hyperparameter tuning. The first estimator of RFR algorithms developed in remote sensing of CHLa, turbidity, and BGA are shown in Figure 3.10.

### 3.5.4 Remotely Sensed Water Quality during HAB Events

In Grand Lake, there were temporally coincident satellite flyovers during four days in 2021 including February 03, March 07, April 08, and July 13. There was a HAB

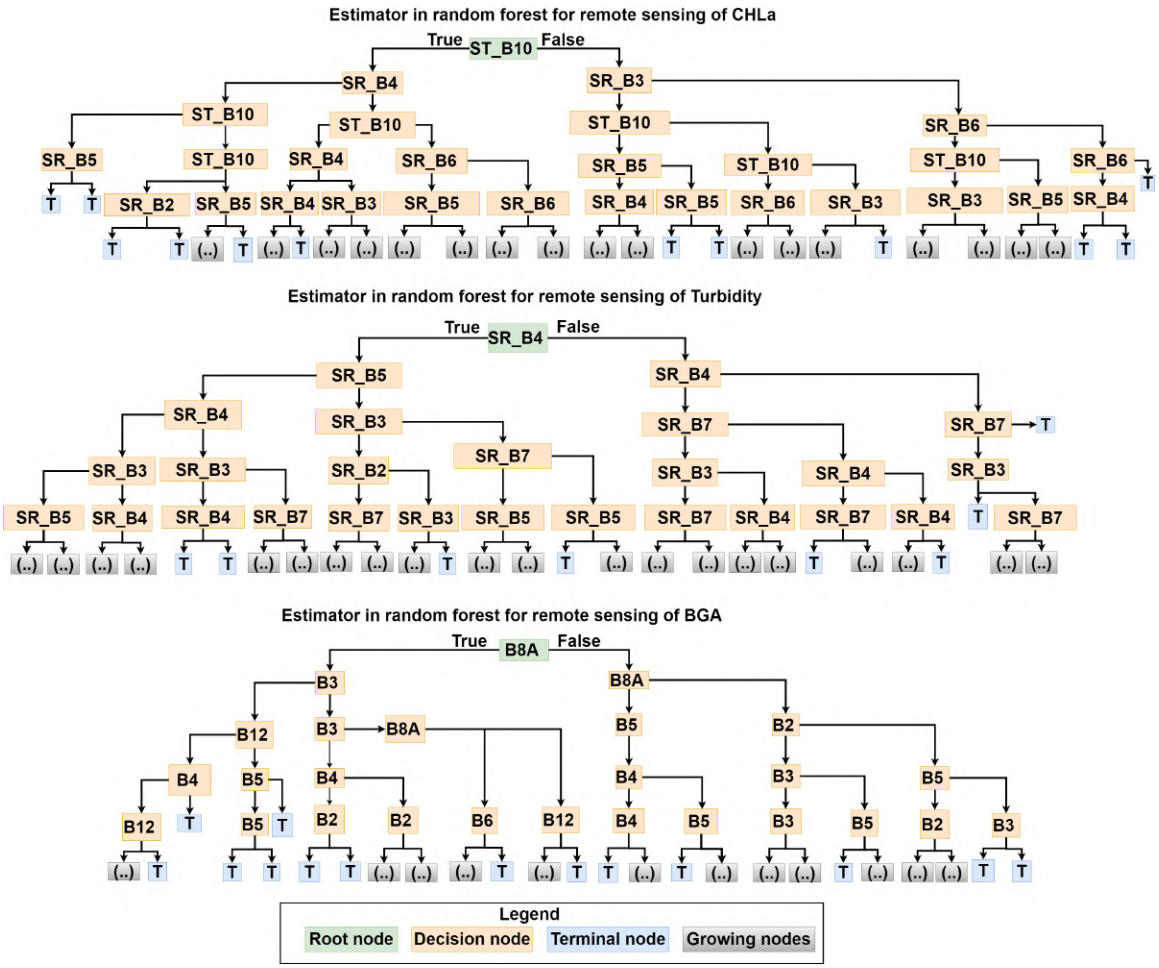


Figure 3.10: Random forest estimators

event ending in July 2021 when the satellite data were used to remotely sense CHLa, and BGA. The concentrations of CHLa are higher on the edges of Horse Creek in the month of July which is confirmed by the field sampling personnel of GRDA. In addition to algal pigments, turbidity was also remotely sensed in Grand Lake using RFR. Figure 3.11 shows that the Landsat-8 remotely sensed turbidity was not affected by the surface CHLa and BGA pigment concentrations.

Finally, since the temporal and spatial data obtained from satellite data is coarse, increasing number of satellite flyover and high resolution in-situ datasets can further improve the developed in-lake water quality model. Besides, there was a significant HAB event during June 2021 but due to significant cloud cover the imagery was not obtained by the remote sensing tool for remote sensing. The remote sensing

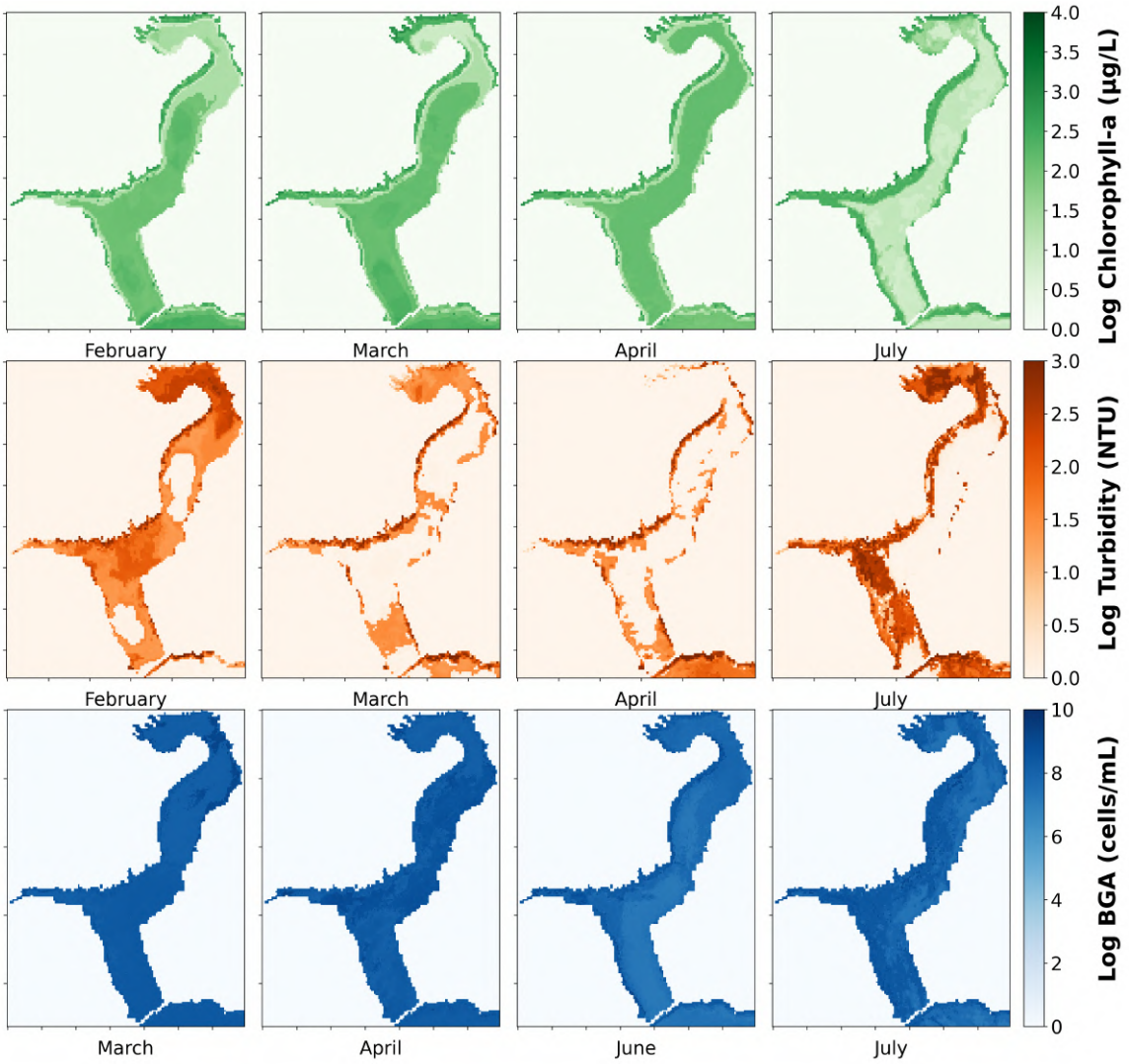


Figure 3.11: Remotely sensed CHLa (µg/L) (top), Turbidity (NTU) (middle), BGA (cells/mL) (bottom) using RFR algorithm developed in building the remote sensing tool

tool developed here has the capability to remotely sense both HABs and turbidity. In addition to remote sensing application the current tool could be used to develop surface water quality models. These water quality models can be integrated with the Environmental Fluid Dynamics Code (EFDC) to understand how sediment and nutrient influxes into lakes and reservoirs can cause HAB events in surface water bodies (Gao and Li, 2015; Xie, 2012).

## 3.6 Conclusions

The goals of this study were to (1) develop remote sensing algorithms for water quality using Landsat and Sentinel imagery and (2) build a real-time monitoring tool to characterize HABs in inland water bodies in the Southern Great Plains. The automated Landsat and Sentinel-based tools for detection of HABs in the surface water bodies using open-source Python facilitate image acquisition and display of the spatial extent of remotely sensed algae and turbidity. The tool provides satisfactory results for monitoring water quality. This remote sensing approach can be extended to new water resources when large datasets of in-situ data water quality data is available to estimate algae concentration on the surface. However, there are certain limitations to the application of this tool. Landsat-8 and Sentinel-2 have moderate spatial and temporal resolutions. A Landsat pixel, which is 900 m2, is susceptible to land-water mixed pixel problems, interference from other objects, and low intensity reflectance from low-concentrations of CHLa and BGA. Additionally, Landsat-8 has temporal resolution of 16 days during which time, several events could occur that are not observed. Moreover, satellites are not helpful for species identification and detecting toxic algal scum formation within hours. Integrating unmanned aerial vehicles (UAVs) and unmanned surface vehicles (USVs) to collect remote sensing and in-situ data would help address these issues. The tools we developed here could be expanded to use the raster datasets collected by USVs and other satellites to remotely sense water quality data. Integration of these datasets and applying machine learning and artificial intelligence could potentially produce more reliable daily HABs advisories.

# CHAPTER IV

## Watershed Modeling

### 4.1 Abstract

Watershed models are useful to analyze water quantity and quality issues. Hydrologic Simulation Program in Fortran (HSPF) is a mathematical model that can be used to analyze sediment and nutrient transport within a watershed. PyHSPF automates the data extraction, model calibration, post-processing of HSPF files. This chapter provides information about building a HSPF model that can be used to simulate Horse Creek Watershed after successful calibration and simulation of flows. In addition to building a watershed model, the capability of PyHSPF to extract data from web has been improved by adding climate data from Oklahoma Mesonet. In addition to watershed modeling this chapter provides information about water quality monitoring of streams using unmanned systems.

Keywords: Watershed modeling, water quality monitoring, unmanned systems, water quantity and quality

### 4.2 Introduction

Nutrient and sediment runoff into water bodies create eutrophic conditions in lakes and reservoirs causing harmful algal blooms. Frequent monitoring of streams and reservoirs provide water quality information that can translate into new and enhanced simulation tools (Bencala, 1983). The influence of climate and nutrient sources in streams can be understood using watershed models and simulations. Quantifying land-use and climate effects on watershed is possible using mathematical models like

Hydrologic Simulation Program in Fortran (HSPF). Large amount of data is needed to construct HSPF models with information including stream reaches, catchment areas, land use percentages, observed stream flows, meteorological and climate timeseries (Singh, 1995).

Graphical User Interface (GUI) tools like BASINS help in extracting and preprocess data, create input files, calibrate, postprocess and visualize the results. However there are limitations in GUI based tools because they lack flexibility to preprocess data and are rigid in nature. Therefore it is necessary to use flexible tools that provides an environment to produce results which are reproducible and transparent. PyHSPF is one of the tools that provides a flexible environment for users to develop a watershed model (Lampert et al., 2015). A watershed model for Horse Creek has been developed to understand nutrient and sediment transport which might be causing HAB events in Grand Lake, Oklahoma.

The Horse Creek Watershed within Grand Lake has two gauging stations installed by United States Geological Survey with stage and discharge records from August 2020. The Horse Creek Watershed majorly contains agricultural land, grasslands, and a wastewater treatment plant. To enhance the data acquisition capabilities of PyHSPF, a Mesonet tool is developed that can extract climate data of all the Mesonet stations within Oklahoma. Currently a baseline watershed model with simulation period from 2019 to 2021 is developed which is further calibrated to simulate Horse Creek Watershed. The HSPEXP hydrology calibration is used to understand and compare the flow statistics between observed and simulated flows. HSPEXP provides the errors in the simulated flow results and helps to improve the watershed simulation. By matching both the observed and simulated flow results, it is expected that the model is performing well. Finally the working watershed model will be used to simulate nutrient and sediment fluxes that might be initiating the HAB formations in Horse Creek and Horse of Grand Lake.

### 4.2.1 Background

Grand Lake has a large watershed with 10,000 square miles of natural grasslands and agricultural lands extending into Oklahoma, Kansas, Arkansas, and Missouri as shown in Figure 4.1. Grand Lake is a primary source of water to several cities, farms, and power plants in northeastern Oklahoma. Also, the location where the Grand Lake is located is a popular tourist destination which provides significant revenue to the local economy. The reservoir experiences regular outbreaks of HABs during summer months that affect the ecosystem provision services locally. Grand Lake is managed by Grand River Dam Authority (GRDA) and monitors the reservoir for HABs regularly using in-situ sensors and field sampling. Unmanned systems including MANUEL, refrigerated auto-sampler, and in-stream monitoring sensors in conjunction with monitoring efforts of GRDA helps in studying HAB formations.

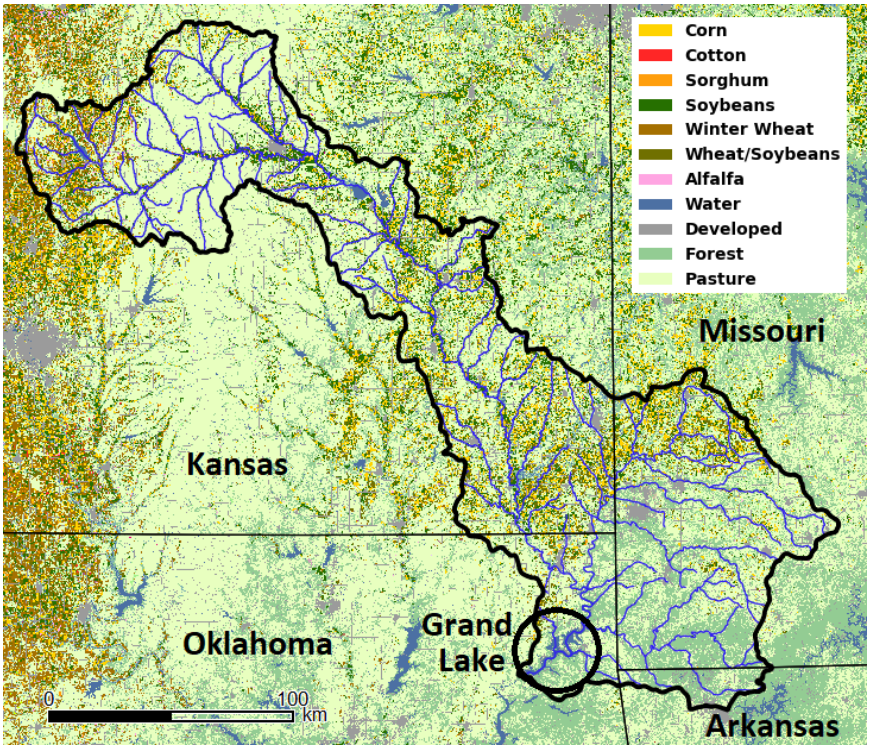


Figure 4.1: Grand Lake watershed

Also the watershed model that is being developed can help the local economy that is negatively affected due to frequent HAB formations in the reservoir due to

summer storms, climatological, and anthropogenic factors. Also, the climate data for the Miami station is available for twenty years combined with in-situ monitoring from GRDA and unmanned systems can provide significant insights in understanding the HAB formations in Grand Lake.

The construction of Pensacola Dam on the Neosho River formed Grand Lake. The primary contributors of water to Grand Lake are Spring and Elk rivers. However there are several small tributaries that feed water to Grand Lake. In the areas around Grand Lake, different land management practices result in varying trophic status. The Horse Creek Watershed has large amount of croplands in the region including corn, soybeans, and wheat as shown in Figure 4.2. The Honey Creek, Sycamore Creek, and Downing Creek watersheds have grasslands. Apart from natural and agricultural lands, wastewater treatment plant in Afton, animal feeding operations, and septic systems are additional sources of nutrients into the surface water systems.

The lack of control in nutrient runoff due to different land management practices resulted in diminished surface water quality in Grand Lake. Based on the previous reports and frequent monitoring from Grand River Dam Authority (GRDA) and Oklahoma Conservation Commission (OCC) both Sycamore Creek and Downing Creek Watersheds are relatively undisturbed and has streams with better water quality compared to streams in Horse Creek Watershed. GRDA has partnered with other state agencies and stakeholders to implement nutrient reduction practices for Honey Creek (Mittelstet et al., 2011). The watershed model in this chapter is simulating the flow that helps to understand sediment, and nutrient dynamics in the Horse Creek Watershed using PyHSPF. The simulated watershed model of Horse Creek can help understand the precursors responsible for HAB formations in Grand Lake.

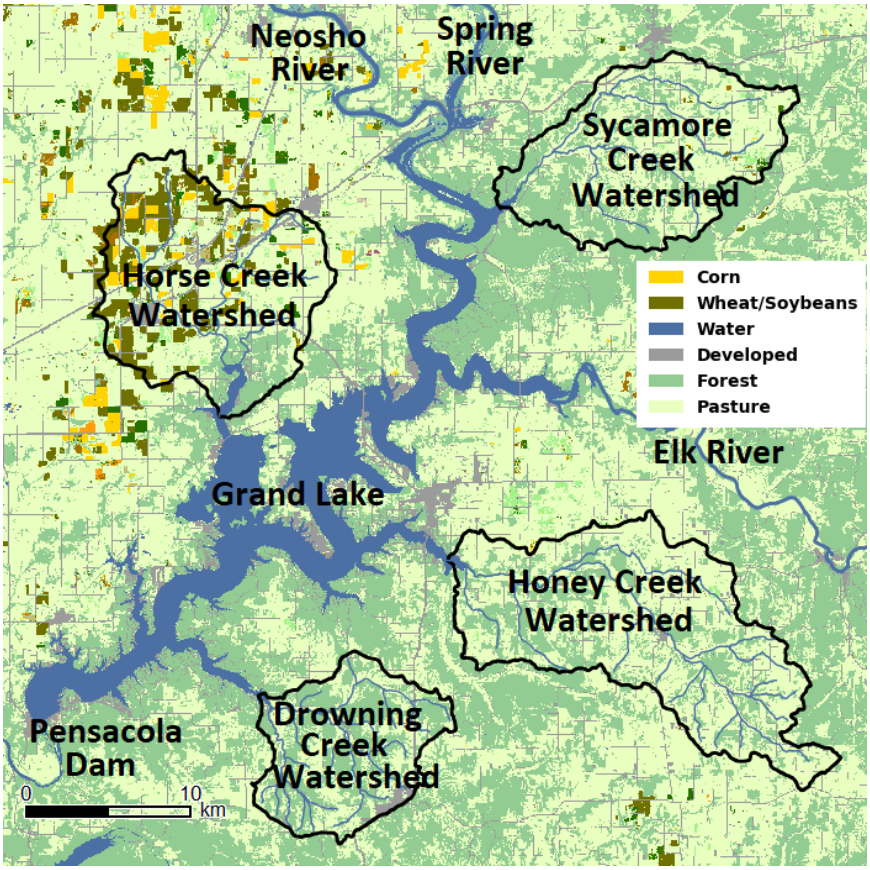


Figure 4.2: Tributaries of Grand Lake

## 4.3 Materials and Methods

### 4.3.1 Water Quality Monitoring in Horse Creek Watershed

The monitoring system in Horse Creek Watershed involves using MeterGroup CTD sensor with turbidity sensors to measure water depth (stage), Temperature, Conductivity, and Temperature. An Arduino based Mayfly board is used to capture the sensor data at a regular intervals of 15 min during the simulation period. The algorithm to program the Mayfly board is made available in appendix. In addition to sensors, a solar powered Teledyne ISCO auto-sampler was used to collect refrigerated samples from Horse creek to measure nutrients including nitrates, dissolved and total phosphorous. The influence of land use and climatology in the Horse Creek Watershed shown in Figure 4.3 can be better understood with a model that simulates runoff

and land management practices.

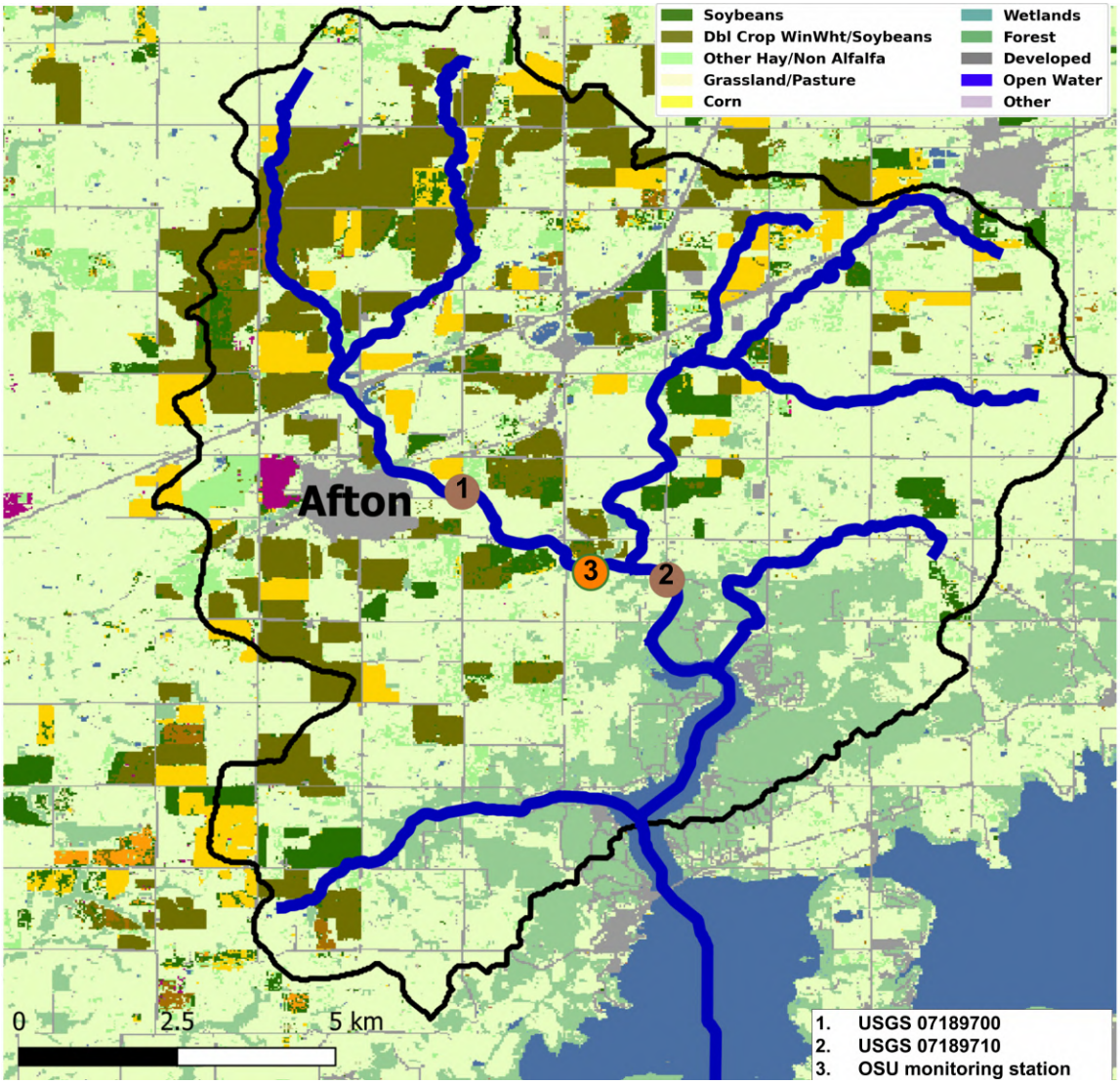


Figure 4.3: Horse Creek Watershed

### 4.3.2 PyHSPF Description and Dependencies

Python is an open-source, interpretative, and high level language for data driven models. PyHSPF use numeric python (NumPy) that builds FORTRAN to python connection as explained in the previous literature. The pre-processor class extracts climate and hydrology information from public datasets. Two input files including User Control Input (UCI) file and Watershed Data Management (WDM) file estab-

lishes communication between the modeler and program. The WDM files contain timeseries information and the UCI files have instructions for the modeler to perform simulations. The watershed information including hydrography information and land use data and a HSPF class which are used to generate input and output UCI and WDM files. The postprocessor and WDMUtil classes take results from watershed simulations and perform statistical analysis. Also the postprocessor plots can generate plots using matplotlib dependency. The process flow diagram of performing a watershed simulation to build a baseline model and calibrate the model is shown in Figure 4.4

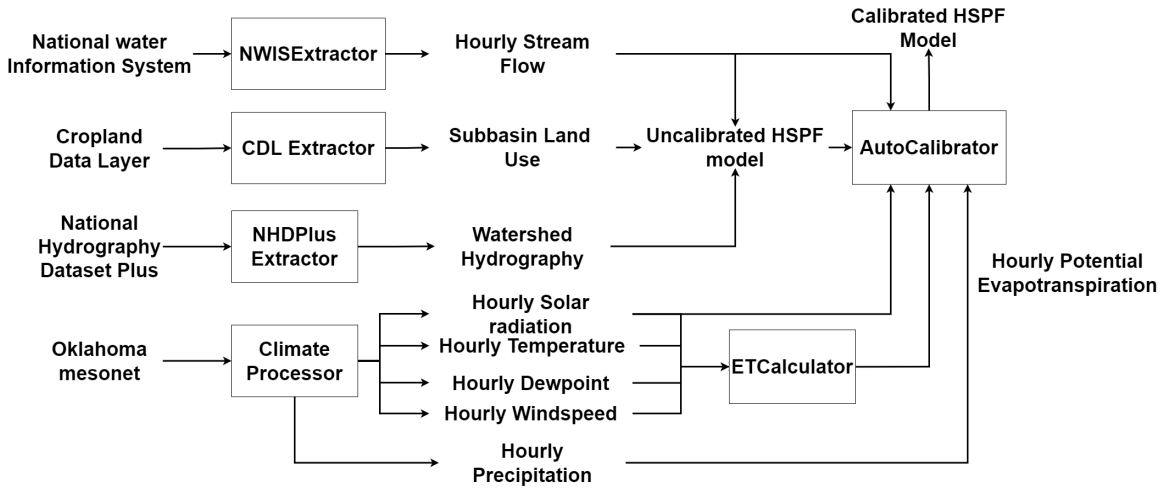


Figure 4.4: Information flow from online sources to calibrated model using PyHSPF (data tools are shown as boxes)

### 4.3.3 Hydrology, Hydrography and Land Use Data Extraction and Processing

Hourly discharge and gauge stage data has been extracted from the USGS National Water Information System (NWIS) website. There are two USGS gauge stations in Horse Creek Watershed such that data from both stations are obtained and processed. The raw data are downloaded and processed as data-frames that can be used to be provided as inputs. The Hydrography and elevation data are obtained from National

Hydrography Dataset Plus (NHDPlus) version 2. The land-use data for Horse Creek watershed is obtained from Cropland Data Layer (CDL). The Flowline and Catchment data in NHDPlus are used to describe the stream connectivity which are explained in detail in Appendix. The land-use categories can be aggregated into several groups based on user preference. For creating a baseline model, three groups (Agriculture, Forest, and Developed) are used initially for the Horse Creek Watershed. The model simulation estimates all the relevant hydrologic parameters based on the input data provided to the HSPF model including slope, reach length, area, latitude, longitude. The CDL data for the year 2021 is used to simulate Horse Creek Watershed. Figure 4.5 gives the information about the land use percentages in the Horse Creek Watershed.

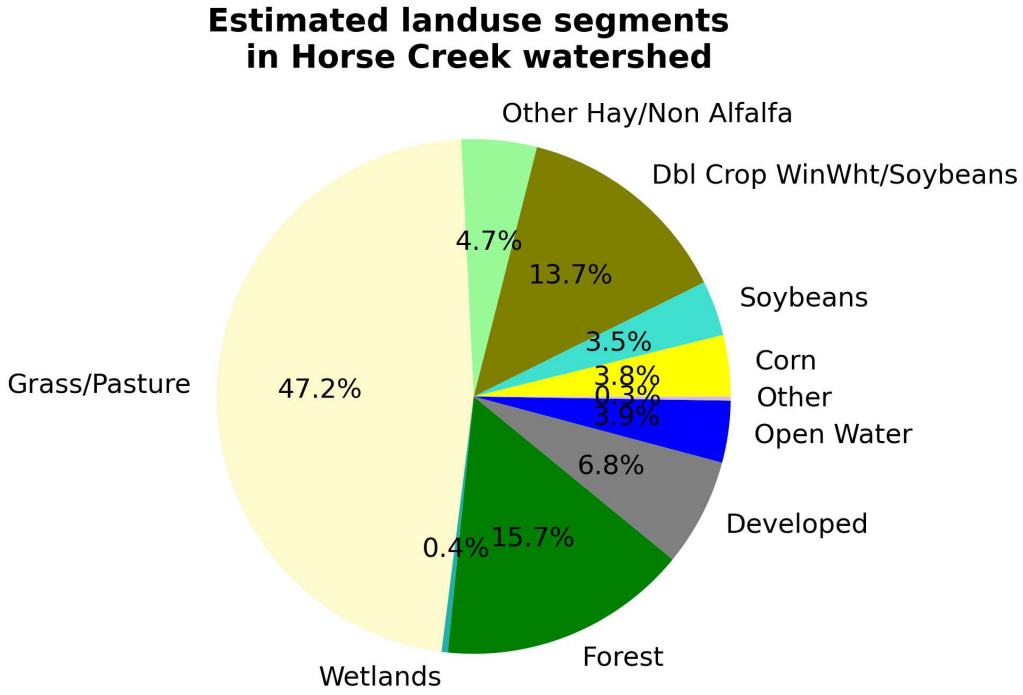


Figure 4.5: Land use in Horse Creek Watershed

### 4.3.4 Climate Data Extraction and Processing

The PyHSPF has classes already that can extract climate information from various sources including National Climate Data Center (NCDC) and National Solar

Radiation Database (NSRDB). New PyHSPF class Mesonet extractor downloads climate data including Hourly Solar Radiation, Hourly Temperature, Hourly Dewpoint, Hourly Windspeed, and Hourly Precipitation from Oklahoma Mesonet. The raw climate data is downloaded from Mesonet site and customized into data structures for further analysis and processing. For Horse Creek Watershed the MIAMI station of Mesonet was selected to obtain the climate timeseries data. The missing climate data from MIAMI station are replaced by the spatially nearest Mesonet station using Mesonet extractor. In HSPF models, precipitation and Potential Evapotranspiration (PET) data are required within the watershed. The climate data was used to estimate the hourly reference Evapotranspiration ($ET_0$) timeseries. The timeseries of $ET_0$ and the other climate parameters are shown in Figure 4.6.

