Assessing CyAN satellite accuracy: A closer look at Kansas' aquatic boundaries
by

Jordan Angot

B.S., Kansas State University, 2022

A REPORT

submitted in partial fulfillment of the requirements for the degree

MASTER OF SCIENCE

Department of Biological and Agricultural Engineering
Carl R. Ice College of Engineering

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2024

Approved by:

Major Professor
Dr. Trisha L. Moore

Copyright

Abstract

Cyanobacteria, commonly known as blue-green algae, pose significant environmental and public health challenges, producing toxins that degrade water quality and affect human health, local economies, and ecosystems. In Kansas, bodies of water like Marion Reservoir, Milford Lake, and Webster Lake are notably affected. Traditional in-situ sampling methods by the Kansas Department of Health and Environment (KDHE) often rely on public reports, potentially missing broader occurrences of harmful algal blooms (HABs). This study evaluates the efficacy of the Cyanobacteria Assessment Network (CyAN) satellite remote sensing technology as a method of bolstering existing efforts, examining its accuracy against localized in-situ measurements at the three eutrophic Kansas reservoirs. Through comparative analyses, this research assesses CyAN's performance in estimating cyanobacteria concentrations at the water's edge as well as when classified these into KDHE's HAB levels: Watch, Warning, and Hazard. Findings from the cell concentration analysis demonstrate that the satellite data's accuracy varies significantly with lake geometry and spatial processing techniques, such as grid cell aggregation and zonal statistic type. However, in the KDHE HAB level classification analysis accuracy improved across all waterbodies. This research supports the potential of integrating remote sensing into existing monitoring frameworks to enhance HAB surveillance and management, providing a more comprehensive understanding of public health and environmental threats along the water's edge in Kansas's freshwater systems.

Table of Contents

List of Figures

List of Tables

List of Equations

Acknowledgements

I would like to express my deepest gratitude to Dr. Trisha Moore, for her patience, support, and guidance throughout my research. Thank you for allowing me the opportunity to continue with my studies even through the start of my professional career.

I am also profoundly grateful to Dr. Aleksey Sheshukov and Dr. Jida Wang for serving on my committee. A special thanks to Dr. Wang for his continued support despite moving to a different state and university.

I would also like to shoutout Dr. Shawn Hutchinson, whose course on Python for GIS sparked a passion for the subject that has significantly influenced my academic and professional pursuits. His enthusiasm and excellent teaching inspired me to delve deeper into the world of GIS and programming.

A huge thank you to my family—my parents and my brother—for their unwavering encouragement, love, and support. I could not have done it without you guys.

Lastly, I wish to express my sincere appreciation to my girlfriend, Michelle, for her constant support and understanding during this journey.

Introduction

Cyanobacteria Background

In Kansas' diverse aquatic ecosystems, cyanobacteria—commonly referred to as blue-green algae—are an inherent but sometimes problematic part of the water's microbial community. These organisms thrive under an influx of phosphorus and nitrogen, stagnant water, and during periods of elevated temperature (J. Alex Elliott, 2010). Certain cyanobacteria species, such as *Microcystis, Aphanizomenon*, and *Anabaena*, are particularly prone to form harmful algal blooms (HABs) in Kansas (A. Dzialowski, V. H. Smith, et al 2011). These blooms pose a significant health hazard, as they produce a variety of toxins—most notably microcystin (KDHE, n.d.). The presence and proliferation of microcystin in Kansas' waterbodies poses a public and environmental health concern threatening drinking source and recreational waters. Exposure to microcystin is known to result in a spectrum of health issues both in humans and animals. In humans, the health effects of microcystin exposure range from mild symptoms, such as skin irritation and gastrointestinal discomfort, to severe conditions like liver and kidney toxicity, and even neurotoxic effects in cases of high or prolonged exposure. Animals, including livestock, pets, and wildlife, are also prone to microcystin poisoning, often exhibiting similar effects. Exposure in animals can lead to immediate health effects such as rapid onset of symptoms leading to death, and chronic exposure may result in reproductive issues, immune system suppression, and tumor promotion (Health Effects, 2015). The structure of microcystins allows them to be highly water-soluble, making them difficult to remove through conventional water treatment processes. (R.M Dawson, 1998, Y. Hu, J. Chen, et al 2016). The environmental and economic threats HABs pose underscore the urgency for ongoing monitoring and management strategies to mitigate adverse effects on Kansas' waterbodies.

KDHE HAB Background

The Kansas Department of Health and Environment (KDHE) is tasked with the oversight of the environmental health of a wide variety of waterbodies across Kansas, ranging from expansive reservoirs to smaller lakes and ponds. Each of these waterbodies has a unique ecological and geological history. A significant aspect of KDHE's responsibilities includes the monitoring and reporting of harmful algal blooms (HABs) across 167 public waterbodies in Kansas (KDHE HAB Viewer, 2024). The KDHE's approach to sampling for HABs involves in-situ or on-site collection of grab samples. This methodology is initiated in response to either visual sightings of algal blooms or reports of illnesses in humans or animals potentially linked to HAB exposure. This monitoring effort is specifically concentrated on the water recreation season, extending from April through October. To address the public health risks posed by HABs, KDHE has established a classification system for HAB advisories, delineated into four levels based on algal cell concentrations per milliliter of water: Below Detection (0-80,000 cells/mL), Watch (80,000-250,000 cells/mL), Warning (250,000-10,000,000 cells/mL), and Hazard (>10,000,000 cells/mL). These categories inform the public about the necessary precautions, ranging from awareness to closure of the waterbody to the public. Sampling of reported waterbodies is carried out in the early morning to ensure a uniform assessment of algal activity (*KDHE Agency Response Plan – 2022*). This timing is crucial as species like Microcystis, known for their positive buoyancy, tend to rise and gather at the water's surface during these hours (Stumpf, R. P. et al, 2012).

Table 1: Waterbody Status Determination sourced from 2022 KDHE HAB Response Plan

Waterbody Status Determination Blue-Green Cell Count and Toxin Exposure Levels		
Condition of Waterbody	Advisory Level	Recommendations
Microcystin toxicity at ≤4 µg/L AND Cell count of ≤80,000 cells/ml	None – Waterbody clear	None
Microcystin toxicity at >4 µg/L to ≤8 µg/L OR Cell count of >80,000 cells/ml to ≤250,000 cells/ml OR Visual confirmation of bloom	**Waterbody will be placed on a Public Health WATCH**	» Post signage » Post on website » Notify public water suppliers » Notify health dept., doctors, vets, health providers » Issue media release
Microcystin toxicity at >8 µg/L to ≤2,000 µg/L OR Cell count of >250,000 cells/ml to ≤10,000,000 cells/ml OR Presence of significant Cyanobacterial surface scum	**Waterbody will be placed on a Public Health WARNING**	» Post signage » Restrict direct contact with water » Post on website » Issue media release » Notify public water suppliers » Notify health dept., doctors, vets, health providers, etc.
Microcystin toxicity at >2,000 µg/L OR Cell count of >10,000,000 cells/ml	**Waterbody will be placed on a Public Health HAZARD**	» Recommend that portions of the lake, the entire lake, or the zone, be closed. If necessary – close adjacent land up to 100 ft from shoreline. » Post signage » Post on website » Issue media release » Notify public water suppliers » Notify health dept., doctors, vets, health providers, etc.

The program's capacity to monitor HABs comprehensively is challenged by financial constraints, as the sampling process is labor intensive and time consuming. Cyanobacteria blooms must first be visually identified and sampled by a trained employee. The samples are then sent to a lab, where they are analyzed to obtain a final cyanobacterial cell count. In order to manage the 167 waterbodies, priorities are set according to the response timeframes delineated by the KDHE. As outlined in the 2022 KDHE HAB response plan, lakes that support public beach access, water supply, or full body contact activities receive the highest priority due to their direct impact on public health and safety. Sampling for these waterbodies is scheduled with the most immediacy, aiming to collect and process reports within a week.

The KDHE's HAB Response Plan reveals financial constraints that critically restrict the frequency and scope of sampling activities, which undermines the program's efficiency in monitoring and mitigating the threats posed by HABs. While waterbodies classified as Priority 1 are tested almost immediately, during times of limited resources, whether labor or funding, waterbodies lacking proximity to major recreational areas may be de-prioritized and overlooked. The KDHE's efforts to monitor and manage HABs through its dedicated program are crucial for protecting public health and the environment. However, the significant challenges facing the program necessitate innovative solutions and sustained support to enhance HAB monitoring effectiveness and reach.

CyAN Background

To monitor and manage cyanobacteria blooms, various surveillance methodologies have been developed. The Cyanobacteria Assessment Network (CyAN) is a collaborative remote sensing project involving the U.S. Environmental Protection Agency (EPA), National Aeronautics and Space Administration (NASA), National Oceanic and Atmospheric Administration (NOAA), United States Geological Survey (USGS), and, as of 2023, the U.S. Army Corps of Engineers (USACE), designed for monitoring harmful algal blooms in freshwater systems across the continental United States and Alaska. Established initially to provide an early warning indicator system, CyAN aims to support the environmental management and public use of U.S. freshwater systems more efficiently and effectively. The Cyanobacteria Index (CI) was originally developed using data collected in western Lake Erie (Wynne et al., 2008) and was later updated and assessed in Ohio, Florida, and eight northeastern U.S. states (Clark et al., 2017; Lunetta et al., 2015). The network's initial goals included creating a standardized method for the early identification of algal blooms, developing an information dissemination system for public health advisories, and understanding the impact of cyanobacterial blooms on health, the economy, and the environment. Since its inception, CyAN has expanded its capabilities to include statewide mapping and providing actionable data to various stakeholders (CyAN, 2023) (NASA Applied Sciences, n.d). The CyAN project is a testament to the power of interdisciplinary collaboration and innovative use of satellite technology in environmental science. Through its comprehensive approach to monitoring algal blooms, CyAN plays a pivotal role in safeguarding water quality and, by extension, protecting public health and aquatic ecosystems. Cyanobacteria, like other phytoplankton, contain photosynthetic pigments that absorb light in specific parts of the spectrum. The key pigments in cyanobacteria are chlorophyll-*a*, phycocyanin, and phycoerythrin. These pigments absorb light primarily in the blue and red regions, reflecting more light in the green part of the spectrum. However, what makes cyanobacteria unique and detectable from space is the presence of phycocyanin, a pigment that absorbs light in the orange to red part of the spectrum (about 620 to 630 nm) and is relatively rare in other types of algae (Kasinak, Jo-Marie E. et al., 2014).

The CyAN project utilizes the CI algorithm, which is a spectral shape algorithm, shown below, tailored to identify cyanobacteria presence through the 620 to 630 nm range. This algorithm is central to the network's capability to detect and quantify algal blooms by analyzing satellite data records. By discerning the unique spectral signatures associated with cyanobacteria, the CI algorithm offers a quantifiable measure of bloom concentration and facilitates the monitoring of water quality on a large scale (CyAN Ocean Color, 2023; NASA Applied Sciences, n.d).

$$SS(\lambda) = \rho_s(\lambda) - \rho_s(\lambda_-) + \{\rho_s(\lambda_-) - \rho_s(\lambda_+)\}\frac{(\lambda - \lambda_-)}{(\lambda_+ - \lambda_-)}, \quad (1)$$

In the above equation, λ denotes the central band's wavelength, with $\lambda-$ and $\lambda+$ representing the adjacent reference bands. $SS(\lambda)$ indicates the spectral shape, while $\rho s(\lambda)$ is the Rayleigh-corrected reflectance. The cyanobacteria index $CIcyano$ is derived from $SS(681)$, where λ=681nm, $\lambda-$=665nm, and $\lambda+$=709nm. A negative $SS(681)$ signifies cyanobacteria presence. More specific identification is possible by recalculating with λ=665nm, $\lambda-$=620nm, and $\lambda+$=681nm. A positive $SS(665)$ indicates phycocyanin, further confirming cyanobacteria.

Since its inception, the CI satellite product available through the Oceancolor portal has undergone a series of validations and evolutions, and as of May 22, 2023 is on Version 5. This version shows on average a 15-20% increase in CI values as well as higher values along the water's edge compared to version 4 (NASA Ocean Color. n.d.). The evolution of the CI algorithm underscores the dynamic nature of remote sensing technologies. As satellite imagery becomes increasingly sophisticated, so do the tools and methods used to analyze these data.

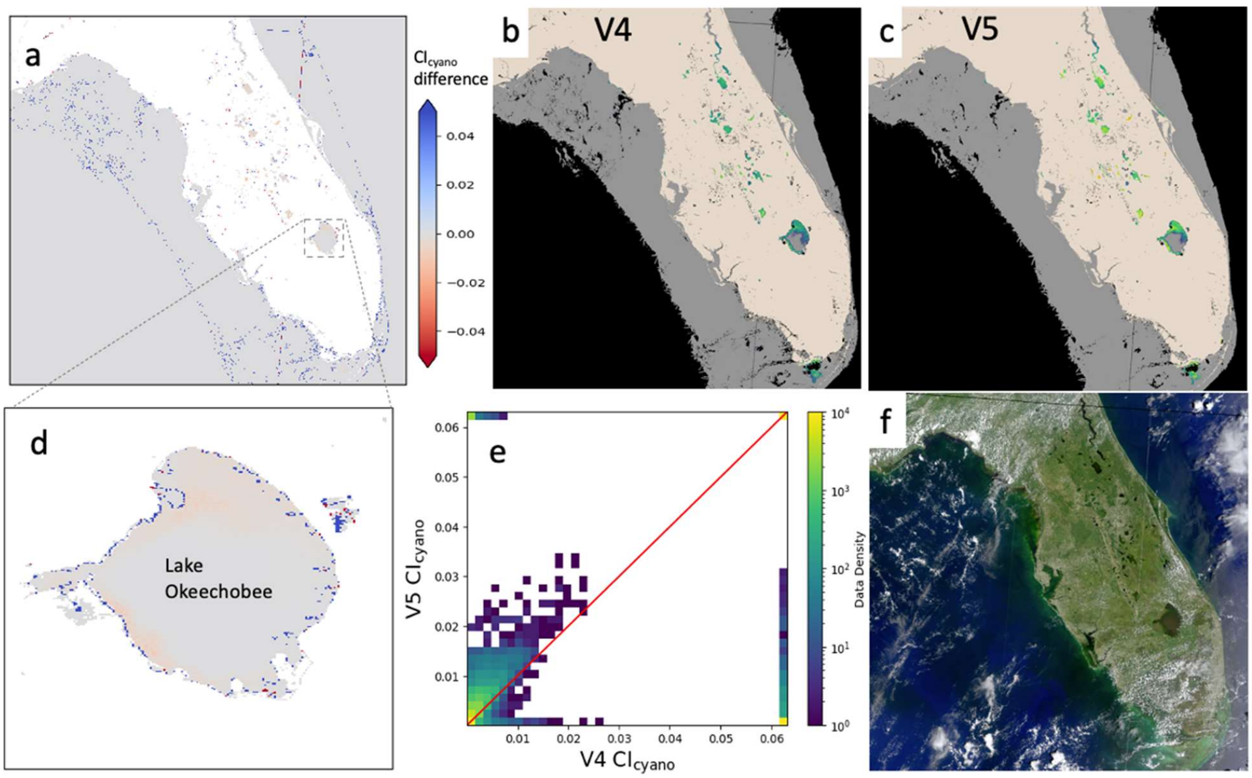

Figure 1: Example of V5 vs. V4 Imagery and Analysis. Sourced from the CyAN Version 5 release notes on the OceanColor domain.

CyAN satellite product utilizes a spatial resolution of 300 meters for detailed waterbody analysis, enabling the detection of algal blooms and identifying cyanobacterial hotspots. Observations are strategically timed around solar noon to ensure consistent lighting for accurate water quality assessments. The project's temporal coverage includes data from 2002 to 2012 using the Envisat MERIS archive and from 2016 onward with Sentinel-3 OLCI, with plans to integrate Sentinel-2 data in 2024 for enhanced resolution (Cyanobacteria Assessment Network, 2023, NASA Applied Sciences, n.d.).