$ET_0$ is the potential evaporation from small grass and is computed using Penman-Monteith equation. Ideally estimates of the $ET_0$ combined with daily crop coefficients generates PET timeseries for each land-use categories within the watershed. The estimates $ET_0$ highly correlates with the observed pan evaporation data (with $r^2 = 0.91$) as shown in Figure 4.7. Daily reference evapotranspiration data for Horse Creek Watershed is estimated from January 2000 to December 2021 where estimation from 2020 and 2021 are shown in Figure 4.6

### 4.3.5 Model Performance and Calibration

The model performance of the watershed model is analyzed using the simulated and observed flows using the root mean square error (RMSE). The RMSE across the n-day period is mathematically represented as follows:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(S_i - O_i)^2} \tag{4.1}$$

where $S_i$ is the simulated flow and $O_i$ is the observed flow. The cumulative observed flow ($\frac{1}{n}\sum_{i=1}^{n} O_i$) can can be used to represent the variance in the daily

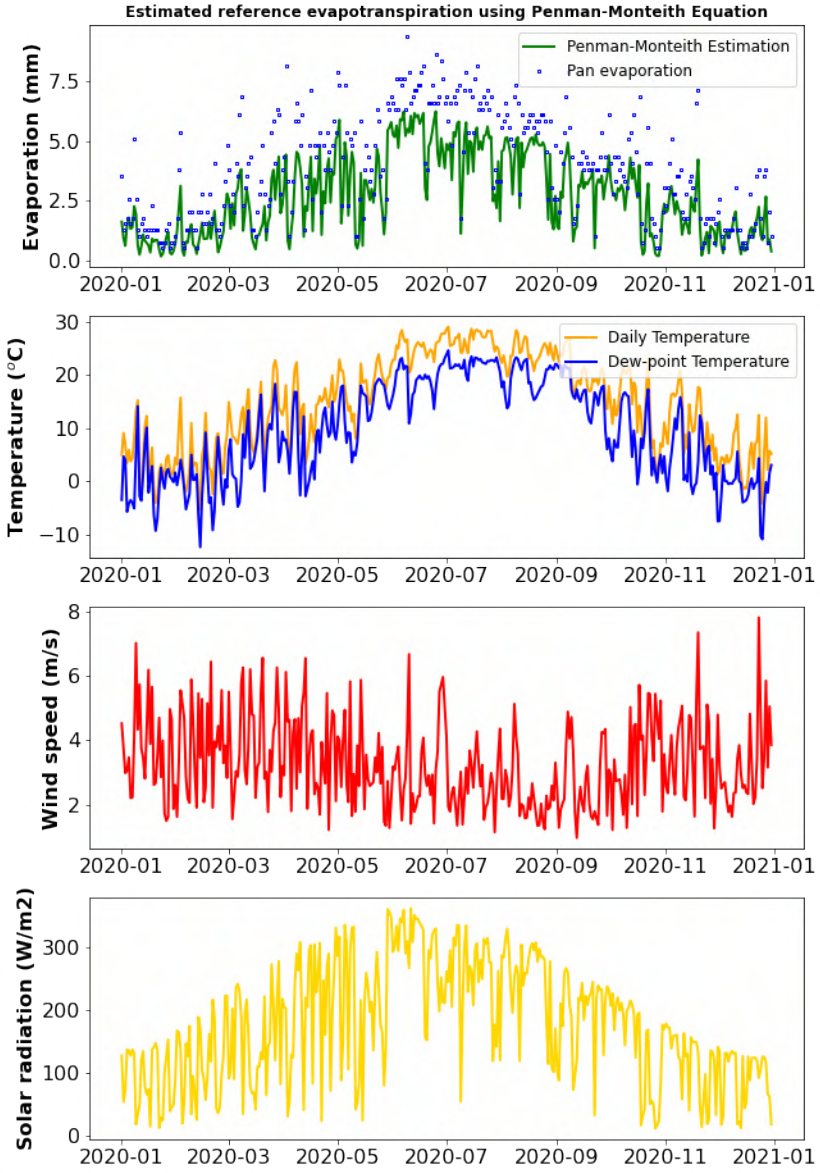


Figure 4.6: Estimates of reference evapotranspiration using hourly Penman-Monteith equation

observed flows. The variance in the daily observed flows is as follows:

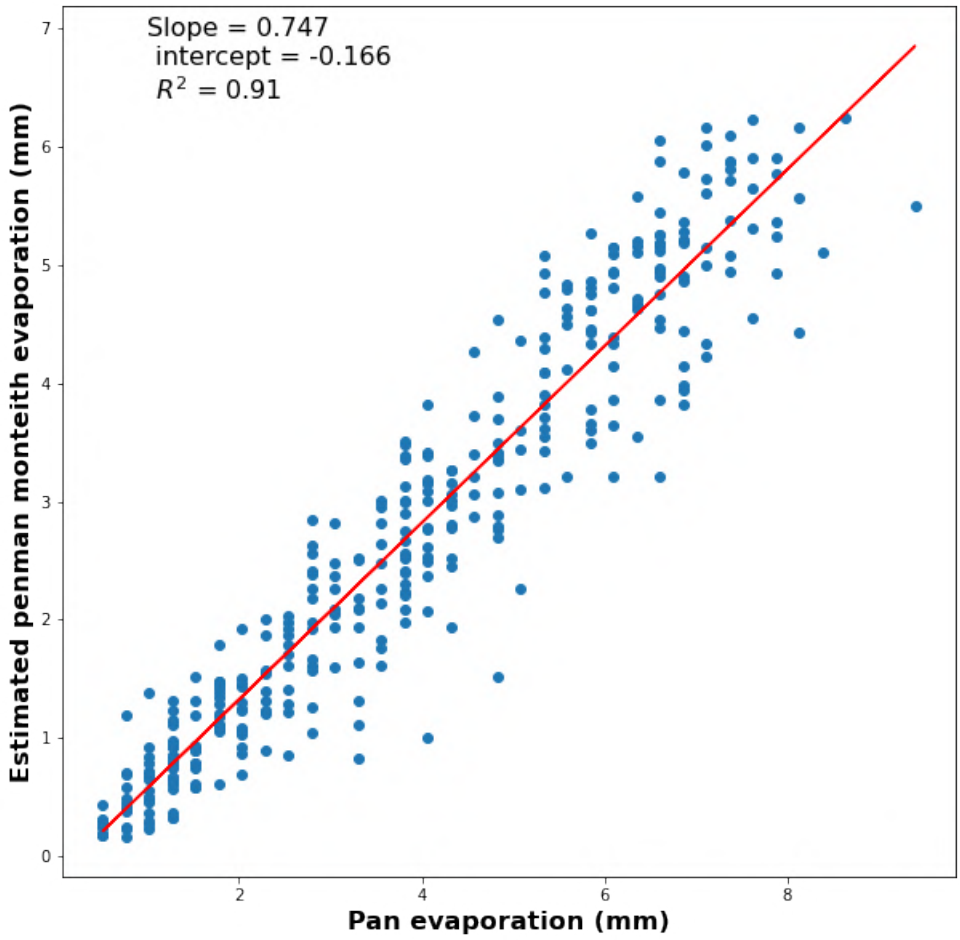


Figure 4.7: Linear fit for the estimated vs observed reference evapotranspiration in Horse Creek Watershed

$$\sigma^2 = \frac{1}{n}\sum_{i=1}^{n}(O_i - \frac{1}{n}\sum_{i=1}^{n}O_i) \tag{4.2}$$

The efficiency of the model is assessed using the Nash and Sutcliffe (NSE) model efficient coefficient. The NSE can be written in terms of RMSE and $\sigma^2$ as follows:

$$NSE = 1 - \frac{\sum_{i=1}^{n}}{\sum_{i=1}^{n}} \tag{4.3}$$

As $\sigma^2$ does not depend on the simulation results, a watershed model which minimizes RMSE will maximize the NSE.

## 4.4 Results and Discussions

The pre-processing steps shown in Figure 4.3 shows the NHDPlus, USGS NWIS, and CDL data to build a watershed model for Horse Creek Watershed. The cross-sectional width of the streams are based on EROM model estimates in NHDPlus. Figure 4.8 shows the non-calibrated fit between the observed flow and simulated flow in Horse Creek Watershed. The storm events captured by the model includes both summer and non-summer storms in the years 2020 and 2021.

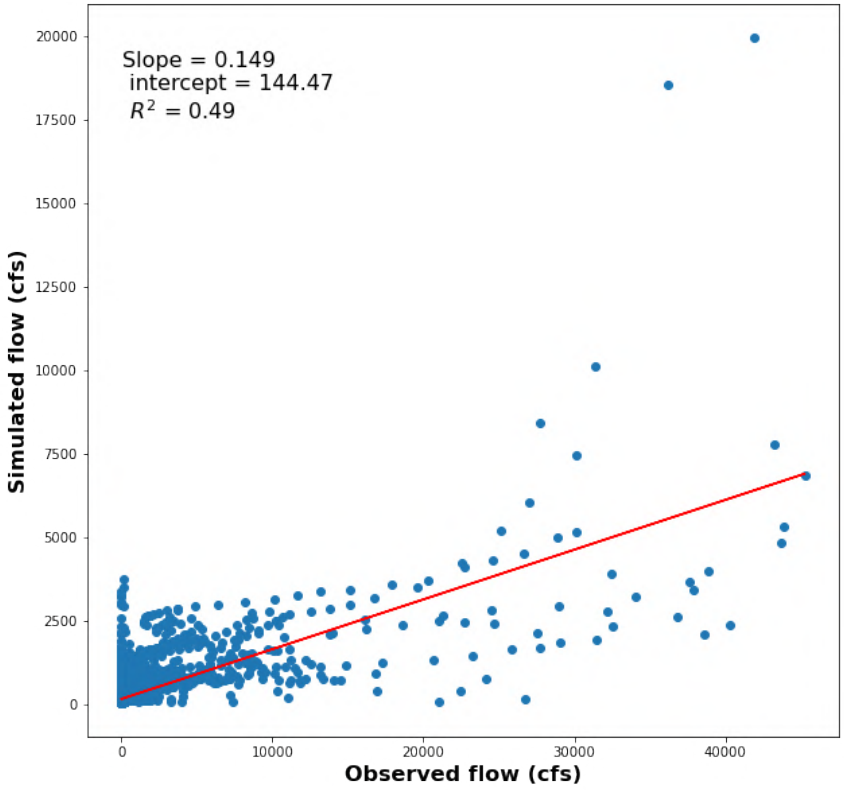


Figure 4.8: Linear fit for the simulated vs observed flows in Horse Creek Watershed (non-calibrated model)

The storms recorded in year 2020 are from August 15, 2020 through September 4,2020 and from October 26, 2020 through November 07, 2020. In the year 2021,

the storms recorded are from May 15, 2021 through June 17, 2021 and June 25, 2021 through July 10, 2021. Several low flow periods during winter months are overestimated due to an absence of snow and ice simulation as shown in Figure 4.9.

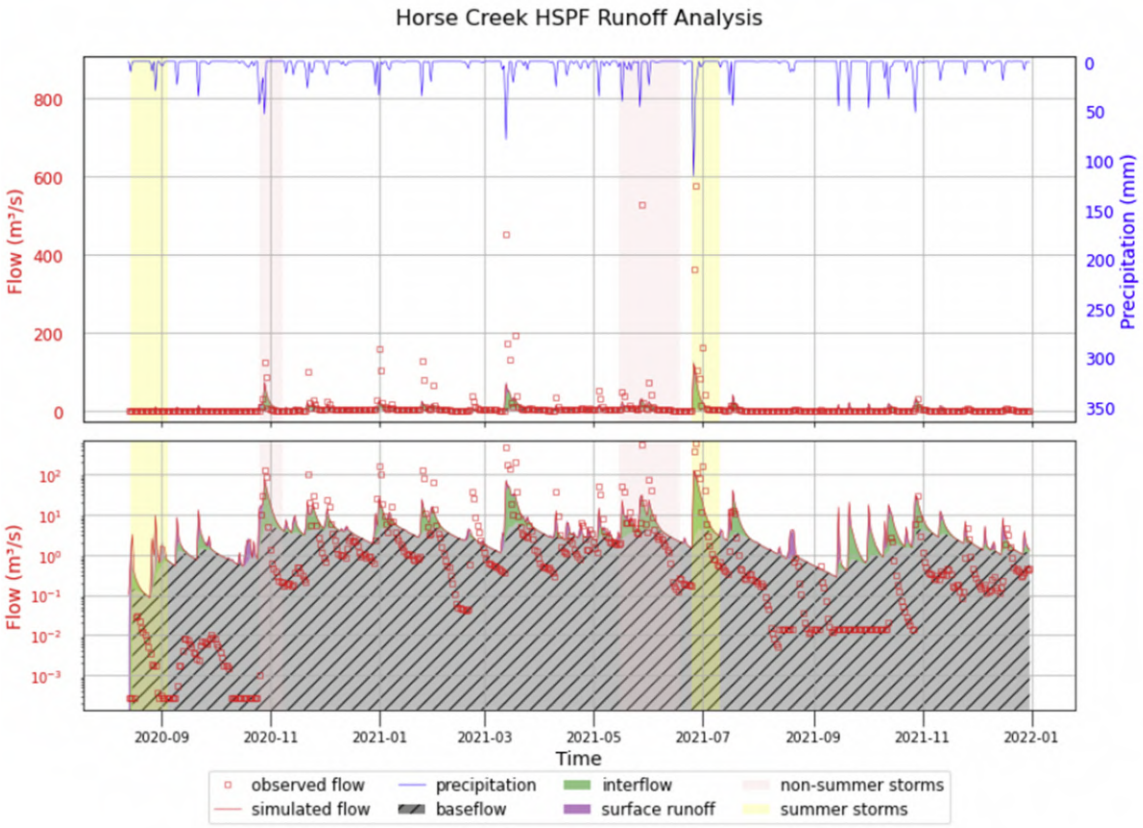


Figure 4.9: Horse Creek Watershed HSPF model non-calibrated Hydrograph

Also to understand the statistics of the simulation a mass balance of water was observed using the postprocessor class of PyHSPF. It was observed that the observed outflow is more than the simulated flow at the Gage station as shown in Table 4.1

To understand how the flow volume is during summer and winter more statistics were estimated using the postprocessor. In both summer and winter simulation, the observed flow is more than the simulated flow as shown in Table 4.2. While the simulated flow is over predicted compared to the observed flow that one of the reason might be that the estimated evaporation data used might be higher than the actual potential evapotranspiration. The potential evpotranspiration of agricultural land-use in Horse Creek Watershed is shown in Figure 4.10

Table 4.1: Mass balance on Horse Creek Watershed during the simulation period

| Parameter | Volume ($Mm^3$) |
|---|---|
| Initial watershed volume | 68 |
| Final watershed volume | 66 |
| Total precipitation volume | 681 |
| Total simulated evapotranspiration | 432 |
| Total simulated outflow | 252 |
| Total groundwater recharge | 0 |
| Total potential evapotranspiration | 554 |
| Total observed outflow | 422 |

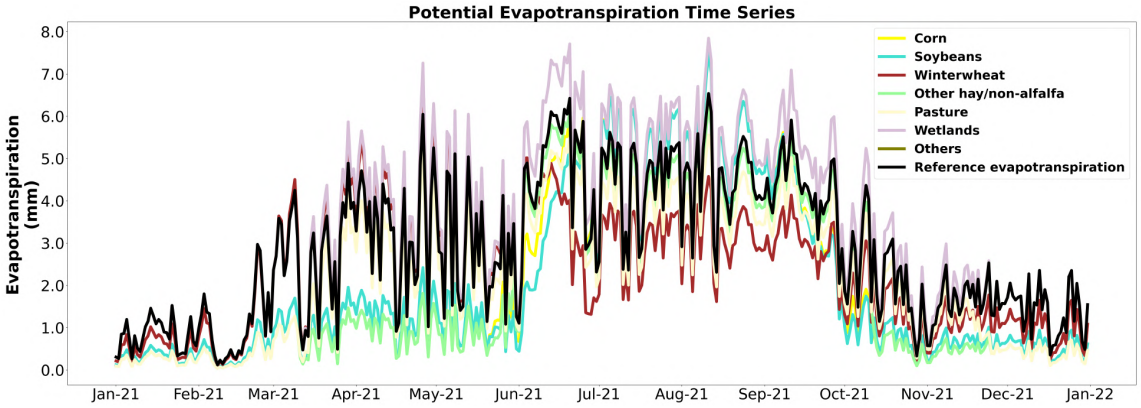


Figure 4.10: Potential evapotranspiration of agricultural crops in Horse Creek Watershed

Table 4.2: HSPF calibration statistics

| Parameter | Observed | Simulated |
|---|---|---|
| Total Precipitation (mm/yr)) | 1200.5 | 1200.5 |
| Total runoff (mm/yr)) | 743.3 | 444.8 |
| Total of highest 10 % flows | 588.7 | 204.5 |
| Total of lowest 50 % flows | 4.5 | 57.6 |
| Total storm volume (mm) | 514.3 | 210.1 |
| Average of storm peaks ($m^3/s$ | 307.3 | 57.6 |
| Baseflow recession rate | 0.923 | 0.951 |
| Summer flow volume (mm/yr) | 233.9 | 114.5 |
| Winter flow volume (mm/yr) | 133.5 | 83.3 |

In addition to evapotranspiration, the clay sediment of Grand Lake might not infiltrate as much water as other soil sediments. Therefore reducing the infiltration rates will bring the simulated flows close to the observed flows.

## 4.5 Conclusions

The simulated watershed model of Horse Creek helps in estimating sediment and nutrient fluxes into the Horse Creek cove from upstream in the watershed. Historical climate data, land-use data with calibrated simulated flows can capture long-term impacts of land-use and climate change affecting sediment and nutrient dynamics within the watershed. The flexible and automated approach of PyHSPF can inform the impact of land use changes associated and climate change on hydrology and water quality within the watershed. Understanding both the nutrient and sediment dynamics within the Horse Creek Watershed provides significant insights on the reasons for HAB formation in Grand Lake.

# CHAPTER V

## Water Quality Monitoring using Unmanned Systems

### 5.1 Abstract

Land use changes, climate change are causing rapid changes to both water quantity and quality. The water quality and ecological status of surface water systems are judged based on physical sampling or single point in-situ measurements using sensors. Due to rapid changes in surface water quality especially due to formation of HABs in freshwater ecosystems, dynamic monitoring technologies are of a critical need. This chapter provides the outcomes of using an unmanned surface vehicle that is used for collect water quality data in Grand Lake, Oklahoma. From the water quality data collected, it is possible to map shallow areas of lake to measure algal pigments including Chlorophyll-a, Phycocyanin and other water quality parameters. This chapter identifies the opportunities of using unmanned vehicles for applications to map water quality in inaccessible areas where humans cannot get in contact with surface water bodies due to the toxic environment especially during Harmful Algal Blooms.

Keywords: Water quality monitoring, Unmanned surface vehicles, Water quality mapping, Water quantity

### 5.2 Introduction

Monitoring HABs is a global concern because of the increase in occurrences of HAB events due to anthropogenic sources of nutrients and climate change. Nutrient pollution has raised the growth of algal biomass by 60% compared to previous conditions

(Schoumans et al., 2014). The limiting nutrients of HAB formations are often either nitrogen or phosphorous. Based on literature and recent research indicates that phosphorous limits HAB growth during summer and nitrogen limits growth during spring. A study conducted at Grand Lake reported that phosphorous released from sediment created nitrogen limiting conditions after thermocline erosion. Therefore to better understand the HAB formations in Grand Lake a monitoring system was used to better understand the HAB formations and control strategies at Grand Lake. It is anticipated that these control strategies can be transferred to other locations. In addition to nutrient influxes, sediment re-suspension is another important phenomenon that can release resting cysts which causes HABs as shown in Figure 5.1. Therefore it is important to understand the nutrient sources and sediment dynamics in coves which are hydrologically isolated during growing season. Apart from HAB proliferation, increasing concentration of suspended solids in surface water have negative impacts on navigation and reduce reservoir capacity. Excessive sedimentation reduces the navigational capacity of surface water bodies and require dredging for restoring navigable channel depths. Also the presence of suspended solids will increase the cost of drinking water due to excessive energy and material used by water treatment plants. Aquatic ecosystems and fish habitats are negatively affected by high concentration of suspended solids in fresh water systems.

Besides suspended solids, heavy metals transported by sediment can increase the levels of bio-accumulation in aquatic species causing detrimental health effects. Besides toxic hydrophobic organic compounds are released into surface water bodies which pose an environmental risk. Therefore it is very important to monitor the surface water quality. Finally increase in human interaction with surface water bodies in conjunction with climate change creates the need to acquire large amounts of data to understand the highly dynamic aquatic environment paving the way for unmanned system to collect high resolution datasets.

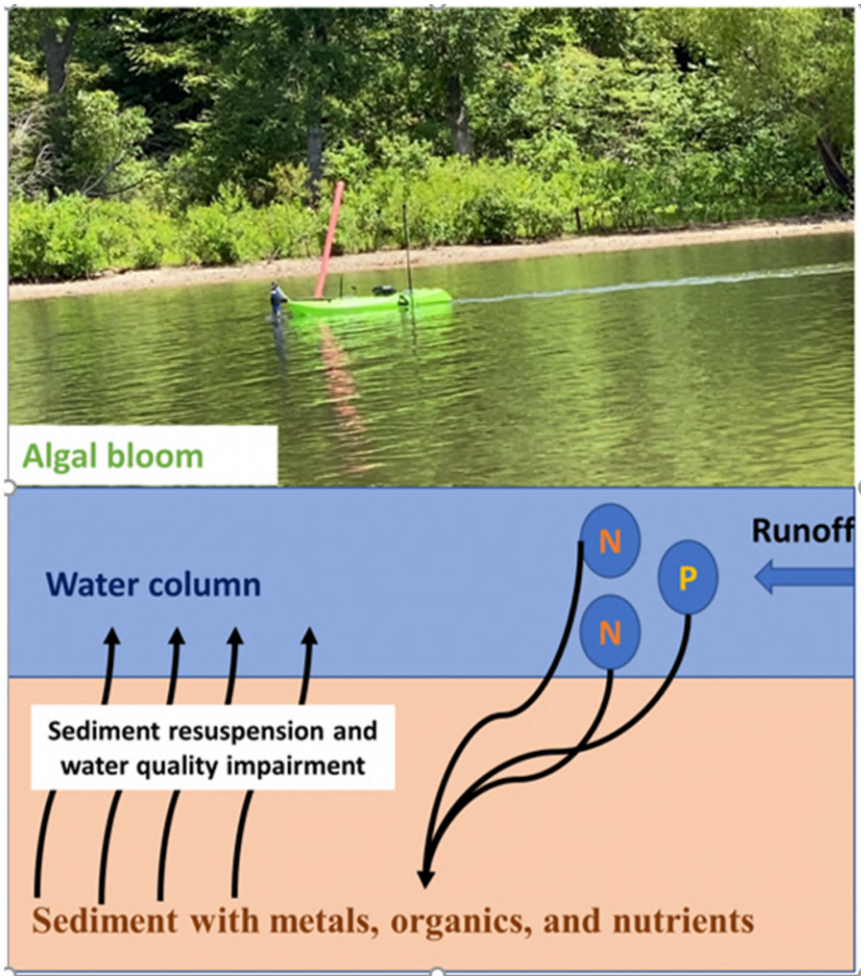


Figure 5.1: Schematic diagram showing the interaction between sediment and water column

## 5.3 Materials and Methods

The sensors on board MANUEL include a ping sonar to collect bathymetry data and a bundled sensor module of YSI EXO-3 sonde. The sensors of EXO-3 sonde have the capability to collect water quality data at a rate of 1 HZ. All of the sensors on board MANUEL requires frequent calibration which is explained in detail further:

### 5.3.1 Ping Sonar Calibration

An open-source GUI, Ping-viewer was used to calibrate the ping sonar. It was known the speed of sound in fresh and salt water is 1435, and 1500 m/s respec-

tively (O'Donnell et al., 1981). The ping sonar is calibrated in freshwater and the speed is set at 1435 m/s because the lakes monitored in Oklahoma are freshwater lakes. When it comes to error, ping sonar has an error percentage of 0.5.

### 5.3.2 EXO-3 Sonde Calibration

The sensor bundle in EXO-3 sonde are calibrated separately using the calibration cup. Before calibration, all the sensors are rinsed with de-ionized water three times and air dried. The clean, dry sensor probes are placed in the calibration cup for calibration. The detailed calibration process of each sensor in EXO-3 sonde is as follows:

1. **pH Calibration**

   The pH sensor was 3-point calibrated using three calibration standards. The calibration standards include pH buffers with pH at 4, 7, and 10. The 3 point calibration ensures maximum accuracy when the pH of the monitoring media cannot be anticipated.

2. **Conductivity Calibration** The conductivity sensor can measure conductivity, specific conductance, salinity, non-linear function conductivity, Total Dissolved Solids (TDS), resistivity, and density. Calibration of the conductivity sensor is available only for specific conductance, conductivity, and salinity. Calibration of one of the specific conductance or conductivity will calibrate the other parameters. YSI recommends calibrating specific conductance because of the ease of use and accuracy. For calibration, a pre-rinsed EXO calibration cup was filled with a conductivity standard of 1000 $\mu$S/cm up to the second marked line. The air dried probes are inserted into the calibration cup and a calibration standard of 1000 is entered. After calibration the QC score is checked to make sure that the calibration is within the factory limits.

3. **Temperature Calibration** The temperature sensor cannot be calibrated by

the user based on the information provided by the user manual of YSI. However, it was mentioned that calibrating conductivity sensor using specific conductance will calibrate the temperature and conductivity parameters. For testing purposes, the temperature sensor was tested with a thermometer periodically before deployment in the field. This periodical check is suggested by the YSI as a best management practice

4. **Turbidity Calibration** The air dried turbidity sensor is rinsed with 0 FNU standard 2 to 3 times before calibration. The calibration cup was filled with 0 FNU standard and the sensors are immersed into the calibration cup. The KOR software is used to enter 0 and calibrate the turbidity sensor. Similarly using two more calibration standards including 124 FNU and 1010 FNU a three point calibration is performed. After performing the 3-point calibration, the turbidity sensor was checked using the QC score to ensure that the sensor is measuring within the factory limits.

5. **Dissolved Oxygen Calibration** Based on the YSI recommendation there are two ways to calibrate the dissolved oxygen sensor. The two ways involve using either water-saturated air or air-saturated water for calibration. For monitoring lakes, the DO sensor was calibrated using water-saturated air where the air dried sensor is placed into the calibration cup about 1/8 inch of water. The calibration cup to the sonde is not sealed during calibration and there is a wait time of 10 to 15 minutes for temperature and oxygen pressure to reach an equilibrium before calibration. In the calibration menu, the current barometric pressure in mm of Hg is entered for DO sensor calibration. Finally the QC score is checked to ensure that the DO readings are within the factory limits.

6. **Total Algae Calibration** YSI provides flexibility for users to prepare their own calibration standards. However, Rhodamine WT is a secondary standard

that can be used to calibrate the total algae sensors that measures Chlorophyll-a and Phycocyanin. The concentration of the Rhodamine dye used is 625 $\mu$g/L. As far as the calibration is concerned, a 2-point calibration is performed with de-ionized water where the the value is set to zero and then with the Rhodamine dye with know fluorescence value of 66. After calibration the QC score is checked to ensure that the algal pigment concentrations are within the factory limits.

### 5.3.3 Data Collection and Processing

An Unmanned surface vehicle (USV) named MANUEL is developed for monitoring lakes and freshwater bodies. The automation system used in MANUEL is an orange cube Pixhwak flight controller with GPS telemetry feedback. The prototype is designed to be compact and water resistant which is explained in detail in literature review section 2.2. The logging system hardware consists of a Teensy 3.6 USB development board with integrated SD card reader, a Venus GPS module with an ANT-555 GPS antenna, and a waterproof DS18B20 temperature sensor. MANUEL is outfitted with a YSI EXO3 Sonde to continuously gather water quality data and a ping sonar to estimate depth. All the data collected from ping sonar and EXO sonde are merged using a MATLAB algorithm developed by the Unmanned Systems Research Institute (USRI). The depth data and water quality data collected by the sonar and EXO-3 sonde is geo-tagged using the GPS coordinates from Pixhawk. The depth data is collected with a frequency of 2 HZ and the water quality data is collected with a frequency of 1 HZ. The merged depth and water quality data is read using open source python extension libraries that can read the way points and plot them as shown in Figure 5.2

## 5.4 Results and Discussions

### 5.4.1 Bathymetry Mapping

A hydrographic survey of Grand Lake was performed in the year 2009 by Oklahoma Water Resources Board (OWRB). Using MANUEL, another hydrographic survey was performed during summer 2020 in Horse Creek Cove to test the capability of the USV in gaining important insights into sediment accumulation. Following Hydrographic survey the observed bathymetry at the individual points are linearly interpolated to create a Digital Elevation Model (DEM) of the bottom of the reservoir as shown in Figure 5.2.

The linear interpolation between the individual observations of depth is performed using the scientific python module which is free and open source.

The DEM can be used to find the accumulation of sediment since 2009 until summer 2020. The lake level used is 226.5 and the LIDAR-based DEM is combined with the survey results to fill the elevations in the inundated areas as shown in Figure 5.2. Sedimentation is assessed pixel by pixel basis such that the raster grid of LIDAR is 2 x 2 m. The data from the hydrographic survey are combined with results shown in Figure 5.2 to estimate the changes in hydrography of the reservoir associated with sedimentation over the period of 2009 to 2020. The difference in DEM resulted in estimating sediment deposition of 42,569 $m^3$. This indicates that there is 42,569 $m^3$ loss of storage capacity in Horse Creek Cove due to sediment deposition. The DEM difference can be observed in Figure 5.3 which shows the excess sediment deposition and dredging close to the dock areas in the cove.

### 5.4.2 Water Quality Mapping

In summer 2020, collecting water quality data in Horse creek cove was successful with MANUEL but no significant HAB development happened. However, there are certain

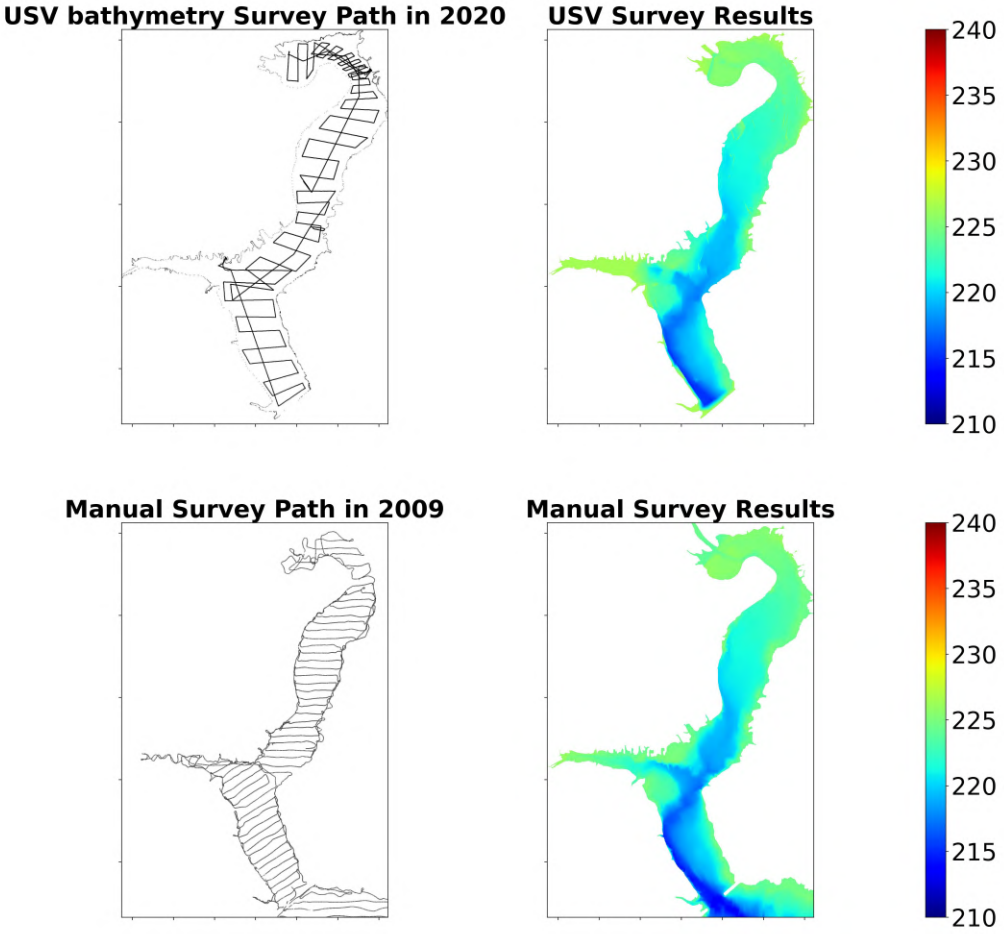


Figure 5.2: Hydrographic survey of Horse Creek Cove by USV (top) compared with previous manual survey from 2009 (bottom)

hot-spots of Chlorophyll-a and Phycocyanin pigments as shown in Figure 5.4

The presence of an algal blooms happening in the Horse Creek Cove can be interpreted by using other water quality parameters such as increasing levels of dissolved oxygen and pH. As the green algal biomass performs photosynthesis in the presence of sunlight, there is increase in dissolve oxygen concentration. Similarly increase levels of algal biomass create basic pH during a bloom event. The sampling process was performed in the afternoon and therefore the temperature was significantly higher

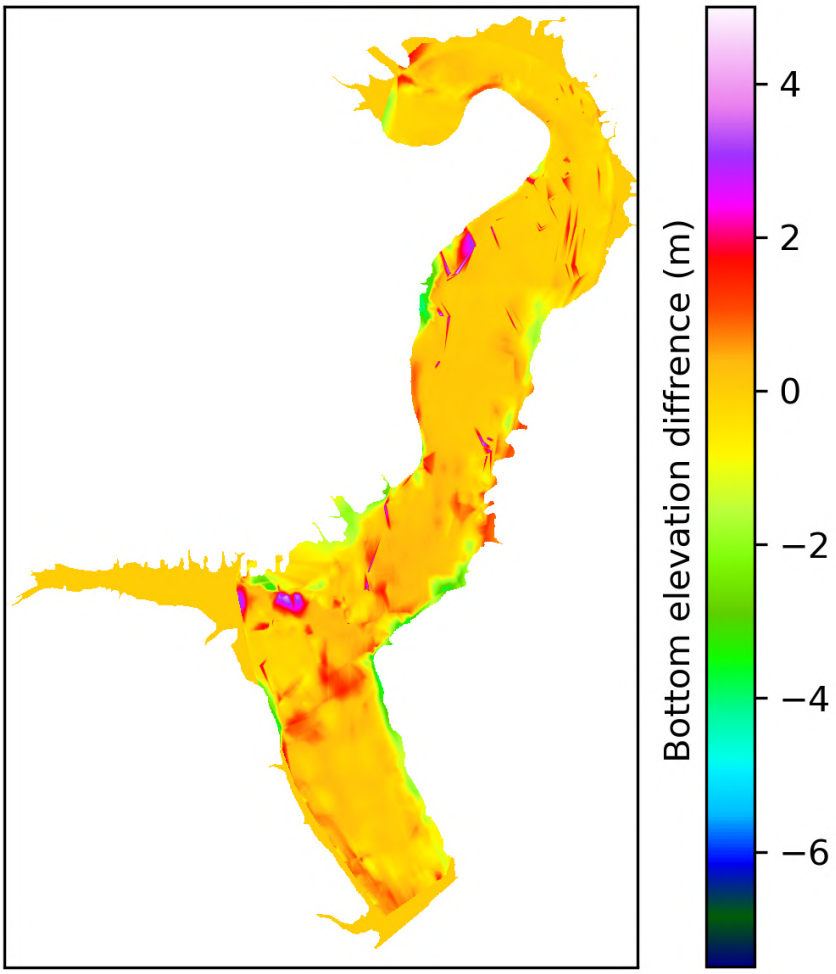


Figure 5.3: Bottom elevation difference of Horse Creek Cove since the manned hydrographic survey of 2009 until summer 2020

than 30 $^oC$. The movement of MANUEL and placement of EXO-3 sonde was placed strategically such that the water quality data is not affected during rapid data collection. This reliability of data collection was previously explored in Oklahoma State University (Jdiobe, 2020).