Validation Studies

The development and application of remote sensing technologies, particularly satellite data, have significantly improved the capacity for environmental monitoring on a global scale. The Cyanobacteria Assessment Network (CyAN) satellite product stands as a significant advancement in this field, providing critical data for the detection and monitoring of cyanobacterial harmful algal blooms (HABs) across expansive waterbodies in the continental United States and Alaska. The validation of such technologies, through rigorous studies and comparisons with field data, is crucial to ensure their reliability, accuracy, and applicability in environmental management and public health advisories. Validation studies serve as a cornerstone for the operational credibility of remote sensing products. For tools like the CyAN satellite product, which offers wide-reaching and potentially impactful data, the need for thorough validation cannot be overstated; in the context of environmental monitoring, the accuracy of data concerning the presence, absence, and concentration of cyanobacteria is vital for informing public health decisions, recreational water use advisories, and environmental policy. While there have been numerous validation studies of the CyAN satellite product, these have predominantly centered around analyses of the presence or absence of cyanobacterial blooms (Whitman et al., 2022, Handler et al., 2023, Mishra et al., 2021). While such studies are valuable in supporting the product's ability to detect the occurrence of HABs, previous validation efforts have yet to thoroughly examine the correlation between classified satellite-derived cell concentrations and the established health advisory levels along the water's edge in Kansas, specifically those defined by the KDHE for HABs. In a study by Lunetta et al. published in 2015, researchers used in-situ cyanobacteria cell concentration data from Ohio, Florida, New York, Vermont, New Hampshire, Delaware, Connecticut, and Rhode Island and classified it into

4 levels: Low (10,000–109,999), Medium (110,000–299,999), High (300,000–1,000,000), and Very High (> 1,000,000). These data were paired with any matching data occurring within a 7-day window before or after the satellite's pass. MERIS satellite data were processed using a modified CI algorithm classified into the same categories, providing a more detailed view than the absence/presence studies and resulted in an R^2 of 0.87 and an RMSE of 225,369. In the extant literature, the nuanced aspects of bloom assessment, particularly the quantitative estimation of cyanobacterial cell concentrations, have not been investigated. The ability to accurately predict cell concentrations as well as the cyanobacteria hazard levels from satellite data is crucial for assessing the severity of blooms and the potential risks they pose to human health and aquatic ecosystems in Kansas. The primary objective of this study is to address this gap in knowledge for waterbodies in Kansas by assessing the accuracy and viability of using the freely accessible CyAN satellite data by comparing it to localized measurements from KDHE at the water's edge, which is a critical point for cyanobacteria concentration and human activity. This analysis also aims to assess the potential of CyAN data to enhance current approaches to managing and predicting HABs in Kansas.

Waterbodies Background

This study evaluated KDHE HAB response database information for all public reservoirs in Kansas. Only three Kansas waterbodies, Milford Lake, Marion Reservoir and Webster Reservoir, had enough pairs of data for consideration in the analysis. The data collected by KDHE resides along the water's edge, an important boundary where the public and animals interact with the aquatic ecosystems. These reservoirs are not just significant landmarks in Kansas, but also critical resources that directly influence and sustain the surrounding communities and environments. In terms of their water quality conditions, all three reservoirs can be categorized as eutrophic, characterized by elevated nutrient levels, chlorophyll-a concentrations exceeding 25 µg/L, and increased demand for oxygen in the hypolimnetic region (Dodds et al., 2006). Each of these reservoirs provides a unique case study to understand remote sensing of cyanobacteria at the water's edge. Through a linear regression analysis of KDHE in-situ data and CyAN satellite raster data from near the water's edge in these Kansas reservoirs, this research contributes to the larger narrative of sustainable water resource management and offers insights that can be extended to reservoirs across the CyAN network's footprint.

Milford Lake, located just north of Junction City, Kansas is managed by the U.S. Army Corps of Engineers in the Kansas City District, stands as Kansas's largest lake, covering 15,700 acres of water and 163 miles of shoreline. Since its establishment in 1967, the lake has primarily served for flood control, reportedly preventing about $165 million in flood damages, a figure substantially more than its operational costs (U.S. Army Corps of Engineers, Kansas City District, n.d.). Besides flood control, Milford Lake supports water supply, navigation, recreation, and fish and wildlife conservation. It has evolved into a biodiverse habitat, gaining recognition as a prime spot for hunting and fishing. Despite its ecological and recreational value, Milford Lake faces significant challenges, especially in water quality. The lake was listed on the 2022 Kansas 303(d) list of impaired waters due to eutrophication, leading to harmful algae blooms that disrupt the ecological balance and recreational utility (U.S. Army Corps of Engineers, n.d.). Efforts like the Watershed Restoration and Protection Strategy (WRAPS) are underway to address these issues, focusing on reducing phosphorus loads and improving water quality (Milford Reservoir WRAPS, 2019). Another concern is sedimentation, which has been occurring at a rate of 1,007 acre-feet per year from 1967 to 2009, affecting the lake's water storage volume and depth (U.S.

Army Corps of Engineers, n.d.). Milford Lake is one of the few reservoirs across Kansas that the KDHE has deemed large and important enough to split into zones for HAB reporting. Advisories for HAB's are issued by zone rather than at the lake level. See **Figure 2**: Milford Lake HAB Monitoring Zones for map.

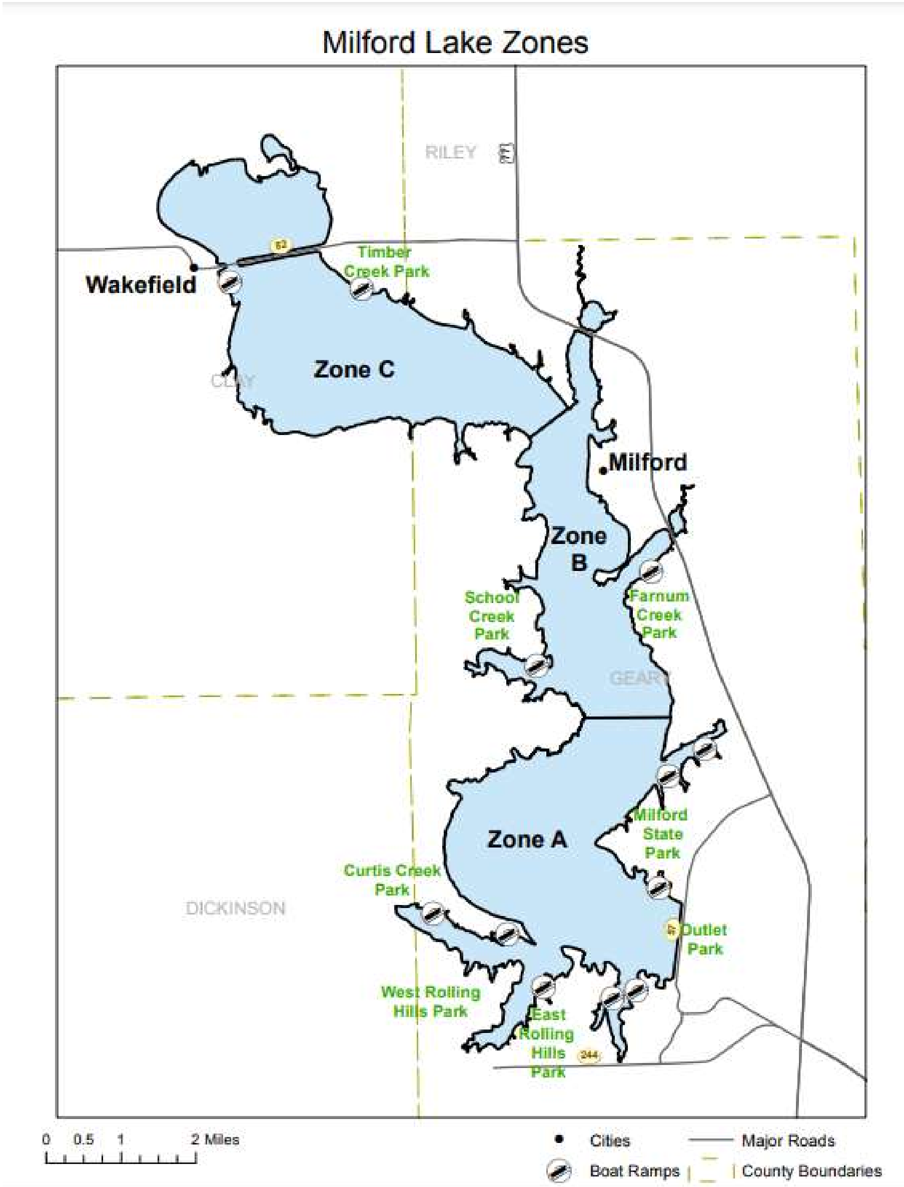

Figure 2: Milford Lake HAB Monitoring Zones. Sourced from the Harmful Algal Bloom page on the KDHE's website.

Marion Reservoir is situated within the Cottonwood River valley of central Kansas between the cities of Marion and Hillsboro. Authorized by the Flood Control Act of 1950 and under the stewardship of the Tulsa District Corps of Engineers, the reservoir was completed in 1968 and cost $13,600,000. Marion Reservoir was designed primarily for flood management and low-flow regulation in the upper Grand (Neosho) River Valley, aligning with other significant projects like the John Redmond Dam and Reservoir and Council Grove Lake (History of Marion Reservoir, n.d.). Marion Reservoir is a multiple-use reservoir constructed by damming the North Cottonwood River to control flooding, provide water supply, and offer recreational opportunities (History of Marion Reservoir, n.d.). The reservoir has a normal pool surface area of 2,509 hectares (6,200 acres) and lies within a 52,836-hectare (204-square mile) watershed that is predominantly cultivated row crop (43%) and grassland (40%) (Marion Lake Fact Sheet, n.d.). Marion Reservoir has been experiencing water quality problems due to anthropogenic eutrophication, threatening algal blooms, low dissolved oxygen, fish kills, and taste and odor issues in drinking water (Marion Lake Fact Sheet, n.d.). According to the Kansas 303(d) list of impaired waters, Marion Reservoir is impaired for aquatic life and water supply uses due to eutrophication, fluoride, and sulfate (KDHE List of Impaired Waters, 2024). The reservoir is also listed as potentially impaired for aquatic life due to dissolved oxygen (KDHE List of Impaired Waters, 2024). These impairments indicate that Marion Reservoir is not meeting the state water quality standards and needs a total maximum daily load (TMDL) to reduce the pollutant loads and restore its designated uses (Marion Lake Fact Sheet, n.d.).

Webster Reservoir, in Rooks County of northwestern Kansas, is a critical infrastructure project that underscores the complex relationship between environmental management and community development. Initiated in response to the devastating Great Flood of 1951, the construction of Webster Reservoir was authorized under the auspices of the Pick-Sloan Missouri Basin Program, highlighting a federal commitment to flood control and water resource management in the Midwest. The reservoir became operational in 1956, following a series of challenges that underscored the project's significant impact on the local community, including the relocation of the town of Webster to accommodate the reservoir's footprint (Jean Lindsey, n.d.). Geographically positioned within the Smoky Hills region of the Great Plains, Webster Reservoir plays a pivotal role in the hydrological dynamics of northwestern Kansas. Its creation marked a significant intervention in the natural landscape, designed to serve multiple functions including flood mitigation, irrigation support, and as a venue for recreational activities. The reservoir, managed by the U.S. Bureau of Reclamation, is complemented by Webster State Park, managed by the Kansas Department of Wildlife, Parks, and Tourism. This park is a testament to the reservoir's recreational value, offering amenities for fishing, camping, and wildlife observation, andcontributing to the region's biodiversity and conservation efforts (Kansas Department of Wildlife, Parks, and Tourism, nd). Despite its utility, Webster Reservoir faces environmental challenges, notably those related to water quality. It is listed on the Kansas 303(d) list of impaired waters for issues includingsiltation and eutrophication (KDHE List of Impaired Waters, 2024). These challenges highlight the ongoing need for management strategies that balance the reservoir's ecological health with its role in supporting agricultural, recreational, and community needs.

Methodology

Data Selection

Cyanobacteria bloom data for the study reservoirs were sourced from the online Kansas Department of Health and Environment Harmful Algal Bloom (KDHE HAB) Test Results Viewer, which publishes freely accessible bloom data collected by the agency through its HAB response program. The data, including sample dates, total algal cell counts in cells/mL, percent blue-green algae, and coordinates, were exported to a CSV file. This study started by looking at all 167 unique waterbodies with data available through the KDHE HAB program. A screenshot of the KDHE HAB viewer map showing the testing locations is shown in **Figure 3**.

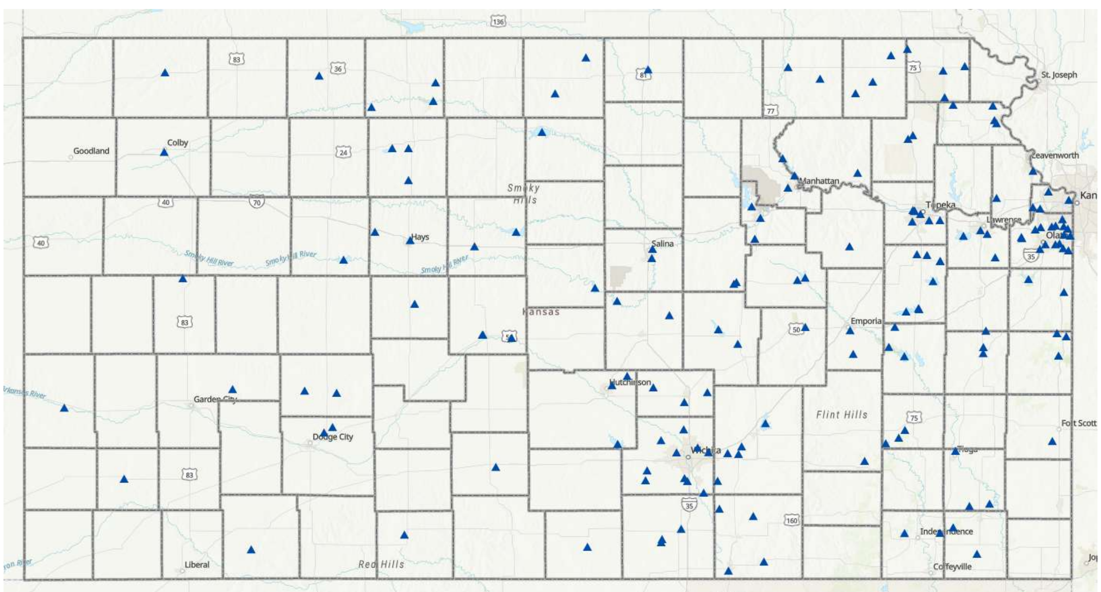

Figure 3: KDHE HAB Response Program Sampling Locations. Sourced from the KDHE HAB Viewer.

This dataset was first filtered to waterbodies with at least 30 data points - the selected waterbodies are published in **Table 2** below. Appendix A contains the all of the waterbodies in the KDHE HAB dataset with their respective number of datapoints and the date range of their samples.

Table 2: Waterbodies with at least 30 samples

Waterbody Name	# of Samples	First sample date	Most recent sample date
CENTRAL PARK LAKE	51	8/18/2010	10/21/2019
CENTRALIA LAKE	43	7/26/2010	10/8/2012
CHENEY LAKE	60	7/28/2011	7/31/2023
CHISHOLM CREEK PARK LAKE	32	7/16/2012	10/26/2015
GATHERING POND	56	6/24/2019	10/31/2022
HARVEY CO. EAST LAKE	47	8/9/2011	10/30/2023
LAKE AFTON	91	8/2/2011	10/30/2023
LAKE MEADE STATE PARK	53	7/27/2010	10/24/2016
LAKE SHAWNEE	51	10/8/2013	10/16/2023
LOVEWELL LAKE	121	7/12/2010	10/30/2023
MARION CO. STATE FISHING LAKE	61	6/27/2011	6/5/2023
MARION RESERVOIR	265	6/27/2011	7/11/2023
MELVERN LAKE	55	8/20/2018	5/23/2023
MELVERN OUTLET POND	62	8/3/2015	9/18/2023
MELVERN OUTLET SWIM POND	33	8/2/2017	9/5/2023
MEMORIAL PARK LAKE	46	8/16/2010	10/30/2017
MILFORD LAKE ZONE A	229	7/18/2011	6/12/2023
MILFORD LAKE ZONE B	136	7/18/2011	6/12/2023
MILFORD LAKE ZONE C	229	7/18/2011	6/12/2023
NORTON LAKE	44	7/7/2014	5/30/2023

POMONA LAKE	87	8/21/2018	6/19/2023
SOUTH PARK LAKE	58	5/30/2012	10/23/2023
WEBSTER LAKE	164	6/30/2015	7/10/2023

The CyAN network offers freely accessible data online for the continental United States through the Oceandata platform. The Cyanobacteria Index daily composite raster data were selected for their temporal similarity to the KDHE grab samples. The data was acquired as a text file of URLs, each corresponding to a TIFF raster for a specific day and indexed by a year and day of year code (e.g., 2020154).

Python Script

To facilitate comparison and ensure repeatability, a series of Python scripts were developed. The main geoprocessing script was designed to integrate the KDHE data with the raster data from the CyAN network, focusing on the date range 2002-2023. A dedicated project folder and a new ArcGIS Pro file were established as part of the initial setup for each waterbody. The dates KDHE sampled were compared to the CyAN list of daily composite raster downloads. The tiles with the matching dates were downloaded and then added to the geodata base. The script then plots the sampling sites from the KDHE dataset on to the CyAN raster and performs a series of buffers and zonal statistics to get the minimum, maximum, and mean values. The buffer distances ranged from 0m, 300m-1000m at 100m increments, & 1500m. These buffer distances – which are defined as the radius of a circle centered on the KDHE sample point – were selected to examine the effects of aggregating more pixels from the CyAN raster in comparison to the in-situ observations at the water's edge similar to the 2015 study by Lunetta et al. The CyAN raster is an 8-bit raster with digital numbers (DN) ranging from 0-255.

This 8-bit raster is related to the calculated Cyanobacteria Index for cell concentrations ranging from ~6500 to ~7,000,000 with the following equation:

$$y = 2 * 10^{-5} * e^{0.0341} \tag{2}$$

where

y = Cyanobacteria index
x = Digital number (0-255)

To convert the DN values into an estimate of cyanobacterial cell concentration, an exponential model given in Error! Reference source not found. was derived (NASA Ocean Color, n.d.).

$$y = 6280 * e^{(0.0277*x)} \tag{3}$$

where

y = Cell concentration (cells/mL)
x = Digital number (0-255)

Raster cells which had a DN equalling 0 (below detection) had their calculated cell concentration values adjusted from 6280 to 0. See Error! Reference source not found. for the CyAN Colorbar and **Figure 5**,**Figure** 6, and **Figure** 7 for location maps.

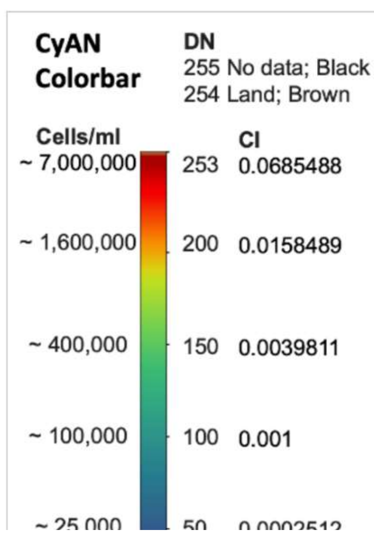

Figure 4: CyAN Colorbar showing relationship between Digital Number (DN), Cyanobacteria Index (CI), and Cell Concentration. Sourced from the CyAN page on the OceanColor domain, 2022.

After running the script for all 24 of the waterbodies with greater than 30 points the output files were inspected to determine whether at least 30 of the days when KDHE collected a sample, CyAN also had a measurement. Of the 24 waterbodies with at least 30 measurements from KDHE, 1 had no points where CyAN also got readings, 4 were too small to be picked up by the 300m satellite resolution, and 12 had fewer than 30 points to compare – leaving just 3 waterbodies – Marion Reservoir, Milford Lake, and Webster Reservoir.

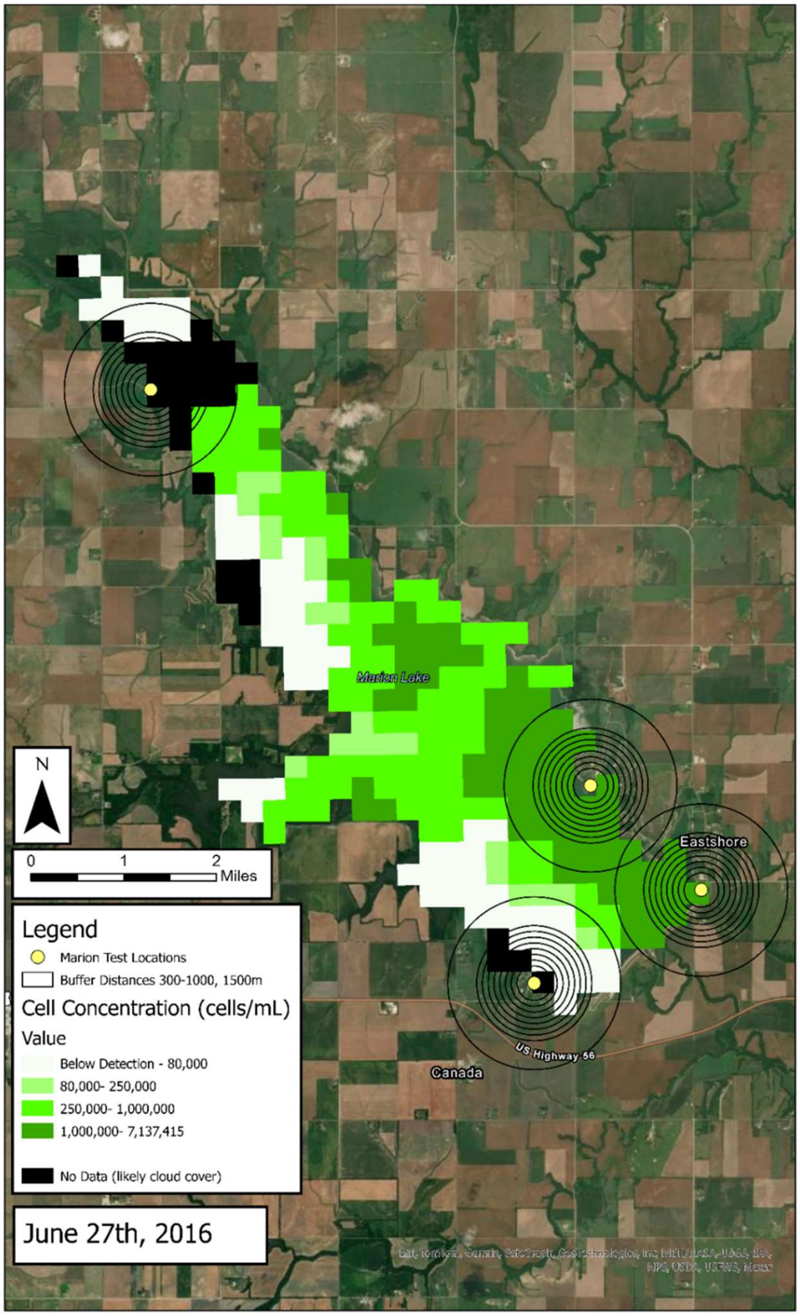

Figure 5: Calculated Cell Concentration Raster for Marion Reservoir - June 27th, 2016. Buffer distances used in the analysis are shown as a series of black concentric circles around KDHE sample points (yellow)

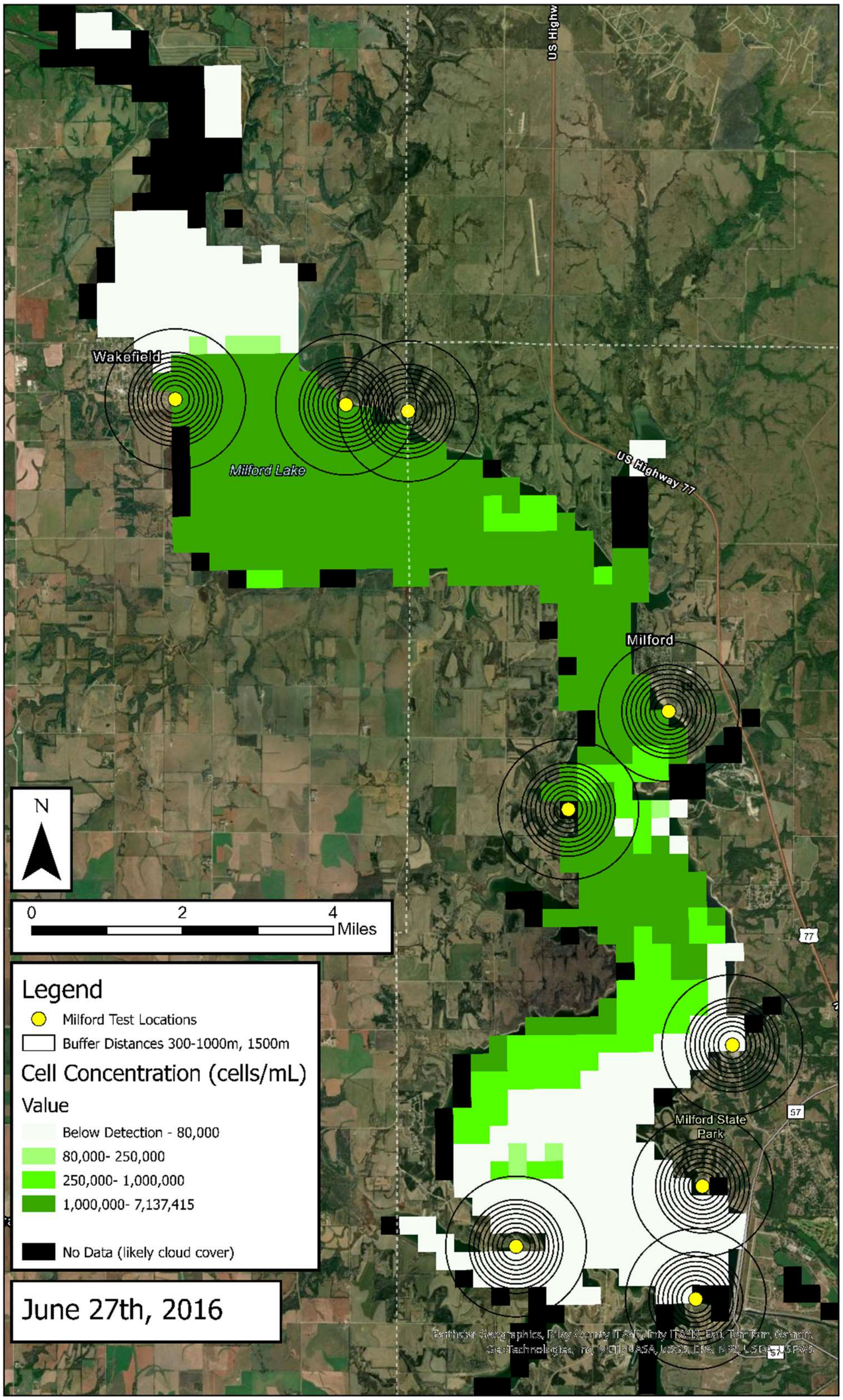

Figure 6: Calculated Cell Concentration Raster for Milford Lake - June 27th, 2016

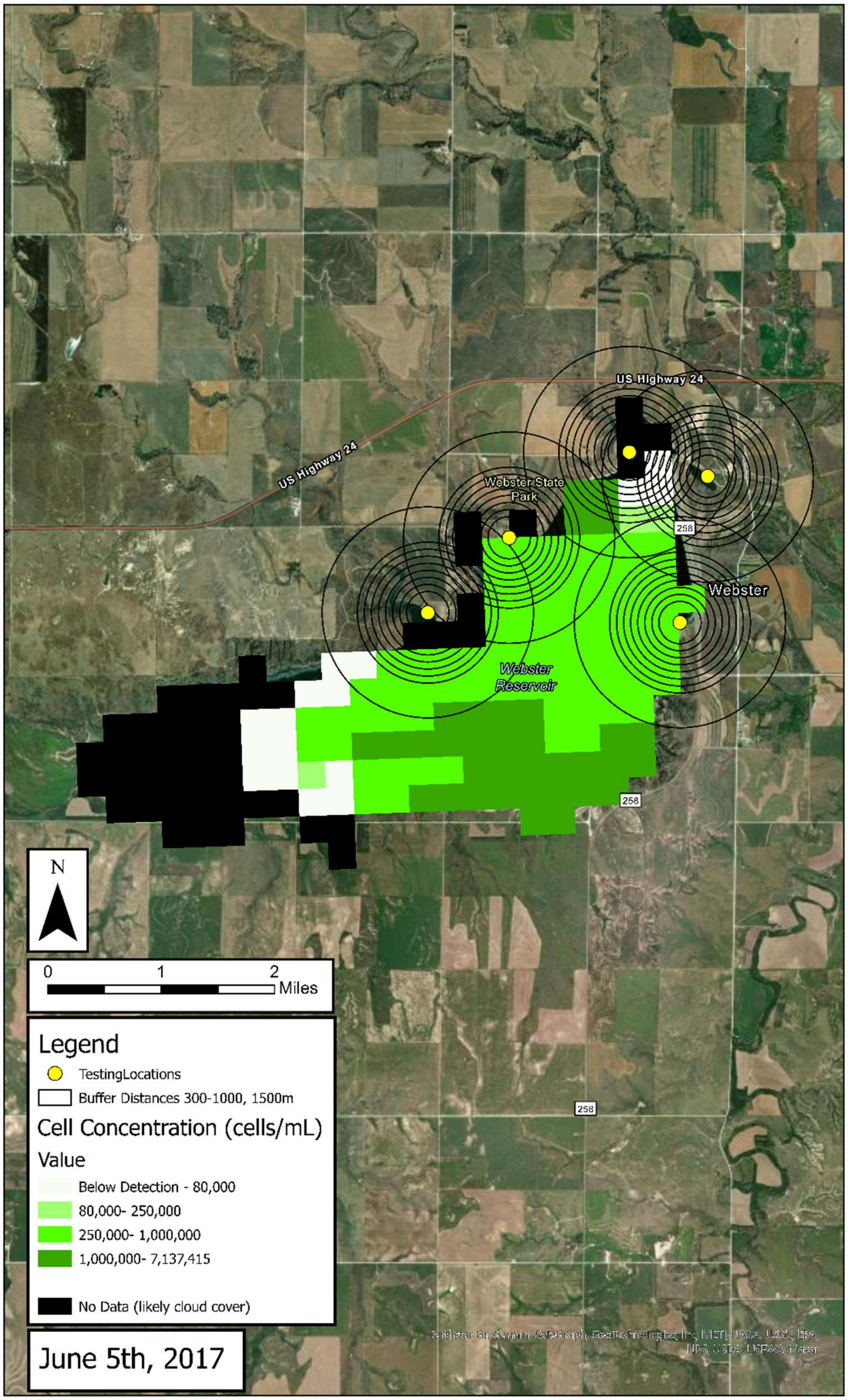

Figure 7: Calculated Cell Concentration Raster for Webster Lake - June 5th, 2017

In addition to the cell concentration analysis the CyAN measurements were then classified into the KDHE HAB classification thresholds: Below Detection (0-80,000 cells/mL), Watch (80,000-250,000 cells/mL), Warning (250,000-10,000,000 cells/mL), and Hazard (>10,000,000 cells/mL) and confusion matrices were created for each buffer distance and zonal statistic type. See Appendix XXX for all matrices. The code can be found on GitHub (https://github.com/Jangot98/KDHE_CyAN_tool/tree/main).

Analysis of CyAN Accuracy

A linear regression analysis was conducted to explore the relationship between zonal statistics' minimum, maximum, and mean values across a series of buffer distances (0m, 300m-1000m, & 1500m) and the in-situ samples collected by the Kansas Department of Health and Environment (KDHE). The study aimed to assess the model's predictive accuracy and systematic bias across various buffer distances and waterbodies. To quantify the model's performance, two key metrics were utilized: Root Mean Square Error (RMSE), which measures the model's predictive accuracy, indicating the average magnitude of the prediction errors. A lower RMSE value signifies higher accuracy. Percent Bias (P-bias), which quantifies the model's systematic bias, indicating any consistent over- or under-estimation in the predictions. A P-bias close to zero indicates minimal bias.

In addition, the following metrics were calculated for each comparison between the zonal statistics' values classified into KDHE HAB levels and KDHE's in-situ samples (also classified into their respective levels) to further evaluate the model's validity:

Accuracy: Indicates the proportion of correctly identified instances, providing a fundamental measure of the model's overall effectiveness in predicting the correct KDHE HAB level.

$$Accuracy = \frac{Number\ of\ Correct\ Predictions}{Total\ Number\ of\ Predictions} \quad (4)$$

Precision: Measures the proportion of true positive results among all positive predictions, highlighting the model's reliability in identifying positive observations.

$$Precision = \frac{1}{N}\sum_{i=1}^{N}\frac{TP_i}{TP_i + FP_i} \quad (5)$$

where N is the number of classes, TP_i is the number of true positives for class i, and FP_i is the number of false positives for class i.

Recall: Assesses the model's ability to identify all relevant instances, ensuring that no significant occurrences are missed.

$$Recall = \frac{1}{N}\sum_{i=1}^{N}\frac{TP_i}{TP_i + FN_i} \quad (6)$$

where N is the number of classes, TP_i is the number of true positives for class i, and FN_i is the number of false negatives for class i.