### 5.4.3 Artifacts in Data

In summer 2021, the focus is to capture a Harmful Algal Bloom (HAB) in Grand Lake. We were able to monitor the study areas shown in Figure 5.5. To avoid the EXO-3 sonde being dragged on the lake bottom and effectively monitor shallow regions, the

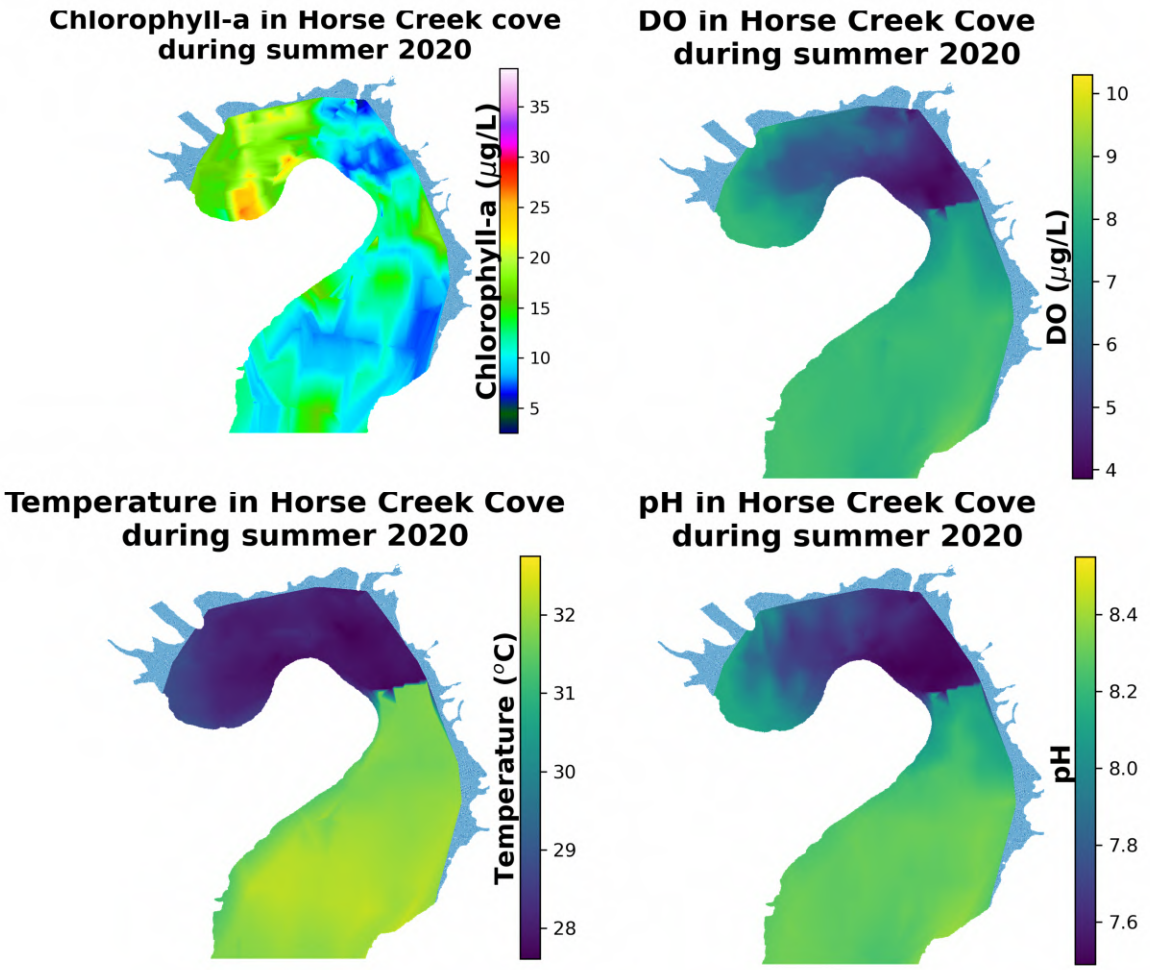


Figure 5.4: Water quality observations mapped as gridded estimates from summer sampling of 2020 in Horse Creek Cove

position of the EXO-3 sonde was moved up by 1 feet. This change in sonde position has created artifacts in the water quality data collection. By revisiting the collected water quality and plotting, it was found that the choppy water during sampling might have made the probes lose contact with water surface causing noise in data collection. Therefore it is very important to keep the position of the sonde at least a foot below the surface of water for smooth interpolation estimates of water quality especially algal pigments including Chlorophyll-a and Phycocyanin. Based on further testing of EXO-3 sonde and data analysis, it was found that the artifacts in the gridded plots are due to the position of the EXO-3 sonde on-board MANUEL which was supposed

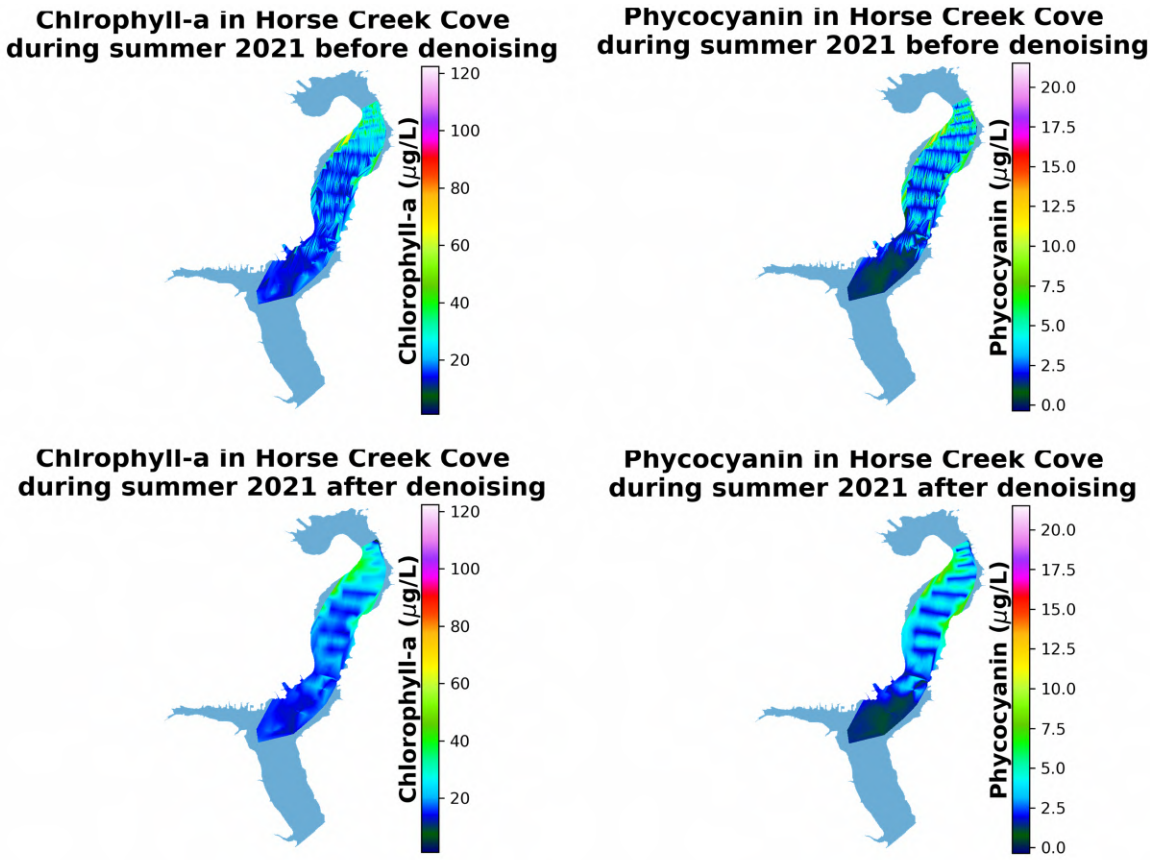


Figure 5.5: Artifacts in gridded estimates corrected using Scientific python digital filter

to be deeper and the USV needs to be programmed to collect more data points near the shores. After diving deeper into data cleaning, the noise in the data has been reduced using scientific python digital filter. However the resultant interpolation cannot be considered as the ground truth of algal pigment concentrations during a HAB outbreak in Grand Lake.

### 5.4.4 Other Water Quality Parameters

The water quality data collection using the EXO-3 sonde probes barely affected the interpolation plots of other water quality parameters. Although several artifacts can be seen in case of mapping algal pigments, water quality parameters including dissolved oxygen, temperature, pH, and turbidity are not affected by the sonde position.

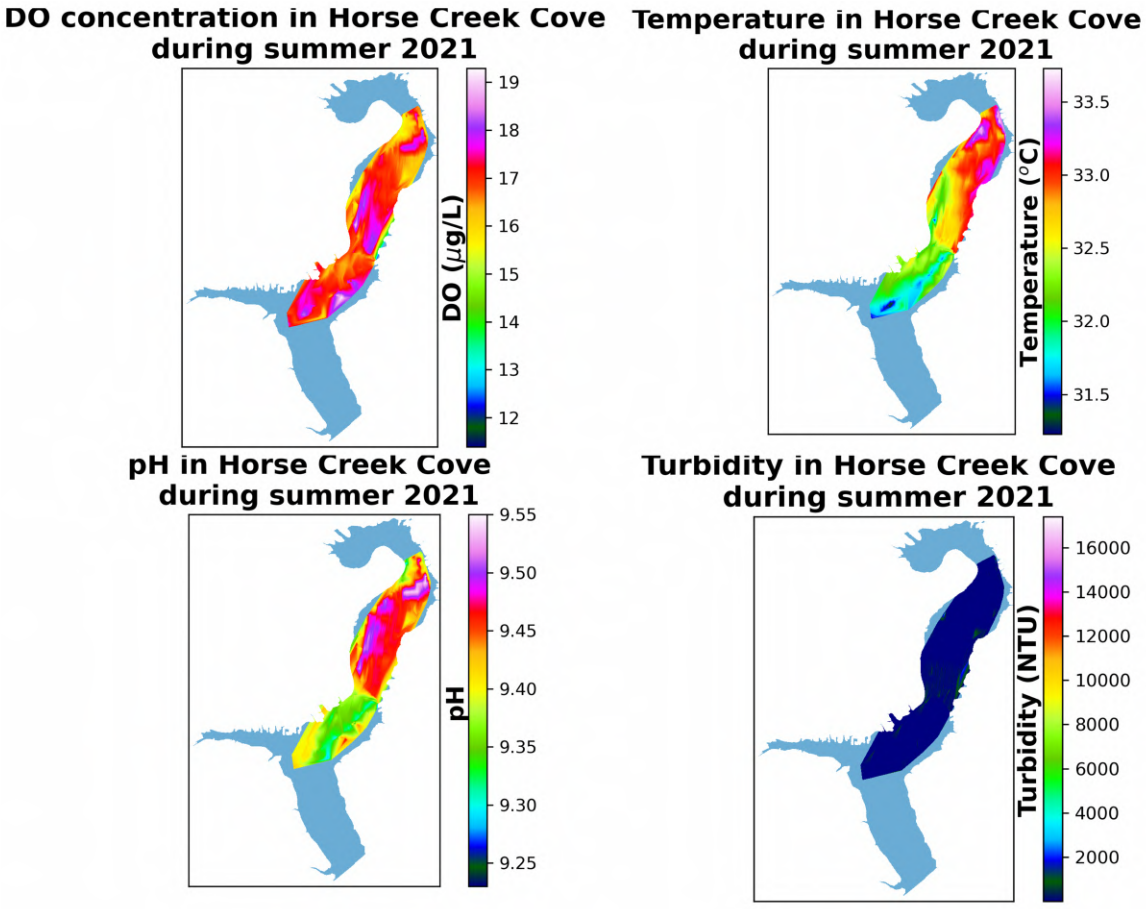


Figure 5.6: Gridded plots of other water quality parameters

This indicates that the dynamic concentrations of algae pigment concentrations that is needed to be captured with excessive data points which can be mapped and considered as ground truth. The high resolution in-situ data collected can help in testing the developed remote sensing tools for model calibration and produce efficient water quality models.

## 5.5 Conclusions

The data collected from USV provides the ground truth of water quality in Grand Lake. The gridded water quality estimates show the capability of USV for rapid data collection without affecting the reliability of data collection process. The bathymetry data collected by the USV provides important insights into the sediment accumulation

from 2009 to 2020 and loss of reservoir capacity. In addition to the water quantity, the coincident satellite flyovers in conjunction with in-situ data and multispectral aerial imagery can help in building in-lake water quality models.

# CHAPTER VI

## Additional work: Well-to-wake Analysis of Switchgrass to Jet Fuel Via a Novel Co-fermentation of Sugars and CO2

### 6.1 Abstract

Lignocellulosic biomass such as switchgrass can be converted to n-butanol using fermentation, which can be further processed into jet fuel. Traditional acetone-butanol-ethanol (ABE) fermentation only converts sugars derived from switchgrass to ABE. Novel co-fermentation processes convert sugars and gas ($CO_2/H_2$) produced during fermentation into butanol, thus increasing ABE yields by 15.5% compared to traditional ABE fermentation. Herein, the environmental impact of a Switchgrass to Jet Fuel (STJ) pathway was assessed using life cycle assessment (LCA) from well-to-wake. LCAs were performed for greenhouse gas (GHG) emissions from jet fuel production via co-fermentation of sugars and gas for ideal and practical cases of ABE fermentation and seven other jet fuel pathways. The ideal case assumes 100% sugar recovery and 95% ABE yield. The practical case assumes 90% sugar recovery and an 80% ABE yield. Results are presented based on 100-year global warming potential (GWP) per MJ of jet fuel. Co-products were allocated using various methods. The increase in butanol yield via the co-fermentation technology reduced GWP-100 for the STJ pathway by 6.5% compared to traditional ABE fermentation. Similarly, the STJ pathway for the practical case with co-fermentation had 14.2%, 47.5%, 73.8%, 44.4% less GWP-100 compared to HRJ, Fischer-Tropsch jet fuel from switchgrass, Fischer-Tropsch jet fuel from coal, and conventional petroleum jet fuel. The results demonstrate that the STJ pathway via co-fermentation has the potential to increase

product yield while reducing GHG emissions compared to other jet fuel production pathways.

Keywords: Life cycle assessment, Global warming potential, Co-fermentation of sugars and gas, Switchgrass, Jet fuel, Greenhouse gas emissions.

## 6.2 Introduction

Jet fuel consumption in the USA was 3 trillion MJ in 2015 and is projected to increase by 40% in 2040 (Conti et al., 2016). Fossil-based jet fuel consumption releases greenhouse gases (GHGs) that contribute to climate change (Barbir et al., 1990). In 2005, aviation fuel usage contributed 1.6% of the total anthropogenic radiative forcing in the atmosphere (Scheelhaase et al., 2018). Aviation accounted for 12% of US transportation energy in 2019, and renewable feedstocks such as lignocellulosic biomass (LB) are expected to make up an increased share of jet fuels in the future (Administration, 2019). Soybeans currently provide most of the feedstock for bio-based jet fuels, but soy production is unlikely to meet future demands for liquid fuels (Scaldaferri and Pasa, 2019; Staples et al., 2014; Zarchin et al., 2015). To meet growing demands and increase sustainability, alternative jet fuel production pathways are needed.

LB like switchgrass provides a potentially more renewable source of carbon and energy to produce jet fuel (Budsberg et al., 2016). Switchgrass is a perennial warm-season crop that has the potential to produce biofuels at a commercial scale with enhanced environmental benefits while potentially avoiding land use change (McLaughlin and Kszos, 2005; Sanderson et al., 2006; Sokhansanj et al., 2009). Switchgrass can be grown on marginal land in the US using dryland farming techniques, where it is expected to have limited effects on food supplies (Arima et al., 2011; Fischer et al., 2010; Graham-Rowe, 2011; Moriarty et al., 2018; Tang et al., 2010; Valentine et al., 2012; Varvel et al., 2008). LB can be used to generate butanol, which can then be

readily converted into jet fuel. Butanol can be produced using traditional acetone-butanol-ethanol (ABE) fermentation of LB including switchgrass, crop stubble, and wood/forest residue, but the conversion efficiencies from these feedstocks are limited (Ezeji et al., 2007; Liu et al., 2015a; Ni and Sun, 2009; Sauer, 2016). A new process has been developed that uses co-fermentation of sugars and off-gases ($CO_2/H_2$) to increase ABE yields from LB, which improves commercial process viability and provides potential environmental benefits (Atiyeh, 2021)). To assess the environmental benefits of this novel switchgrass to jet fuel (STJ) pathway, a life cycle assessment (LCA) of the GHG emissions is required.

Conventional and emerging pathways to produce jet fuel include corn to jet (CTJ), Fischer-Tropsch jet fuel (FTJ) from LB and coal, hydro processed renewable jet fuel (HRJ) from oils of algae, palm, jatropha, soybean, rapeseed, camelina and conventional petroleum jet fuel (CPJ) (Enright, 2011; Liu et al., 2015a,b). Many of these alternative pathways have been investigated using LCA (Bailis and Baka, 2010; Cox et al., 2014; Elgowainy et al., 2012; Han et al., 2013a; Shonnard et al., 2010; Staples et al., 2014; Stratton et al., 2011). However, there are no previous LCA studies for the proposed switchgrass to jet fuel (STJ) pathway.

The goal of this study was to estimate the life cycle GHG emissions from the STJ pathway using the material and energy inputs and compare it with other alternative pathways. Data used to assess the jet fuel pathways were obtained from recent laboratory studies, literature and Greenhouse gases, Regulated Emissions, and Energy Use in Transportation (GREET) model. The novel STJ pathway via co-fermentation was compared with STJ via traditional ABE and other well-characterized pathways, including CTJ via ABE fermentation, HRJ from soybean oil, FTJ from coal, FTJ from LB, and CPJ from crude oil. A sensitivity analysis was performed to examine the effects of ABE yield, conversion technologies, and other key assumptions for the STJ pathway via the co-fermentation route.

## 6.3 Methods

### 6.3.1 Goal, Scope, and Assumptions

LCA analysis was conducted as per ISO 14044 guidelines. The life cycle emissions and associated global warming potential (GWP) for the new STJ pathway were compared with other conventional and emerging alternatives. For the STJ pathway, both a "practical" and an "ideal" case were considered. The practical case assumes 90% sugar yield from hydrolysis and 80% yield from ABE fermentation (Gao et al., 2014; Liu et al., 2015a,b; Sun et al., 2020). The ideal case assumes 100% sugar and 95% ABE of theoretical yields based on stoichiometry. The functional unit used for the analysis was 1 MJ of jet fuel, since the combustion efficiency was assumed to be similar for all the target fuels. All results were presented and compared using 100-year Global Warming Potential (GWP-100) expressed as $CO_2$ equivalent ($CO_2e$) based on the data obtained from the 5th Assessment Report of the Intergovernmental Panel on Climate Change (IPCC) (IPCC, 2014).

### 6.3.2 System Boundary

The LCA encompasses processes from well-to-wake (WTWa), including feedstock production, conversion, and fuel combustion. Jet fuel production pathways evaluated include STJ (via co-fermentation and traditional ABE fermentation), CTJ (biochemical pathway), FTJ (coal and switchgrass gasification) pathway, hydro-processed renewable jet HRJ (soybean bio-oil) pathway, and CPJ from crude oil. The system boundary of biofuel pathways (STJ, CTJ, and HRJ) includes farming, conversion, and combustion (Figure 6.1). These pathways generate many co-products such as electricity, acetone, ethanol, distiller's dried grains with solubles (DDGS), gasoline, diesel, soybean meal, propane, and naphtha that require allocation. The alternative jet fuel pathways have been analyzed previously, including the effects of different co-

product allocation methods except for the new STJ pathway (Alves et al., 2017; Han et al., 2013b; Pierobon et al., 2018; Wang et al., 2011).

Major inputs and emissions from the STJ pathway were analyzed in detail as shown in Figure 6.2. The life cycle for STJ pathway begins with the manufacturing of fertilizers. Fertilizer production requires energy, which was assumed to come from natural gas (NG). Switchgrass farming generates $N_2O$ emissions due to the nitrogen fertilizer application. Although it is not shown in the system boundary, $CO_2$ emissions from farm equipment and other indirect emissions for these processes were accounted for based on data from the GREET model. The switchgrass pretreatment-hydrolysis step to generate sugars requires NG, electricity, sulfuric acid, ammonia, and cellulase, which generates $CO_2$ emissions. In the pretreatment-hydrolysis step, lignin, being a valuable co-product, was assumed to be combusted to produce electricity. Lignin separated from the cellulose during pretreatment was assumed to be combusted to generate stream used in the distillation step, which displaces energy and reduces GHG emissions. Sugars from switchgrass and $H_2$ and $CO_2$ generated during fermentation were co-fermented to an ABE mixture using microorganisms. $CO_2$ emissions from co-fermentation are generated from excess gases not used in the process. Additional emissions from energy inputs for the fermentation process are generated from the usage of natural gas for heating and electricity mixing. Distillation was assumed to separate butanol from the ABE mixture, which consumes NG and produces $CO_2$. Finally, butanol was considered to be converted to jet fuel using a chemical catalytic process. Finally, the produced jet fuel is assumed to be combusted in a Boeing-747-400 aircraft. Estimates of the material and energy requirements for the steps in the full-scale conversion process were based on an Aspen Plus simulation of a corn ABE refinery (Wu et al., 2007). The "well-to-product" GHG emissions for resources available in GREET were then used to estimate emissions from production of most inputs used in each of the pathways as described below. Emissions from preparation

and handling of microbes used in the ABE fermentation were assumed to be small compared to other units in the system boundary and were excluded from the analysis (Dunn et al., 2015).

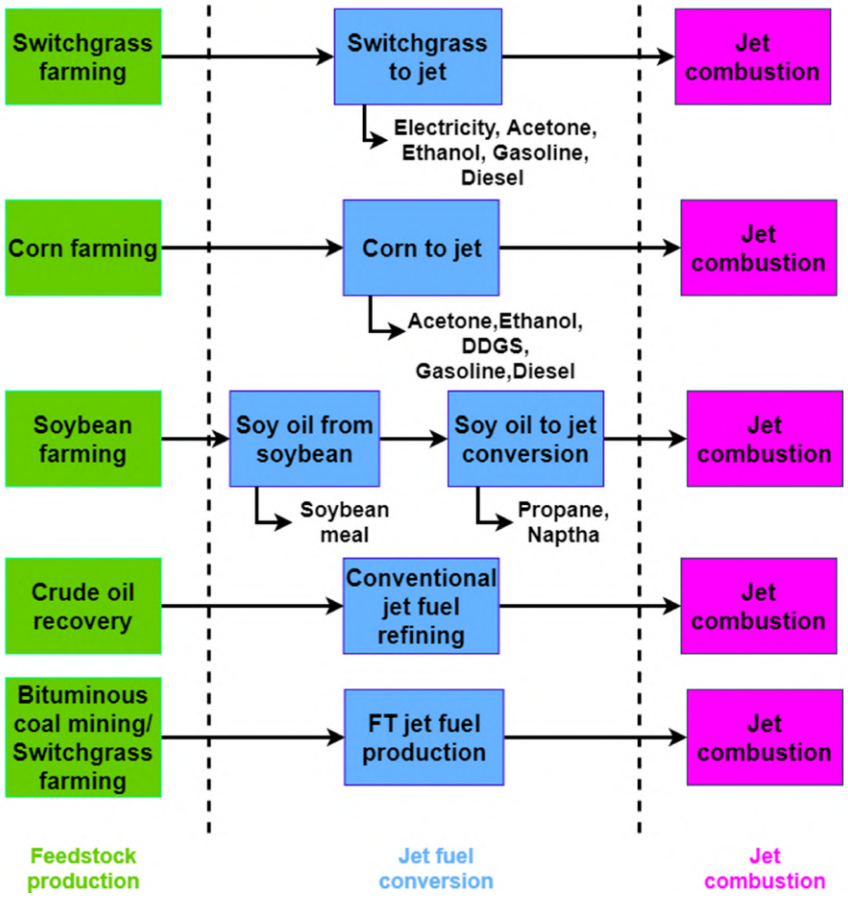


Figure 6.1: System boundary of jet fuel production pathways.

Emissions from the construction of infrastructure were assumed to be similar for each pathway and relatively small compared to emissions from the fuel cycle (Lucas et al., 2012), so they were excluded from the analysis. LB was assumed to come from marginal lands that do not directly affect food supplies (Gopalakrishnan et al., 2011; Tang et al., 2010). The effects of temporary sequestration and releases of carbon associated with growing biomass feedstocks were assumed to be in a steady-state condition, and were not considered; i.e., $CO_2$ pulled from the atmosphere during feedstock production was assumed to be instantaneously released during fuel combustion despite the actual time lag between these two events. A brief assessment

of the importance of direct land use change on life cycle emissions was performed for the STJ pathway, since direct land use changes can have an important effect of increased bioenergy production. The impacts of indirect land use change were considered beyond the scope of this analysis. The estimation of carbon sequestration is difficult for bio-energy crops because of variability of many factors such as soil types, land management practices (Post and Kwon, 2000; McLaughlin and Walsh, 1998). In addition, the application of nitrogen fertilizers to improve crop yields impact the carbon sequestered in soil, which varies with farming locations (Schuman et al., 2002). Information on the time the bio-energy crops take to establish and attain equilibrium or how much carbon can be sequestered is debatable and hypothetical (Vaughan et al., 2018). Therefore, carbon sequestration in bio-energy crops used was beyond the scope of this analysis. Further research is required to thoroughly investigate carbon sequestration in lignocellulosic biomass.

## 6.4 Data Collection and Inventory Analysis

### 6.4.1 Agricultural Inputs

Data used to estimate the emissions from fertilizer production, herbicide, and diesel used during switchgrass farming are shown in Table 6.1. Fertilizers used in switchgrass farming include nitrogen (N), phosphorus ($P_2O_5$), and potassium ($K_2O$). N inputs and life cycle emissions were assumed to be the US average nitrogen mix in GREET (Plevin, 2009). $P_2O_5$ was assumed to come from mono-ammonium and diammonium phosphate mix derived from phosphorous mining (Johnson et al., 2013). $K_2O$ was assumed to be produced from potassium mining operations (Wang et al., 1999). Apart from fertilizers, other agricultural inputs such as lime, herbicides, and high-density polyethylene were included in the estimation of emissions, consistent with the GREET model (Cai et al., 2016).

### 6.4.2 Switchgrass Farming

Switchgrass can be grown in various geographical, climatic conditions as different varieties adapt to changing weather conditions in different seasons (Guretzky et al., 2011). Most of the southern lowland ecotypes of switchgrass include Alamo, Kanlow, Blackwell, and Caddo (Lemus et al., 2002). Alamo switchgrass was assumed to be the feedstock for the STJ pathway from various locations in the USA. Alamo switchgrass has a one- or two-year establishment period after planting where no fertilizer is added after seeding due to the potential growth of weeds (Parrish and Fike, 2005). After the establishment year or first cutting, the stand life of switchgrass is at least ten years (Fike et al., 2006). Based on these assumptions, the yield of switchgrass was assumed to be 4.9 dry tons/acre based on the 10-year life cycle (Dunn et al., 2012). Irrigation of switchgrass is needed in dry months when precipitation is much lower than evapotranspiration (Giannoulis et al., 2016). In the current LCA, emissions from energy use in irrigation were not accounted because switchgrass was assumed to come from rain-fed Conservation Reserve Program (CRP) lands of southern, southeast, and Midwestern USA (Elgowainy et al., 2014).

The co-fermentation of sugars and gases requires 348 g and 247 g of switchgrass in practical and ideal cases, respectively, to produce 1 MJ of jet fuel. Similarly, the switchgrass requirement was 402 g and 288 g of from the traditional ABE fermentation process to produce 1 MJ of jet fuel. The feedstock requirement for both traditional and co-fermentation technology routes to produce 1 MJ of jet fuel was estimated based on the stoichiometry of ABE produced from sugars (including both glucose and xylose). Usage of fertilizers and diesel generates $N_2O$ and $CO_2$ emissions in the farming process (Cai et al., 2016).

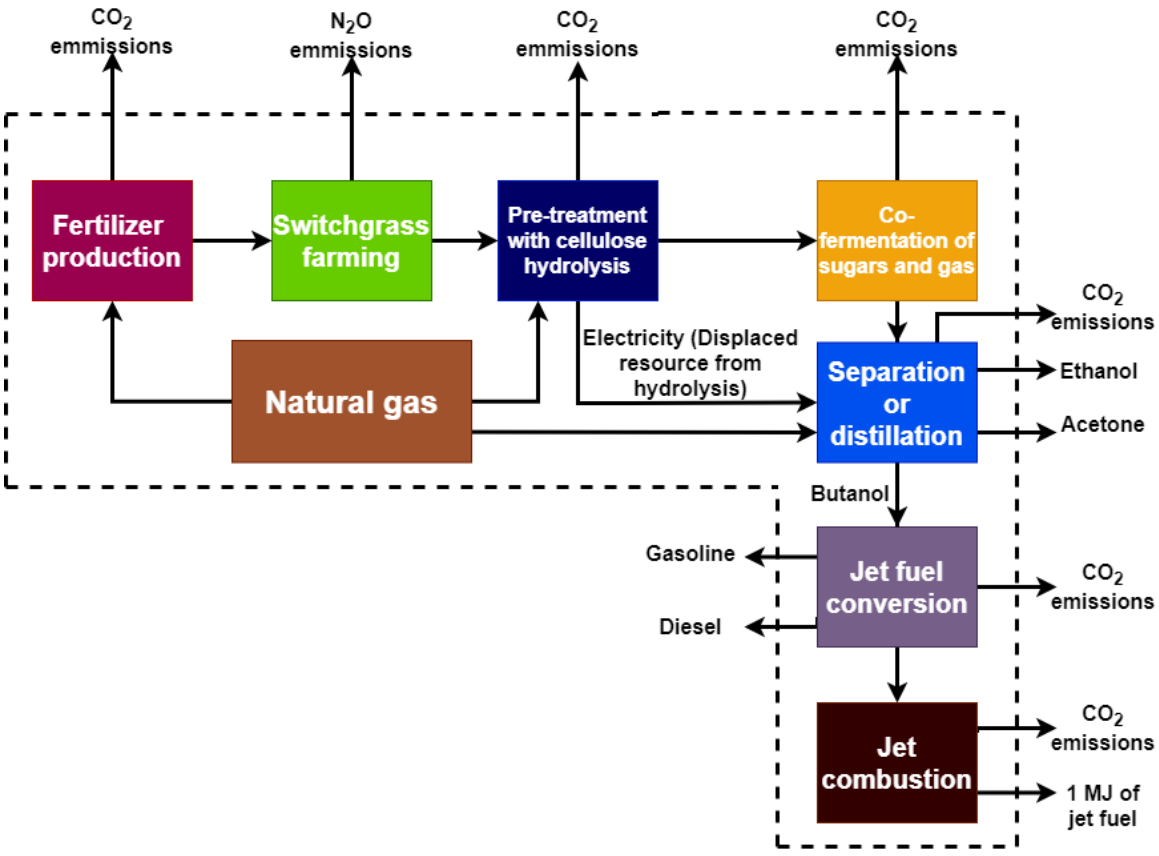


Figure 6.2: Major process inputs and emissions for the switchgrass to jet fuel pathway

$N_2O$ emissions were estimated from nitrogen input, assuming 1.2% of the applied N was converted to $N_2O$ (Wang et al., 2012). The nitrogen fertilizer and diesel fuel requirements to produce 1 MJ of jet fuel under various STJ assumptions are shown in Table 6.1. Co-fermentation decreases fertilizer and diesel requirement per unit output by improving ABE yields.

### 6.4.3 Dilute Acid Pretreatment and Hydrolysis

After switchgrass is harvested, it is assumed to be pretreated with sulfuric acid followed by hydrolysis with enzymes to break down the feedstock into sugars, including glucose and some xylose. A metric ton of switchgrass can produce 690 kg and 620 kg of sugars for the ideal case (Dunn et al., 2015).