F1 Score: A balanced mean of precision and recall, offers a singular metric that encapsulates both the reliability of positive predictions and the model's sensitivity in detecting relevant cases.

$$F1\ Score = \frac{2 * Precision * Recall}{Precsion + Recall} \quad (7)$$

These metrics were chosen to offer a comprehensive evaluation of the model's performance when predicting within KDHE's harmful algal bloom (HAB) levels. The methodological approach is designed to assess the concordance between CyAN and KDHE data at different buffer distances and when categorized into KDHE HAB classification levels, providing a multidimensional analysis of predictive accuracy and reliability under varying conditions.

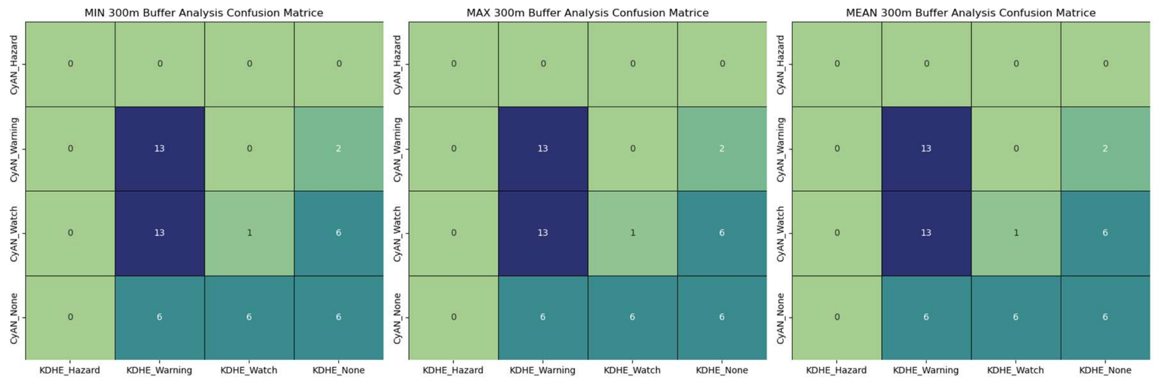

Figure 8: 300m Buffer Analysis Confusion Matrices and Heat Map. The darker colors denote a higher number of classified pairs in that bin.

Results

This section presents the outcomes of two comparative analyses between in-situ cell concentration measurements of cyanobacteria cell concentrations, as reported by the Kansas Department of Health and Environment (KDHE), and remote sensing-derived cell concentration data provided by the Cyanobacteria Assessment Network (CyAN). The first is between the quantities of measured cells per mL and the second is between the categorized KDHE harmful algal bloom level: below detection (0-80,000 cells/mL), watch (80,000-250,000 cells/mL), warning (250,000-10,000,000 cells/mL), and hazard (>10,000,000 cells/mL). The focal points of the analysis are Marion Reservoir, Milford Lake, and Webster Lake but the study started with the 167 waterbodies monitored by KDHE. The primary objective of this research is to assess the validity of the CyAN satellite data by comparing it against the localized measurements from KDHE at the water's edge, which is a critical point for cyanobacteria concentration and human activity.

The results that follow provide insights into the spatial and quantitative relationships between the KDHE and CyAN datasets, covering various distances around KDHE sampling locations. Through this analysis, the study also aims to determine the potential of CyAN data in supplementing the current approaches to HAB management and prediction in Kansas.

Linear regression analyses were conducted to discern the relationship between KDHE's in-situ measurements and CyAN raster data, with a focus on varying buffer distances and statistic type: Minimum, Maximum, and Mean – and spanning from 0m to 1500m. These analyses were done individually for each water body or water body zone to offer a detailed perspective on the correlations and potential disparities at the water body level. The regression results are summarized in **Table 3-**.

In addition to a linear regression analysis between the cell concentrations measured by KDHE and the cell concentrations measured by CyAN, the values were reclassified into the harmful algal bloom levels determined by KHDE. The accuracy, precision, recall, and F1 score were calculated for each confusion matrix and comprehensive graphs for each water body are below. See Appendix B – F for complete results and matrices.

Marion Reservoir

Linear regression analyses exploring the relationship between KDHE's in-situ measurements and CyAN raster data were performed across buffer distances ranging from 0m to 1500m, encompassing Minimum, Maximum, and Mean zonal statistics. For the 0m or point analysis, the correlation was weak, with an R-squared value of just 0.042 for all zonal statistics (MIN, MEAN, MAX), indicating that only a small fraction of the variance in KDHE measurements could be explained by the CyAN data. Despite a consistent slope of 0.76 across these analyses, the high p-values suggested that the relationships were not statistically significant. As buffer distances increased, the analyses generally showed improvements in the R-squared values at shorter distances, indicating a better model fit. For instance, at the 300m buffer, the R-squared values ranged between 0.599 and 0.611 with slopes at 0.54, demonstrating moderate to good explanatory power of the models and strong statistical significance with p-values well below the threshold of 0.05. However, despite these improvements, the percent bias indicated a moderate underestimation by the models, and the root mean square errors (RMSE) suggested substantial prediction errors were still prevalent. At medium distances, such as 400m to 1000m buffers, the R^2 began to decline. While the statistical significance of the models remained high, the decrease in R^2 values was evident, particularly by the 1000m buffer where it hovered around 0.304 to 0.528. The increasing RMSE and percent bias across these distances further indicated growing inaccuracies in the predictions as the distance from the sampling points increased. The trend continued at the 1500m buffer where the model performance significantly dropped, especially for the minimum pixel value analysis, where the R-squared value fell to 0.145. Although the maximum pixel value analysis maintained a relatively higher R-squared of 0.527, the overall prediction errors and biases notably increased, reflecting larger discrepancies in model accuracy at this extended buffer distance. See **Table 3** for cell concentration analysis results and **Figure 10**: Marion R2 Values by Buffer Distance. Additionally, it is important to note the presence of multiple data points where CyAN values are zero despite KDHE indicating bloom conditions. This discrepancy highlights potential limitations or anomalies in detecting or reporting bloom presence under certain measurement or environmental conditions. See **Figure 9**: Marion 400m Buffer Analysis Cell Concentration Results.

For the analysis of reclassified CyAN data into KDHE's harmful algal bloom categories at Marion Reservoir, variations in performance metrics—accuracy, precision, recall, and F1 score—were noted across various buffer distances for MAX, MEAN, and MIN statistics. Accuracy for all statistics began around 0.389 for the point analysis and improved at 300m and 400m, peaking at approximately 0.449 for MAX and MEAN. The MIN statistic showed a distinctive trend, achieving higher accuracy up to 0.5 at 1000m with MAX and MEAN, which showing lesser accuracy at extended distances. Precision for the MAX statistic reached its highest at 400m with a value of 0.473 but then declined sharply to 0.192 by 1500m. Both MEAN and MIN also peaked early in precision, yet MIN demonstrated a lesser decline, maintaining higher precision values up to 1000m, reaching 0.396. This indicates that MIN measures were more stable over distance, providing a more consistent positive predictive value. Recall measurements for all statistics slightly increased up to 500m, with MIN achieving the highest recall, peaking at 0.397 at 1000m. This trend suggests that MIN was more adept at detecting all relevant instances, particularly at further distances. The F1 Score, which balances precision and recall, reflected these early peaks in precision and recall, with a notable peak for MIN at 1000m of 0.353, marking significant improvement over MAX and MEAN. The latter two peaked at 0.292 at 400m and 0.320 at 600m, respectively. MIN consistently outperformed in balancing precision and recall, especially at extended distances. This analysis illustrates that while MAX and MEAN provided stronger predictions at closer ranges, their performance diminished at greater distances. Conversely, MIN maintained more consistent and higher scores across most metrics at longer distances. See **Figure 10**: Marion R2 Values by Buffer Distance, **Figure 11:** Marion KDHE HAB Level Classification - Accuracy, **Figure 12**: Marion KDHE HAB Level Classification - Precision, **Figure 13**: Marion KDHE HAB Level Classification - Recall, **Figure 14**: Marion KDHE HAB Level Classification - F1 Score

Marion Reservoir Results

Table 3: Marion Reservoir Cell Concentration Analysis Results

Analysis Type	Analysis	Slope	Intercept	R-squared	P-value	Standard Error	PBIAS	RMSE	# of CyAN pixels	# of KDHE observa tions
Point Analysis	MIN vs KDHE_CC	0.76	399628	0.042	2.3E-01	0.618	-358.2	779307	36	36
Point Analysis	MAX vs KDHE_CC	0.76	399628	0.042	2.3E-01	0.618	-358.2	779307	36	36
Point Analysis	MEAN vs KDHE_CC	0.76	399628	0.042	2.3E-01	0.618	-358.2	779307	36	36
300m Buffer	MIN vs KDHE_CC	0.54	448789	0.611	9.0E-11	0.064	-51.9	1063866	54	47
300m Buffer	MAX vs KDHE_CC	0.54	459683	0.599	1.8E-10	0.066	-54.5	1072362	54	47
300m Buffer	MEAN vs KDHE_CC	0.54	454236	0.605	1.3E-10	0.065	-53.2	1068055	54	47
400m Buffer	MIN vs KDHE_CC	0.36	477342	0.403	9.5E-07	0.064	-27.2	1204759	67	49
400m Buffer	MAX vs KDHE_CC	0.55	442046	0.605	4.8E-11	0.064	-50.3	1041473	67	49
400m Buffer	MEAN vs KDHE_CC	0.45	459770	0.524	4.2E-09	0.063	-38.2	1126390	67	49
500m Buffer	MIN vs KDHE_CC	0.36	421794	0.402	2.7E-07	0.060	-21.7	1152932	139	54
500m Buffer	MAX vs KDHE_CC	0.58	470620	0.576	2.9E-11	0.069	-75.0	1016784	139	54
500m Buffer	MEAN vs KDHE_CC	0.46	444390	0.509	1.4E-09	0.063	-45.4	1078523	139	54
600m Buffer	MIN vs KDHE_CC	0.36	389121	0.451	1.1E-08	0.054	-15.9	1113772	189	57
600m Buffer	MAX vs KDHE_CC	0.58	470970	0.569	1.3E-11	0.068	-77.6	999494	189	57
600m Buffer	MEAN vs KDHE_CC	0.47	426395	0.533	1.2E-10	0.059	-43.1	1043090	189	57
700m Buffer	MIN vs KDHE_CC	0.29	354475	0.466	4.8E-09	0.042	2.3	1157460	274	57
700m Buffer	MAX vs KDHE_CC	0.63	492883	0.587	3.8E-12	0.071	-95.4	998453	274	57
700m Buffer	MEAN vs KDHE_CC	0.49	409212	0.592	2.7E-12	0.055	-41.9	1015619	274	57
800m Buffer	MIN vs KDHE_CC	0.21	365825	0.291	1.5E-05	0.044	9.6	1208632	330	57
800m Buffer	MAX vs KDHE_CC	0.63	517198	0.573	9.4E-12	0.073	-103.3	1023506	330	57
800m Buffer	MEAN vs KDHE_CC	0.43	426889	0.533	1.1E-10	0.055	-37.7	1064195	330	57
900m Buffer	MIN vs KDHE_CC	0.20	351442	0.286	1.5E-05	0.042	14.0	1205441	464	58
900m Buffer	MAX vs KDHE_CC	0.63	554340	0.534	7.4E-11	0.079	-115.7	1065700	464	58
900m Buffer	MEAN vs KDHE_CC	0.41	438298	0.498	6.0E-10	0.056	-38.1	1074885	464	58
1000m Buffer	MIN vs KDHE_CC	0.20	340497	0.304	7.2E-06	0.040	16.7	1206789	576	58
1000m Buffer	MAX vs KDHE_CC	0.62	576506	0.526	1.2E-10	0.079	-121.5	1088267	576	58
1000m Buffer	MEAN vs KDHE_CC	0.42	428236	0.528	1.0E-10	0.054	-36.9	1063861	576	58
1500m Buffer	MIN vs KDHE_CC	0.11	294622	0.145	2.6E-03	0.036	35.0	1228095	1128	60
1500m Buffer	MAX vs KDHE_CC	0.62	646603	0.527	5.3E-11	0.077	-142.2	1143597	1128	60
1500m Buffer	MEAN vs KDHE_CC	0.37	425573	0.481	8.3E-10	0.050	-26.5	1087214	1128	60

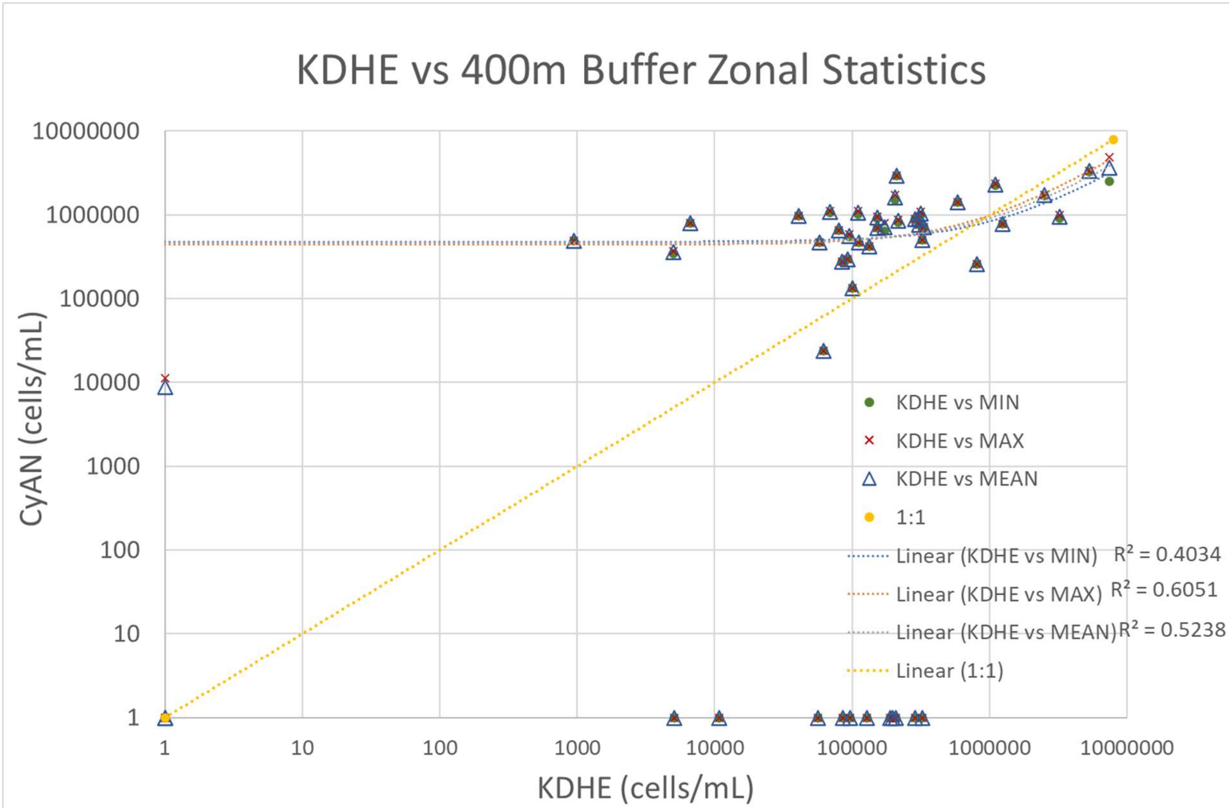

Figure 9: Marion 400m Buffer Analysis Cell Concentration Results

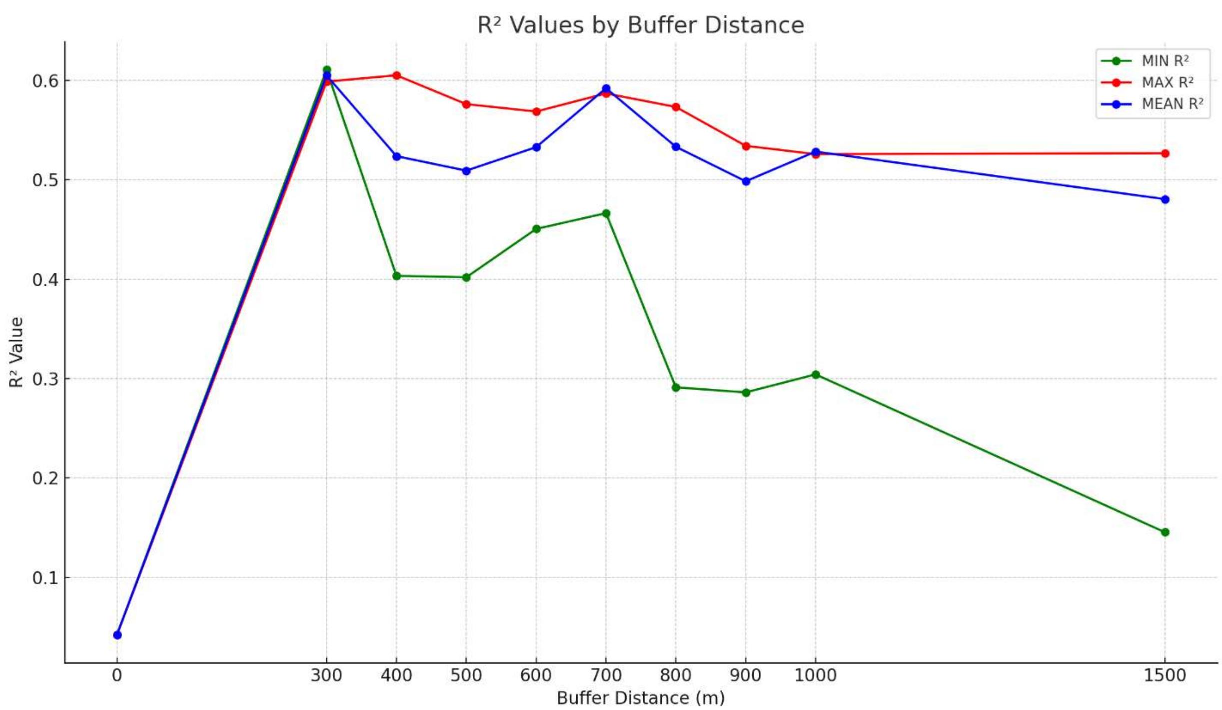

Figure 10: Marion R^2 Values by Buffer Distance

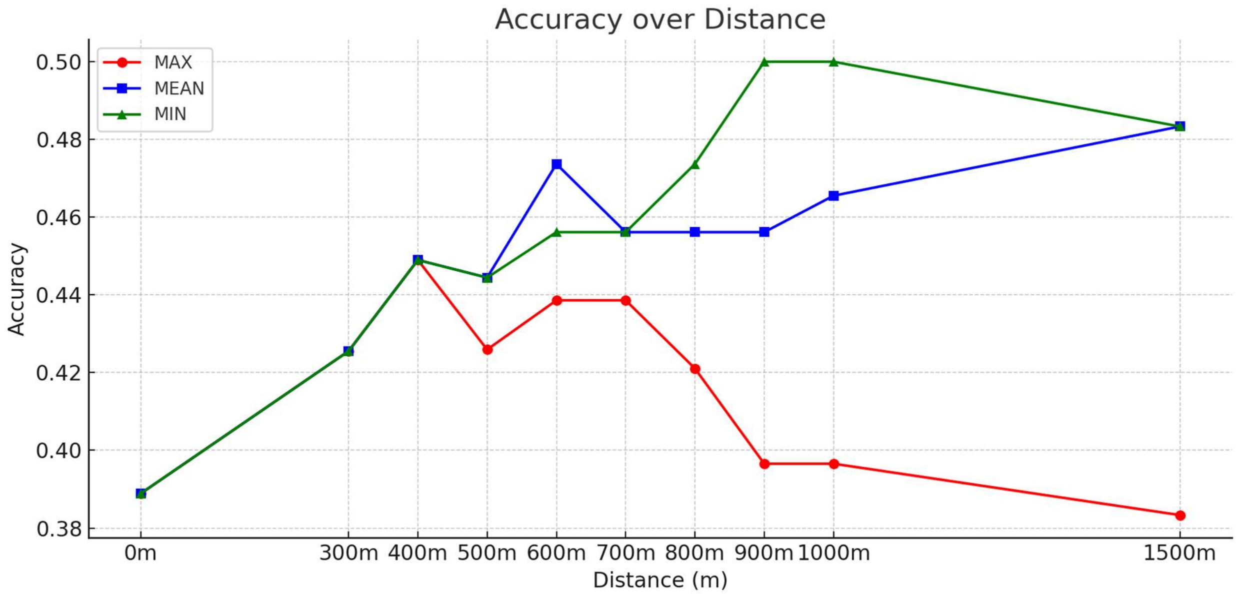

Figure 11: Marion KDHE HAB Level Classification - Accuracy

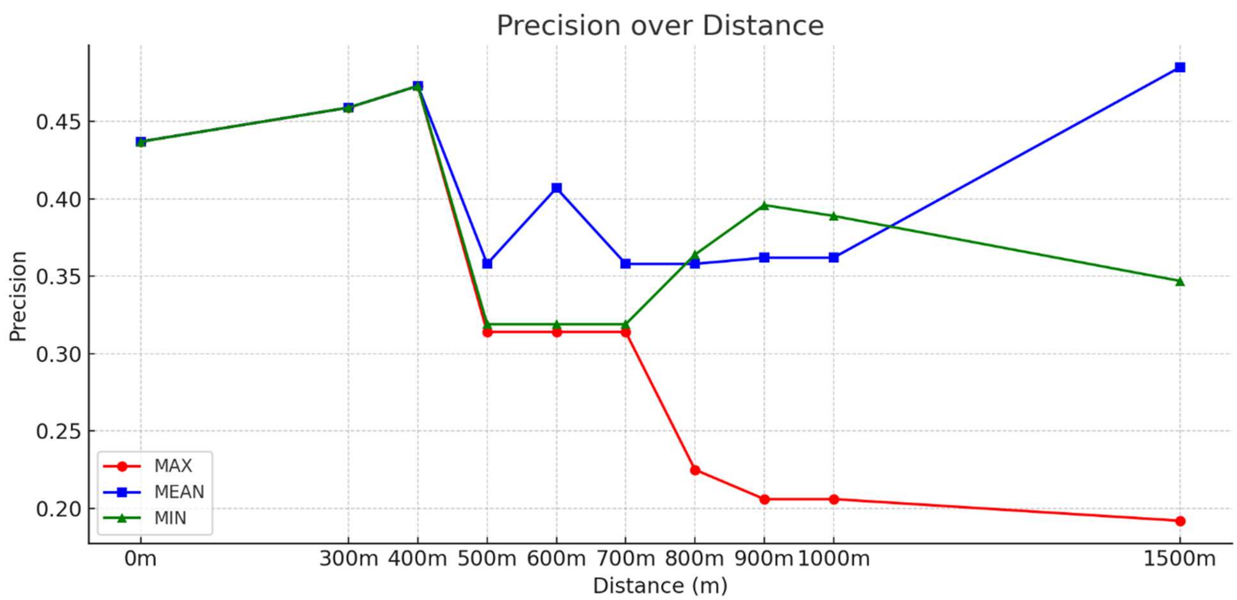

Figure 12: Marion KDHE HAB Level Classification - Precision

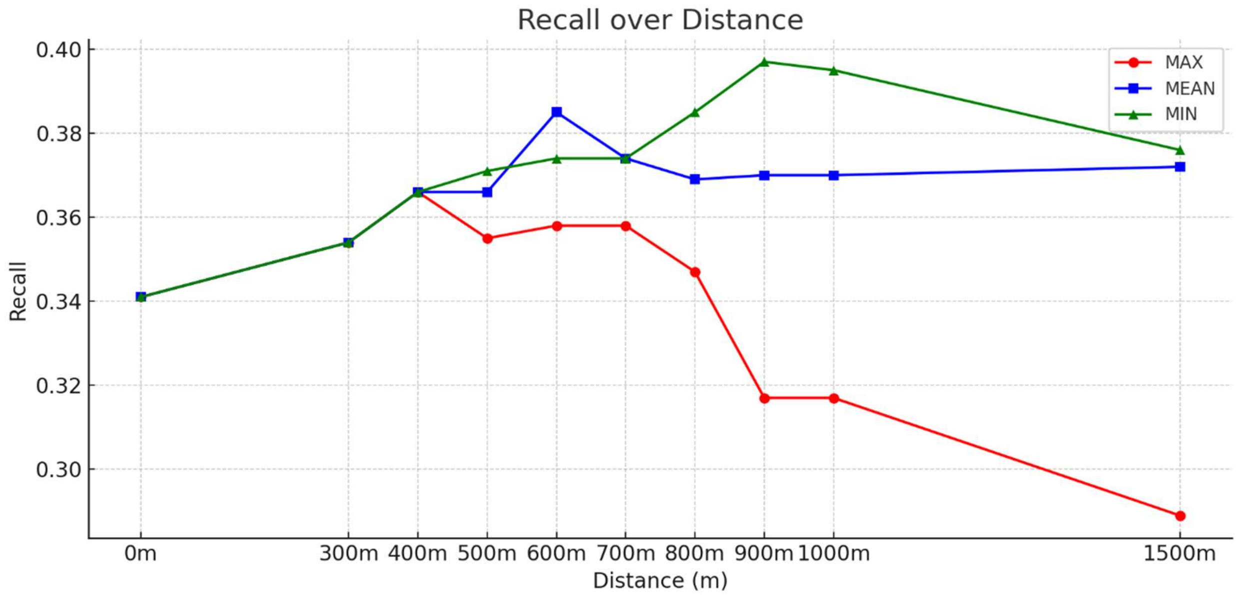

Figure 13: Marion KDHE HAB Level Classification - Recall

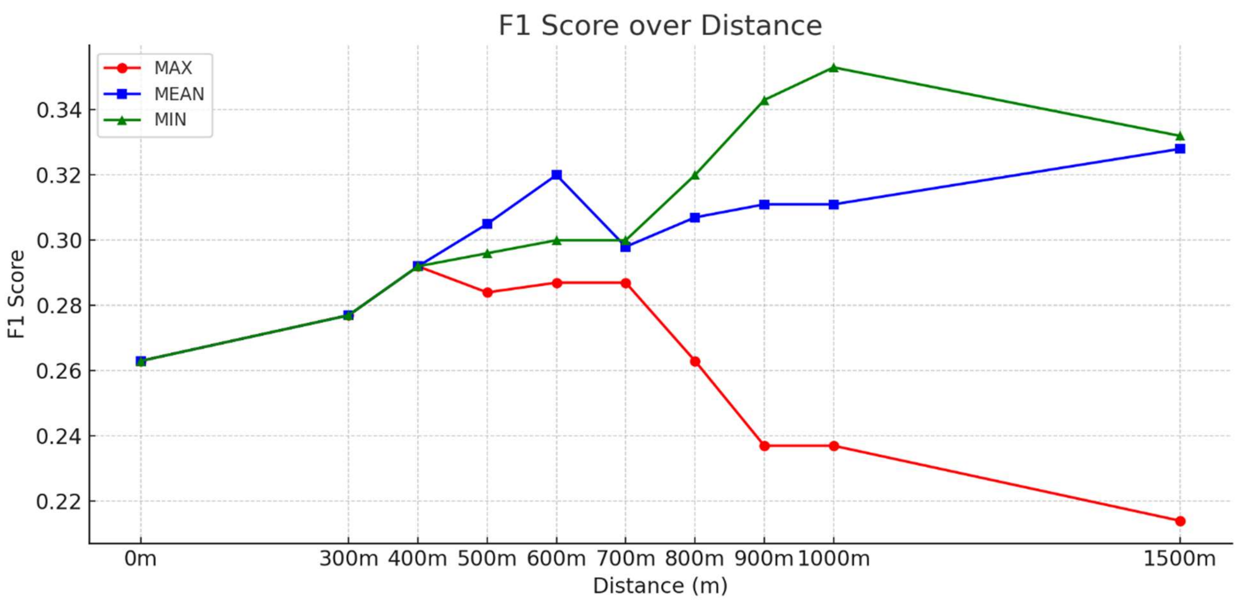

Figure 14: Marion KDHE HAB Level Classification - F1 Score

Milford Lake Zone A

For 0m or point linear regression analysis, the relationship was extremely weak, with an R-squared value of 0.005 across MIN, MEAN, and MAX zonal statistics. The consistent negative slope of -0.31 in these analyses and high p-values indicated a lack of statistical significance in these relationships. As buffer distances increased, the results showed some improvements in model performance at shorter ranges. Notably, at the 300m buffer, R-squared values improved to around 0.431 for all analyses, reflecting moderate explanatory power and substantial statistical significance with very low p-values. However, despite these enhancements, the models still suggested notable prediction errors, as evidenced by substantial RMSE values, and a percent bias that suggested moderate overestimations. Moving to intermediate buffer distances from 400m to 1000m, there was a noticeable decrease in the accuracy of the models. Although the models remained statistically significant, the reduction in R-squared values was apparent, particularly at the 1000m buffer, which showed R-squared values ranging between 0.098 and 0.221. The rise in RMSE and percent bias across these distances highlighted increasing inaccuracies in the predictions as the buffer increased. The diminishing trend in model accuracy was even more pronounced at the 1500m buffer. At this distance, the minimum zonal statistics analysis showed a further reduced R-squared of 0.021. While the maximum zonal statistics analysis maintained a higher R-squared of 0.185, the general trend was towards larger errors and biases, indicating significant discrepancies in model accuracy at extended distances. See **Table 4**: Milford Lake Zone A Reservoir Cell Concentration Analysis Results and **Figure 16**: Milford Lake Zone A R2 Values by Buffer Distance. It is important to note the presence of multiple data points where CyAN values are zero despite KDHE indicating bloom conditions. Again, this discrepancy highlights potential limitations or anomalies in detecting or reporting bloom presence under certain measurement or environmental conditions. See **Figure 15**: Milford Lake Zone A 400m Buffer Analysis Cell Concentration Results,

For the analysis of reclassified CyAN data into KDHE's harmful algal bloom categories at Milford Lake Zone A, variations in performance metrics—accuracy, precision, recall, and F1 score—were noted across various buffer distances for MAX, MEAN, and MIN statistics. Accuracy for all metrics began at approximately 0.438 for the point analysis and showed incremental improvements, particularly noticeable in MIN, which consistently demonstrated the highest accuracy across all distances, peaking at 0.510 from 500m to 1500m. This trend suggests a robust ability of the MIN statistic to maintain accuracy over greater distances compared to MAX and MEAN, which displayed lesser accuracy at extended distances. Precision for the MAX statistic reached its highest at 500m with a value of 0.323 but then declined to 0.311 by 1500m. Both MEAN and MIN also peaked in precision at 500m; however, MIN demonstrated a less significant decline, maintaining higher precision values up to 500m, reaching 0.291. This indicates that MIN measures were more consistent over distance, providing a more reliable positive predictive value. Recall measurements for all statistics gradually increased up to 500m, with MIN achieving the highest recall, peaking at 0.267 from 600m to 900m. This upward trend in recall for MIN, particularly at mid to long-range distances, suggests that it was more adept at detecting all relevant instances, particularly at further distances. The F1 Score, which balances precision and recall, reflected these trends in precision and recall, with MIN consistently showing improvements at longer distances. Notably, MIN achieved an F1 score of 0.257 at 500m and maintained relatively high scores, indicating a balanced performance between precision and recall over extended ranges. This analysis illustrates that while MAX and MEAN provided stronger predictions at closer ranges, their performance tended to diminish at greater distances.

Conversely, MIN maintained more consistent and higher scores across most metrics at longer distances. See **Figure 17**: Milford Lake Zone A KDHE HAB Level Classification - Accuracy, **Figure 18**: Milford Lake Zone A KDHE HAB Level Classification - Precision, **Figure 19**: Milford Lake Zone A KDHE HAB Level Classification - Recall, and **Figure 20**: Milford Lake Zone A KDHE HAB Level Classification - F1 Score.