Table 6.1: Resources requirements and associated GWP-100 to produce 1 MJ of jet fuel via the STJ pathway with co-fermentation

| Production resource | Resource quantity (practical case)[a] | Life cycle emissions (g) (practical case)[a] | Resource quantity (ideal case)[b] | Life cycle emissions (g) (ideal case)[b] | Source of data |
|---|---|---|---|---|---|
| Nitrogen (N), g | 1.87 | 7.09 | 1.32 | 5.02 | Cai et al., 2016 |
| Phosphoric acid ($P_2O_5$), g | 0.88 | 0.72 | 0.63 | 0.51 | Cai et al., 2016 |
| Potassium oxide ($K_2O$), g | 1.23 | 0.71 | 0.87 | 0.50 | Cai et al., 2016 |
| $CaCO_3$, g | 2.24 | 0.99 | 1.60 | 0.70 | Cai et al., 2016 |
| Herbicides, g | 0.02 | 0.39 | 0.02 | 0.28 | Cai et al., 2016 |
| High-density polyethylene, g | 0.13 | 0.38 | 0.09 | 0.27 | Cai et al., 2016 |
| Diesel in farm equipment, mL | 0.69 | 4.53 | 0.51 | 3.22 | Cai et al., 2016 |
| Natural gas in pretreatment and hydrolysis, g | 25 | 6.83 | 19.81 | 5.39 | Dunn et al., 2015 |
| Electricity in pretreatment and hydrolysis, Wh | 16.66 | 7.96 | 13.33 | 6.28 | Dunn et al., 2015 |
| Ammonia, g | 4.32 | 10.10 | 3.4 | 7.96 | Dunn et al., 2015 |
| Cellulase, g | 2.16 | 3.70 | 1.70 | 2.92 | Dunn et al., 2015 |
| Sulfuric acid, g | 6.48 | 0.28 | 5.11 | 0.22 | Dunn et al., 2015 |
| Natural gas in evaporator, co-fermentation and gas stripper, g | 15.71 | 7.87 | 13.23 | 6.63 | Atiyeh, 2020; Wu et al., 2007, |
| Electricity in evaporator, co-fermentation and gas stripper, Wh | 7.42 | 3.33 | 7.39 | 3.32 | Atiyeh., 2020; Wu et al., 2007, |
| Natural gas in distillation, g | 6.66 | 1.81 | 5.25 | 1.42 | Baral et al., 2016 |
| Electricity in butanol to jet fuel conversion, Wh | 9.29 | 4.42 | 9.29 | 4.42 | Wang et al., 2016 |
| Hydrogen, g | 0.67 | 6.95 | 0.67 | 6.95 | Wang et al., 2016 |
| ZSM-5 catalyst, g | 0.11 | 0.72 | 0.11 | 0.72 | Wang et al., 2016 |

[a] Practical case assumes yields of 90% of the theoretical value for hydrolysis and 80% of the theoretical value for ABE fermentation

[b] Ideal case assumes yields of 100% of the theoretical value for hydrolysis and 95% of the theoretical value for ABE fermentation

For traditional ABE fermentation and co-fermentation processes in the ideal case, it was assumed that 95% of the sugars (glucose and xylose) produced from switchgrass were converted to ABE. In the practical case, the assumption of a 10% sugar loss was based on previous literature (Liu et al., 2015a,b; Sun et al., 2020). Sugars are lost during pretreatment and hydrolysis when enzymatic process slows near equilibrium. Generating 1 MJ of jet fuel using the co-fermentation process in the STJ pathway requires 216 g and 170 g of sugars in the practical and ideal cases, respectively. The fermented sugars generate $H_2$ and $CO_2$ that are converted to butanol by gas fermenting microorganisms (Atiyeh, 2021). However, 250 g and 200 g of sugars are needed in the practical and ideal cases with the traditional ABE fermentation route for 1 MJ of jet fuel, which are estimated based on the data from literature and proposed stoichiometry in Section 2.3.4 (Dunn et al., 2015). The inputs for switchgrass pretreatment and hydrolysis include NG, electricity, ammonia, cellulase, and sulfuric acid, as shown in Table 6.1. Emissions from NG and electricity production were based on US averages from GREET (Cai et al., 2012; Clark et al., 2012). Electricity is consumed for pumping and mixing and produced from lignin combustion in the pretreatment process. Since the electricity production from lignin combustion substantially exceeds the process requirements, the excess was assumed to displace emissions from U.S average electricity. Apart from energy inputs, ammonia was added to the pretreated slurry to raise the pH from 1 to 6 for enzymatic hydrolysis (Humbird et al., 2011). Emissions from ammonia production are important from the life cycle perspective because it contributes to most of the emissions in the pretreatment and hydrolysis. Emissions from ammonia production, in the pretreatment and hydrolysis are based on the average nitrogen fertilizer produced in the United States (Johnson et al., 2013). Cellulase is an enzyme used to break the cellulose into sugars. The emissions from cellulase production were estimated from previous studies (Schneider et al., 2018; Wang et al., 2012). Sulfuric acid is required for dilute acid pretreatment

(DAP). Emissions for sulfuric acid production, which is a co-product of petroleum refining, were taken from GREET (Dunn et al., 2015; Lampert et al., 2015).

### 6.4.4 Co-fermentation of Sugars and Gas

Sugars released during pretreatment and hydrolysis are converted to ABE through evaporation, fermentation and gas stripping. In the fermentation step, both sugars and gas ($H_2$, $CO_2$) are co-fermented to produce the ABE mixture. The efficiency of the ABE fermentation process was assumed to be at 80% and 95% in the practical and ideal cases, respectively, based on previous literature (Sun et al., 2020). Yields of 0.321 g ABE/g sugar in the practical case and 0.374 g ABE/g sugar in the ideal case were estimated for the co-fermentation process based on the stoichiometry in Equations 6.1 and 6.2. However, traditional ABE fermentation yields only 0.278 g ABE/g sugar and 0.321 g ABE/g sugar for both the practical and ideal cases, respectively, based on Equation 6.1. This result is close to previous estimates for ABE fermentation yields (Kumar et al., 2012; Qureshi et al., 2010). The molar ratio of acetone: butanol: ethanol is 3:6:1, as reported by many researchers (Baral et al., 2016; De Jong and Kersten, 2013). However, the molar ratio of acetone: butanol: ethanol for the co-fermentation process increases to 3:7.3:1 by utilizing the off gases as shown in Equation 6.2.

$$\begin{aligned} 69C_6H_{12}O_6 \rightarrow 36C_4H_{10}O + 18C_3H_6O + 6C_2H_4O_2 + 12C_2H_4O_2 + \\ 6C_4H_8O_2 + H_2O + 156CO_2 + 108H_2 \end{aligned} \tag{6.1}$$

$$108H_2 + 36CO_2 \rightarrow 9C_4H_{10}O + 36H_2O \tag{6.2}$$

The co-fermentation process requires additional energy for heating and mixing, which leads to additional indirect emissions. The fuel requirements and emissions for

the fermentation steps were assumed to be similar to those for a corn ABE biorefinery, which has previously been analyzed using an ASPEN Plus model (Wu et al., 2007). Data for these requirements were taken from the corn ABE analysis and GREET as shown in Table 6.1 (Atiyeh, 2021; Wu et al., 2007).

### 6.4.5 Distillation, Jet Fuel Conversion, and Combustion

The ABE mixture produced from fermentation is separated using distillation. The main product, butanol, can be subsequently converted to jet fuel (Figure 6.2). The acetone and ethanol are valuable co-products requiring allocation. Butanol is converted to jet fuel by a series of steps that involve dehydration using silane-modified-$\gamma$-alumina catalyst, oligomerization with Ziegler-Natta catalyst, and hydrogenation (Wang and Tao, 2016). GHG emissions for the fuels used in distillation are shown in Table 6.1, based on a previous study that included heat recovery (Baral et al., 2016). The butanol to jet fuel conversion step requires external $H_2$ gas, a ZSM-5 catalyst, and electricity (Harvey and Meylemans, 2011). Conversion requires external $H_2$, electricity, and ZSM-5 zeolite, leading to indirect GHG emissions as shown in Table 6.1. Finally, the emissions from the combustion of 1 MJ of jet fuel are based on a consumption rate of 11.7 L/km (5 gal/mile) in a Boeing 747-400 (Park and O'Kelly, 2014). The Boeing 747-400 was selected because it has been used with traditional petroleum jet fuel and with several emerging biofuels (Marsh, 2008; Mazlan et al., 2015; Rahmes et al., 2009). Virgin Atlantic has been running their firm flights with one engine on low carbon jet fuel manufactured from industrial wastes (Brooks et al., 2016b). Also, Air New Zealand has been flying Boeing 747-400 on a jatropha-derived HRJ with (50%) on one engine and kerosene-type fuel on the other (50%). KLM has also been flying with a blend of 50% FTJ and 50% conventional jet fuel in Boeing 747-400 (Johansson et al., 2012).

### 6.4.6 Life Cycle GHG Emissions from Other Jet Fuel Pathways

The GWP-100 of conventional and other emerging jet fuel pathways including CPJ, CTJ, HRJ, and FTJ were compiled to assess the potential impact of the new STJ pathway. Results from previous analyses compiled within GREET were used in each case, but with slight modifications for consistency with the assumptions and system boundaries used for the STJ pathway in this study. The estimates of GHG emissions for CPJ were based on energy consumption in jet fuel refined from crude oil using data from 43 refineries representing 70% of the U.S. refining sources Elgowainy et al. (2014), For the CTJ pathway, the GHG emissions were estimated from a previous study on ABE fermentation Wu et al. (2007) with an assumption that the corn butanol to jet fuel conversion and combustion steps were similar to STJ pathway. For the HRJ pathway, estimated GWP-100 were based on a previous analysis of various alternative aviation fuels, including various bio-based oil feedstocks (Elgowainy et al., 2012). For FTJ, both coal and switchgrass, were considered as feedstocks using results from previous analysis (Xie et al., 2011). A detailed description of the assumptions used to estimate the life cycle emissions from the CPJ, CTJ, HRJ, and FTJ pathways is provided in the supplemental data file.

### 6.4.7 Co-product Allocation

Co-product handling and assumptions about biogenic carbon are important factors when estimating GHG emissions from bioenergy pathways. The biogenic jet fuel pathways (STJ, CTJ, and HRJ) generate many co-products in the conversion step, such as alcohols and fuel products (Figure 6.1) that require allocation of the GHG emission burden. Co-product allocation methods used include displacement, energy allocation, mass allocation, and market value allocation. In the STJ pathway, there are five co-products (electricity, acetone, ethanol, gasoline, and diesel), as shown in Figure 6.1. The CTJ pathway has ethanol, acetone, and DDGS as co-products in

the distillation and gasoline blendstock and low-sulfur diesel in the final butanol to jet fuel conversion step. Co-products in the HRJ pathway include soy meal, which displaces soybeans as an animal feed (Knudsen, 1997), and propane and naphtha produced during conversion.

### 6.4.8 Inventory Analysis and Impact Assessment

Data from laboratory studies, GREET, and other literature described above were compiled into a Microsoft Excel spreadsheet that was used to estimate life cycle GWP-100 for the various pathways. $CO_2e$ emissions were summed for all unit processes, excluding biogenic $CO_2$, and co-products were allocated using a variety of methods to estimate GWP-100 for the entire pathway. For HRJ, the estimated GWP-100 from a previous report were used (Elgowainy et al., 2012). Emissions estimates for FTJ pathways were taken from previous research (Elgowainy et al., 2012; Xie et al., 2011). GWP-100 for CPJ were taken from previous studies (Elgowainy et al., 2012; Xie et al., 2011). For each pathway, biogenic carbon, co-product allocation, and emissions from fuel combustion were handled consistently with the approach used for the goal and scope of the STJ pathway.

## 6.5 Results and Discussion

### 6.5.1 Emissions in STJ Pathway

The various jet fuel production pathways were quantitatively assessed and compared based on the GWP-100 from well-to-wake (WTWa) as summarized in Table 6.2 and Table 6.3. The STJ pathway includes farming, hydrolysis, evaporation, fermentation, gas stripping, separation, distillation, and combustion. This pathway was analyzed for four different alternative assumptions: with and without co-fermentation of the off-gases for both the theoretical and practical cases.GWP-100 from WTWa is often divided from well to pump (WTP) and pump to wake (PTW) (Bicer and Dincer,

2017). For biogenic sources of carbon, the PTW emissions were neglected as per standard LCA practices, and because no credit was provided for removal of $CO_2$ from the atmosphere during feedstock production (Finkbeiner, 2014; de Normalización, 2006; Wu et al., 2015). The GWP-100 of different unit processes for the STJ pathway via the co-fermentation route for the practical and ideal cases are illustrated in Figure 6.3 and Figure B.8 in the supplementary data, respectively. The STJ pathway begins with switchgrass farming, which requires agricultural chemicals that contribute to 9.3% of the WTP GWP-100 as illustrated in Figure 6.3. Fertilizer production and usage were the third-largest contributors to the total GWP-100 in the STJ pathway. Most of these emissions come from nitrogen fertilizer usage while farming. GWP-100 contributions from switchgrass farming were 3.5% for WTP, primarily from diesel and electricity usage. Agricultural chemicals contribute to 5.8% of GWP-100, significantly more than the farm equipment, which illustrates the importance of improved nitrogen fertilizer management for biofuel sustainability

Table 6.2: Pre-allocated GWP-100 (g $CO_2e$/MJ) in bio-jet fuel production pathways from WTP

| Processes | STJ (co-fermentation) practical case)[a] | STJ(co-fermentation) ideal case) [b] | CTJ (practical case)[c] | CTJ (ideal case)[d] | HRJ |
|---|---|---|---|---|---|
| Farming | 23.6 | 16.8 | 30.1 | 25.1 | 51.4 |
| Feedstock to jet fuel conversion | 138.8 | 115.5 | 67.0 | 65.8 | 67.3 |
| Combustion | 74.3 | 74.3 | 74.3 | 74.3 | 74.3 |

Pretreatment and hydrolysis are the largest contributors to GWP- 100 in the STJ pathway as shown in Figure 6.3. The majority of emissions in pretreatment and hydrolysis result from the combustion of natural gas and ammonia input. Additional significant sources of emissions include natural gas for evaporation, co-fermentation and gas stripping. The emissions from co-fermentation step represents the excess $CO_2$ not used to produce ABE with the $H_2$ gas. The co-fermentation process contributes

only 1.0% of the total GWP-100 from WTP, which is the lowest $CO_2$ contributor in the STJ pathway. Traditional ABE fermentation contributes to 16.8% more $CO_2$ emissions than co-fermentation. More importantly, the new process improves yields, which saves emissions from upstream processes in the supply chain.

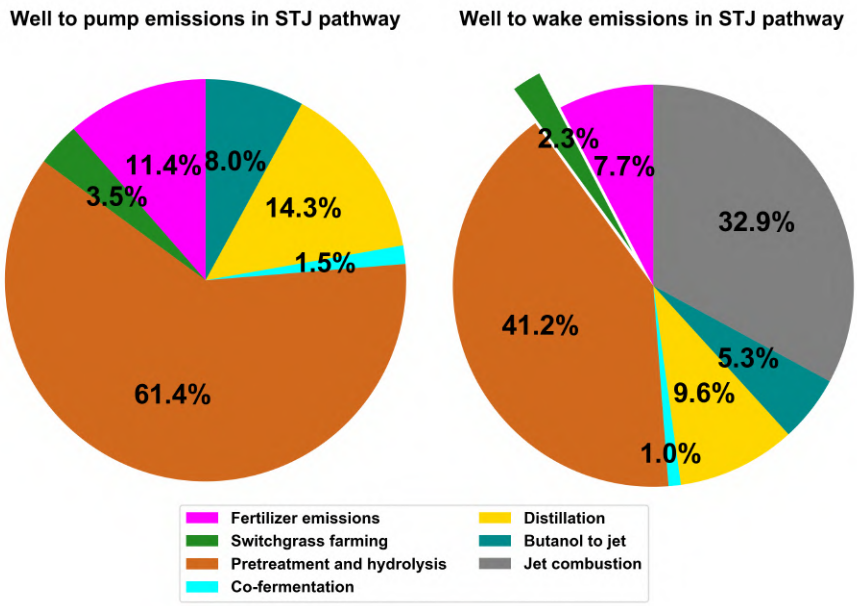


Figure 6.3: WTP and WTWa emissions from processes in the STJ pathway with co-fermentation

The ABE separation is energy-intensive and represents the second largest contribution to GWP-100 at 12.4% from as shown in Figure 6.3. The butanol to jet fuel conversion step contributes 7.4% of WTP GWP-100, since a significant amount of external $H_2$ fuel is required to reduce the butanol to hydrocarbons. The WTP GWP-100 breakdown for the STJ pathway shown in Figure 6.3 illustrates that most of the emissions result from fuel usage for switchgrass hydrolysis pretreatment, distillation, and fertilizer production. Therefore, future research should investigate approaches to reduce the energy intensity of these processes. Moreover, estimations of the emissions for all the notable alternative jet fuel production pathways compared in the current study are provided in Sections 6.5.7, 6.5.8 and and Table B.14 in the supplementary document.

Table 6.3: GWP-100 (g $CO_2e$/MJ) or STJ, CTJ, and HRJ pathways with energy allocation method

| Allocated resource | STJ (practical case)[a] | STJ (ideal case)[b] | CTJ (practical case)[c] | CTJ (ideal case)[d] | HRJ |
|---|---|---|---|---|---|
| Jet fuel | 48.3 | 40.1 | 44.0 | 41.2 | 56.2 |
| Ethanol | 12.0 | 10.0 | 0.007 | 0.006 | 0 |
| Acetone | 23.6 | 19.6 | 12.3 | 11.5 | 0 |
| DDGS | 0 | 0 | 26.2 | 24.5 | 0 |
| Propane | 0 | 0 | 0 | 0 | 13 |
| Naphtha | 0 | 0 | 0 | 0 | 9.5 |
| Soybean meal | 0 | 0 | 0 | 0 | 40 |
| Displaced electricity | 62.6 | 49.5 | 0 | 0 | 0 |
| Other products | 15.9 | 13.2 | 14.5 | 13.5 | 0 |
| $CO_2$ biogenic[e] | 74.3 | 74.3 | 74.3 | 74.3 | 74.3 |

[a] STJ-Practical case assumes yields of 90% of the theoretical value for hydrolysis and 80% of the theoretical value for ABE fermentation (0.321 g ABE/g sugar)
[b] STJ-Ideal case assumes yields of 100% of the theoretical value for hydrolysis and 95% of the theoretical value for ABE fermentation (0.374 g ABE/g sugar)
[c] CTJ-Practical case assumes yield of 80% of the theoretical value for ABE fermentation (0.372 g ABE/g sugar)
[d] CTJ-Ideal case assumes yield of 100% of the theoretical value for ABE fermentation (0.465 g ABE/g sugar)
[e] $CO_2$ biogenic refers to the $CO_2$ emissions from the combustion of jet fuel which is assumed to be sequestered by the feedstock

### 6.5.2 Sensitivity of Results to Switchgrass Yield and Nitrogen Fertilizer Application

Soil properties, weather patterns, nutrient availability in different geographical locations affect switchgrass yields, which have important impacts on the STJ life cycle (Bradford et al., 2006; Wullschleger et al., 2010). Spatial variability in switchgrass yield and nitrogen fertilizer application was used to assess the impacts of farming on the STJ pathway life cycle emissions for the practical case as shown in Figure 6.4. A similar comparison plot affecting GWP-100 of STJ pathway (ideal case) by different farming techniques across these regions were shown in Figure B.11 in the supplementary data. Quantities of applied nitrogen fertilizer and switchgrass yields were obtained from previous studies (De et al., 2003; Epplin, 1996; Hallam et al., 2001; Khanna et al., 2008; Turhollow, 2000). These data were used to assess GWP-100 variability across various regions of the United States, including Oklahoma, the Corn Belt, Maryland, Iowa, and Illinois. Data representing the Corn Belt include an average of switchgrass yields across Illinois, Indiana, Iowa, Kansas, Kentucky, Michigan, Minnesota, Missouri, Nebraska, North Dakota, South Dakota, Ohio, and Wisconsin.

Life cycle emissions were estimated during the establishment year and after the first cut which were compared on the weight (kg) basis of switchgrass harvested. Even without the application of nitrogen fertilizer, there was significant variation in the biomass yield in the establishment year, as illustrated in Figure 6.4. Native nitrogen available in the soil can affect the biomass yield in the establishment year (Owens et al., 2013). From Figure 6.4, Oklahoma has the highest switchgrass yield at 2.3 metric ton/acre, while Iowa has a comparatively low yield of 1.1 metric ton/acre in establishment year. After the first year of establishment, nitrogen fertilizer is applied to improve the yield of switchgrass. Illinois and Iowa have relatively high annual nitrogen fertilizer requirements of 17.8 and 13.6 g/kg of switchgrass, respectively, and comparatively low yields of 2.6 and 4.2 metric ton/acre (Khanna et al., 2008;

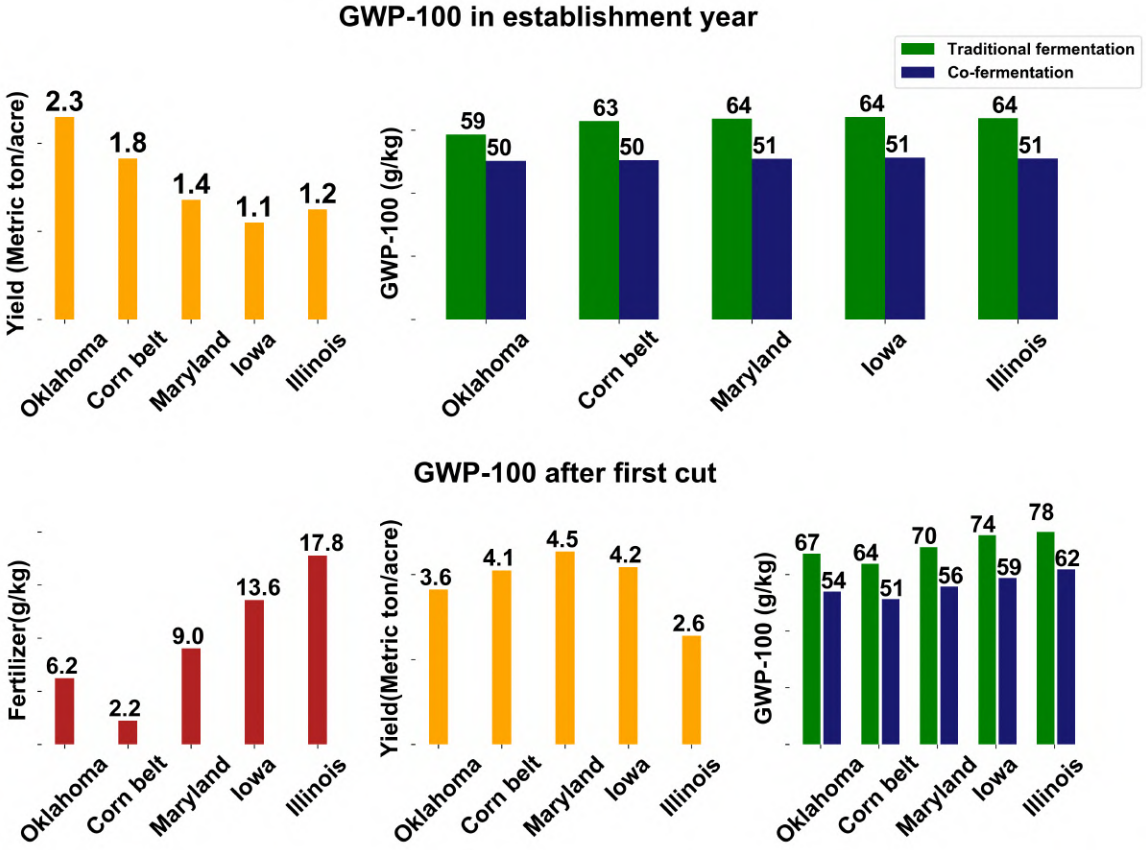


Figure 6.4: WTP and WTWa emissions from processes in the STJ pathway with co-fermentation

Hallam et al., 2001) as shown in Figure 6.4. Illinois and Iowa switchgrass have high GWP-100 emissions from fertilizer application after establishment, while the Corn Belt has comparatively low GWP-100 emissions, as illustrated in Figure 6.4. These results demonstrate the importance of controlling nitrogen fertilizer applications to reduce GHG emissions in STJ pathways. As shown in Figure 6.4, cofermentation reduces the life cycle emissions by 19.5% in Illinois and 19.7% in Iowa, respectively, which demonstrates the significance of the co-fermentation technology on the life cycle emissions.

### 6.5.3 Effects of Process Technology Selection on STJ Pathway

There are multiple alternatives for some of these processes in the STJ pathway, as explained in Section 6.3.2. Figure 6.3 shows that pretreatmenthydrolysis and dis-

tillation technologies contributed the most to the GWP-100 for this pathway. The effects of alternative assumptions for some of the most significant unit processes on the GWP-100 were assessed. $CO_2e$ emissions were estimated for the STJ pathway prior to co-product allocation for both the ideal and practical cases. As shown in Figure 6.5, the STJ pathway via the co-fermentation route has unit processes including farming, DAP, co-fermentation, distillation (heat recovery), butanol to jet fuel conversion, and jet combustion. DAP is a leading pretreatment technology to produce fermentable sugars (Chandra et al., 2007; Wyman et al., 2009; Yang and Wyman, 2008). As an alternative, ammonia fiber expansion (AFEX) can also be used to pretreat switchgrass, enabling cellulase to break the biomass into sugars. AFEX can reduce the $CO_2e$ from WTWa in both the practical and ideal cases by 12.9% and 5.2% compared to DAP in STJ pathways via co-fermentation (Dunn et al., 2015). Distillation is used to separate ABE, and it can be performed with or without heat recovery. Distillation with heat recovery was used to estimate GWP-100 for the STJ pathway. Recovering heat in distillation reduces the GWP-100 by 7.5% and 7.8% in the practical and ideal cases, respectively. In addition, the $CO_2e$ emissions for both the practical and ideal cases of the STJ pathway with co-fermentation are reduced by 12.3% and 6.2%, respectively, versus traditional ABE fermentation due to its low ABE yield

### 6.5.4 Impacts of Co-product Handling

The biofuel pathways of STJ, CTJ, and HRJ generate multiple coproducts, including acetone, ethanol, DDGS, soybean meal, gasoline blendstock, and low-sulfur diesel. Allocation of GHG emissions to these co-products plays a significant role in the estimation of GWP-100. Four co-product allocation methods (energy, mass, market value allocation, and displacement) were used in the biofuel production pathways of STJ, CTJ, and HRJ. Displacement or system boundary expansion was chosen to allo-

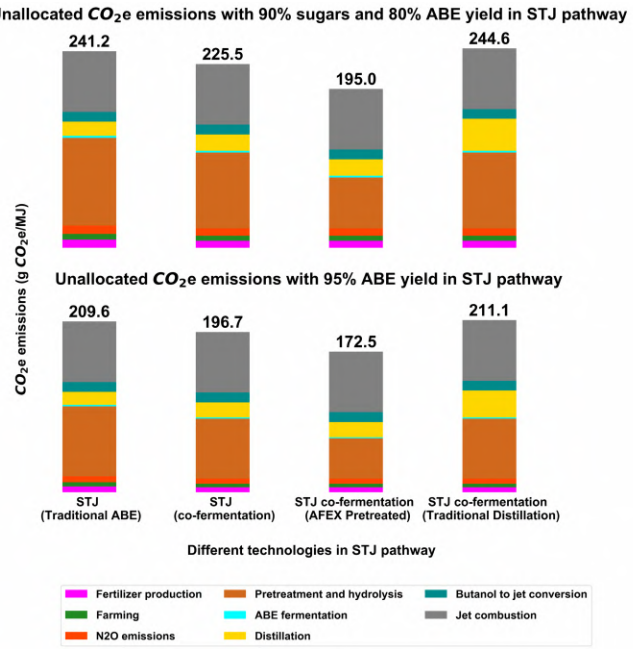


Figure 6.5: $CO_2e$ emissions variability based on different unit processes in STJ pathways (practical and ideal cases); STJ (traditional ABE) unit processes include DAP, traditional ABE fermentation, distillation (heat recovery); STJ (co-fermentation) unit processes include DAP, co-fermentation, distillation (heat recovery); STJ (AFEX pretreated) unit processes include AFEX, co-fermentation, distillation (heat recovery); STJ (traditional distillation) unit processes include DAP, co-fermentation, distillation (without heat recovery)

cate emission burdens to co-produced electricity in each case, since it is the preferred method when possible (Iso, 2006). Electricity produced during lignin combustion was assumed to displace U.S. average emissions in four allocation methods in the STJ pathway. Allocation of all co-products in the STJ (practical case), CTJ (80% ABE),and HRJ pathways with various approaches are illustrated in Figure 6.6. Similarly, co-product allocation of ideal cases for the STJ and CTJ pathways are shown in Figure B.12 in the supplementary data. The non-electricity co-products receive only a small allocation compared to electricity ($< 20\%$), but they are significant on the STJ life cycle. The most notable co-product allocations are given to ethanol and acetone, as shown in Figure 6.6. The carbon in the switchgrass LB is biogenic (taken from the atmosphere), so a credit of 74.3 g/$CO_2e$/MJ is provided to life cycle emissions to get the final GWP-100 value for the STJ pathway, as shown in Table 6.3. There are at least 24 potential ways of allocating emissions to the co-products in the STJ pathways that were considered (mass, energy, and market basis for each co-product). The GPW-100 for the STJ, CTJ, and HRJ pathways with energy, mass, and market allocations are shown in Figure 6.6. Out of all the allocation scenarios, market value allocation has the least contribution to GWP-100 at 43.1% and 31.9% less than the energy and mass allocations. Since jet fuel and the co-products are all combustible and potential energy sources, the energy allocation approach was selected for final GWP-100 comparison with other jet fuel pathways. Similarly, the co-product allocation for both the ideal and practical cases of CTJ pathways are shown in Figure 6.6 and Figure B.12. In the CTJ pathways, the energy, mass, and market basis were used to allocate emissions between five co-products including ethanol, acetone, DDGS, gasoline blendstock, and low-sulfur diesel. System boundary expansion was not possible for these co-products, since there is no substitute product with life cycle emissions data readily available. As in the STJ pathway, the market allocation provides the least burden to the jet fuel (22.1%) for the CTJ pathways, since the price of

both acetone and DDGS are relatively high compared to other co-products. For the CTJ pathways, corn grain does not co-produce electricity, whereas corn stover does (Wang et al., 2009b). The allocation of emissions to different co-products is shown in Table 6.3

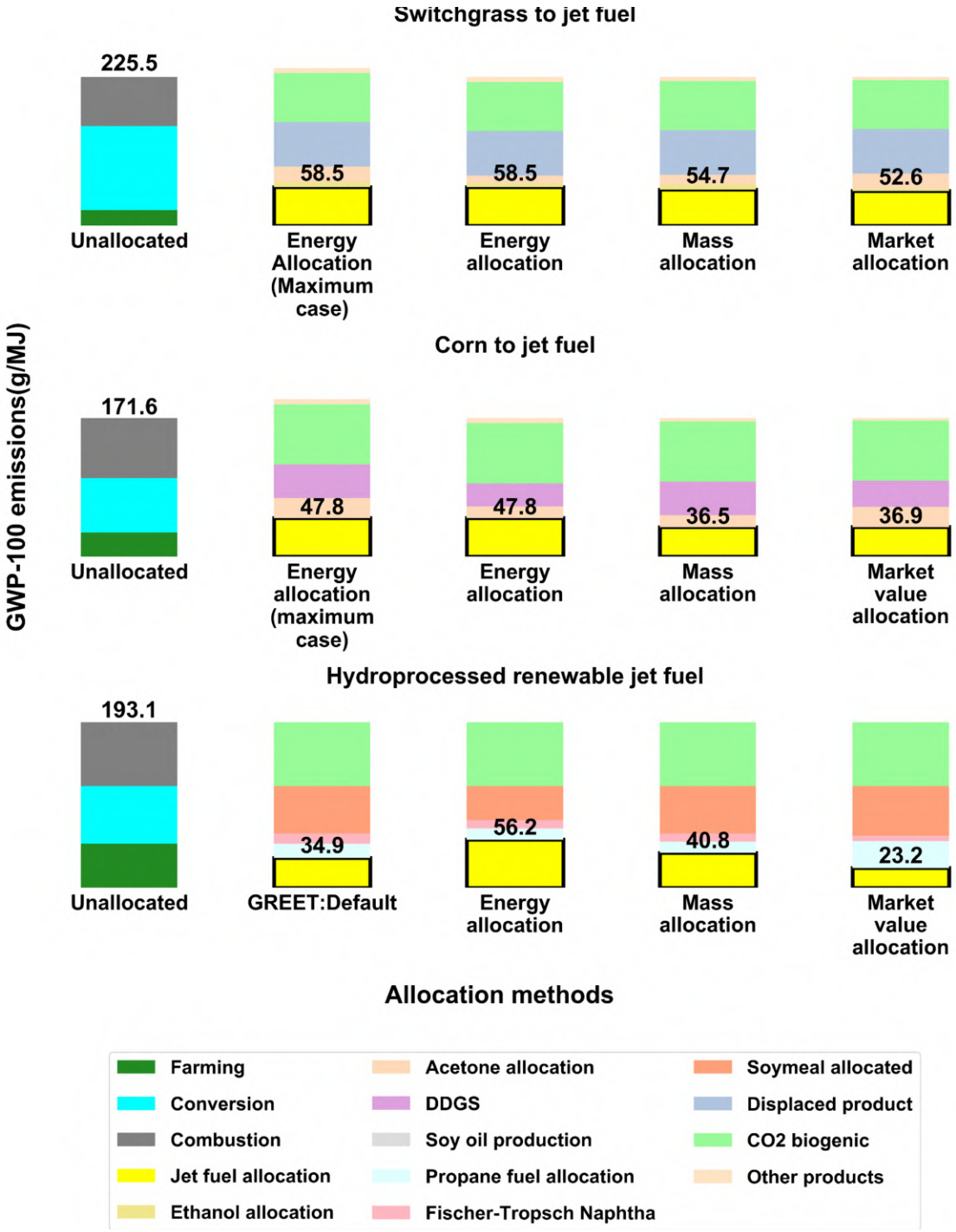


Figure 6.6: Co-product allocation methods in biofuel pathways

In the case of HRJ, the jet fuel conversion step from soy oil generates multiple co-products including propane and naphtha. The GREET default method (energy allocation) was used to estimate the GWP-100 of the HRJ pathway. All allocation scenarios shown in Figure 6.6 and Figure B.12 for HRJ pathway were assessed using GREET. During allocation, soymeal displaces most of the life cycle emissions, making it an essential co-product in the HRJ pathway. The allocation method therefore has a significant effect for this pathway. Energy allocation was used to compare the STJ pathway with other potential jet fuel production pathways, since most of the products can be used as fuels. Since market prices are not stable, the market allocation approach is subject to changes in the future, unlike the energy basis. Figure 6.6 shows that co-product allocation plays an important role in assessing the GWP-100 when alternative jet fuel production pathways are compared

### 6.5.5 Sensitivity Analysis of GWP-100 in STJ Pathway

A sensitivity analysis was conducted to examine the importance of the ABE yield on GWP-100 from both the traditional ABE and the ABE with co-fermentation for the STJ pathway practical case. Assumptions that significantly affect the GWP-100 by changing ABE yields are shown in Figure 6.7. The co-fermentation ABE yield was changed from the base values of 0.321 g ABE/g sugars by ±20% to assess effects of yield on the life cycle. Other assumptions affecting the life cycle emissions by more than 1% for the traditional ABE process with a base yield of 0.278 g ABE/g sugars are illustrated in Figure 6.7. Changes in input assumptions that modified the life cycle GWP-100 by more than 3% from the baseline of 48.3 g $CO_2e$/MJ of jet fuel for the co-fermentation pathway and 51.7 g $CO_2e$/MJ of jet fuel for traditional ABE were determined. For pretreatment and hydrolysis, the sensitivity of results to the NG, electricity, ammonia, and enzyme requirements were analyzed. As shown in Figure 6.7, the NG requirement for hydrolysis for sugar is the most significant contributor to

the life cycle GWP-100. NG combustion during distillation is the second most energy-intensive process in the STJ pathway via both co-fermentation and traditional ABE routes.