Milford Lake Zone A Results

Table 4: Milford Lake Zone A Reservoir Cell Concentration Analysis Results

Analysis Type	Analysis	Slope	Intercept	R-squared	P-value	StandardError	PBIAS	RMSE	# of CyAN pixels	# of KDHE observations
Point Analysis	MIN vs KDHE_CC	-0.31	123839	0.005	8.0E-01	1.193	-51.7	121122	16	16
Point Analysis	MAX vs KDHE_CC	-0.31	123839	0.005	8.0E-01	1.193	-51.7	121122	16	16
Point Analysis	MEAN vs KDHE_CC	-0.31	123839	0.005	8.0E-01	1.193	-51.7	121122	16	16
300m Buffer	MIN vs KDHE_CC	0.18	119336	0.431	2.0E-04	0.041	59.3	1395321	42	27
300m Buffer	MAX vs KDHE_CC	0.18	133247	0.427	2.2E-04	0.041	55.0	1393764	42	27
300m Buffer	MEAN vs KDHE_CC	0.18	126413	0.431	2.0E-04	0.041	57.1	1394499	42	27
400m Buffer	MIN vs KDHE_CC	0.15	76196	0.391	5.7E-06	0.029	74.9	1240991	60	44
400m Buffer	MAX vs KDHE_CC	0.15	94719	0.311	8.4E-05	0.034	69.6	1237782	60	44
400m Buffer	MEAN vs KDHE_CC	0.15	85533	0.352	2.2E-05	0.031	72.2	1239321	60	44
500m Buffer	MIN vs KDHE_CC	0.04	78677	0.055	1.1E-01	0.025	76.9	1215011	115	48
500m Buffer	MAX vs KDHE_CC	0.15	113478	0.266	1.7E-04	0.036	61.4	1181721	115	48
500m Buffer	MEAN vs KDHE_CC	0.09	98281	0.170	3.6E-03	0.031	69.0	1200816	115	48
600m Buffer	MIN vs KDHE_CC	0.02	69898	0.014	4.2E-01	0.023	79.6	1206644	204	49
600m Buffer	MAX vs KDHE_CC	0.15	117129	0.258	1.9E-04	0.036	59.5	1169245	204	49
600m Buffer	MEAN vs KDHE_CC	0.08	93979	0.144	7.1E-03	0.029	70.2	1192443	204	49
700m Buffer	MIN vs KDHE_CC	0.02	69898	0.014	4.2E-01	0.023	79.6	1206644	254	49
700m Buffer	MAX vs KDHE_CC	0.14	122995	0.254	2.2E-04	0.036	57.6	1168350	254	49
700m Buffer	MEAN vs KDHE_CC	0.08	96082	0.124	1.3E-02	0.029	69.9	1193218	254	49
800m Buffer	MIN vs KDHE_CC	0.02	66941	0.017	3.8E-01	0.021	80.4	1207309	302	49
800m Buffer	MAX vs KDHE_CC	0.14	135657	0.232	4.5E-04	0.038	53.6	1166774	302	49
800m Buffer	MEAN vs KDHE_CC	0.08	97153	0.145	7.0E-03	0.028	69.3	1191992	302	49
900m Buffer	MIN vs KDHE_CC	0.02	63899	0.017	3.7E-01	0.021	81.3	1208029	461	49
900m Buffer	MAX vs KDHE_CC	0.14	147981	0.224	5.9E-04	0.038	49.7	1165054	461	49
900m Buffer	MEAN vs KDHE_CC	0.07	97499	0.116	1.7E-02	0.028	69.7	1194185	461	49
1000m Buffer	MIN vs KDHE_CC	0.01	63938	0.010	4.8E-01	0.021	81.4	1208255	538	49
1000m Buffer	MAX vs KDHE_CC	0.14	150824	0.221	6.5E-04	0.038	48.8	1164781	538	49
1000m Buffer	MEAN vs KDHE_CC	0.06	97868	0.098	2.9E-02	0.028	69.8	1195212	538	49
1500m Buffer	MIN vs KDHE_CC	0.02	52128	0.021	3.2E-01	0.017	84.8	1211008	1055	49
1500m Buffer	MAX vs KDHE_CC	0.15	199141	0.185	2.0E-03	0.045	32.5	1156247	1055	49
1500m Buffer	MEAN vs KDHE_CC	0.06	95873	0.109	2.0E-02	0.026	70.5	1195677	1055	49

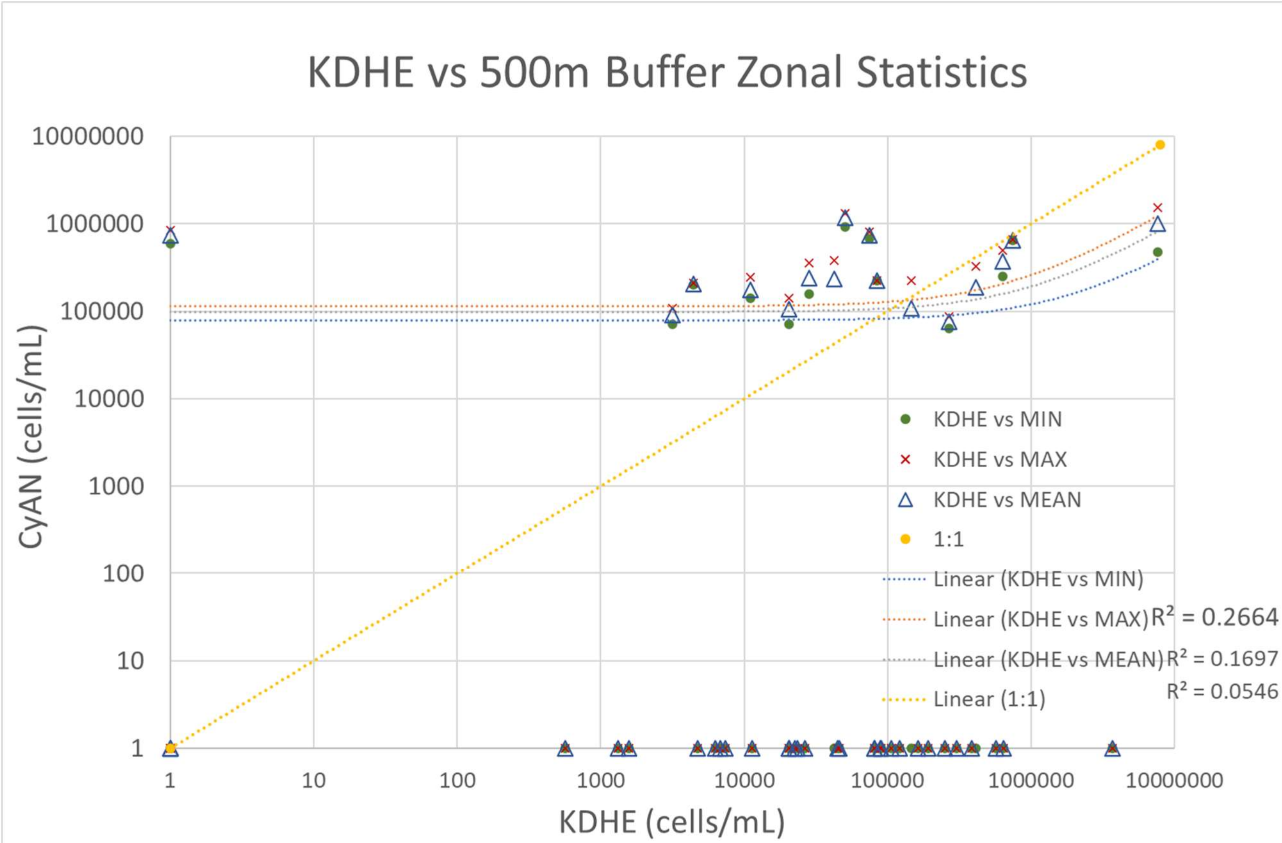

Figure 15: Milford Lake Zone A 400m Buffer Analysis Cell Concentration Results

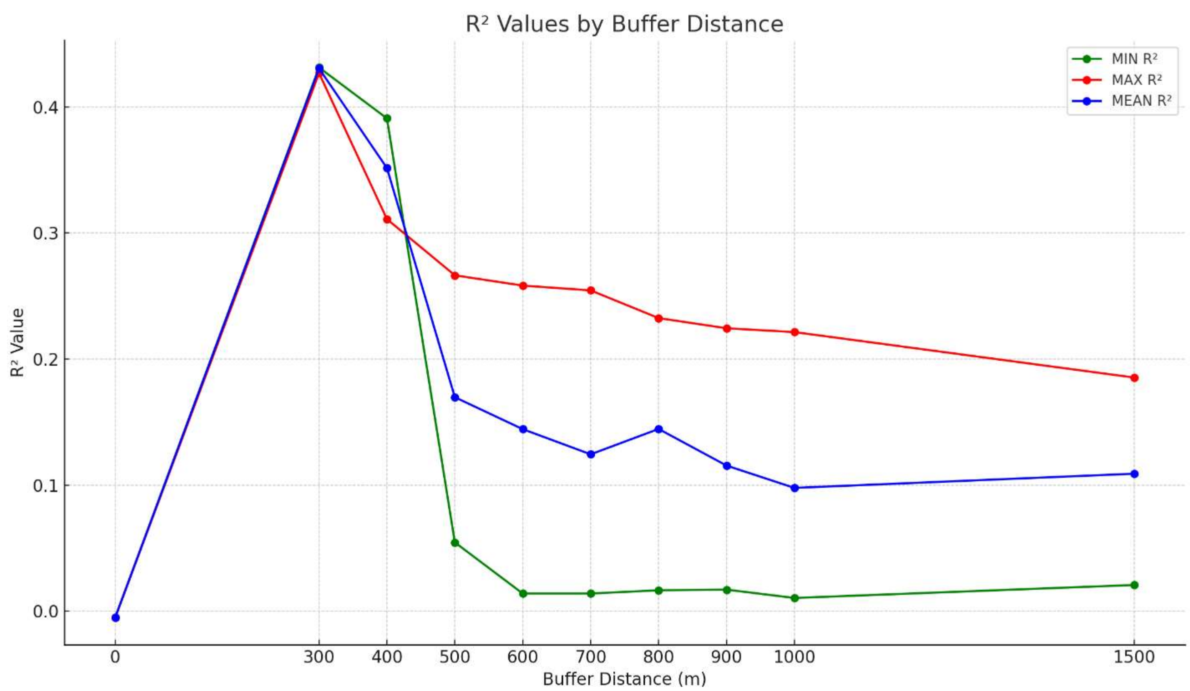

Figure 16: Milford Lake Zone A R^2 Values by Buffer Distance

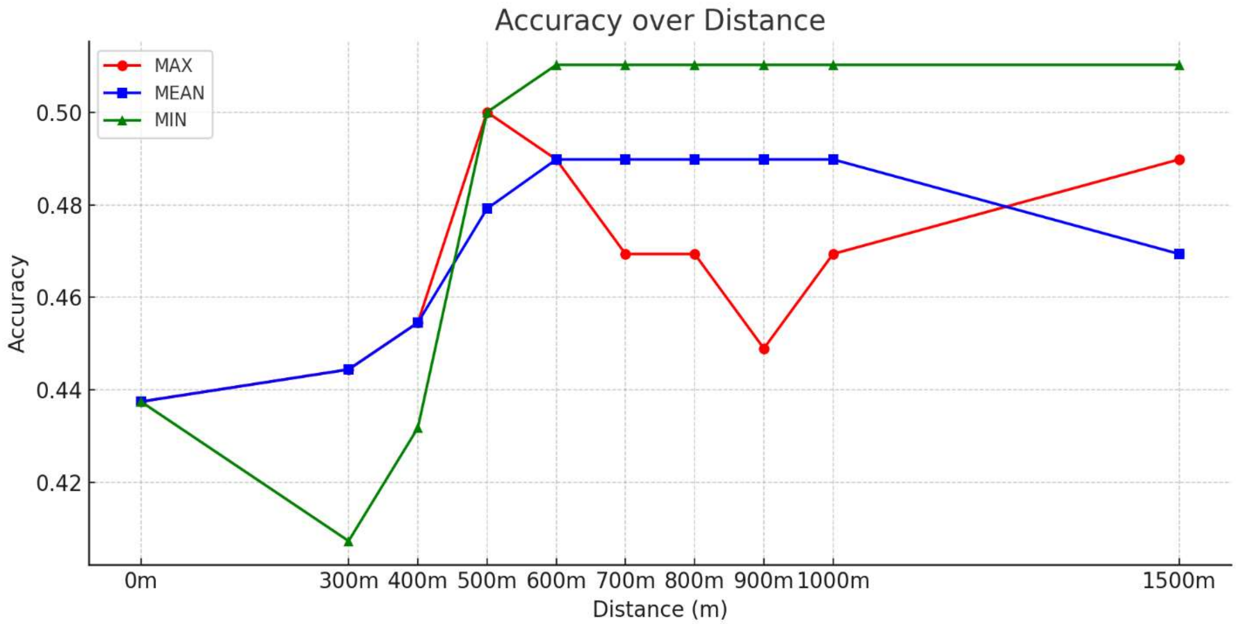

Figure 17: Milford Lake Zone A KDHE HAB Level Classification - Accuracy

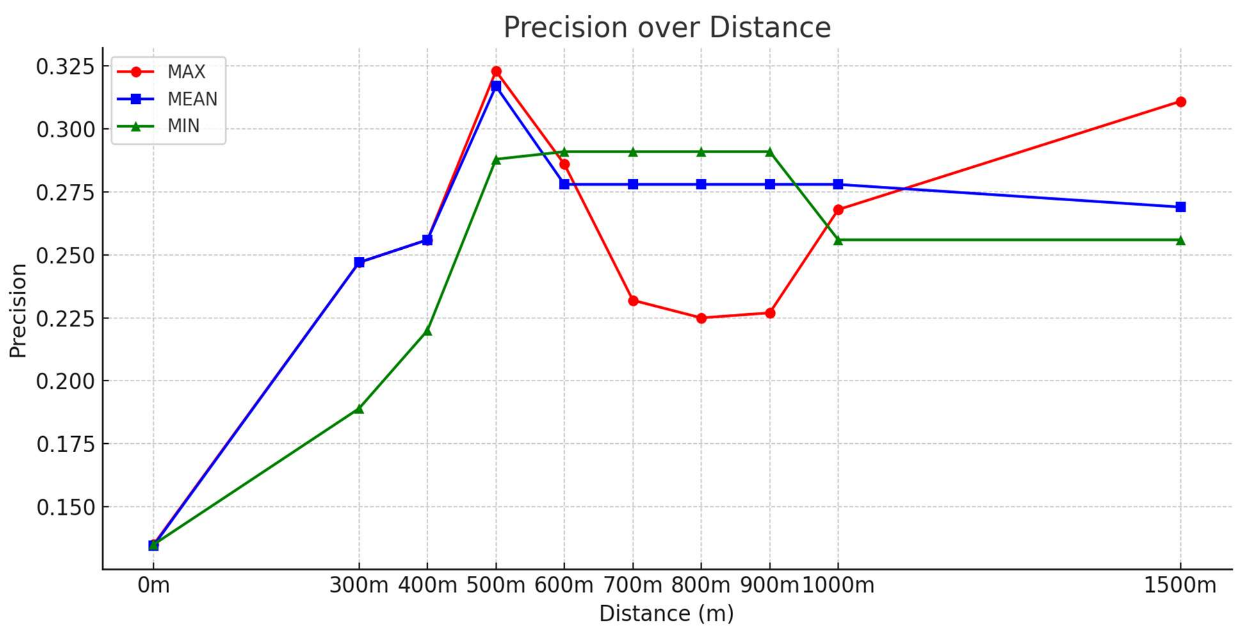

Figure 18: Milford Lake Zone A KDHE HAB Level Classification - Precision

Figure 19: Milford Lake Zone A KDHE HAB Level Classification - Recall

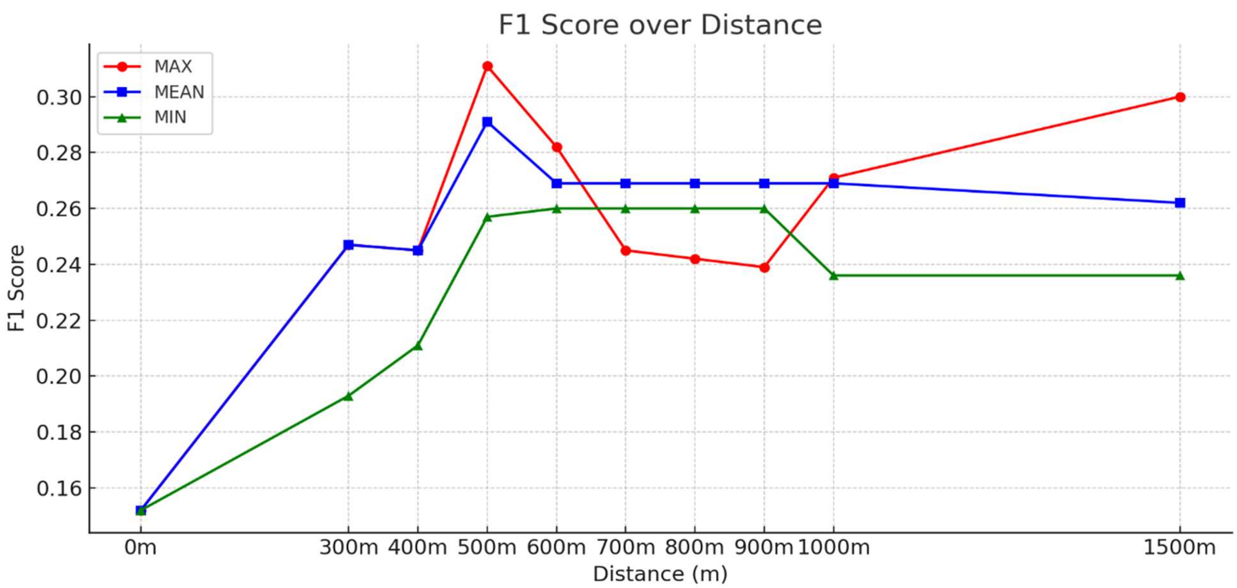

Figure 20: Milford Lake Zone A KDHE HAB Level Classification - F1 Score

Milford Lake Zone B

The point analysis demonstrated a moderate explanatory power, with an R-squared value of 0.241 for MIN, MEAN, and MAX evaluations, suggesting that about 24.1% of the variance in KDHE measurements was accounted for by the satellite data. Each model indicated a positive slope of 0.11, and while the p-values hovered around 0.11, they signaled that the relationships, though present, were not robustly significant. As the analysis extended to larger buffer distances, initial findings showed some consistency but also revealed limitations. At the 300m buffer, the R-squared values dropped to 0.072, indicating a reduction in the models' ability to explain the variability in KDHE measurements effectively. This trend was somewhat stable across the MIN, MEAN, and MAX analyses, with each maintaining similar p-values and slight variations in standard errors and biases, indicating mild prediction errors compared to the initial point analysis. Progressing to intermediate distances, from 400m to 1000m, the data exhibited a decline in the predictive accuracy of the models. For instance, at the 400m and 500m buffers, R-squared values remained low, around 0.030 to 0.047, with p-values rising, suggesting diminishing statistical significance. By the 600m to 900m buffers, while some improvements were observed in R-squared values for specific analyses, such as a 0.075 at 600m for the MAX analysis, the general pattern of increasing RMSE and shifting biases highlighted a growing discrepancy in prediction accuracy. The models' efficacy continued to decrease noticeably at the 1000m buffer and became more pronounced at the 1500m buffer, where the R-squared values dropped further, reaching 0.011 for the MIN analysis and peaking at 0.070 for the MAX analysis, which emphasized substantial errors and inconsistencies. See **Table 5**: Milford Lake Zone B Reservoir Cell Concentration Analysis Results and **Figure 22**: Milford Lake Zone B R2 Values by Buffer Distance. It is important to note the presence of multiple data points where

CyAN values are zero despite KDHE indicating bloom conditions. Again, this discrepancy highlights potential limitations or anomalies in detecting or reporting bloom presence under certain measurement or environmental conditions. See **Figure 21**: Milford Lake Zone B 400m Buffer Analysis Cell Concentration Results.

For the analysis of reclassified CyAN data into KDHE's harmful algal bloom categories at Milford Lake Zone B, performance metrics—accuracy, precision, recall, and F1 score—showed variation across buffer distances for MAX, MEAN, and MIN statistics. Accuracy for all statistics started strong at 0.583 at the point of measurement and showed a general decline across distances, with a notable stability in MEAN, which maintained higher accuracy at mid-ranges peaking at 0.543 at 500m. MIN statistics demonstrated resilience over distance, achieving a peak of 0.556 at both 800m and 1000m. This suggests that the MIN measure provides consistent accuracy, maintaining a reliable detection rate over extended distances. Precision for MAX and MEAN was highest at 300m, recording 0.442, and generally decreased towards 1500m. However, MIN showed an increasing trend in precision, peaking at 0.489 at 500m, and maintaining relatively high precision up to 1000m. This indicates that MIN was more effective in correctly identifying positive cases, even at greater distances. Recall measurements for all statistics slightly varied with MIN displaying a stronger performance, especially at 800m and 1000m where it peaked at 0.472. This upward trend in recall for MIN suggests its capability to identify relevant instances across a broader range of distances. The F1 Score, which balances precision and recall, was highest initially for all statistics but showed notable resilience in MIN. MIN achieved a higher F1 Score peaking at 0.412 at 500m and maintained good scores up to 1000m. This indicates that MIN effectively balances precision and recall over a range of distances, showing less decline compared to MAX and MEAN, which displayed more significant drops past 500m. These metrics suggest that while MAX and MEAN provided effective predictions at closer ranges, their performance in precision and recall diminished notably at greater distances. Conversely, MIN demonstrated consistent and higher scores across most metrics at longer distances. See **Figure 23**: Milford Lake Zone B KDHE HAB Level Classification - Accuracy, **Figure 24:** Milford Lake Zone B KDHE HAB Level Classification - Precision, **Figure 25**: Milford Lake Zone B KDHE HAB Level Classification - Recall, and **Figure 26**: Milford Lake Zone B KDHE HAB Level Classification - F1 Score.

Milford Lake Zone B Results

Table 5: Milford Lake Zone B Reservoir Cell Concentration Analysis Results

Analysis Type	Analysis	Slope	Intercept	R-squared	P-value	StandardError	PBIAS	RMSE	# of CyAN pixels	# of KDHE observations
Point Analysis	MIN vs KDHE_CC	0.11	399621	0.241	1.1E-01	0.061	47.8	2388135	12	12
Point Analysis	MAX vs KDHE_CC	0.11	399621	0.241	1.1E-01	0.061	47.8	2388135	12	12
Point Analysis	MEAN vs KDHE_CC	0.11	399621	0.241	1.1E-01	0.061	47.8	2388135	12	12
300m Buffer	MIN vs KDHE_CC	0.08	498143	0.072	1.3E-01	0.049	4.6	1634283	34	34
300m Buffer	MAX vs KDHE_CC	0.08	498143	0.072	1.3E-01	0.049	4.6	1634283	34	34
300m Buffer	MEAN vs KDHE_CC	0.08	498143	0.072	1.3E-01	0.049	4.6	1634283	34	34
400m Buffer	MIN vs KDHE_CC	0.05	500124	0.030	3.2E-01	0.049	7.0	1640064	58	34
400m Buffer	MAX vs KDHE_CC	0.06	583793	0.046	2.2E-01	0.052	-10.3	1637838	58	34
400m Buffer	MEAN vs KDHE_CC	0.06	537118	0.047	2.2E-01	0.049	-1.2	1637367	58	34
500m Buffer	MIN vs KDHE_CC	0.05	500124	0.030	3.2E-01	0.049	7.0	1640064	58	34
500m Buffer	MAX vs KDHE_CC	0.06	583793	0.046	2.2E-01	0.052	-10.3	1637838	58	34
500m Buffer	MEAN vs KDHE_CC	0.06	537118	0.047	2.2E-01	0.049	-1.2	1637367	58	34
600m Buffer	MIN vs KDHE_CC	0.03	442613	0.018	4.4E-01	0.045	16.8	1622650	130	35
600m Buffer	MAX vs KDHE_CC	0.09	629443	0.075	1.1E-01	0.057	-25.5	1614942	130	35
600m Buffer	MEAN vs KDHE_CC	0.06	525118	0.051	1.9E-01	0.048	-1.7	1615494	130	35
700m Buffer	MIN vs KDHE_CC	0.04	394406	0.026	3.5E-01	0.042	24.3	1603081	177	36
700m Buffer	MAX vs KDHE_CC	0.08	676446	0.062	1.4E-01	0.057	-36.0	1601169	177	36
700m Buffer	MEAN vs KDHE_CC	0.06	519778	0.055	1.7E-01	0.046	-2.4	1593713	177	36
800m Buffer	MIN vs KDHE_CC	0.04	374259	0.032	3.0E-01	0.040	27.9	1604516	235	36
800m Buffer	MAX vs KDHE_CC	0.11	682471	0.096	6.6E-02	0.058	-40.9	1597184	235	36
800m Buffer	MEAN vs KDHE_CC	0.07	506963	0.065	1.3E-01	0.047	-0.6	1592203	235	36
900m Buffer	MIN vs KDHE_CC	0.02	364230	0.012	5.2E-01	0.038	31.1	1608041	304	36
900m Buffer	MAX vs KDHE_CC	0.11	702234	0.090	7.6E-02	0.059	-44.7	1600791	304	36
900m Buffer	MEAN vs KDHE_CC	0.06	522612	0.050	1.9E-01	0.047	-2.7	1594138	304	36
1000m Buffer	MIN vs KDHE_CC	0.03	329579	0.015	4.7E-01	0.036	37.7	1590555	398	37
1000m Buffer	MAX vs KDHE_CC	0.11	707896	0.080	9.0E-02	0.062	-45.3	1580610	398	37
1000m Buffer	MEAN vs KDHE_CC	0.06	501663	0.049	1.9E-01	0.047	1.6	1572319	398	37
1500m Buffer	MIN vs KDHE_CC	0.02	249208	0.011	5.2E-01	0.031	51.0	1563035	871	39
1500m Buffer	MAX vs KDHE_CC	0.11	830009	0.070	1.0E-01	0.064	-78.0	1577741	871	39
1500m Buffer	MEAN vs KDHE_CC	0.05	498838	0.037	2.4E-01	0.045	-1.5	1536720	871	39