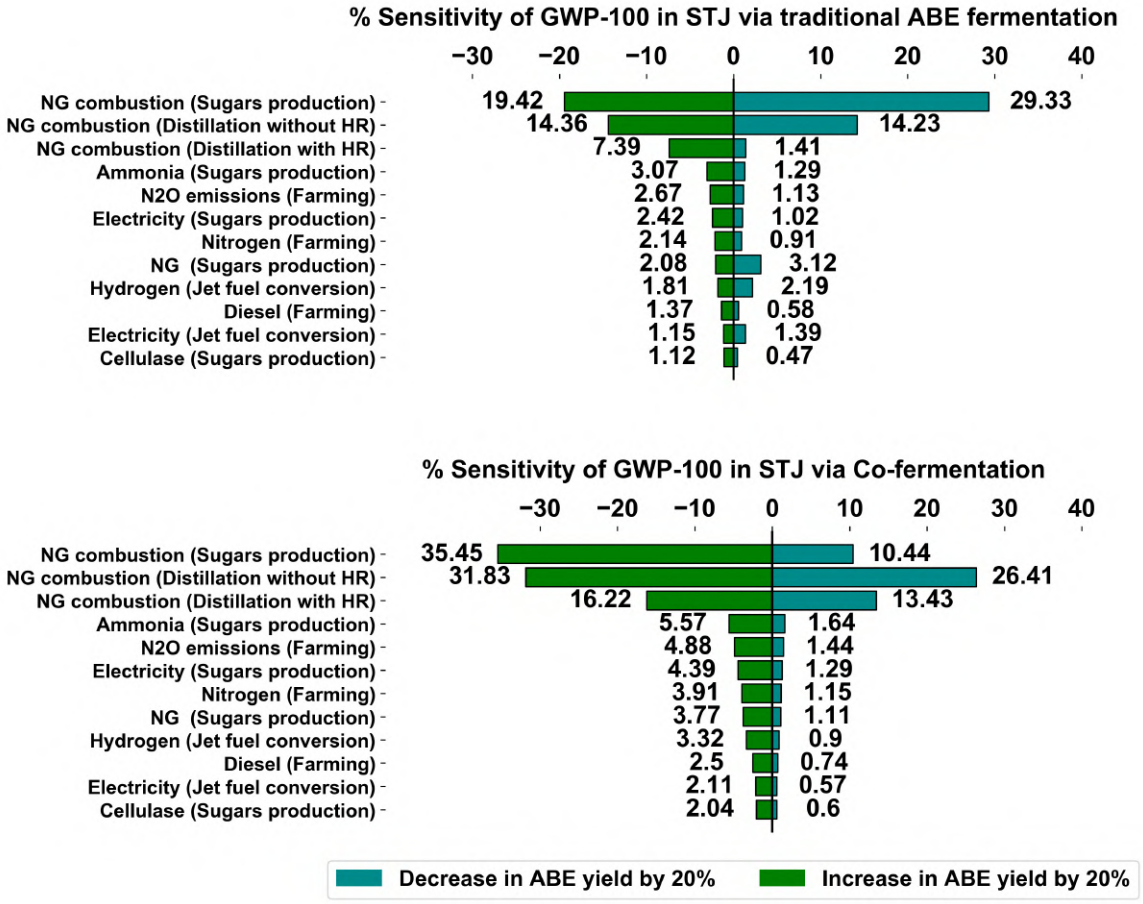


Figure 6.7: GWP-100 sensitivity to ±20% changes in ABE yield for the STJ practical case with both co-fermentation and traditional ABE routes. The GWP-100 for a base yield of 0.321 g ABE/g sugars for STJ with co-fermentation was 48.3 g $CO_2e$/MJ of jet fuel, while the emissions from a base yield of 0.278 g ABE/g sugars for STJ via traditional ABE was 51.7 g $CO_2e$/MJ of jet fuel

After the NG requirements, the ammonia requirement for sugars hydrolysis is the most significant contributor of GWP-100. The ammonia requirements for farming also contribute significantly to the GWP-100 of the STJ pathway, as shown in Table 6.1. In addition to emissions from production of ammonia, $N_2O$ emissions generated during farming contribute significantly to life cycle GHG emissions for the STJ pathway via both cofermentation and traditional ABE routes. Based on the results from

the sensitivity analysis, NG and ammonia inputs alter the life cycle GWP-100 most significantly, and therefore strategies to reduce emissions from these aspects of the process should be explored in more detail. The NG requirements for sugar production and distillation steps are the most sensitive variables in the production life cycle. Therefore, the 20% increase in ABE yield using co-fermentation significantly reduces the energy requirement and associated GHG emissions. Similar effects are apparent for upstream inputs like ammonia requirements for farming and pretreatment hydrolysis. Figure 6.7 and Figure B.13 in the supplementary data clearly show that the use of co-fermentation technology to produce 1 MJ of jet fuel reduces resource consumption and GHG emissions due to higher ABE yields relative to traditional ABE fermentation

### 6.5.6 Effects of Direct Land Use Change on the STJ Pathway

Large-scale land use changes to farm switchgrass commercially could affect food supplies and deforestation both directly and indirectly. When switchgrass farming was assumed to take place on existing farmland, this direct land use change can affect the life cycle GWP-100 from $-3.9$ to 13 g $CO_2e$/MJ for both the ideal and practical cases (Dunn et al., 2013). The life cycle GWP-100 for the STJ pathway for the practical case can range from 44.4 g $CO_2e$/MJ to 61.3 g $CO_2e$/MJ of jet fuel, depending on land use change effects. The direct land use change worst case land use change scenario increases GWP-100 by 26.9% for the STJ practical case. Even in this scenario, the life cycle GWP-100 is 33.4%, 66.7%, and 29.4% less than the FTJ (switchgrass), FTJ (coal), and CPJ pathways. These results indicate that the STJ pathway can reduce GWP-100 impact for producing jet fuel when compared with fossil-based pathways like FTJ and CPJ. The increase in ABE yield using the novel cofermentation of sugars and gas ($CO_2, H_2$) has the potential to produce sustainable aviation fuels with further reduction in emissions. Future efforts should focus on strategies to improve

performance of energy intensive unit processes like pretreatment, hydrolysis, and distillation. Additional research should focus on different pretreatment technologies and separation methods to produce butanol from ABE

### 6.5.7 Net Energy Ratio and Economic Feasibility of STJ Pathway

The net energy ratio in an important sustainability metric for alternative fuel supplies (King et al., 2015; Yang and Chen, 2012; Zhang and Colosi, 2013). The net energy cost for the STJ pathway was estimated by adding all the non-renewable energy inputs required to produce 1 MJ of jet fuel via the co-fermentation route. The nonrenewable energy requirements were estimated using data available in GREET. Co-products were allocated a share of the energy requirements as shown in Table B.13. The net energy ratio was then estimated using the energy output relative to the energy requirements. The primary inputs into the system shown in Table 6.1 were used to estimate the non-renewable energy requirements.

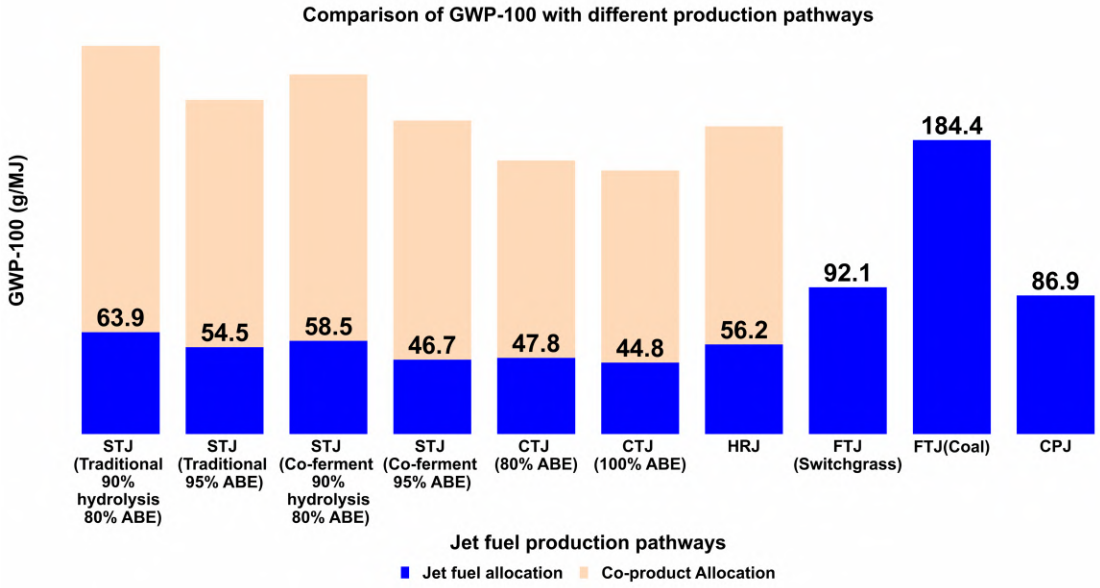


Figure 6.8: Comparison of GWP-100 in alternative jet fuel production pathways

After estimating energy cost for all the inputs, co-products including ethanol, and

acetone were allocated using the energy basis. Based on the estimations, the net energy ratio of the STJ pathway (practical case) is 1.213 MJ of non-renewable energy expended per MJ of jet fuel produced. The net energy ratio for CTJ pathway is 0.994 MJ per MJ of jet fuel, which is slightly better than STJ as in the case of life cycle GWP-100. Therefore, more research is needed to determine strategies to reduce the non-renewable energy costs for aviation biofuels. Estimates of net energy cost for all the jet fuel production pathways used in the current LCA are reported in the supplemental document.

### 6.5.8 Comparison of GWP-100 of Alternative Jet Fuel Pathways

The life cycle GWP-100 for the STJ pathway for each of the four scenarios (traditional ABE, practical case with 90% efficient hydrolysis and 80% ABE conversion), STJ (traditional ABE, ideal case with 95% ABE conversion), STJ (co-fermentation, practical case with 90% efficient hydrolysis and 80% ABE conversion), STJ (co-fermentation, ideal case with 95% ABE conversion) are shown in Figure 6.8. The results are compared with the other jet fuel pathways, CTJ (80% ABE), CTJ, HRJ, FTJ (switchgrass), FTJ (coal), and CPJ. Results are shown using the energy allocation method for co-products for the biofuel pathways. Jet fuel produced from switchgrass via co-fermentation in both the ideal and practical cases has a 6.5% and 6% reduction in GWP-100 compared to STJ via traditional ABE fermentation for both the practical and ideal cases. The GWP-100 for the STJ pathway with co-fermentation in the ideal case was 2.6% lower than for the CTJ pathway (0.465 g ABE/g sugar, 100% ABE) primarily due to the high energy requirements for pretreatment and higher ABE yield for corn grain relative to LB. The 15.8% and 24% increases in ABE yields for both CTJ (assuming 80% ABE conversion) and CTJ (assuming 100% ABE conversion) compared to STJ with co-fermentation for both the practical and ideal cases indicate that corn-based feedstock performs slightly better than LB with respect to

GWP-100. The STJ pathway outperforms the HRJ, FTJ, and CPJ pathways for all cases. These results show that the biochemical conversion of switchgrass to jet fuel is potentially more environmentally friendly than thermochemical conversion.In addition, the STJ pathway via co-fermentation route results in lower GWP-100 when compared to traditional ABE pathway.

### 6.5.9 Recommendations for Further Research

Lignocellulosic biomass has lower energy density when compared to fossil fuel sources. Therefore, technologies that improve the energy density such as pelletization, torrefaction, and pyrolysis need to be investigated further. In addition, water requirements in biofuel pathways justify the need for further research because water usage in the fermentation step can increase the energy required for distillation causing more emissions. In addition, Commercial switchgrass farming might need irrigated water, and the energy required for pumping paves the way for future studies. Although the ABE co-fermentation process is a promising technology, the economic feasibility of the STJ pathway should be assessed by a detailed techno-economic analysis in the further studies.

## 6.6 Conclusions

The novel co-fermentation process for switchgrass presents potential environmental benefits when compared with conventional and emerging bio-jet fuel production pathways. The estimated GWP-100 for switchgrass to jet fuel STJ pathway for the practical case (90% hydrolysis and 80% ABE conversion) with co-fermentation and an ABE yield of 0.321 g ABE/g of sugars was estimated to be 6.5%, 14.2%, 47.5%, 73.8%, and 44.4% lower than STJ with traditional ABE fermentation, HRJ, FTJ (switchgrass), FTJ (coal) and CPJ pathways. An increase in ABE yield by 15.5% with co-fermentation reduced the GWP-100 in the fermentation step by 14.4%. Pathways

that involve biomass as feedstocks including STJ pathway for the ideal case (100% hydrolysis and 95% ABE conversion) with co-fermentation and the STJ pathway ideal case (100% hydrolysis and 95% ABE conversion) using traditional ABE fermentation have 53.8% and 50.8% less GWP-100 than CPJ. Therefore, irrespective of the fermentation route, the STJ pathway results in less GWP-100 impact than CPJ, although the net energy ratio indicates that more research is needed. The FTJ pathway for switchgrass has 5.9% more GWP-100 than CPJ. The co-fermentation process using lignocellulosic feedstock is less energy intensive for production of jet fuel with lower GWP-100 compared to other alternative jet fuel pathways. The incorporation of direct land use change in the current LCA results in lower GWP-100 from STJ pathway compared to FTJ (switchgrass), FTJ (coal), and CPJ pathways. The most sensitive input parameters to the life cycle emissions of STJ pathway include natural gas for distillation and pretreatment and ammonia requirements for farming and conversion. Future research should focus on reducing the energy intensity of these processes.

# CHAPTER VII

## Summary, Conclusions and Future Work

### 7.1 Research Objectives

The present work has developed tools for water quality monitoring and modeling using publicly available satellite datasets, in-situ sensing, and field sampling. In addition to tools developed, in-situ sensing methods by unmanned vehicles demonstrated their capability to collect water quality data for water quality mapping. Both statistical and watershed models are used to analyze surface water quality. This dissertation had the following purposes regarding the current state of water quality monitoring and modeling:

1. Implement unmanned surface vehicles for in-situ monitoring of bathymetry, algal pigments, and other water quality parameters

2. Implement satellite remote sensing of algal levels, turbidity, in freshwater bodies

3. Utilize the data collected using unmanned systems including unmanned surface vehicles, refrigerated auto-sampler, and in-stream sensors to measure the formation of HABs, sedimentation, and nutrients

4. Interpret observed nutrient and sediment loadings using a watershed model

To achieve the research objectives laboratory, field, and computational/modeling exercises were performed including field assessment of Horse Creek Cove and Horse Creek Watershed of Grand Lake in Oklahoma.

## 7.2 Research Accomplishments

The key conclusion of this dissertation is that in-situ data collection in conjunction with multi-spectral imagery produce in-lake water quality models for real time monitoring of lakes and reservoirs. Rapid data collection using unmanned surface vehicles in conjunction with satellite remote sensing and other unmanned systems helps in understanding sedimentation and HAB formations in freshwater bodies. A summary of individual conclusions that can be drawn from each chapter of this dissertation are as follows:

1. Chapter 3: Satellite remote sensing provides multispectral imagery to produce remote sensing tools for near real-time monitoring of Harmful Algal blooms and sediment. Random forest models with multi-spectral satellite imagery can estimate Chlorophyll-a concentrations, Phycocyanin concentrations, and turbidity in reservoirs that has coefficient of determination ($R^2$) of 0.75, 79, 0.82 respectively, with validation data. During tool calibration and data validation it was found that random forest models provide superior performance in estimating in-lake water quality parameters compared to other machine learning models used including multiple regression, and support vector regression.

2. Chapter 4: HSPF integrated with python provided a baseline watershed model for Horse Creek Watershed that explains that observed total runoff is 40% more than the simulated runoff. During the simulation time-frame i.e., from late August 2020 to the end of 2021, all the peak flows during summer and winter are 50% and 37% more than the simulated flows. Incorporation of potential evapotranspiration using crop coefficients and calibrating the watershed model can help understand sediment dynamics and nutrient fluxes responsible for HAB formations in Grand Lake.

3. Chapter 5: Demonstrates the capability of an unmanned surface vehicle to

collect water quality data in Grand Lake Oklahoma. The reliable data collected at the frequency of 1 HZ are used to produce water quality maps on a LIDAR raster with 1m spatial grid. These spatial estimates is ground truth when integrated with satellite images and multi-spectral imagery can produce high-resolution real-time gridded estimates of algal pigments and other water quality parameters.

4. Chapter 6: Additional work that estimates the life cycle global warming potential of a novel switchgrass to jet fuel (STJ) pathway via a novel-cofermentation of sugars and $CO_2$. The estimated GWP-100 for switchgrass to jet fuel STJ pathway for the practical case (90% hydrolysis and 80% ABE conversion) with co-fermentation and an ABE yield of 0.321 g ABE/g of sugars was estimated to be 6.5%, 14.2%, 47.5%, 73.8%, and 44.4% lower than STJ with traditional ABE fermentation, HRJ, FTJ (switchgrass), FTJ (coal) and CPJ pathways.STJ pathway outperforms all the other jet fuel pathways including HRJ, FTJ, and CPJ. Therefore bio-chemical conversion of STJ is potentially environmental friendly than conventional jet fuel production pathways.

## 7.3 Outstanding Research Needs and Future Work

The data obtained from unmanned systems provide reliable water quality data for modeling HABs. However it is important to connect unmanned monitoring systems with microbial molecular analysis to determine the early signs associated with HAB formations. Exploring both unmanned systems for modeling and molecular methods (DNA/RNA extraction) help identify the connection between existing monitoring methods and early molecular signatures for HAB formations. Integrating molecular methods with data intensive water quality monitoring is critical in the longer-term development of mass-deployable monitoring and prediction of HAB formation. Also the multi-pronged approach of monitoring HABs is expected to provide a new technology

for predicting likelihood of HAB outbreaks, and enhanced assessment of the interconnections between nutrients, HAB biomarkers, and HAB formations in freshwater ecosystems.

## References


Administration, U. E. I. (2019). Monthly energy review. *Report No. DOE/EIA-0035 (2020/8)*.

Afroz, R., Masud, M. M., Akhtar, R., and Duasa, J. B. (2014). Water pollution: Challenges and future direction for water resource management policies in malaysia. *Environment and urbanization ASIA*, 5(1):63–81.

Agarwal, S. et al. (2011). Weighted support vector regression approach for remote healthcare monitoring. In *2011 International Conference on Recent Trends in Information Technology (ICRTIT)*, pages 969–974. IEEE.

Aggarwal, S. (2004). Principles of remote sensing. *Satellite remote sensing and GIS applications in agricultural meteorology*, 23(2):23–28.

Ahn, C.-Y., Joung, S.-H., Yoon, S.-K., and Oh, H.-M. (2007). Alternative alert system for cyanobacterial bloom, using phycocyanin as a level determinant. *Journal of microbiology*, 45(2):98–104.

Ahn, Y.-H. and Shanmugam, P. (2006). Detecting the red tide algal blooms from satellite ocean color observations in optically complex northeast-asia coastal waters. *Remote Sensing of Environment*, 103(4):419–437.

Alipour, M., Tavallaey, S. S., Andersson, A. M., and Brandell, D. (2022). Improved battery cycle life prediction using a hybrid data-driven model incorporating linear support vector regression and gaussian. *ChemPhysChem*, 23(7):e202100829.

Allan, M. G., Hamilton, D. P., Hicks, B. J., and Brabyn, L. (2011). Landsat remote

sensing of chlorophyll a concentrations in central north island lakes of new zealand. *International Journal of Remote Sensing*, 32(7):2037–2055.

Alves, C. M., Valk, M., de Jong, S., Bonomi, A., van der Wielen, L. A., and Mussatto, S. I. (2017). Techno-economic assessment of biorefinery technologies for aviation biofuels supply chains in brazil. *Biofuels, Bioproducts and Biorefining*, 11(1):67–91.

Amejecor, K. Y. (2021). *Water Quality and Remote Sensing: a Preliminary Study on the Effects of Harmful Algal Blooms (HABs) on Ponds and Lakes in Upstate New York and Pennsylvania*. PhD thesis, State University of New York at Binghamton.

Anderson, D. M. (2009). Approaches to monitoring, control and management of harmful algal blooms (habs). *Ocean & coastal management*, 52(7):342–347.

Antoniou, M. G., De La Cruz, A. A., and Dionysiou, D. D. (2005). Cyanotoxins: New generation of water contaminants.

Arima, E. Y., Richards, P., Walker, R., and Caldas, M. M. (2011). Statistical confirmation of indirect land use change in the brazilian amazon. *Environmental Research Letters*, 6(2):024010.

Atiyeh, H. K. (2021). System and method of biocatalytic conversion for production of alcohols, ketones, and organic acids. US Patent 11,180,779.

Authority, G. R. D. (2009). Lake hudson shoreline management plan. Technical report, Grand River Dam Authority.

Babin, M., Morel, A., and Gentili, B. (1996). Remote sensing of sea surface sun-induced chlorophyll fluorescence: consequences of natural variations in the optical characteristics of phytoplankton and the quantum yield of chlorophyll a fluorescence. *International Journal of Remote Sensing*, 17(12):2417–2448.

Bailis, R. E. and Baka, J. E. (2010). Greenhouse gas emissions and land use change from jatropha curcas-based jet fuel in brazil. *Environmental science & technology*, 44(22):8684–8691.

Banerjee, B. P., Raval, S., Maslin, T. J., and Timms, W. (2020). Development of a uav-mounted system for remotely collecting mine water samples. *International Journal of Mining, Reclamation and Environment*, 34(6):385–396.

Baral, N. R., Slutzky, L., Shah, A., Ezeji, T. C., Cornish, K., and Christy, A. (2016). Acetone-butanol-ethanol fermentation of corn stover: current production methods, economic viability and commercial use. *FEMS microbiology letters*, 363(6):fnw033.

Barbir, F., Veziroğlu, T., and Plass Jr, H. (1990). Environmental damage due to fossil fuels use. *International journal of hydrogen energy*, 15(10):739–749.

Barrett, D. C. and Frazier, A. E. (2016). Automated method for monitoring water quality using landsat imagery. *Water*, 8(6):257.

Bates, S., Hastie, T., and Tibshirani, R. (2021). Cross-validation: what does it estimate and how well does it do it? *arXiv preprint arXiv:2104.00673*.

Belgiu, M. and Drăguţ, L. (2016). Random forest in remote sensing: A review of applications and future directions. *ISPRS journal of photogrammetry and remote sensing*, 114:24–31.

Bell, T. E. (1995). Remote sensing. *IEEE spectrum*, 32(3):24–31.

Benaman, J., Shoemaker, C. A., and Haith, D. A. (2001). Modeling non-point source pollution using a distributed watershed model for the cannonsville reservoir basin, delaware county, new york. In *Bridging the Gap: Meeting the World's Water and Environmental Resources Challenges*, pages 1–9.

Bencala, K. E. (1983). Simulation of solute transport in a mountain pool-and-riffle stream with a kinetic mass transfer model for sorption. *Water resources research*, 19(3):732–738.

Bennett, L. (2017). Algae, cyanobacteria blooms, and climate change. *The Climate Institute*.

Bicer, Y. and Dincer, I. (2017). Life cycle evaluation of hydrogen and other potential fuels for aircrafts. *International Journal of Hydrogen Energy*, 42(16):10722–10738.

Bicknell, B. R., Imhoff, J. C., Kittle Jr, J. L., Donigian Jr, A. S., and Johanson, R. C. (1997). Hydrological simulation program—fortran user's manual for version 11. *Environmental Protection Agency Report No. EPA/600/R-97/080. US Environmental Protection Agency, Athens, Ga*.

Blackwell, R. J. (1979). *Trophic Classification of Selected Colorado Lakes: Lake Classification Through the Amalgamation of Contact-sensed Data and Digitally Processed Multispectral Scanner Data Acquired by Satellite and Aircraft*, volume 78. Jet Propulsion Laboratory, California Institute of Technology.

Board, O. W. R. (2008). Hydrographic survey of hudson lake. Technical report, Oklahoma Water Resources Board.

Bradford, J. B., Lauenroth, W. K., Burke, I. C., and Paruelo, J. M. (2006). The influence of climate, soils, weather, and land use on primary production and biomass seasonality in the us great plains. *Ecosystems*, 9(6):934–950.

Brezonik, P., Menken, K. D., and Bauer, M. (2005). Landsat-based remote sensing of lake water quality characteristics, including chlorophyll and colored dissolved organic matter (cdom). *Lake and Reservoir Management*, 21(4):373–382.

Brooks, B. W., Lazorchak, J. M., Howard, M. D., Johnson, M.-V. V., Morton, S. L., Perkins, D. A., Reavie, E. D., Scott, G. I., Smith, S. A., and Steevens, J. A.

(2016a). Are harmful algal blooms becoming the greatest inland water quality threat to public health and aquatic ecosystems? *Environmental toxicology and chemistry*, 35(1):6–13.

Brooks, K., Snowden-Swan, L., Jones, S., Butcher, M., Lee, G.-S., Anderson, D., Frye, J., Holladay, J., Owen, J., Harmon, L., et al. (2016b). Low-carbon aviation fuel through the alcohol to jet pathway. In *Biofuels for aviation*, pages 109–150. Elsevier.

Budsberg, E., Crawford, J. T., Morgan, H., Chin, W. S., Bura, R., and Gustafson, R. (2016). Hydrocarbon bio-jet fuel from bioconversion of poplar biomass: life cycle assessment. *Biotechnology for biofuels*, 9(1):1–13.

Cai, H., Canter, C. E., Dunn, J. B., Tan, E., Biddy, M., Talmadge, M., Hartley, D., Searcy, E., and Snowden-Swan, L. (2016). Supply chain sustainability analysis of indirect liquefaction of blended biomass to produce high octane gasoline. Technical report, Argonne National Lab.(ANL), Argonne, IL (United States).

Cai, H., Wang, M., Elgowainy, A., and Han, J. (2012). Updated greenhouse gas and criteria air pollutant emission factors and their probability distribution functions for electricity generating units. Technical report, Argonne National Lab.(ANL), Argonne, IL (United States).

Cao, C., Hu, J., Li, S., Wilcox, W., and Wang, Y. (2009). Intensified fischer–tropsch synthesis process with microchannel catalytic reactors. *Catalysis Today*, 140(3-4):149–156.

Carpenter, D. and Carpenter, S. (1983). Modeling inland water quality using landsat data. *Remote Sensing of Environment*, 13(4):345–352.

Casci, J. L., Lok, C. M., and Shannon, M. D. (2009). Fischer–tropsch catalysis: The

basis for an emerging industry with origins in the early 20th century. *Catalysis today*, 145(1-2):38–44.

Chandra, R. P., Bura, R., Mabee, W., Berlin, d. A., Pan, X., and Saddler, J. (2007). Substrate pretreatment: the key to effective enzymatic hydrolysis of lignocellulosics? *Biofuels*, pages 67–93.

Chen, C., Zhu, J., Beardsley, R. C., and Franks, P. J. (2003). Physical-biological sources for dense algal blooms near the changjiang river. *Geophysical Research Letters*, 30(10).

Chen, J., Zhang, M., Cui, T., and Wen, Z. (2013). A review of some important technical problems in respect of satellite remote sensing of chlorophyll-a concentration in coastal waters. *IEEE journal of selected topics in applied earth observations and remote sensing*, 6(5):2275–2289.

Chen, R., Qin, Z., Han, J., Wang, M., Taheripour, F., Tyner, W., O'Connor, D., and Duffield, J. (2018). Life cycle energy and greenhouse gas emission effects of biodiesel in the united states with induced land use change impacts. *Bioresource technology*, 251:249–258.

Chipman, J. W., Lillesand, T. M., Schmaltz, J. E., Leale, J. E., and Nordheim, M. J. (2004). Mapping lake water clarity with landsat images in wisconsin, usa. *Canadian journal of remote sensing*, 30(1):1–7.

Chislock, M. F., Doster, E., Zitomer, R. A., and Wilson, A. E. (2013). Eutrophication: causes, consequences, and controls in aquatic ecosystems. *Nature Education Knowledge*, 4(4):10.

Chitralekha, S. B. and Shah, S. L. (2010). Application of support vector regression for developing soft sensors for nonlinear processes. *The Canadian Journal of Chemical Engineering*, 88(5):696–709.

Clark, C., Han, J., Burnham, A., Dunn, J., and Wang, M. (2012). Life-cycle analysis of shale gas and natural gas. Technical report, Argonne National Lab.(ANL), Argonne, IL (United States).

Claverie, M., Ju, J., Masek, J. G., Dungan, J. L., Vermote, E. F., Roger, J.-C., Skakun, S. V., and Justice, C. (2018). The harmonized landsat and sentinel-2 surface reflectance data set. *Remote Sensing of Environment*, 219:145–161.

Conti, J., Holtberg, P., Diefenderfer, J., LaRose, A., Turnure, J. T., and Westfall, L. (2016). International energy outlook 2016 with projections to 2040. Technical report, USDOE Energy Information Administration (EIA), Washington, DC (United States ....

Cowx, I. G., Arlinghaus, R., and Cooke, S. J. (2010). Harmonizing recreational fisheries and conservation objectives for aquatic biodiversity in inland waters. *Journal of Fish Biology*, 76(9):2194–2215.

Cox, K., Renouf, M., Dargan, A., Turner, C., and Klein-Marcuschamer, D. (2014). Environmental life cycle assessment (lca) of aviation biofuel from microalgae, pongamia pinnata, and sugarcane molasses. *Biofuels, Bioproducts and Biorefining*, 8(4):579–593.

Crosta, A. P. and MOORE, J. M. (1989). Geological mapping using landsat thematic mapper imagery in almeria province, south-east spain. *International Journal of Remote Sensing*, 10(3):505–514.

Daoud, J. I. (2017). Multicollinearity and regression analysis. In *Journal of Physics: Conference Series*, volume 949, page 012009. IOP Publishing.

Das, S. P. and Padhy, S. (2012). Support vector machines for prediction of futures prices in indian stock market. *International Journal of Computer Applications*, 41(3).

Datt, B. (1998). Remote sensing of chlorophyll a, chlorophyll b, chlorophyll a+ b, and total carotenoid content in eucalyptus leaves. *Remote Sensing of Environment*, 66(2):111–121.

De, D. G., La Torre Ugarte, M. E. W., Shapouri, H., Slinsky, S. P., and De, G. (2003). The economic impacts of bioenergy crop production on us agriculture. *USDA*.

De Jong, W. and Kersten, S. (2013). Biomass, a sustainable energy source for the future-fundamentals of conversion processes. 2013. *Netherlands: Wiley STM*.

de Normalización, O. I. (2006). *ISO 14044: Environmental Management, Life Cycle Assessment, Requirements and Guidelines*. ISO.

De Souza, L. L. P., Lora, E. E. S., Palacio, J. C. E., Rocha, M. H., Renó, M. L. G., and Venturini, O. J. (2018). Comparative environmental life cycle assessment of conventional vehicles with different fuel options, plug-in hybrid and electric vehicles for a sustainable transportation system in brazil. *Journal of cleaner production*, 203:444–468.

Deng, Y., Wu, M., Zhang, H., Zheng, L., Acosta, Y., and Hsu, T.-T. D. (2017). Addressing harmful algal blooms (habs) impacts with ferrate (vi): Simultaneous removal of algal cells and toxins for drinking water treatment. *Chemosphere*, 186:757–761.

Drucker, H., Burges, C. J., Kaufman, L., Smola, A., and Vapnik, V. (1996). Support vector regression machines. *Advances in neural information processing systems*, 9.

Dunbabin, M. and Grinham, A. (2010). Experimental evaluation of an autonomous surface vehicle for water quality and greenhouse gas emission monitoring. In *2010 IEEE International Conference on Robotics and Automation*, pages 5268–5274. IEEE.

Dunbabin, M., Grinham, A., and Udy, J. (2009). An autonomous surface vehicle for water quality monitoring. In *Australasian conference on robotics and automation (ACRA)*, pages 2–4. Citeseer.

Dunn, J. B., Adom, F., Sather, N., Han, J., Snyder, S., He, C., Gong, J., Yue, D., and You, F. (2015). Life-cycle analysis of bioproducts and their conventional counterparts in greet. Technical report, Argonne National Lab.(ANL), Argonne, IL (United States).

Dunn, J. B., Mueller, S., Kwon, H.-y., and Wang, M. Q. (2013). Land-use change and greenhouse gas emissions from corn and cellulosic ethanol. *Biotechnology for biofuels*, 6(1):1–13.

Dunn, J. B., Mueller, S., Wang, M., and Han, J. (2012). Energy consumption and greenhouse gas emissions from enzyme and yeast manufacture for corn and cellulosic ethanol production. *Biotechnology letters*, 34(12):2259–2263.

Elgowainy, A., Han, J., Cai, H., Wang, M., Forman, G. S., and DiVita, V. B. (2014). Energy efficiency and greenhouse gas emission intensity of petroleum products at us refineries. *Environmental science & technology*, 48(13):7612–7624.

Elgowainy, A., Han, J., Wang, M., Carter, N., Stratton, R., Hileman, J., Malwitz, A., Balasubramanian, S. N., et al. (2012). Life-cycle analysis of alternative aviation fuels in greet. Technical report, Argonne National Laboratory.

Ellingsen, L. A.-W., Singh, B., and Strømman, A. H. (2016). The size and range effect: lifecycle greenhouse gas emissions of electric vehicles. *Environmental Research Letters*, 11(5):054010.

Enright, C. (2011). Aviation fuel standard takes flight. *ASTM standardization news*, 39(5).

Epplin, F. M. (1996). Cost to produce and deliver switchgrass biomass to an ethanol-conversion facility in the southern plains of the united states. *Biomass and Bioenergy*, 11(6):459–467.

Esteban, J., McRoberts, R. E., Fernández-Landa, A., Tomé, J. L., and Nsset, E. (2019). Estimating forest volume and biomass and their changes using random forests and remotely sensed data. *Remote Sensing*, 11(16):1944.

Ezeji, T., Qureshi, N., and Blaschek, H. P. (2007). Butanol production from agricultural residues: impact of degradation products on clostridium beijerinckii growth and butanol fermentation. *Biotechnology and bioengineering*, 97(6):1460–1469.

Ferreira, P., Le, D. C., and Zincir-Heywood, N. (2019). Exploring feature normalization and temporal information for machine learning based insider threat detection. In *2019 15th International Conference on Network and Service Management (CNSM)*, pages 1–7. IEEE.

Fike, J. H., Parrish, D. J., Wolf, D. D., Balasko, J. A., Green Jr, J. T., Rasnake, M., and Reynolds, J. H. (2006). Long-term yield potential of switchgrass-for-biofuel systems. *Biomass and bioenergy*, 30(3):198–206.

Finkbeiner, M. (2014). The international standards as the constitution of life cycle assessment: the iso 14040 series and its offspring. In *Background and future prospects in life cycle assessment*, pages 85–106. Springer.

Fischer, G., Prieler, S., van Velthuizen, H., Berndes, G., Faaij, A., Londo, M., and de Wit, M. (2010). Biofuel production potentials in europe: Sustainable use of cultivated land and pastures, part ii: Land use scenarios. *Biomass and bioenergy*, 34(2):173–187.

Flores-Anderson, A. I., Griffin, R., Dix, M., Romero-Oliva, C. S., Ochaeta, G., Skinner-Alvarado, J., Ramirez Moran, M. V., Hernandez, B., Cherrington, E.,

Page, B., et al. (2020). Hyperspectral satellite remote sensing of water quality in lake atitlán, guatemala. *Frontiers in Environmental Science*, 8:7.

Fraser, R. (1998). Hyperspectral remote sensing of turbidity and chlorophyll a among nebraska sand hills lakes. *International journal of remote sensing*, 19(8):1579–1589.

Gao, K., Boiano, S., Marzocchella, A., and Rehmann, L. (2014). Cellulosic butanol production from alkali-pretreated switchgrass (panicum virgatum) and phragmites (phragmites australis). *Bioresource technology*, 174:176–181.

Gao, L. and Li, D. (2015). A review of hydrological/water-quality models. *Frontiers of Agricultural Science and Engineering*, 1(4):267–276.

Garreta, R. and Moncecchi, G. (2013). *Learning scikit-learn: machine learning in python*. Packt Publishing Ltd.

Giannoulis, K., Karyotis, T., Sakellariou-Makrantonaki, M., Bastiaans, L., Struik, P., and Danalatos, N. (2016). Switchgrass biomass partitioning and growth characteristics under different management practices. *NJAS-Wageningen Journal of Life Sciences*, 78:61–67.

González-Márquez, L. C., Torres-Bejarano, F. M., Torregroza-Espinosa, A. C., Hansen-Rodríguez, I. R., and Rodríguez-Gallegos, H. B. (2018). Use of landsat 8 images for depth and water quality assessment of el guájaro reservoir, colombia. *Journal of South American Earth Sciences*, 82:231–238.

Gopalakrishnan, G., Cristina Negri, M., and Snyder, S. W. (2011). A novel framework to classify marginal land for sustainable biomass feedstock production. *Journal of environmental quality*, 40(5):1593–1600.

Graham-Rowe, D. (2011). Agriculture: beyond food versus fuel. *Nature*, 474(7352):S6–S8.

Guretzky, J. A., Biermacher, J. T., Cook, B. J., Kering, M. K., and Mosali, J. (2011). Switchgrass for forage and bioenergy: harvest and nitrogen rate effects on biomass yields and nutrient composition. *Plant and soil*, 339(1):69–81.

Haas, M. J., McAloon, A. J., Yee, W. C., and Foglia, T. A. (2006). A process model to estimate biodiesel production costs. *Bioresource technology*, 97(4):671–678.

Hallam, A., Anderson, I. C., and Buxton, D. R. (2001). Comparative economic analysis of perennial, annual, and intercrops for biomass production. *Biomass and Bioenergy*, 21(6):407–424.

Han, J., Elgowainy, A., Cai, H., and Wang, M. Q. (2013a). Life-cycle analysis of bio-based aviation fuels. *Bioresource Technology*, 150:447–456.

Han, J., Elgowainy, A., Cai, H., and Wang, M. Q. (2013b). Life-cycle analysis of bio-based aviation fuels. *Bioresource technology*, 150:447–456.

Han, J., Tao, L., and Wang, M. (2017). Well-to-wake analysis of ethanol-to-jet and sugar-to-jet pathways. *Biotechnology for Biofuels*, 10(1):1–15.

Harod, R., Eswar, R., and Bhattacharya, B. K. (2021). Effect of surface emissivity and retrieval algorithms on the accuracy of land surface temperature retrieved from landsat data. *Remote Sensing Letters*, 12(10):983–993.

Harrington Jr, J. A., Schiebe, F. R., and Nix, J. F. (1992). Remote sensing of lake chicot, arkansas: Monitoring suspended sediments, turbidity, and secchi depth with landsat mss data. *Remote Sensing of Environment*, 39(1):15–27.

Harvey, B. G. and Meylemans, H. A. (2011). The role of butanol in the development of sustainable fuel technologies. *Journal of Chemical Technology & Biotechnology*, 86(1):2–9.

Higinbotham, J. R., Moisan, J. R., Schirtzinger, C., Linkswiler, M., Yungel, J., and Orton, P. (2008). Update on the development and testing of a new long duration solar powered autonomous surface vehicle. In *OCEANS 2008*, pages 1–10. IEEE.

Hong, W.-C., Dong, Y., Zheng, F., and Lai, C.-Y. (2011). Forecasting urban traffic flow by svr with continuous aco. *Applied Mathematical Modelling*, 35(3):1282–1291.

Humbird, D., Davis, R., Tao, L., Kinchin, C., Hsu, D., Aden, A., Schoen, P., Lukas, J., Olthof, B., Worley, M., et al. (2011). Process design and economics for biochemical conversion of lignocellulosic biomass to ethanol: dilute-acid pretreatment and enzymatic hydrolysis of corn stover. Technical report, National Renewable Energy Lab.(NREL), Golden, CO (United States).

Hunter, S. L., Trevisan, A. R., Villa, J., and Smith, K. A. (2020). Bathymetric map, surface area, and capacity of grand lake o'the cherokees, northeastern oklahoma, 2019. Technical report, US Geological Survey.

Huo, H., Wang, M., Bloyd, C., and Putsche, V. (2009). Life-cycle assessment of energy use and greenhouse gas emissions of soybean-derived biodiesel and renewable fuels. *Environmental science & technology*, 43(3):750–756.

IPCC, A. (2014). Ar5 synthesis report: climate change 2014. *Climate change*.

Iso, E. (2006). 14040: 2006. *Environmental management–Life cycle assessment–Principles and framework*, 2.

Izydorczyk, K., Carpentier, C., Mrówczyński, J., Wagenvoort, A., Jurczak, T., and Tarczyńska, M. (2009). Establishment of an alert level framework for cyanobacteria in drinking water resources by using the algae online analyser for monitoring cyanobacterial chlorophyll a. *Water Research*, 43(4):989–996.

Jdiobe, M. (2020). *Design and Development of a Water Observatory: An Autonomous Environmental Sampling System for In-Situ Sensing of Lakes and Rivers*. PhD thesis, Oklahoma State University.

Johansson, T. B., Patwardhan, A. P., Nakićenović, N., and Gomez-Echeverri, L. (2012). *Global energy assessment: toward a sustainable future*. Cambridge University Press.

Johnson, M. C., Palou-Rivera, I., and Frank, E. D. (2013). Energy consumption during the manufacture of nutrients for algae cultivation. *Algal Research*, 2(4):426–436.

Joyce, S. (2000). The dead zones: oxygen-starved coastal waters. *Environmental health perspectives*, 108(3):A120–A125.

Kalnes, T. N., McCall, M. M., and Shonnard, D. R. (2010). Renewable diesel and jet-fuel production from fats and oils. *Thermochemical conversion of biomass to liquid fuels and chemicals*, pages 468–495.

Khan, F. M. and Zubek, V. B. (2008). Support vector regression for censored data (svrc): a novel tool for survival analysis. In *2008 Eighth IEEE International Conference on Data Mining*, pages 863–868. IEEE.

Khanna, M., Dhungana, B., and Clifton-Brown, J. (2008). Costs of producing miscanthus and switchgrass for bioenergy in illinois. *Biomass and bioenergy*, 32(6):482–493.

King, C. W., Maxwell, J. P., and Donovan, A. (2015). Comparing world economic and net energy metrics, part 2: total economy expenditure perspective. *Energies*, 8(11):12975–12996.

Knudsen, K. E. B. (1997). Carbohydrate and lignin contents of plant materials used in animal feeding. *Animal feed science and technology*, 67(4):319–338.

Koparan, C., Koc, A. B., Privette, C. V., Sawyer, C. B., and Sharp, J. L. (2018). Evaluation of a uav-assisted autonomous water sampling. *Water*, 10(5):655.

Krstajic, D., Buturovic, L. J., Leahy, D. E., and Thomas, S. (2014). Cross-validation pitfalls when selecting and assessing regression and classification models. *Journal of cheminformatics*, 6(1):1–15.

Kumabe, K., Sato, T., Matsumoto, K., Ishida, Y., and Hasegawa, T. (2010). Production of hydrocarbons in fischer–tropsch synthesis with fe-based catalyst: Investigations of primary kerosene yield and carbon mass balance. *Fuel*, 89(8):2088–2095.

Kumar, M., Goyal, Y., Sarkar, A., and Gayen, K. (2012). Comparative economic assessment of abe fermentation based on cellulosic and non-cellulosic feedstocks. *Applied Energy*, 93:193–204.

Kutser, T., Metsamaa, L., Strömbeck, N., and Vahtmäe, E. (2006). Monitoring cyanobacterial blooms by satellite remote sensing. *Estuarine, Coastal and Shelf Science*, 67(1-2):303–312.

Kwiatkowski, J. R., McAloon, A. J., Taylor, F., and Johnston, D. B. (2006). Modeling the process and costs of fuel ethanol production by the corn dry-grind process. *Industrial crops and products*, 23(3):288–296.

Lampert, D. J., Cai, H., Wang, Z., Keisman, J., Wu, M., Han, J., Dunn, J., Sullivan, J. L., Elgowainy, A., and Wang, M. (2015). Development of a life cycle inventory of water consumption associated with the production of transportation fuels. Technical report, Argonne National Lab.(ANL), Argonne, IL (United States).

Lee, H., Wang, J., and Leblon, B. (2020). Using linear regression, random forests, and support vector machine with unmanned aerial vehicle multispectral images to predict canopy nitrogen weight in corn. *Remote Sensing*, 12(13):2071.

Lemus, R., Brummer, E. C., Moore, K. J., Molstad, N. E., Burras, C. L., and Barker, M. F. (2002). Biomass yield and quality of 20 switchgrass populations in southern iowa, usa. *Biomass and Bioenergy*, 23(6):433–442.

Liang, S., Zheng, D., Standley, D. M., Yao, B., Zacharias, M., and Zhang, C. (2010). Epsvr and epmeta: prediction of antigenic epitopes using support vector regression and multiple server results. *BMC bioinformatics*, 11(1):1–6.

Liu, G., Yan, B., and Chen, G. (2013). Technical review on jet fuel production. *Renewable and Sustainable Energy Reviews*, 25:59–70.

Liu, J., Wu, M., and Wang, M. (2009). Simulation of the process for producing butanol from corn fermentation. *Industrial & engineering chemistry research*, 48(11):5551–5557.

Liu, K., Atiyeh, H. K., Pardo-Planas, O., Ezeji, T. C., Ujor, V., Overton, J. C., Berning, K., Wilkins, M. R., and Tanner, R. S. (2015a). Butanol production from hydrothermolysis-pretreated switchgrass: quantification of inhibitors and detoxification of hydrolyzate. *Bioresource technology*, 189:292–301.

Liu, S., Zhu, Q., Guan, Q., He, L., and Li, W. (2015b). Bio-aviation fuel production from hydroprocessing castor oil promoted by the nickel-based bifunctional catalysts. *Bioresource technology*, 183:93–100.

Lobo, P., Whitefield, P. D., Hagen, D. E., Rye, L. J., Blakey, S., Wilson, C. W., Williams, P. I., Christie, S., Raper, D., Uryga-Bugajska, I., et al. (2011). Sae e31 methodology development and association pm emissions characteristics of aircraft apus burning conventional and alternative aviation fuels. *mit*.

Lorenzen, C. J. (1966). A method for the continuous measurement of in vivo chlorophyll concentration. In *Deep Sea Research and Oceanographic Abstracts*, volume 13, pages 223–227. Elsevier.

Lucas, A., Silva, C. A., and Neto, R. C. (2012). Life cycle analysis of energy supply infrastructure for conventional and electric vehicles. *Energy Policy*, 41:537–547.

Lutz, M. (2013). *Learning python: Powerful object-oriented programming.* " O'Reilly Media, Inc.".

Ma, J., Jin, S., Li, J., He, Y., and Shang, W. (2021). Spatio-temporal variations and driving forces of harmful algal blooms in chaohu lake: a multi-source remote sensing approach. *Remote Sensing*, 13(3):427.

Main-Knorn, M., Pflug, B., Louis, J., Debaecker, V., Müller-Wilm, U., and Gascon, F. (2017). Sen2cor for sentinel-2. In *Image and Signal Processing for Remote Sensing XXIII*, volume 10427, page 1042704. International Society for Optics and Photonics.

Mansaray, A. S., Dzialowski, A. R., Martin, M. E., Wagner, K. L., Gholizadeh, H., and Stoodley, S. H. (2021). Comparing planetscope to landsat-8 and sentinel-2 for sensing water quality in reservoirs in agricultural watersheds. *Remote Sensing*, 13(9):1847.

Marsh, G. (2008). Biofuels: aviation alternative? *Renewable Energy Focus*, 9(4):48–51.

Masek, J. G., Wulder, M. A., Markham, B., McCorkel, J., Crawford, C. J., Storey, J., and Jenstrom, D. T. (2020). Landsat 9: Empowering open science and applications through continuity. *Remote Sensing of Environment*, 248:111968.

Mason, A. M. and Triplett, J. R. (2016). Controlling environmental crisis messages in uncontrollable media environments: the 2011 case of blue-green algae on grand lake o'the cherokees, ok. In *Communicating Climate-Change and Natural Hazard Risk and Cultivating Resilience*, pages 189–204. Springer.

Mazlan, N. M., Savill, M., and Kipouros, T. (2015). Effects of biofuels properties on aircraft engine performance. *Aircraft Engineering and Aerospace Technology: An International Journal*.

McLaughlin, S. B. and Kszos, L. A. (2005). Development of switchgrass (panicum virgatum) as a bioenergy feedstock in the united states. *Biomass and bioenergy*, 28(6):515–535.

McLaughlin, S. B. and Walsh, M. (1998). Evaluating environmental consequences of producing herbaceous crops for bioenergy. *Biomass and Bioenergy*, 14(4):317–324.

Millman, K. J. and Aivazis, M. (2011). Python for scientists and engineers. *Computing in Science & Engineering*, 13(2):9–12.

Mittelstet, A. R., Heeren, D. M., Fox, G. A., Storm, D. E., White, M. J., and Miller, R. B. (2011). Comparison of subsurface and surface runoff phosphorus transport rates in alluvial floodplains. *Agriculture, ecosystems & environment*, 141(3-4):417–425.

Moges, M. A., Schmitter, P., Tilahun, S. A., and Steenhuis, T. S. (2018). Watershed modeling for reducing future non-point source sediment and phosphorus load in the lake tana basin, ethiopia. *Journal of soils and sediments*, 18(1):309–322.

Moriarty, K. L., Milbrandt, A. R., Warner, E., Lewis, J. E., and Schwab, A. A. (2018). 2016 bioenergy industry status report. Technical report, National Renewable Energy Lab.(NREL), Golden, CO (United States).

Murphy, R., Tolhurst, T., Chapman, M., and Underwood, A. (2005). Estimation of surface chlorophyll-a on an emersed mudflat using field spectrometry: accuracy of ratios and derivative-based approaches. *International journal of remote sensing*, 26(9):1835–1859.

Nadadoor, V. R., De la Hoz Siegler, H., Shah, S. L., McCaffrey, W. C., and Ben-Zvi, A. (2012). Online sensor for monitoring a microalgal bioreactor system using support vector regression. *Chemometrics and Intelligent Laboratory Systems*, 110(1):38–48.

Nagol, J. R., Sexton, J. O., Anand, A., Sahajpal, R., and Edwards, T. C. (2018). Isolating type-specific phenologies through spectral unmixing of satellite time series. *International Journal of Digital Earth*, 11(3):233–245.

Ni, Y. and Sun, Z. (2009). Recent progress on industrial fermentative production of acetone–butanol–ethanol by clostridium acetobutylicum in china. *Applied microbiology and biotechnology*, 83(3):415–423.

Odetti, A., Bruzzone, G., Altosole, M., Viviani, M., and Caccia, M. (2020). Swamp, an autonomous surface vehicle expressly designed for extremely shallow waters. *Ocean Engineering*, 216:108205.

Ogashawara, I., Mishra, D. R., Mishra, S., Curtarelli, M. P., and Stech, J. L. (2013). A performance review of reflectance based algorithms for predicting phycocyanin concentrations in inland waters. *Remote Sensing*, 5(10):4774–4798.

Olmanson, L. G., Brezonik, P. L., Finlay, J. C., and Bauer, M. E. (2016). Comparison of landsat 8 and landsat 7 for regional measurements of cdom and water clarity in lakes. *Remote Sensing of Environment*, 185:119–128.

O'Shea, R. E., Pahlevan, N., Smith, B., Bresciani, M., Egerton, T., Giardino, C., Li, L., Moore, T., Ruiz-Verdu, A., Ruberg, S., et al. (2021). Advancing cyanobacteria biomass estimation from hyperspectral observations: Demonstrations with hico and prisma imagery. *Remote Sensing of Environment*, 266:112693.

Owens, V., Viands, D., Mayton, H., Fike, J., Farris, R., Heaton, E., Bransby, D., and

Hong, C. (2013). Nitrogen use in switchgrass grown for bioenergy across the usa. *biomass and bioenergy*, 58:286–293.

Oyama, Y., Matsushita, B., and Fukushima, T. (2015). Distinguishing surface cyanobacterial blooms and aquatic macrophytes using landsat/tm and etm+ shortwave infrared bands. *Remote Sensing of Environment*, 157:35–47.

O'Donnell, M., Jaynes, E., and Miller, J. (1981). Kramers–kronig relationship between ultrasonic attenuation and phase velocity. *The Journal of the Acoustical Society of America*, 69(3):696–701.

Paerl, H. W., Gardner, W. S., Havens, K. E., Joyner, A. R., McCarthy, M. J., Newell, S. E., Qin, B., and Scott, J. T. (2016). Mitigating cyanobacterial harmful algal blooms in aquatic ecosystems impacted by climate change and anthropogenic nutrients. *Harmful Algae*, 54:213–222.

Palou-Rivera, I. and Wang, M. Q. (2010). Updated estimation of energy efficiencies of us petroleum refineries. Technical report, Argonne National Lab.(ANL), Argonne, IL (United States).

Park, J., Church, J., Son, Y., Kim, K.-T., and Lee, W. H. (2017). Recent advances in ultrasonic treatment: challenges and field applications for controlling harmful algal blooms (habs). *Ultrasonics sonochemistry*, 38:326–334.

Park, Y. and O'Kelly, M. E. (2014). Fuel burn rates of commercial passenger aircraft: variations by seat configuration and stage distance. *Journal of transport geography*, 41:137–147.

Parrish, D. J. and Fike, J. H. (2005). The biology and agronomy of switchgrass for biofuels. *BPTS*, 24(5-6):423–459.

Pielke Jr, R. A., Landsea, C., Mayfield, M., Layer, J., and Pasch, R. (2005). Hurricanes and global warming. *Bulletin of the American Meteorological Society*, 86(11):1571–1576.

Pierobon, F., Eastin, I. L., and Ganguly, I. (2018). Life cycle assessment of residual lignocellulosic biomass-based jet fuel with activated carbon and lignosulfonate as co-products. *Biotechnology for biofuels*, 11(1):1–18.

Plevin, R. J. (2009). Modeling corn ethanol and climate: a critical comparison of the bess and greet models. *Journal of Industrial Ecology*, 13(4):495–507.

Porcar-Castell, A., Tyystjärvi, E., Atherton, J., Van der Tol, C., Flexas, J., Pfündel, E. E., Moreno, J., Frankenberg, C., and Berry, J. A. (2014). Linking chlorophyll a fluorescence to photosynthesis for remote sensing applications: mechanisms and challenges. *Journal of experimental botany*, 65(15):4065–4095.

Post, W. M. and Kwon, K. C. (2000). Soil carbon sequestration and land-use change: processes and potential. *Global change biology*, 6(3):317–327.

Potes, M., Costa, M. J., and Salgado, R. (2012). Satellite remote sensing of water turbidity in alqueva reservoir and implications on lake modelling. *Hydrology and Earth System Sciences*, 16(6):1623–1633.

Powers, C. W., Hanlon, R., Grothe, H., Prussin, A. J., Marr, L. C., and Schmale III, D. G. (2018). Coordinated sampling of microorganisms over freshwater and saltwater environments using an unmanned surface vehicle (usv) and a small unmanned aircraft system (suas). *Frontiers in Microbiology*, 9:1668.

Prempraneerach, P. and Kulvanit, P. (2010). Autonomous robot boat for water sampling and environmental data acquisition tasks. In *Proceedings of 2010 International Conference on Test and Measurement (ICTM2010)*.

Publications, U. N. (2019). *World Population Prospects 2019: Highlights*. UN.

Qureshi, N., Saha, B. C., Hector, R. E., Dien, B., Hughes, S., Liu, S., Iten, L., Bowman, M. J., Sarath, G., and Cotta, M. A. (2010). Production of butanol (a biofuel) from agricultural residues: Part ii–use of corn stover and switchgrass hydrolysates. *Biomass and bioenergy*, 34(4):566–571.

Rabalais, N. N., Turner, R. E., and Wiseman Jr, W. J. (2002). Gulf of mexico hypoxia, aka" the dead zone". *Annual Review of ecology and Systematics*, pages 235–263.

Rahmes, T., Kinder, J., Crenfeldt, G., LeDuc, G., Abe, Y., McCall, M., Henry, T., Zombanakis, G., Lambert, D., Lewis, C., et al. (2009). Sustainable bio-derived synthetic paraffinic kerosene (bio-spk) jet fuel flights and engine tests program results. In *9th AIAA aviation technology, integration, and operations conference (ATIO) and aircraft noise and emissions reduction symposium (ANERS)*, page 7002.

Ritchie, J. C., Zimba, P. V., and Everitt, J. H. (2003). Remote sensing techniques to assess water quality. *Photogrammetric engineering & remote sensing*, 69(6):695–704.

Sanderson, M. A., Adler, P. R., Boateng, A. A., Casler, M. D., and Sarath, G. (2006). Switchgrass as a biofuels feedstock in the usa. *Canadian journal of plant science*, 86(Special Issue):1315–1325.

Sauer, M. (2016). Industrial production of acetone and butanol by fermentation—100 years later. *FEMS microbiology letters*, 363(13):fnw134.

Scaldaferri, C. A. and Pasa, V. M. D. (2019). Production of jet fuel and green diesel range biohydrocarbons by hydroprocessing of soybean oil over niobium phosphate catalyst. *Fuel*, 245:458–466.

Scheelhaase, J., Maertens, S., Grimme, W., and Jung, M. (2018). Eu ets versus corsia–a critical assessment of two approaches to limit air transport's co2 emissions by market-based measures. *Journal of Air Transport Management*, 67:55–62.

Schneider, W. D. H., dos Reis, L., Fontana, R. C., Dillon, A. J. P., and Camassola, M. (2018). Exploring strategies for the use of glycerol in the production of cellulases and xylanases, and the use of these enzymes in the hydrolysis of lignocellulosic biomass. *Industrial Crops and Products*, 122:114–118.

Schoumans, O., Chardon, W., Bechmann, M., Gascuel-Odoux, C., Hofman, G., Kronvang, B., Rubæk, G., Ulén, B., and Dorioz, J.-M. (2014). Mitigation options to reduce phosphorus losses from the agricultural sector and improve surface water quality: a review. *Science of the Total Environment*, 468:1255–1266.

Schratz, P., Muenchow, J., Iturritxa, E., Richter, J., and Brenning, A. (2019). Hyperparameter tuning and performance assessment of statistical and machine-learning algorithms using spatial data. *Ecological Modelling*, 406:109–120.

Schuman, G., Janzen, H., and Herrick, J. (2002). Soil carbon dynamics and potential carbon sequestration by rangelands. *Environmental pollution*, 116(3):391–396.

Shen, L., Xu, H., and Guo, X. (2012). Satellite remote sensing of harmful algal blooms (habs) and a potential synthesized framework. *Sensors*, 12(6):7778–7803.

Shonnard, D. R., Williams, L., and Kalnes, T. N. (2010). Camelina-derived jet fuel and diesel: Sustainable advanced biofuels. *Environmental progress & sustainable energy*, 29(3):382–392.

Silva, B. B. d., Braga, A. C., Braga, C. C., de Oliveira, L. M., Montenegro, S. M., and Barbosa Junior, B. (2016). Procedures for calculation of the albedo with oli-landsat 8 images: Application to the brazilian semi-arid. *Revista Brasileira de Engenharia Agrícola e Ambiental*, 20:3–8.

Singh, V. P. (1995). *Computer models of watershed hydrology*. Water Resources Publications.

Singh, V. P. and Frevert, D. K. (2010). *Watershed models*. CRC press.

Sokhansanj, S., Mani, S., Turhollow, A., Kumar, A., Bransby, D., Lynd, L., and Laser, M. (2009). Large-scale production, harvest and logistics of switchgrass (panicum virgatum l.)–current technology and envisioning a mature technology. *Biofuels, Bioproducts and Biorefining*, 3(2):124–141.

Song, J., Sciubba, M., and Kam, J. (2021). Risk and impact assessment of dams in the contiguous united states using the 2018 national inventory of dams database. *Water*, 13(8):1066.

Sòria-Perpinyà, X., Vicente, E., Urrego, P., Pereira-Sandoval, M., Ruíz-Verdú, A., Delegido, J., Soria, J. M., and Moreno, J. (2020). Remote sensing of cyanobacterial blooms in a hypertrophic lagoon (albufera of València, eastern iberian peninsula) using multitemporal sentinel-2 images. *Science of the Total Environment*, 698:134305.

Srinivas, R., Singh, A. P., Dhadse, K., and Garg, C. (2020). An evidence based integrated watershed modelling system to assess the impact of non-point source pollution in the riverine ecosystem. *Journal of Cleaner Production*, 246:118963.

Staples, M. D., Malina, R., Olcay, H., Pearlson, M. N., Hileman, J. I., Boies, A., and Barrett, S. R. (2014). Lifecycle greenhouse gas footprint and minimum selling price of renewable diesel and jet fuel from fermentation and advanced fermentation production technologies. *Energy & Environmental Science*, 7(5):1545–1554.

Stratton, R. W., Wong, H. M., and Hileman, J. I. (2011). Quantifying variability in life cycle greenhouse gas inventories of alternative middle distillate transportation fuels. *Environmental science & technology*, 45(10):4637–4644.

Stumpf, R., Culver, M., Tester, P., Tomlinson, M., Kirkpatrick, G., Pederson, B., Truby, E., Ransibrahmanakul, V., and Soracco, M. (2003). Monitoring karenia brevis blooms in the gulf of mexico using satellite ocean color imagery and other data. *Harmful Algae*, 2(2):147–160.

Stumpf, R. P. (2001). Applications of satellite ocean color sensors for monitoring and predicting harmful algal blooms. *Human and Ecological Risk Assessment: An International Journal*, 7(5):1363–1368.

Sun, X., Atiyeh, H. K., Adesanya, Y., Okonkwo, C., Zhang, H., Huhnke, R. L., and Ezeji, T. (2020). Feasibility of using biochar as buffer and mineral nutrients replacement for acetone-butanol-ethanol production from non-detoxified switchgrass hydrolysate. *Bioresource Technology*, 298:122569.

Tanaka, A., Kishino, M., Doerffer, R., Schiller, H., Oishi, T., and Kubota, T. (2004). Development of a neural network algorithm for retrieving concentrations of chlorophyll, suspended matter and yellow substance from radiance data of the ocean color and temperature scanner. *Journal of Oceanography*, 60(3):519–530.

Tang, D. L., Kawamura, H., Doan-Nhu, H., and Takahashi, W. (2004). Remote sensing oceanography of a harmful algal bloom off the coast of southeastern vietnam. *Journal of geophysical research: oceans*, 109(C3).

Tang, Y., Xie, J.-S., and Geng, S. (2010). Marginal land-based biomass energy production in china. *Journal of Integrative Plant Biology*, 52(1):112–121.

Trinh, R. C., Fichot, C. G., Gierach, M. M., Holt, B., Malakar, N. K., Hulley, G., and Smith, J. (2017). Application of landsat 8 for monitoring impacts of wastewater discharge on coastal water quality. *Frontiers in Marine Science*, 4:329.

Tseng, Y.-F., Lin, J., Dai, M., and Kao, S.-J. (2014). Joint effect of freshwater

plume and coastal upwelling on phytoplankton growth off the changjiang river. *Biogeosciences*, 11(2):409–423.

Turhollow, A. (2000). Costs of producing biomass from riparian buffer strips. *ORNL*, 27:4–00.

USEPA (2019). Recommended human health recreational ambient water quality criteria or swimming advisories for microcystins and cylindrospermopsin (epa 822-r-19-001). *USEPA*.

Vaesen, K., Gilliams, S., Nackaerts, K., and Coppin, P. (2001). Ground-measured spectral signatures as indicators of ground cover and leaf area index: the case of paddy rice. *Field Crops Research*, 69(1):13–25.

Valentine, J., Clifton-Brown, J., Hastings, A., Robson, P., Allison, G., and Smith, P. (2012). Food vs. fuel: the use of land for lignocellulosic 'next generation'energy crops that minimize competition with primary food production. *Gcb Bioenergy*, 4(1):1–19.

Vapnik, V. (1999). *The nature of statistical learning theory*. Springer science & business media.

Vapnik, V., Guyon, I., and Hastie, T. (1995). Support vector machines. *Mach. Learn*, 20(3):273–297.

Varvel, G. E., Vogel, K. P., Mitchell, R. B., Follett, R., and Kimble, J. (2008). Comparison of corn and switchgrass on marginal soils for bioenergy. *Biomass and bioenergy*, 32(1):18–21.

Vaughan, N. E., Gough, C., Mander, S., Littleton, E. W., Welfle, A., Gernaat, D. E., and Van Vuuren, D. P. (2018). Evaluating the use of biomass energy with carbon capture and storage in low emission scenarios. *Environmental Research Letters*, 13(4):044014.

Wang, J., Gu, W., and Zhu, J. (2009a). Design of an autonomous surface vehicle used for marine environment monitoring. In *2009 International Conference on Advanced Computer Control*, pages 405–409. IEEE.

Wang, M., Han, J., Dunn, J. B., Cai, H., and Elgowainy, A. (2012). Well-to-wheels energy use and greenhouse gas emissions of ethanol from corn, sugarcane and cellulosic biomass for us use. *Environmental research letters*, 7(4):045905.

Wang, M., Huo, H., and Arora, S. (2011). Methods of dealing with co-products of biofuels in life-cycle analysis and consequent results within the us context. *Energy Policy*, 39(10):5726–5736.

Wang, M., Wu, M., and Huo, H. (2007). Life-cycle energy and greenhouse gas emission impacts of different corn ethanol plant types. *Environmental Research Letters*, 2(2):024001.

Wang, M. Q. et al. (1999). Greet 1.5: Transportation fuel-cycle model. vol. 1: Methodology, development, use, and results. Technical report, Argonne National Laboratory.

Wang, W.-C. and Tao, L. (2016). Bio-jet fuel conversion technologies. *Renewable and Sustainable Energy Reviews*, 53:801–822.

Wang, X., Feng, Y., Wang, H., Qu, Y., Yu, Y., Ren, N., Li, N., Wang, E., Lee, H., and Logan, B. E. (2009b). Bioaugmentation for electricity generation from corn stover biomass using microbial fuel cells. *Environmental science & technology*, 43(15):6088–6093.

Ward, F. A., Roach, B. A., and Henderson, J. E. (1996). The economic value of water in recreation: Evidence from the california drought. *Water Resources Research*, 32(4):1075–1081.

Watanabe, F. S. Y., Alcântara, E., Rodrigues, T. W. P., Imai, N. N., Barbosa, C. C. F., and Rotta, L. H. d. S. (2015). Estimation of chlorophyll-a concentration and the trophic state of the barra bonita hydroelectric reservoir using oli/landsat-8 images. *International journal of environmental research and public health*, 12(9):10391–10417.

Weng, C., Cui, Y., Sun, J., and Peng, H. (2013). On-board state of health monitoring of lithium-ion batteries using incremental capacity analysis with support vector regression. *Journal of Power Sources*, 235:36–44.

Witharana, C. and Civco, D. L. (2014). Optimizing multi-resolution segmentation scale using empirical methods: Exploring the sensitivity of the supervised discrepancy measure euclidean distance 2 (ed2). *ISPRS Journal of Photogrammetry and Remote Sensing*, 87:108–121.

Wu, J., Liu, H.-J., Yan, X., Zhou, Y.-J., Lin, Z.-N., Mi, S., Cheng, K.-K., and Zhang, J.-A. (2019). Efficient catalytic dehydration of high-concentration 1-butanol with zn-mn-co modified $\gamma$-al2o3 in jet fuel production. *Catalysts*, 9(1):93.

Wu, M., Wang, M., Liu, J., Huo, H., et al. (2007). Life-cycle assessment of corn-based butanol as a potential transportation fuel. Technical report, Argonne National Lab.(ANL), Argonne, IL (United States).

Wu, P., Xia, B., and Wang, X. (2015). The contribution of iso 14067 to the evolution of global greenhouse gas standards—a review. *Renewable and Sustainable Energy Reviews*, 47:142–150.

Wullschleger, S. D., Davis, E. B., Borsuk, M. E., Gunderson, C. A., and Lynd, L. (2010). Biomass production in switchgrass across the united states: Database description and determinants of yield. *Agronomy Journal*, 102(4):1158–1168.

Wurtsbaugh, W. A., Paerl, H. W., and Dodds, W. K. (2019). Nutrients, eutrophication and harmful algal blooms along the freshwater to marine continuum. *Wiley Interdisciplinary Reviews: Water*, 6(5):e1373.

Wyman, C. E., Dale, B. E., Elander, R. T., Holtzapple, M., Ladisch, M. R., Lee, Y., Mitchinson, C., and Saddler, J. N. (2009). Comparative sugar recovery and fermentation data following pretreatment of poplar wood by leading technologies. *Biotechnology progress*, 25(2):333–339.

Xie, S. (2012). *Remote Sensing Data Assimilation in Water Quality Numerical Models for Simulation of Water Column Temperature*. PhD thesis, Indiana University.

Xie, X., Wang, M., and Han, J. (2011). Assessment of fuel-cycle energy use and greenhouse gas emissions for fischer- tropsch diesel from coal and cellulosic biomass. *Environmental science & technology*, 45(7):3047–3053.

Xie, Z., Lou, I., Ung, W. K., and Mok, K. M. (2012). Freshwater algal bloom prediction by support vector machine in macau storage reservoirs. *Mathematical problems in engineering*, 2012.

Xu, D., Pu, Y., Zhu, M., Luan, Z., and Shi, K. (2021). Automatic detection of algal blooms using sentinel-2 msi and landsat oli images. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 14:8497–8511.

Xu, Y., Ma, C., Liu, Q., Xi, B., Qian, G., Zhang, D., and Huo, S. (2015). Method to predict key factors affecting lake eutrophication–a new approach based on support vector regression model. *International Biodeterioration & Biodegradation*, 102:308–315.

Xu, Y., Wang, B., Kong, Q.-J., Liu, Y., Wang, F.-Y., Xu, Y., and Wang, F. (2014). Spatio-temporal variable selection based support vector regression for urban traffic

flow prediction. In *Proceedings of the 93rd Annual Meeting of the Transportation Research Board, Washington, DC, USA*, volume 14.

Yacobi, Y., Köhler, J., Leunert, F., and Gitelson, A. (2015). Phycocyanin-specific absorption coefficient: Eliminating the effect of chlorophylls absorption. *Limnology and Oceanography: Methods*, 13(4):157–168.

Yajima, H. and Derot, J. (2018). Application of the random forest model for chlorophyll-a forecasts in fresh and brackish water bodies in japan, using multivariate long-term databases. *Journal of Hydroinformatics*, 20(1):206–220.

Yan, Y., Bao, Z., and Shao, J. (2018). Phycocyanin concentration retrieval in inland waters: A comparative review of the remote sensing techniques and algorithms. *Journal of Great Lakes Research*, 44(4):748–755.

Yang, B. and Wyman, C. E. (2008). Pretreatment: the key to unlocking low-cost cellulosic ethanol. *Biofuels, Bioproducts and Biorefining: Innovation for a sustainable economy*, 2(1):26–40.

Yang, Q. and Chen, G. (2012). Nonrenewable energy cost of corn-ethanol in china. *Energy Policy*, 41:340–347.

YSI (2017). *EXO User Manual*. YSI.

Yu, L., Chen, H., Wang, S., and Lai, K. K. (2008). Evolving least squares support vector machines for stock market trend mining. *IEEE Transactions on evolutionary computation*, 13(1):87–102.

Zarchin, R., Rabaev, M., Vidruk-Nehemya, R., Landau, M. V., and Herskowitz, M. (2015). Hydroprocessing of soybean oil on nickel-phosphide supported catalysts. *Fuel*, 139:684–691.

Zhang, Y. and Colosi, L. M. (2013). Practical ambiguities during calculation of energy ratios and their impacts on life cycle assessment calculations. *Energy Policy*, 57:630–633.

Zhang, Y., Liang, J., Zeng, G., Tang, W., Lu, Y., Luo, Y., Xing, W., Tang, N., Ye, S., Li, X., et al. (2020). How climate change and eutrophication interact with microplastic pollution and sediment resuspension in shallow lakes: A review. *Science of the Total Environment*, 705:135979.

Zohdi, E. and Abbaspour, M. (2019). Harmful algal blooms (red tide): a review of causes, impacts and approaches to monitoring and prediction. *International Journal of Environmental Science and Technology*, 16(3):1789–1806.

# APPENDIX A

## Supplemental Information for "A Remote Sensing Tool for Near Real-Time Monitoring of Harmful Algal Blooms and Turbidity in Reservoirs

Table A.1: Level-2 spectral reflectance band properties of Landsat-8

| Band Information | Data range | Properties |
|---|---|---|
| SR_B1 (Ultra-Blue, Coastal): Aerosol 0.435-0.512 $\mu$m | 1-65455 | Mostly used for imaging shallow water; previously used to monitor CHLa concentrations and suspended solids |
| SR_B2 (Blue): 0.452-0.512 $\mu$m | 1-65455 | Used for bathymetry mapping; spectrum at which light penetrates water the deepest |
| SR_B3 (Green): 0.533-0.590 $\mu$m | 1-65455 | Used for estimating vegetation density; previously used to estimate mangrove chlorophyll concentration |
| SR_B4 (Red): 0.636-0.673 $\mu$m | 1-65455 | Used to distinguish between the anthropogenic and natural environments; absorbed by chlorophyll, which enables quantification of vegetation |
| SR_B5 (Near Infrared): 0.851-0.879 $\mu$m | 1-65455 | Exhibits the high contrast between water and non-water objects |
| SR_B6 (Short Wave Infrared 1): 1.566-1.651 $\mu$m | 1-65455 | Discriminates moisture content of soil and vegetation |
| SR_B7 (Short Wave Infrared 2): 2.107-2.294 $\mu$m | 1-65455 | Detects increased moisture content of soil and vegetation |

Table A.2: Level-2 spectral computed band properties of Landsat-8

| Band Information | Data range | Properties |
|---|---|---|
| ST_B10 (Surface Temperature): Aerosol 10.60-11.19 $\mu$m | 0-65535 | Used for thermal mapping of the surface and estimating soil moisture content |
| SR_QA_AEROSOL (Aerosol Attributes) | NA | Documents cloud and surface state flags |
| ST_ATRAN (Atmospheric Transmittance) | 0-10000 | Indicates the ratio of the radiation transmitted to the total radiation incident upon the atmosphere |
| ST_CDIST (Pixel Distance to Cloud) | 0-24000 | Indicates the distance between a pixel and the nearest cloud pixel in kilometers; can be used with emissivity standard deviation to estimate surface temperature |
| ST_DRAD (Downwelled Radiance) | 0-28000 | Indicates the thermal energy emitted by the atmosphere which reaches the earth's surface and reflects back to the sensor |
| ST_EMIS (Emissivity of band 10) | 0-10000 | Indicates the ratio of the energy radiated from the surface of a material to the energy radiated from a blackbody |
| ST_TRAD (Thermal Surface Radiance) | 0-22000 | Indicates thermal surface radiance, used to detect changes in surface temperature |
| ST_URAD (Upwelled Radiance) | 0-28000 | Indicates the amount of electromagnetic radiation reflected from the earth's surface to the atmosphere |
| QA_PIXEL | NA | Bit combinations that define certain quality conditions of the pixels |
| QA_RADSAT | NA | Indicates which sensor bands are saturated |

Table A.3: Level-2 spectral band properties of Sentinel-2

| **Band Information** | **Wavelength (nm)** | **Band properties** |
|---|---|---|
| Aerosol Optical Thickness (AOT) | NA | Provides the visual transparency of the atmosphere |
| B2 (Green) | 496.6 (S2A) / 492.1 (S2B) | Remotely-senses the vegetation and CHLa concentration in surface water bodies |
| B3 (Blue) | 560 (S2A) / 559 (S2B) | Used to detect the presence of water bodies; affected by atmospheric scattering |
| B4 (Red) | 664.5 (S2A) / 665 (S2B) | Used for remotely sensing turbidity of lakes and reservoirs |
| B5, B6, B7, B8A (Red Edge 1, Red Edge 2, Red Edge 3, Red Edge 4) | 703.9 (S2A) / 703.8 (S2B)<br>740.2 (S2A) / 739.1 (S2B)<br>782.5 (S2A) / 779.7 (S2B) | Used for detecting both CHLa and phycocyanin in surface water bodies |
| B11 (Short Wave Infrared 1) | 1613.7 (S2A) / 1610.4 (S2B) | Differentiates cyanobacterial blooms in water compared to aquatic plants; not affected by scattering in turbid waters |
| B12 (Short Wave Infrared 2) | 2202.4 (S2A) / 2185.7 (S2B) | |
| SCL (Scene Classification Map) | NA | Detects clouds and their shadows, snow to generate a classification map |
| TCI (True Color Image) | NA | Used to detect vegetation, algal blooms but the band relies heavily on solar reflectance which makes it not usable if the image is taken during night |
| WVP (Water Vapor Pressure) | NA | Used to track lower-level moisture, identify regions with turbulence, highlight volcanic plumes |

# APPENDIX B

## Supplementary Data: Well-to-wake Analysis of Switchgrass to Jet Fuel Via a Novel Co-fermentation of Sugars and CO2

### 2.1 Introduction

As air traffic grows steadily, the aviation industry needs alternative jet fuel production technologies which can provide lower life cycle emissions compared to petroleum-based fuels. This study presents a life cycle assessment (LCA) for switchgrass to jet fuel (STJ) production pathway and compares its environmental impact with other emerging jet fuel production pathways including corn to jet fuel (CTJ), hydroprocessed renewable jet fuel (HRJ), Fischer Tropsch jet fuel using both lignocellulosic biomass and coal (FTJ), and conventional petroleum jet fuel (CPJ) using 100-year global warming potential (GWP-100). This supplemental data file provides a more detailed description of the assumptions used for both the STJ pathway and other alternative fuel supplies described in the manuscript.

### 2.2 Overview of the co-fermentation process of sugars/off-gas

Butanol is produced by traditional acetone-butanol-ethanol (ABE) fermentation. However, butanol can also be produced via co-fermentation of sugars and gas ($CO_2/H_2$) produced during fermentation as explained in Figure B.1. (Atiyeh, 2021). The novel co-fermentation can be performed in two stages, one for sugar fermentation and one for gas fermentation or in one stage with a co-culture. The use of Clostridium beijerinckii ($C_b$) and Clostridium carboxidivorans ($C_c$) has shown to increase butanol yield from sugars and off-gas (Atiyeh, 2021). A detailed LCA has been performed to account for the energy and life cycle emissions of lignocellulosic biomass to jet fuel pathway using data from laboratory studies, GREET, and literature. Energy and emissions required in the major unit processes of switchgrass to jet fuel (STJ) has been assessed using GREET. While preliminary data of ABE yield is obtained from laboratory results, energy requirements of cofermentation are estimated based on reasonable assumptions from the literature

### 2.3 Data used in process level inputs/outputs

#### 2.3.1 Data used in fertilizer production

Process level data used in fertilizer production for switchgrass include the input quantities of nitrogen (N), phosphorus pentoxide ($P_2O_5$) and potassium oxide ($K2O$). Nitrogen data were taken from nitrogen average mix in GREET which represents US

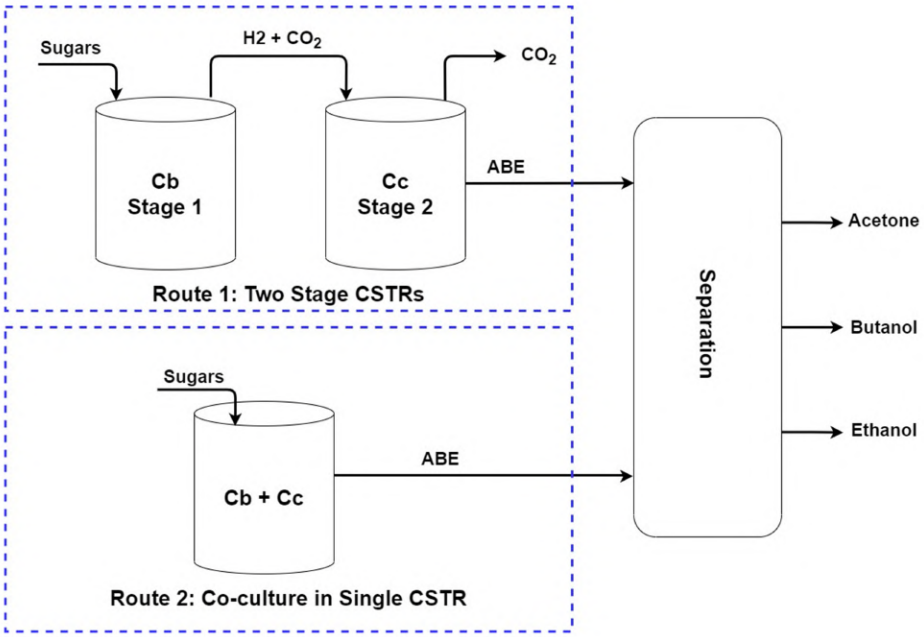


Figure B.1: Routes for butanol production using co-fermentation of sugars/off-gas (Atiyeh, 2021)

average data. Similarly, $P_2O_5$ data was from mono-ammonium phosphate and di-ammonium phosphate mix in GREET. $K_2O$ was from Potassium oxide production pathway in GREET. The output type and output needed as a fertilizer in switchgrass farming is shown in Table B.1. In fertilizer production, $CO_2$ equivalent for inputs in switchgrass farming are taken as GWP-100 in the fertilizer production process

Table B.1: Fertilizer output needed for 1 US ton of switchgrass farming

| Input type | Output type | Output (g) |
|---|---|---|
| Natural gas | Nitrogen (N) | *4877.35 |
| Natural gas | phosphorus pentoxide (P2O5) | *2307.81 |
| Natural gas | Potassium Oxide ($K_2O$) | *3200.42 |

*Data taken from (Cai et al., 2016)

### 2.3.2 Data used in switchgrass farming and hydrolysis

In switchgrass farming all $N_2O$ emissions are assumed to be coming from nitrogen input and N conversion i.e. %kg $N_2O$ of N was obtained from GREET excel. To estimate $N_2O$ emissions from nitrogen input law of conservation of mass is used. After harvesting switchgrass, it was pretreated with sulfuric acid. The pretreated switchgrass is hydrolyzed using cellulase as the enzyme to produce sugars. Main sugars produced from hydrolysis were glucose and xylose. Data used in hydrolysis to produce 1 US ton of sugar from switchgrass is shown in Table B.2, where $CO_2e$ equivalent for all inputs in the production process were taken from GREET.Net 2018.

Table B.2: Data used in switchgrass hydrolysis (per 1 US ton of sugars) *

| Input type | Input | Energy inputs (MJ) | GWP-100 (kg) |
|---|---|---|---|
| Switchgrass, kg | 1306 kg | *21949 | - |
| Natural gas | 105.5 kg | *5302 | 298.47 |
| Electricity, kWh | 70.34 kWh | *525 | 33.44 |
| Ammonia, kg | 18.14 kg | *754 | 42.41 |
| Cellulase, kg | 9.071 kg | *721 | 15.54 |
| Sulfuric acid, kg | 27.22 kg | *17 | 1.16 |

*Data taken from (Dunn et al., 2015)

### 2.3.3 Data and Equations used in co-fermentation of sugars and H2/CO2 gas

Sugars are fermented using solvent producing strains such as Clostridium acetobutylicum and produced $H_2/CO_2$ is co-fermented by gas fermenting strains such as Clostridium carboxidivorans. The co-fermentation yields acetone, butanol and ethanol in a molar ratio of 3:7.3:1. Therefore, butanol was produced both from sugars and from $H_2, CO_2$. Due to more $CO_2$ production than $H_2$ during sugar fermentation, 7.7 moles of $CO_2$ are wasted from each 10 moles of $CO_2$ produced from sugars based on the stoichiometry in Equations 6.1 and 6.2.

### 2.3.4 Data used in distillation

Distillation is required to separate the ABE (acetone-butanol-ethanol) mixture produced from the ABE fermentation. To estimate butanol production from 1 US ton of switchgrass input, the law of conservation of mass was used and the required data needed is the energy required for distillation and density of the products and co-products produced as shown in Table B.3.

Table B.3: Data used in distillation with and without heat recovery

| Input type | Energy Input | Reference |
|---|---|---|
| Natural Gas | 9.4 MJ/kg of ABE produced (without heat recovery) | (Baral et al., 2016) |
| Natural Gas | 4.8 MJ/kg of ABE produced (with heat recovery) | (Baral et al., 2016) |
| Density of Butanol | 0.81 kg/L | (Wu et al., 2007) |
| Density of Acetone | 0.791 kg/L | (Wu et al., 2007) |
| Density of Ethanol | 0.789 kg/L | (Wu et al., 2007) |

### 2.3.5 Data variability

- In switchgrass hydrolysis data variability with co-product (electricity) and without coproduct displacement is used to estimate $CO_2$ equivalent emissions.
- In the ABE fermentation process, both cases based on butanol produced by cofermentation using $H_2$ and $CO_2$ and with traditional ABE fermentation without use of $H_2, CO_2$ are used to estimate $CO_2$ emissions.
- In the distillation process, the energy required to separate acetone, butanol and ethanol is used to estimate $CO_2$ emissions with and without heat recovery as shown in Figure B.6.

## 2.4 Other jet fuel production pathways

### 2.4.1 Other jet fuel production pathways

The corn to jet fuel (CTJ) pathway includes corn farming, conversion of the corn grain to butanol, conversion of butanol to jet fuel, and fuel consumption. Butanol was produced using ABE fermentation and converted to jet fuel using dehydration, isomerization, oligomerization, and hydrogenation steps (Wu et al., 2019). Emissions from corn farming were generated from energy inputs, including diesel for farming operations, electricity, and agricultural chemicals (Han et al., 2017; Wang et al., 2007). Transportation of harvested corn grain to the bio-refinery was neglected for consistency with the other pathways, and because it is likely to be small compared to other emissions (De Souza et al., 2018; Ellingsen et al., 2016; Lucas et al., 2012). Energy for corn processing (dry milling) to generate butanol was assumed to come from NG boilers, and the electricity comes from the US grid. Data for corn farming and processing to produce sugars was taken from an Aspen Plus simulation model for a corn dry mill (Kwiatkowski et al., 2006). Similarly, data for processing sugars to butanol, with an ABE yield of 0.465 g ABE/g sugars, were obtained from the Aspen simulation model to analyze life cycle emissions from well to pump (Wu et al., 2007). Similarly, a practical case of CTJ where the ABE yield was assumed to be 0.374 g ABE/g sugars was assessed in the final comparison of GWP100 based on the yield data from previous literature (Liu et al., 2009). Most of the emissions from processing corn to butanol come from the consumption of NG. GWP-100 from farming to butanol and jet fuel production were estimated from GREET. GWP-100 in jet fuel combustion steps were calculated based on the fuel consumption rate of Boeing 747-400.

### 2.4.2 Hydro-processed renewable (HRJ) jet fuel

Bio-oils, including soy, canola, camelina, jatropha, palm, algae, rapeseed, waste cooking oil, and tallow, can be converted into jet fuels via hydro-processing (Han et al., 2013b). In the present study, only the soybean feedstock was considered for HRJ because it is the primary source for biodiesel and jet fuel production in the US (Haas et al., 2006). The steps in the HRJ pathway include farming, soybean oil extraction,

conversion to jet fuel, and fuel consumption (i.e., from WTWa). Soybean farming requires chemical inputs that generate emissions that were estimated from GREET. GREET was also used to estimate emissions from NG, coal, and electricity consumption for extracting soy oil (Chen et al., 2018; Huo et al., 2009). Soy oil to jet fuel conversion involves hydrodeoxygenation, hydro-isomerization, and hydrocracking to meet jet fuel specifications (Kalnes et al., 2010). Emissions were accounted for from soy oil to jet fuel conversion using GREET. GWP-100 data from soybean farming to jet fuel production steps were obtained from GREET, and in the final step, combustion emissions were estimated similarly to STJ pathway.

### 2.4.3 Fischer-Tropsch jet (FTJ) fuel from coal and lignocellulosic biomass

Fischer-Tropsch technology converts synthetic gas (syngas) generated from feedstocks such as coal and lignocellulosic biomass to liquid transportation fuels (Liu et al., 2013). In FTJ from coal, the feedstock was assumed to be a mixture of bituminous, subbituminous, lignite, and anthracite. The lignocellulosic biomass used in the current LCA is switchgrass that can produce jet fuel via the FTJ pathway. Both feedstocks, including coal and switchgrass, were considered individually in the current LCA with the data used to assess and estimate emissions in the FTJ pathway were obtained from the literature (Xie et al., 2011). Jet fuel production via the FTJ pathway involves the production of syngas, removal of impurities from syngas streams, and fuel refining to produce jet fuel (Casci et al., 2009). These feedstocks can be fed to a gasifier to produce syngas (a mixture of CO, $H_2$, and $CO_2$). The syngas was assumed to be converted to jet fuel using Co, Fe, and $\gamma$ alumina-based catalysts (Cao et al., 2009; Casci et al., 2009; Kumabe et al., 2010). FTJ fuels do not produce sulfur dioxide and emit less particulate matter than conventional or petroleum jet (CPJ), which make them attractive as an alternative fuel source (Lobo et al., 2011). Due to coal usage as a feedstock and fuel processing, FTJ fuel from coal involves energy inputs in mining, such as diesel, electricity, NG, that contribute to significant GHG emissions. Similarly, emissions from LB gasification to FTJ fuel results from farming, diesel, electricity, and NG usage. In the FTJ from the coal pathway, $CO_2$ produced from gasification can be released to the atmosphere or sequestered. In the present study, the case where $CO_2$ was released to the atmosphere during the FTJ fuel processing from syngas was considered because none of the other pathways analyzed potential carbon sequestration strategies. Biogenic $CO_2$ is credited in FTJ from the switchgrass pathway in the current LCA as done in previous studies (Elgowainy et al., 2012; Xie et al., 2011). Emission data for FTJ from coal and switchgrass pathways were obtained from GREET. However, emissions from the final combustion step were estimated, similar to STJ pathway.

### 2.4.4 Conventional petroleum Jet (CPJ) fuel pathway

Conventional or petroleum jet (CPJ) fuel involves crude oil recovery, fuel processing, and fuel consumption. Emissions from the conventional crude oil jet fuel pathway were generated during recovery and refining processes (Palou-Rivera and Wang, 2010).

Crude oil for traditional jet fuel was assumed to be recovered from domestic sources in the US and international sources such as Latin America, Canada, and the Middle East. Among the foreign sources, the production of crude oil from Canadian oil sands contributes substantially to emissions compared to conventional crude recovery because of the GHG emissions involved in the production technologies that include oil sand recovery and refining (Cai et al., 2016). The estimates of GHG emissions were based on energy consumption in jet fuel refined from crude oil using data from 43 refineries representing 70% of the US refining sources (Elgowainy et al., 2014). The authors estimated emissions from the material, energy inputs, and outputs of CPJ pathway using GREET.

## 2.5 Results

### 2.5.1 CO2 emissions in fertilizer production

Based on the system boundary in Figure 6.2 of the main manuscript, $CO_2$ emissions in STJ pathway via co-fermentation from each unit processes are calculated. In fertilizer production based on the output required from 1 US ton of switchgrass farming, $CO_2$ equivalent emissions are taken from GREET and shown in the Table B.4. Comparison of $CO_2$ emissions due to fertilizer inputs in switchgrass farming are shown in Figure B.2.

Table B.4: $CO_2$ emissions from fertilizer production

| Input type | Output type | Output (g) | GWP-100 (kg) |
|---|---|---|---|
| Natural Gas | Nitrogen (N) | 4877.35 | 18.5 |
| Natural Gas | Phosphorus (pentoxide (P2O5)) | 2307.81 | 1.89 |
| Density of Butanol | Potassium (Oxide (K2O) | (K2O) | 1.84 |

Total $CO_2$ emissions from all the fertilizers = 18.5+1.89+1.84 = 22.23 kg $CO_2$/US ton of switchgrass

### 2.5.2 N2O emissions in switchgrass farming

In switchgrass farming $N_2O$ emissions (kg)= $Fertilizerinput(kg) \times \frac{(14gN)}{(17g(NH_3)}$ $\times\ Nconversion \times \frac{(44gN_2O)}{(28gN_2)}$. Where percentage N conversion to N2O is 1.2% (Wang et al., 2012). $N_2O$ emissions = 0.077 kg = 77 g

### 2.5.3 CO2 emissions in switchgrass hydrolysis

Data taken from GREET.Net and Dunn et al. (2015) are for 1 US ton of sugar production from switchgrass. In this study 1 US ton of switchgrass produces 0.69 US ton of total sugar, which is assumed to be all glucose. Based on the glucose yield,

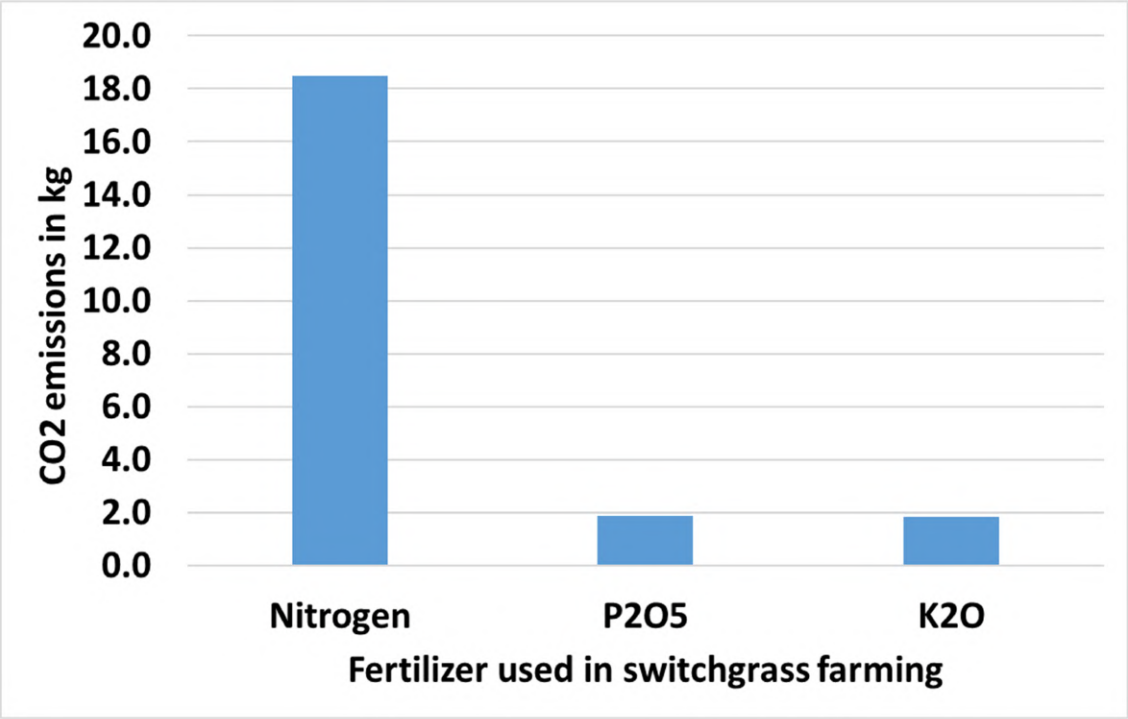


Figure B.2: Comparison of $CO_2$ from different inputs in fertilizer production of STJ pathway via co-fermentation

$CO_2$ emissions are estimated as shown in Figure B.3. Major inputs that contribute to $CO_2$ emissions in switchgrass hydrolysis are shown in Table B.6.

Table B.5: $CO_2$ emissions from switchgrass hydrolysis

| Input type | Input | Energy for inputs (MJ) | GWP-100 (kg) |
|---|---|---|---|
| Switchgrass | 1 US Ton | - | - |
| Natural Gas | 72.8 kg | 3774.15 | 289.57 |
| Electricity | 174.71 MJ | 525 | 23.07 |
| Ammonia | 12.52 kg | 754 | 29.26 |
| Cellulase | 6.26 kg | 721 | 10.72 |
| Sulfuric acid | 18.78 kg | 17 | 0.8 |

Total $CO_2$ emissions from switchgrass hydrolysis without co-product (electricity) displacement = 269.8 kg

Based on previous study, 1 US ton of sugar production displace 2 MMBtu of electricity (Dunn et al., 2015). Based on the calculations, electricity displaced in production of 0.69 US ton of sugar is 1.38 MMBtu. As electricity is considered as a displaced resource, $CO_2$ emissions associated with the electricity displaced are 191.37 kg. If the electricity is not considered as a displaced resource, then $CO_2$ emissions would be more as shown in Figure B.4. In this study, emissions with co-product displacement are considered. Emissions with co-product displacement in switchgrass

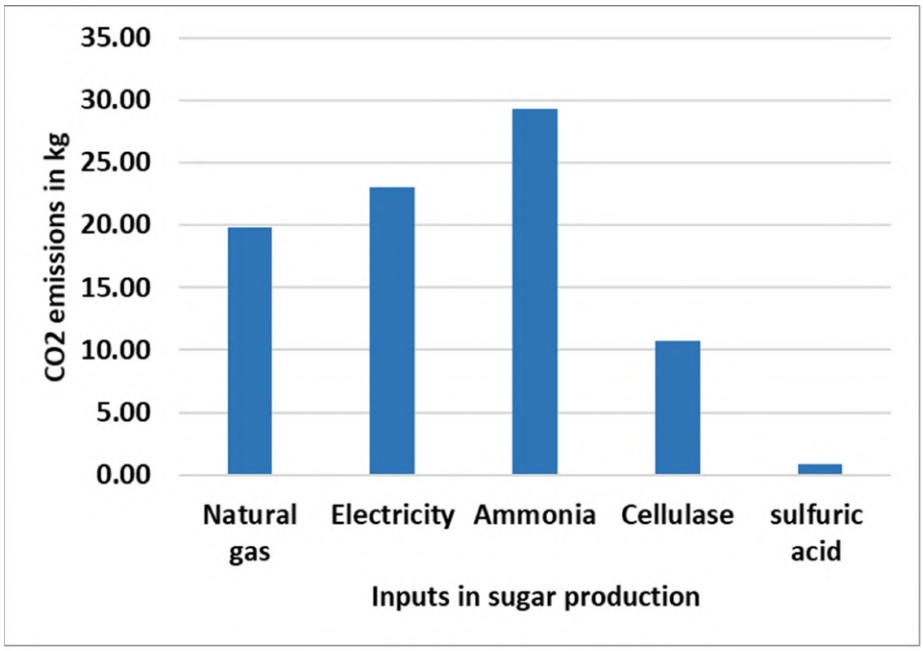


Figure B.3: Comparison of $CO_2$ from different inputs in switchgrass hydrolysis of STJ via co-fermentation

hydrolysis = 269.8-191.37= 78.43 kg.

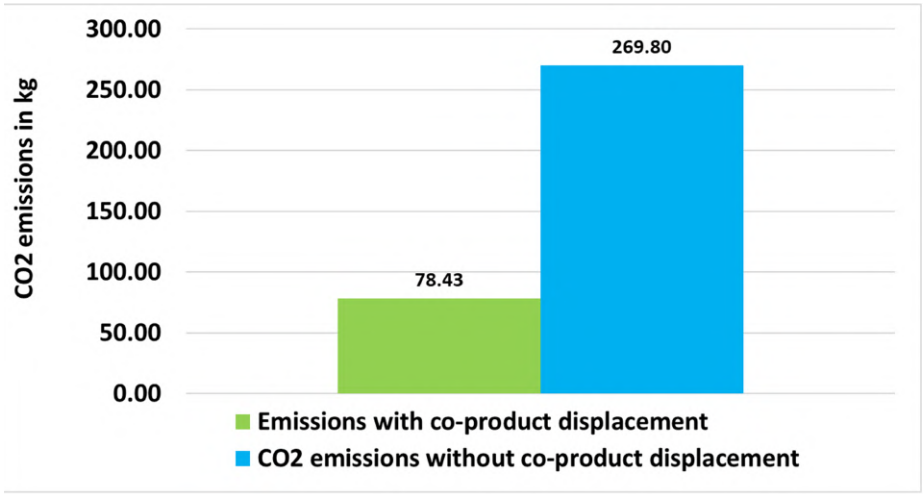


Figure B.4: Comparison of $CO_2$ from switchgrass hydrolysis with and without co-product (electricity) displacement in STJ pathway via co-fermentation

### 2.5.4 CO2 emissions in both traditional ABE and co-fermentation of sugars, H2/CO2

From stoichiometry in equations 6.1 and 6.2 used for the co-fermentation and traditional ABE fermentation, the ABE yield with and without $H_2$ and $CO_2$ conversion are estimated. The use of $H_2$ and $CO_2$ for butanol production generates less emissions

as seen in Table B.7. and Figure B.5. In this study, the $CO_2$ emissions resulting from a yield of 0.374 g ABE per g of sugar are considered.

$$69C_6H_{12}O_6 \rightarrow 36C_4H_{10}O + 18C_3H_6O + 6C_2H_4O_2 + 12C_2H_4O_2 + 6C_4H_8O_2 + H_2O + 156CO_2 + 108H_2 \quad (2.1)$$

$$108H_2 + 36CO_2 \rightarrow 9C_4H_{10}O + 36H_2O \quad (2.2)$$

Table B.6: $CO_2$ process level emissions from ABE fermentation with and without $H_2$ and $CO_2$ conversion.

| ABE yield (g ABE/g sugar) | ABE produced (from 0.69 US ton of sugar) | GWP-100 (kg) |
|---|---|---|
| 0.374 (with $H_2$ and $CO_2$ use) | 222.8 kg | 5.28 |
| 0.321 (without $H_2$ and $CO_2$ use) | 191.3 kg | 6.52 |

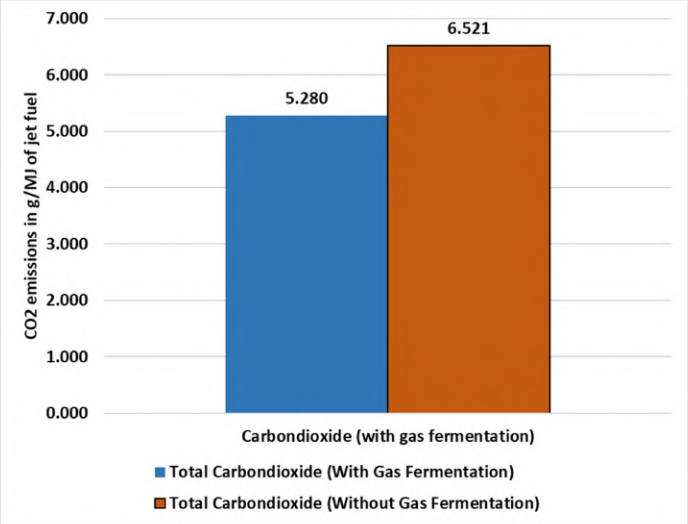


Figure B.5: Comparison of $CO_2$ with and without $H_2$ and $CO_2$ conversion to butanol in STJ pathway

### 2.5.5 CO2 emissions in distillation process

In the distillation process to separate ABE mixture, the $CO_2$ generated is estimated based on the natural gas usage. $CO_2$ emissions without heat recovery are 115.22 kg and with heat recovery are around 58.9 kg as shown in Figure B.6. In this study, $CO_2$ emissions both with and without heat recovery are considered.

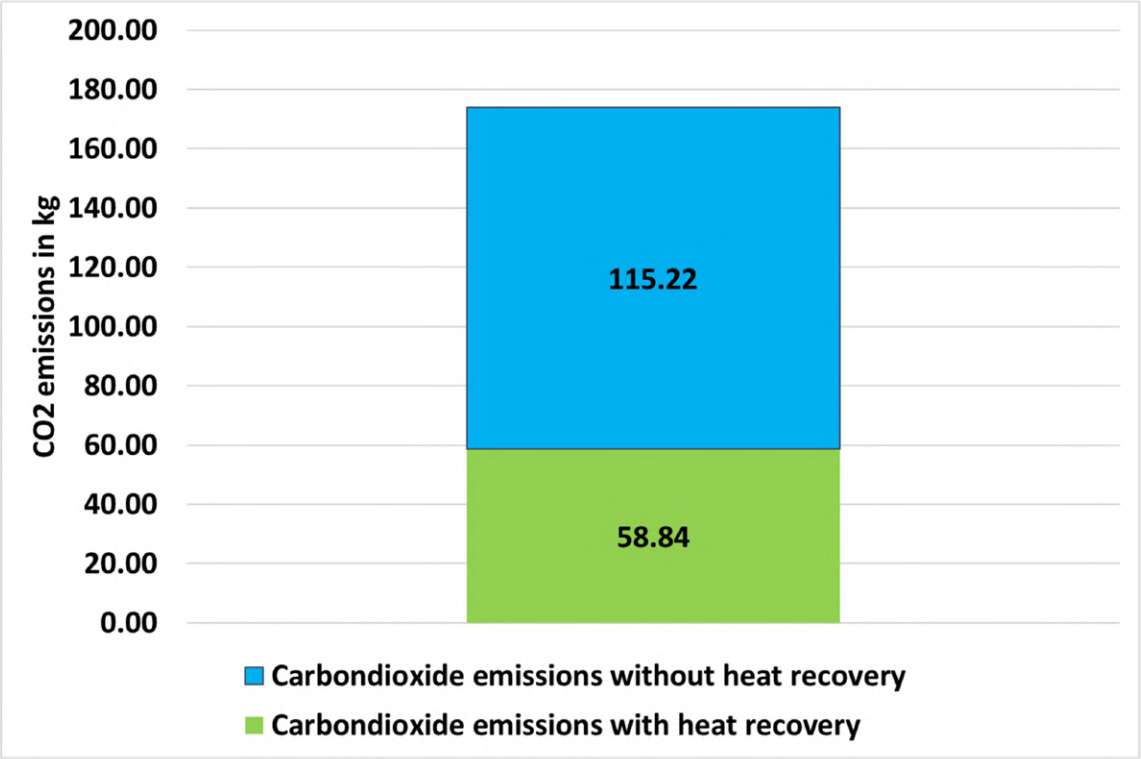


Figure B.6: Comparison of $CO_2$ emissions in distillation with and without heat recovery in STJ pathway via co-fermentation

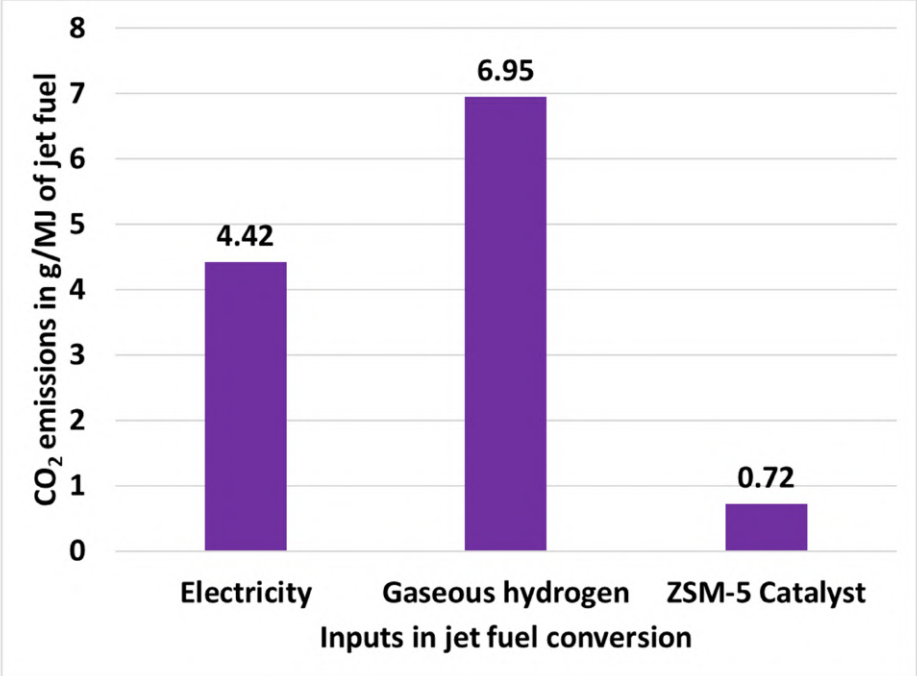


Figure B.7: $CO_2$ emissions in butanol to jet fuel conversion in STJ pathway via co-fermentation

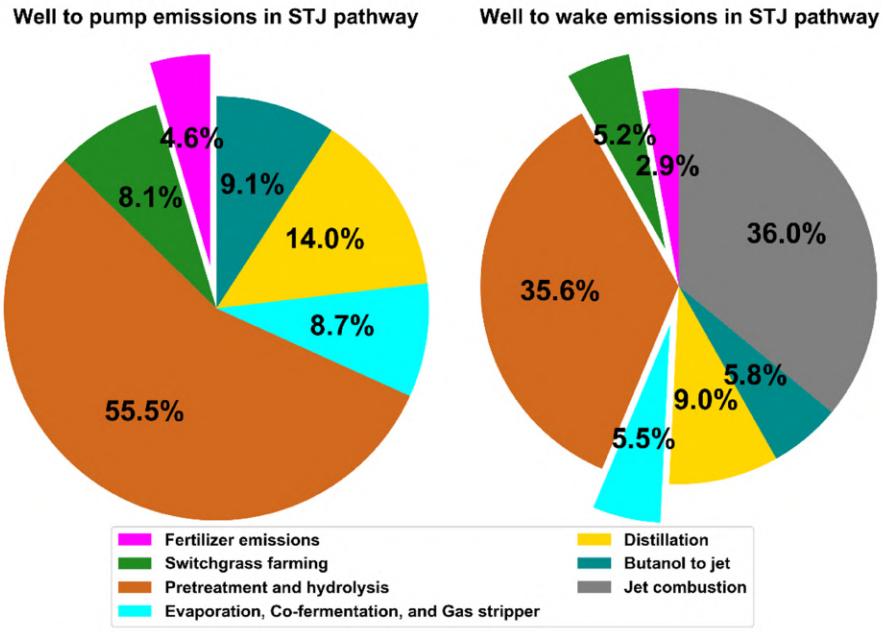


Figure B.8: Emissions in different unit processes of STJ pathway (ideal case) for both WTP and WTWa scenarios

### 2.5.6 Emissions in CTJ pathway

The pre-allocated GWP-100 in corn farming and feedstock to jet fuel conversion from WTP of CTJ pathway were 25.1 g $CO2e$/MJ and 65.8 g $CO2e$/MJ of jet fuel, respectively, as shown in Table 6.2. Similarly, in CTJ with 80% ABE of the simulated yield of 0.465, the WTP emissions were 30.1 g $CO2e$/MJ and 67 g $CO2e$/MJ of jet fuel. Therefore, 72.3% and 69% of GWP-100 from WTP in the CTJ and CTJ (80% ABE, practical case) pathways come from corn to jet fuel conversion. Most of the fossil emissions in the corn to jet fuel conversion step were from the NG. Jet fuel was assumed to be produced from butanol with 1 MJ of jet fuel requiring 1.49 MJ of butanol. Energy inputs needed to produce 1.49 MJ of butanol in the processing of jet fuel from butanol were NG and electricity. Energy contribution in the form of NG was 95.5% and 4.5% in the form of electricity to produce butanol from corn (Wu et al., 2007). For STJ and CTJ pathways, $CO_2$ emissions from nitrogen fertilizer production and $N_2O$ emissions in farming play a significant role in the GWP-100. Corn farming contributes to 15.1% in CTJ and 17.5% in CTJ (80% ABE, practical case) of the total life cycle GHG emissions from WTWa primarily due to the production and application of nitrogen fertilizer. The preallocated GWP-100 from switchgrass farming practical case was 21.5% and 54% lower than for corn in CTJ and soybean farming, respectively (Table 6.2). The main reason for less emission contribution was due to less nitrogen fertilizer application in STJ. Therefore, improved fertilizer application efficiency is a crucial strategy for reducing the environmental impact of bioenergy pathways. However, the pre-allocated GWP-100 for feedstock to jet fuel conversion in the STJ pathway was over 110.9%, 106.2% higher than in CTJ and HRJ, respectively, due to the large GWP-100 due to pretreatment and hydrolysis

process that contributes to almost 57.4% of total emissions in the STJ pathway as shown in Figure B.3.

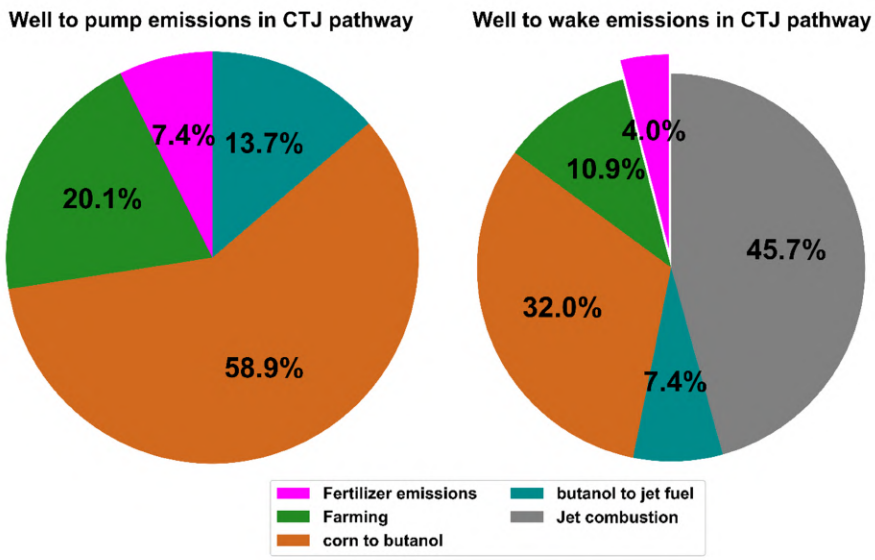


Figure B.9: Emissions in different unit processes of CTJ pathway for both WTP and WTWa scenarios

### 2.5.7 Emissions in HRJ pathway

Emissions in the HRJ pathway result from farming, oil extraction, fuel processing, and combustion steps from WTWa. The WTP pre-allocated GWP-100 in the HRJ pathway includes farming and soybean to jet fuel conversion steps, as shown in Table 6.2. As previously discussed, GWP-100 from farming was mainly from nitrogen fertilizer production and usage, similar to STJ and CTJ pathways. The pre-allocated GWP-100 from soybean farming was 51.4 g $CO_2e$/MJ of jet fuel, which contributed 43.3% of total emission from WTP, as shown in Table 6.2. Emissions from oil extraction mainly depend on the process efficiency and oil content available in the seeds. Emissions in the oil extraction process were 15.8 g $CO_2e$/MJ, while energy-based allocation was used for the soy meal co-product. $CO_2$ released in fuel processing or soy oil to jet fuel conversion were 67.3 g $CO_2e$/MJ. Emissions from the oil extraction were from energy consumption, including coal or NG.

### 2.5.8 Biogenic CO2 credit in biofuel pathways

We accounted for the biogenic $CO_2$, which is identical to the amount of $CO_2$ combusted in the jet,which cancels each other. The biogenic $CO_2$ credit is shown in Figure B.14. Based on the literature,the estimation of carbon sequestration in bioenergy crops is difficult due to large variability of soil types and land management practices Post and Kwon (2000) and the application of variable amount of nitrogen fertilizers at different farming locations (Schuman et al., 2002). Further research is

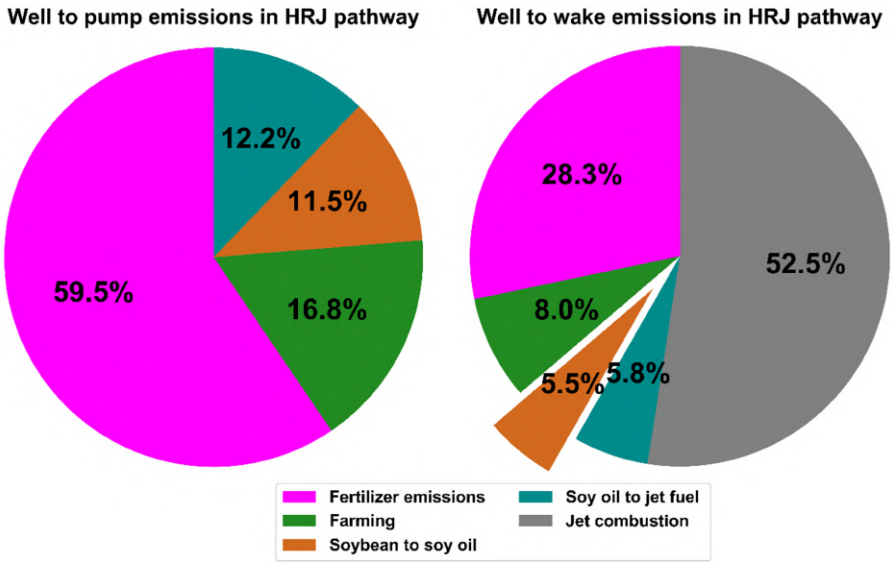


Figure B.10: Emissions in different unit processes of HRJ pathway for both WTP and WTWa scenarios

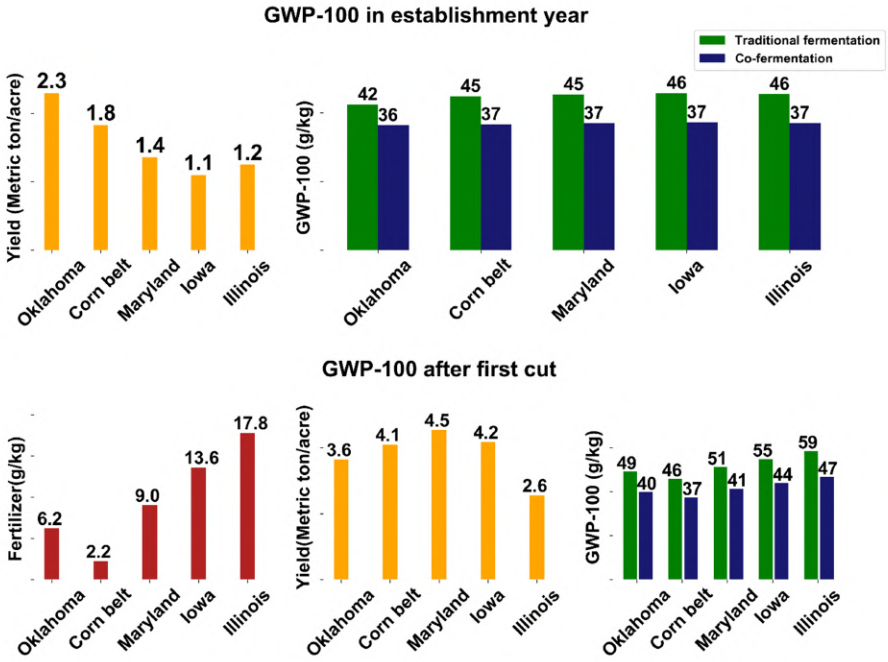


Figure B.11: Impact of different farming techniques on GWP-100 across different locations

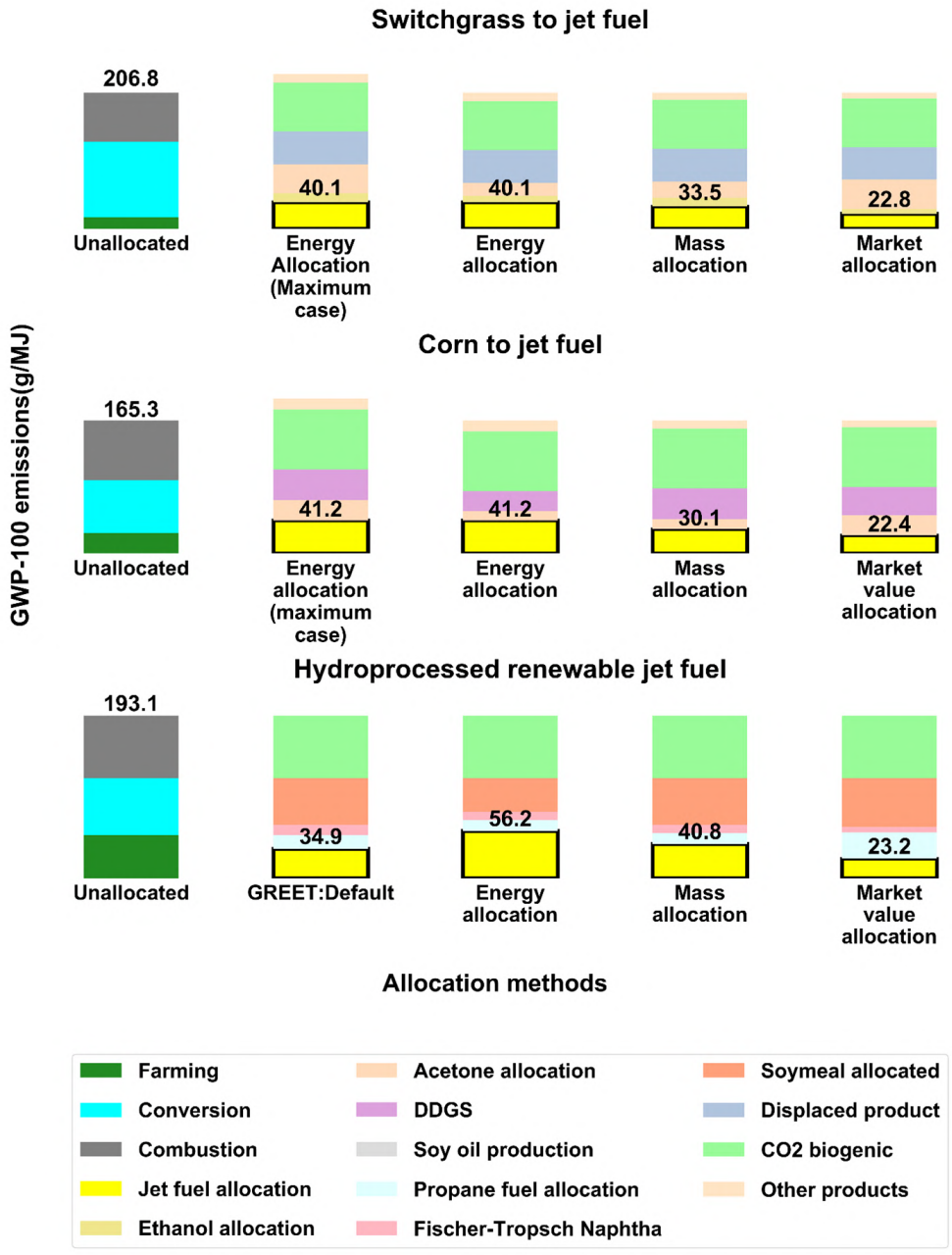


Figure B.12: Co-product allocation scenarios of biofuel pathways (STJ (ideal case), CTJ, and HRJ)

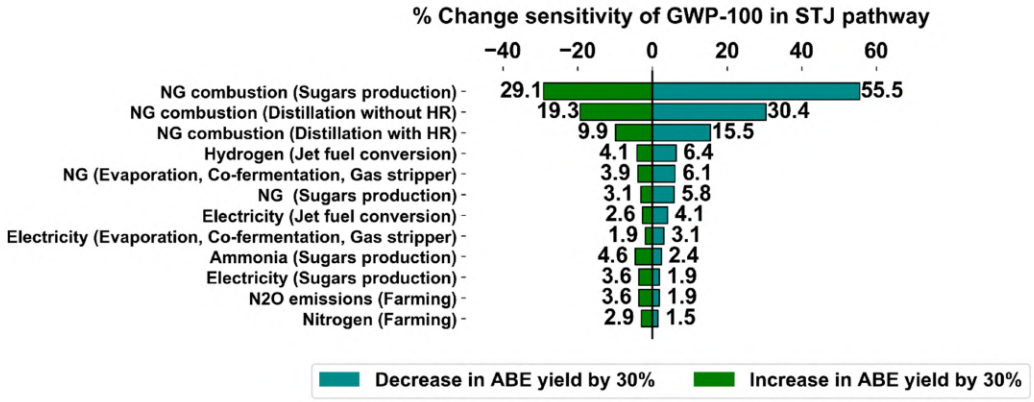


Figure B.13: Tornado chart of STJ (ideal case) sensitivity with changes in ABE yield by ±30% via co-fermentation. GWP-100 for base yield of 0.372 g ABE/g glucose is 43.7 g $CO_2e$/MJ of jet fuel

required to understand carbon sequestration in lignocellulosic biomass, because the duration until bioenergy crops reach equilibrium is debatable (Vaughan et al., 2018).

Table B.7: Energy, mass and market values used in co-product allocations of STJ pathway via co-fermentation

| STJ | Energy content (MJ/kg) | Mass (kg) | Density (kg/L) | Market value ($/gal) | Market value ($/kg) |
|---|---|---|---|---|---|
| Butanol | 36.05 | 0.0413 | 0.810 | 3.750 | 0.0506 |
| Ethanol | 26.80 | 0.0093 | 0.810 | 1.446 | 0.0045 |
| Acetone | 28.55 | 0.0172 | 0.810 | 4.990 | 0.0286 |
| Gasoline blendstock | 44.00 | 0.0048 | 2.210 | 2.210 | 0.0037 |
| Low-sulfur diesel | 42.00 | 0.0029 | 0.810 | 1.930 | 0.0018 |
| Jet fuel | 43.15 | 0.0232 | 0.810 | 1.910 | 0.0146 |

Table B.8: Energy, mass and market values used in co-product allocations of CTJ pathway

| CTJ | Energy content (MJ/kg) | Mass (kg) | Market value ($/gal) | Market value ($/kg) |
|---|---|---|---|---|
| Butanol | 36.049 | 0.0413 | 3.750 | 0.0506 |
| Ethanol | 27.000 | 0.000006 | 1.408 | 0.000003 |
| Acetone | 28.550 | 0.0098 | 4.990 | 0.0163 |
| DDGS | 20.250 | 0.0294 | 0.780 | 0.0230 |
| Gasoline blendstock | 44.000 | 0.0048 | 2.210 | 0.0037 |
| Low-sulfur diesel | 42.000 | 0.0029 | 1.930 | 0.0018 |
| Jet fuel | 43.150 | 0.0232 | 1.910 | 0.0146 |

Table B.9: Market values used in co-product allocations of HRJ pathway

| HRJ | Market value | Mass (kg) | Density (kg/L) | Market value ($/MJ) |
|---|---|---|---|---|
| Jet fuel ($/gal) | 1.910 | 0.0232 | 0.80 | 0.0146 |
| Propane ($/gal) | 1.918 | 0.0199 | 0.51 | 0.0197 |
| Naphtha ($/gal) | 0.480 | 0.0225 | 0.74 | 0.0039 |
| Soy meal ($/kg) | 0.319 | 0.1148 | NA | 0.0366 |

Table B.10: Co-product allocations (g $CO_2e$/MJ) in STJ pathway (practical case) via cofermentation

| Co-products allocated | Energy allocation (maximum case) | Energy allocation | Mass allocation | Market allocation |
|---|---|---|---|---|
| Jet fuel | 48.3 | 48.3 | 40.4 | 27.4 |
| Ethanol | 16.2 | 12.0 | 16.2 | 8.4 |
| Acetone | 53.7 | 23.6 | 29.9 | 53.7 |
| Displaced electricity | 62.6 | 62.6 | 62.6 | 62.6 |
| Gasoline | 10.1 | 10.1 | 8.3 | 6.9 |
| Diesel | 5.8 | 5.8 | 4.9 | 3.3 |

Table B.11: Co-product allocations (g $CO_2e$/MJ) in CTJ pathway (80% ABE)

| Co-products allocated | Energy allocation (maximum case) | Energy allocation | Mass allocation | Market allocation |
|---|---|---|---|---|
| Jet fuel | 44.050 | 44.050 | 32.15058 | 23.913 |
| Ethanol | 0.0078 | 0.007 | 0.0078 | 0.004 |
| Acetone | 26.739 | 12.334 | 13.606 | 26.739 |
| DDGS | 40.831 | 26.254 | 40.831 | 37.566 |
| Gasoline | 9.250 | 9.250 | 6.621 | 6.094 |
| Diesel | 5.286 | 5.286 | 3.964 | 2.864 |

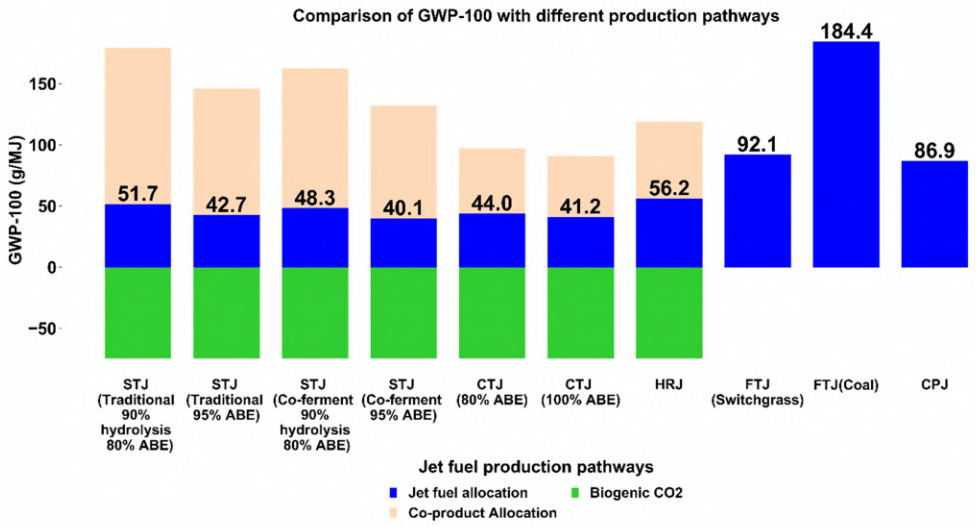


Figure B.14: Comparison of GWP-100 in alternative jet fuel production pathways

Table B.12: Co-product allocations (g $CO_2e$/MJ) in HRJ pathway

| Co-products allocated | GREET:default | Energy allocation | Mass allocation | Market allocation |
|---|---|---|---|---|
| Jet fuel | 34.8 | 56.2 | 40.8 | 23.913 |
| Propane | 16.5 | 13.0 | 13.06 | 0.004 |
| Naphtha | 12.0 | 9.5 | 9.5 | 26.739 |
| Soymeal | 55.4 | 40.0 | 55.4 | 58.1 |

Table B.13: Net energy values for STJ pathway practical case via co-fermentation

| Inputs | Quantity | Unit | Energy intensity coefficients (KJ/Unit) | Total NE cost KJ/MJ of jet fuel |
|---|---|---|---|---|
| Nitrogen | 1.87 | g | 60.52 | 104.09 |
| Phosphoric acid | 0.88 | g | 8.19 | 6.63 |
| Potassium oxide | 1.23 | g | 7.86 | 8.88 |
| $CaCO_3$ | 2.24 | g | 0.12 | 0.24 |
| Herbicides | 2.24 | g | 0.12 | 0.24 |
| High-density polyethylene | 0.13 | g | 79.74 | 9.57 |
| Diesel in farm equipment | 0.69 | mL | 42.96 | 30.07 |
| Natural Gas in pretreatment and hydrolysis | 25 | g | 51.93 | 1194.39 |
| Electricity in pretreatment and hydrolysis | 16.66 | Wh | 6.81 | 103.98 |
| Ammonia | 4.32 | g | 41.45 | 164.55 |
| Cellulase | 2.16 | g | 93.60 | 186.26 |
| Sulfuric acid | 6.48 | g | 0.59 | 3.51 |
| Natural gas in evaporator, cofermentation and gas stripper | 15.71 | g | 51.93 | 815.82 |
| Electricity in evaporator, cofermentation and gas stripper | 7.42 | Wh | 6.81 | 50.53 |
| Natural gas in distillation and hydrolysis | 6.66 | g | 51.93 | 345.85 |
| Electricity in butanol to jet fuel conversion | 9.29 | Wh | 6.81 | 63.26 |
| Hydrogen | 1.23 | g | 7.86 | 8.88 |
| ZSM-5 catalyst | 1.23 | g | 7.86 | 8.88 |
| Electricity displaced in hydrolysis | 139.5 | Wh | 6.81 | -949.95 |
| Ethanol | NA | NA | NA | -248 |
| Acetone | NA | NA | NA | -497 |
| Gasoline blendstock | NA | NA | NA | -210 |
| Low-sulfur diesel | NA | NA | NA | -120 |

Table B.14: Life cycle GHG emissions from well to wake in different jet fuel production pathways

| Jet fuel pathways | Co-products | Co-product handling | GWP-100 |
|---|---|---|---|
| Switchgrass to jet (via co-fermentation (ABE route) yield 0.374 g ABE/g sugars) | Electricity, Ethanol, Acetone | Displacement, Energy-allocation | 42.7 |
| Switchgrass to jet (via traditional ABE (ABE route) yield 0.321 g ABE/g sugars) | Electricity, Ethanol, Acetone | Displacement, Energy-allocation | 51.7 |
| Corn to jet | Gasoline blendstock, Low sulfur diesel | Energy allocation | 41.2 |
| Hydro-processed renewable jet | Soymeal, Propane fuel mix, Naphtha | Displacement, Energy-allocation | 56.2 |
| Fischer-Tropsch jet fuel from switchgrass | NA | NA | 92.1 |
| Fischer-Tropsch jet from coal | NA | NA | 184.4 |
| Conventional jet from crude oil | NA | NA | 86.9 |

- NE for STJ is 1213 KJ/MJ of jet fuel
- NE for CTJ is 994 KJ/MJ of jet fuel
- NE for HRJ 427.96 KJ/MJ of jet fuel
- NE for FTJ (Switchgrass) 130 KJ/MJ of jet fuel
- NE for FTJ (Coal) is 2048.64 KJ/MJ of jet fuel

VITA

ABHIRAM SIVA PRASAD PAMULA

Candidate for the Degree of

Doctor of Philosophy

Dissertation: APPLYING SATELLITE REMOTE SENSING, UNMANNED SYSTEMS, AND MODELS FOR WATER QUALITY ANALYSIS

Major Field: Civil Engineering

Biographical:

Education:

Completed the requirements for the degree of Doctor of Philosophy with a major in Civil Engineering at Oklahoma State University in July 2022.
Received a Masters of Science in Environmental/Environmental Health Engineering at Cleveland State University, Cleveland in May 2018.
Received a Bachelors of Technology in Civil Engineering at Indian Institute of Technology, Bhubaneswar in July 2014.

Experience:

Handling Editor for International Journal of Environment and Waste Management and International Journal of Environmental Engineering.