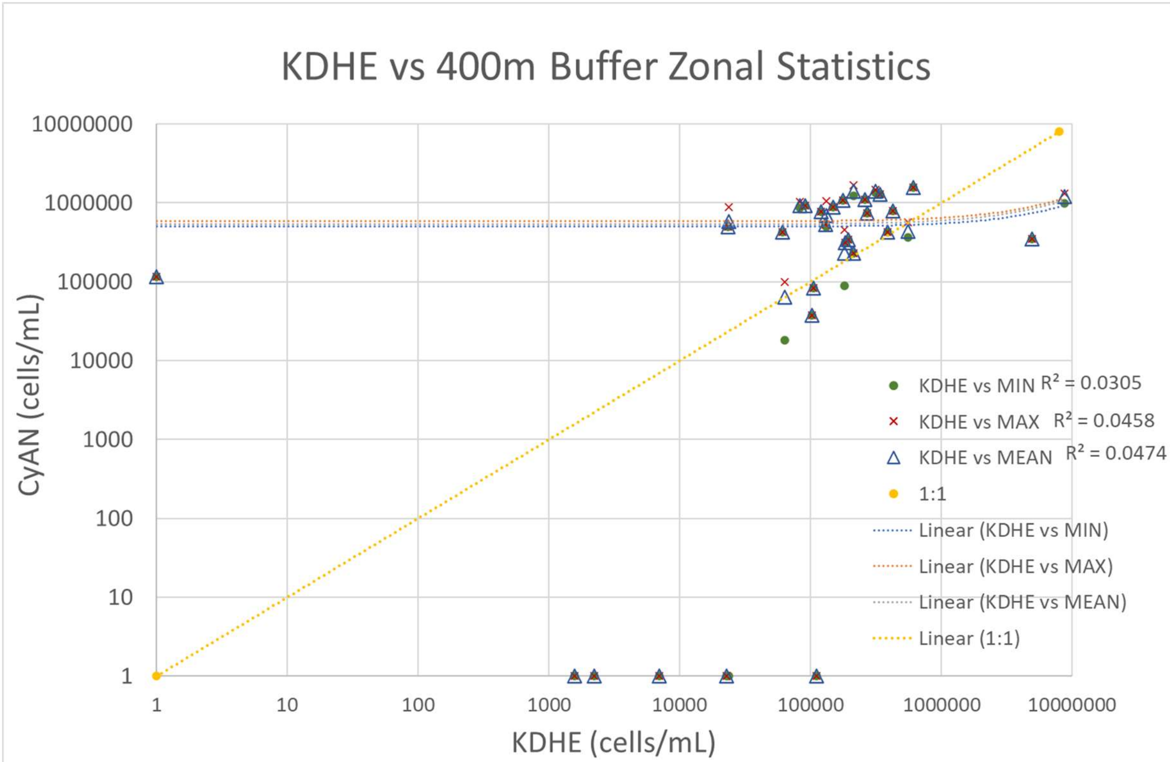

Figure 21: Milford Lake Zone B 400m Buffer Analysis Cell Concentration Results

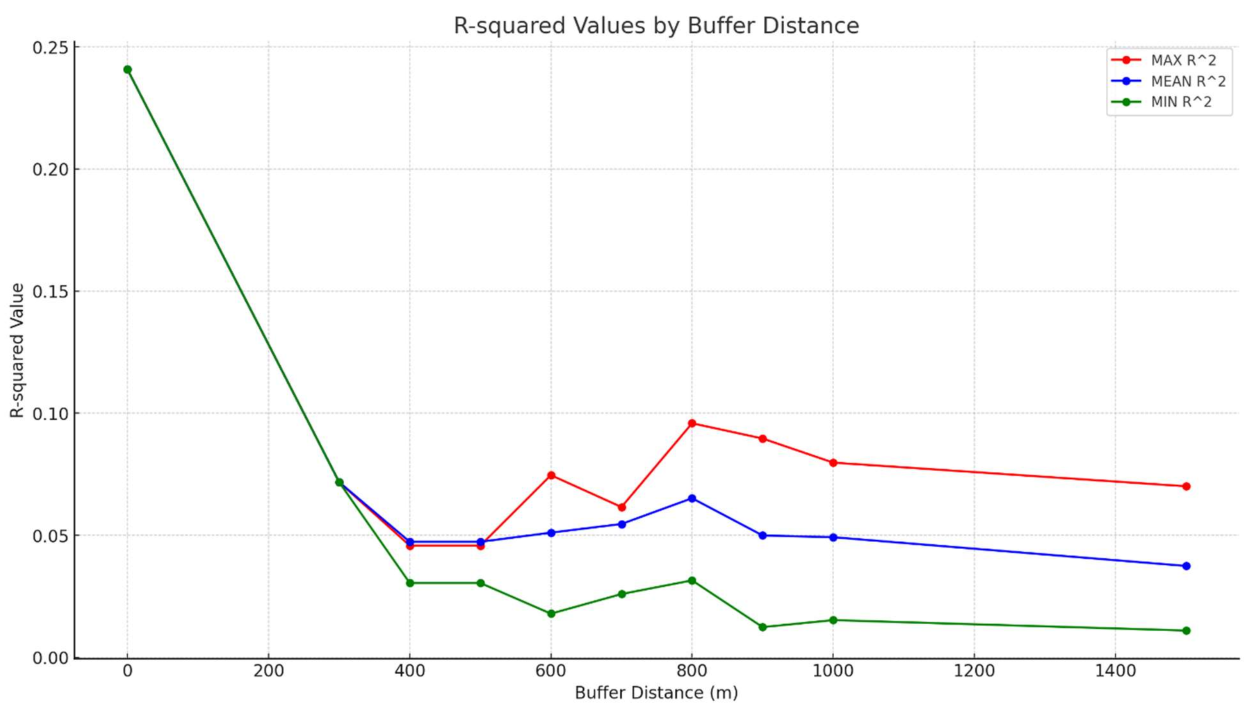

Figure 22: Milford Lake Zone B R^2 Values by Buffer Distance

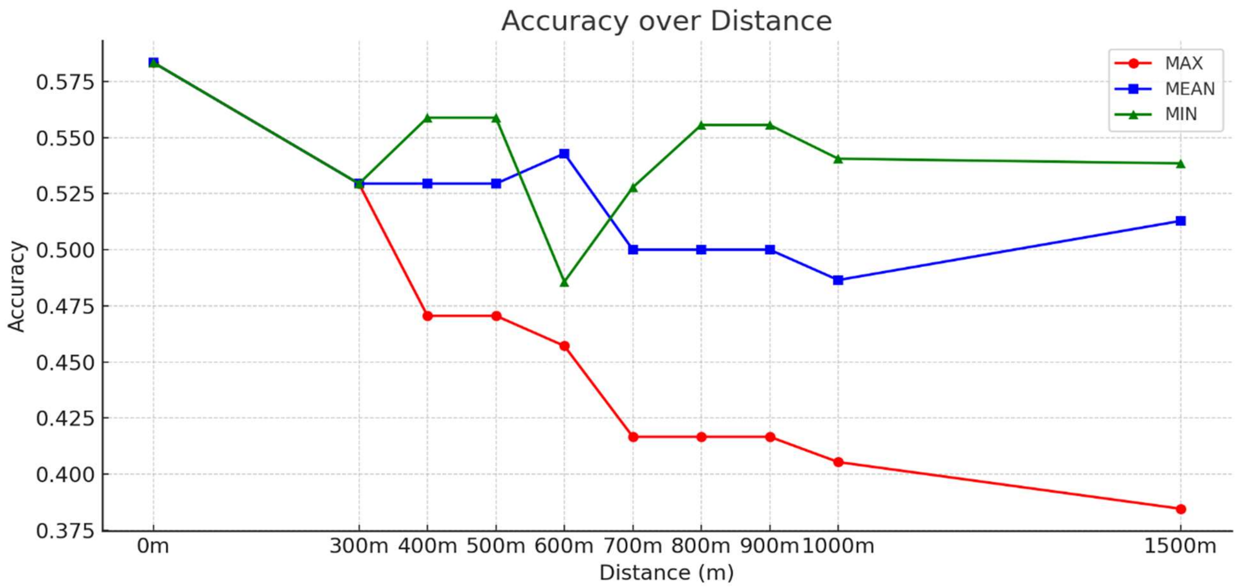

Figure 23: Milford Lake Zone B KDHE HAB Level Classification - Accuracy

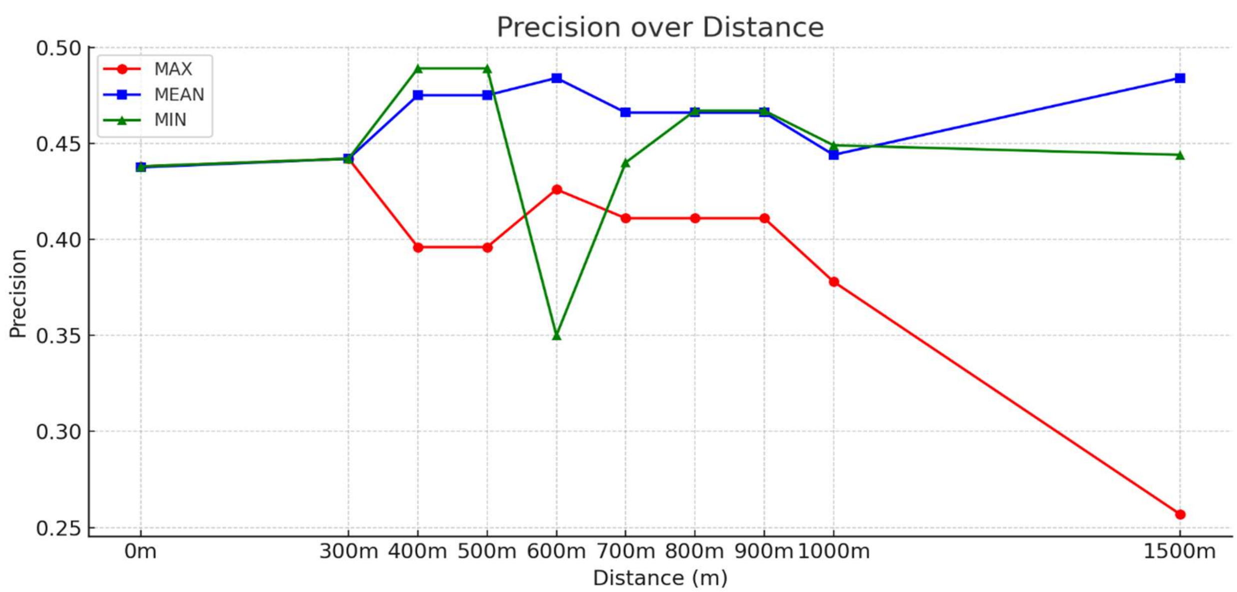

Figure 24: Milford Lake Zone B KDHE HAB Level Classification - Precision

Figure 25: Milford Lake Zone B KDHE HAB Level Classification - Recall

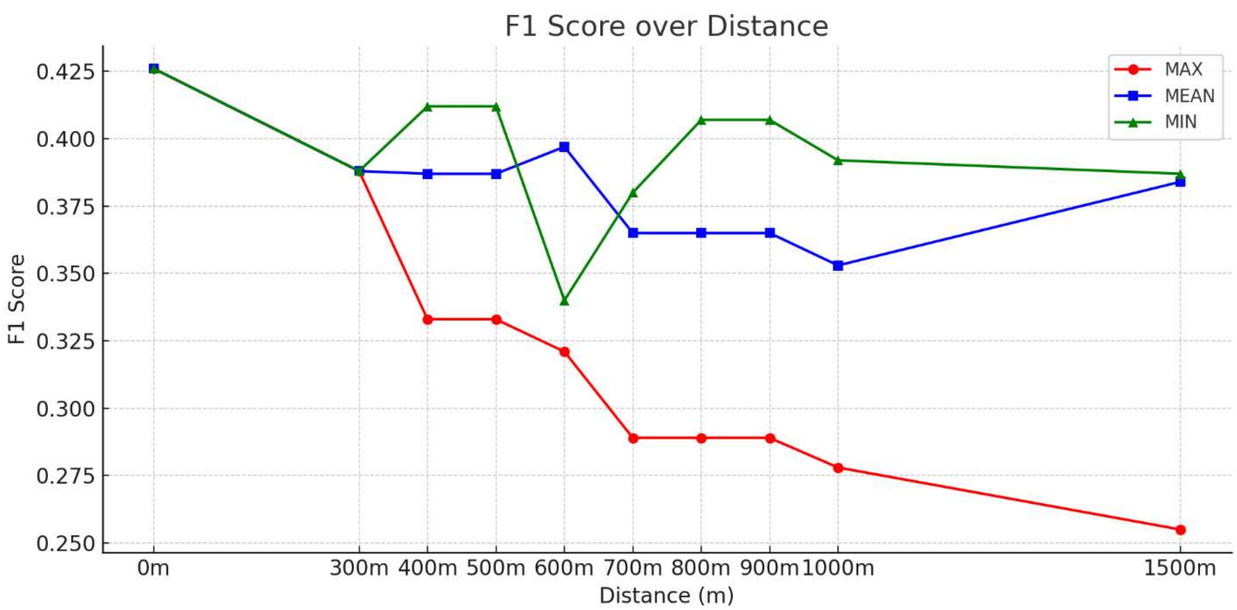

Figure 26: Milford Lake Zone B KDHE HAB Level Classification - F1 Score

Milford Lake Zone C

For the point analysis, the model's explanatory power was limited, with an R-squared value of 0.068 across MIN, MEAN, and MAX evaluations. This indicates that only 6.8% of the variance in KDHE measurements could be explained by the CyAN data, reflecting a rather weak predictive capability. Furthermore, the high p-values around 0.27 in these analyses suggested that the statistical significance of these relationships was minimal, despite a relatively high positive slope of 0.61. As the analysis incorporated increasing buffer distances, improvements in model performance were observed but remained moderate. At the 300m buffer, the R-squared values increased to around 0.357, indicating a better but still modest ability to explain variance in KDHE measurements. The statistical significance of these models improved markedly, as indicated by very low p-values (1.2E-05), suggesting stronger relationships at this buffer distance compared to the point analysis. Progressing to intermediate buffer distances between 400m and 1000m, the models exhibited fluctuating efficacy. R-squared values, such as 0.369 at 400m, suggested a reasonable level of explanatory power, but this metric began to decrease with further increased distances. While the models maintained statistical significance (low p-values), the increasing RMSE indicated growing inaccuracies in the predictions. At larger buffer distances, specifically the 1500m buffer, the diminishing trend in model accuracy was stark. Here, the R-squared values decreased significantly, though some remained above 0.3, indicating inconsistent effectiveness. The RMSE continued to be high, and percentage biases showed substantial deviations, which underscored significant discrepancies in model accuracy at these extended ranges. See **Table 6**: Milford Lake Zone C Reservoir Cell Concentration Analysis Results and **Figure 28**: Milford Lake Zone C R2 Values by Buffer Distance. Like the other waterbodies, it is important to note the presence of multiple data points where CyAN values are zero despite

KDHE indicating bloom conditions. Again, this discrepancy highlights potential limitations or anomalies in detecting or reporting bloom presence under certain measurement or environmental conditions. See **Figure 27**: Milford Lake Zone C 300m Buffer Analysis Cell Concentration Results.

In the analysis of reclassified CyAN data into KDHE's harmful algal bloom categories at Milford Lake Zone C, performance metrics—accuracy, precision, recall, and F1 score—present nuanced relationships across buffer distances for MAX, MEAN, and MIN statistics. Accuracy for all statistics started at 0.6 at the point of measurement, with a noticeable increase to approximately 0.696 at 300m for MAX, MEAN, and MIN, showing a unified peak in the early range. Interestingly, accuracy for MEAN and MIN maintained stronger levels up to 600m, with MEAN peaking at 0.708 at 500m and MIN maintaining over 0.66 up to 1000m. This suggests that MIN, in particular, sustains its accuracy over a broader range of distances, possibly indicating a robustness in MIN's data handling or sensitivity to detect HAB levels accurately over larger spatial extents. Precision saw its highest values at 300m across all statistics, with MAX and MEAN around 0.484 and MIN at 0.489. Notably, MIN maintained higher precision than MAX and MEAN beyond 500m, particularly peaking again at 0.508 at 1000m. This higher precision at extended distances for MIN implies a better consistency in positively identifying true HAB cases. Recall for all statistics showed modest increases up to 300m, with MIN displaying superior performance by maintaining higher recall rates up to 1000m, peaking at 0.466. This higher recall rate for MIN at extended distances suggests an enhanced ability to detect all relevant HAB instances. F1 Score analysis reflects a balance of precision and recall, with MIN consistently achieving higher scores across distances. It peaked at 0.449 at 1000m, showcasing its balanced performance even at extended ranges. The stable and high F1 Score of

MIN compared to the other statistics points to its effectiveness in maintaining a reliable balance between identifying true positives and the ability to detect all relevant cases. The overall analysis demonstrates that while MAX and MEAN provide effective initial predictions at closer ranges, their performance diminishes at greater distances. In contrast, MIN not only maintains higher accuracy and precision but also exhibits sustained recall and F1 Scores at longer distances. See **Figure 29**: Milford Lake Zone C KDHE HAB Level Classification - Accuracy, **Figure 30**: Milford Lake Zone C KDHE HAB Level Classification - Precision, **Figure 31**: Milford Lake Zone C KDHE HAB Level Classification - Recall, and **Figure 32**: Milford Lake Zone C KDHE HAB Level Classification - F1 Score.

Milford Lake Zone C Results

Table 6: Milford Lake Zone C Reservoir Cell Concentration Analysis Results

Analysis Type	Analysis	Slope	Intercept	R-squared	P-value	Standard Error	PBIAS	RMSE	# of CyAN pixels	# of KDHE observations
Point Analysis	MIN vs KDHE_CC	0.61	657065	0.068	2.7E-01	0.534	-335.7	1031272	20	20
Point Analysis	MAX vs KDHE_CC	0.61	657065	0.068	2.7E-01	0.534	-335.7	1031272	20	20
Point Analysis	MEAN vs KDHE_CC	0.61	657065	0.068	2.7E-01	0.534	-335.7	1031272	20	20
300m Buffer	MIN vs KDHE_CC	0.31	654327	0.357	1.2E-05	0.063	7.2	1638597	46	46
300m Buffer	MAX vs KDHE_CC	0.31	654327	0.357	1.2E-05	0.063	7.2	1638597	46	46
300m Buffer	MEAN vs KDHE_CC	0.31	654327	0.357	1.2E-05	0.063	7.2	1638597	46	46
400m Buffer	MIN vs KDHE_CC	0.32	590303	0.369	5.9E-06	0.063	12.9	1615814	67	47
400m Buffer	MAX vs KDHE_CC	0.31	651338	0.346	1.4E-05	0.064	6.3	1622740	67	47
400m Buffer	MEAN vs KDHE_CC	0.32	620983	0.359	8.4E-06	0.063	9.5	1618853	67	47
500m Buffer	MIN vs KDHE_CC	0.26	554193	0.297	6.3E-05	0.059	23.6	1669364	206	48
500m Buffer	MAX vs KDHE_CC	0.33	827650	0.317	3.1E-05	0.072	-20.8	1606750	206	48
500m Buffer	MEAN vs KDHE_CC	0.30	676771	0.342	1.3E-05	0.061	3.8	1621290	206	48
600m Buffer	MIN vs KDHE_CC	0.22	602830	0.395	4.7E-07	0.037	41.2	2548258	211	53
600m Buffer	MAX vs KDHE_CC	0.23	897135	0.325	8.3E-06	0.046	13.2	2480574	211	53
600m Buffer	MEAN vs KDHE_CC	0.22	736430	0.394	5.0E-07	0.039	28.4	2508264	211	53
700m Buffer	MIN vs KDHE_CC	0.22	572682	0.422	1.1E-07	0.035	42.7	2530839	280	54
700m Buffer	MAX vs KDHE_CC	0.23	902833	0.329	5.8E-06	0.046	11.0	2459219	280	54
700m Buffer	MEAN vs KDHE_CC	0.23	725578	0.400	2.9E-07	0.038	28.0	2487021	280	54
800m Buffer	MIN vs KDHE_CC	0.20	441507	0.438	3.8E-08	0.031	55.3	2570498	405	55
800m Buffer	MAX vs KDHE_CC	0.24	981438	0.321	6.5E-06	0.048	0.7	2427107	405	55
800m Buffer	MEAN vs KDHE_CC	0.22	710811	0.397	2.5E-07	0.037	28.8	2477375	405	55
900m Buffer	MIN vs KDHE_CC	0.21	370771	0.480	3.4E-09	0.030	60.8	2563000	486	56
900m Buffer	MAX vs KDHE_CC	0.27	948851	0.373	5.7E-07	0.048	-1.1	2374788	486	56
900m Buffer	MEAN vs KDHE_CC	0.23	663225	0.441	2.4E-08	0.036	30.8	2446610	486	56
1000m Buffer	MIN vs KDHE_CC	0.20	348120	0.491	1.9E-09	0.028	63.3	2578167	634	56
1000m Buffer	MAX vs KDHE_CC	0.28	1003014	0.395	2.1E-07	0.047	-7.7	2364405	634	56
1000m Buffer	MEAN vs KDHE_CC	0.23	682856	0.439	2.7E-08	0.036	29.2	2446201	634	56
1500m Buffer	MIN vs KDHE_CC	0.17	225546	0.490	1.3E-09	0.023	76.4	2644651	1413	57
1500m Buffer	MAX vs KDHE_CC	0.30	1222640	0.276	2.8E-05	0.066	-35.1	2372779	1413	57
1500m Buffer	MEAN vs KDHE_CC	0.22	660103	0.388	2.3E-07	0.037	31.5	2447548	1413	57

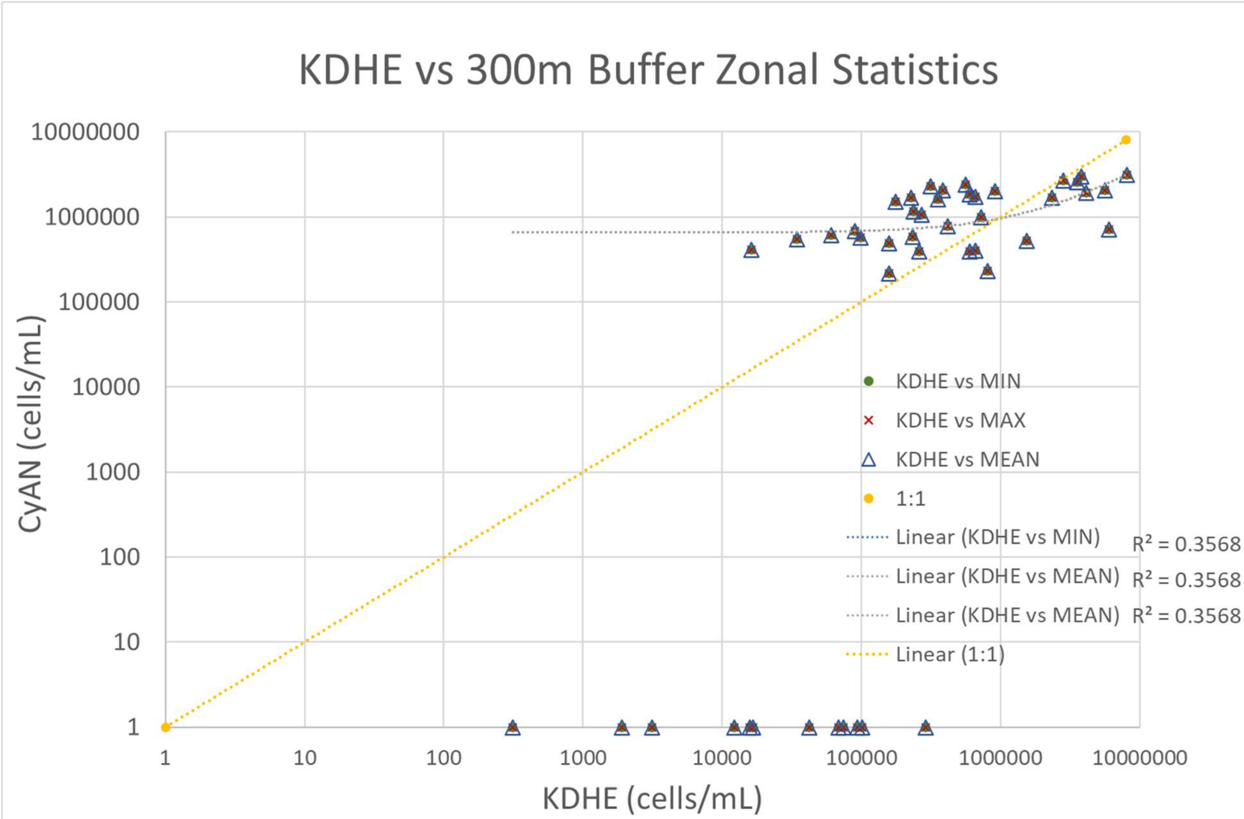

Figure 27: Milford Lake Zone C 300m Buffer Analysis Cell Concentration Results

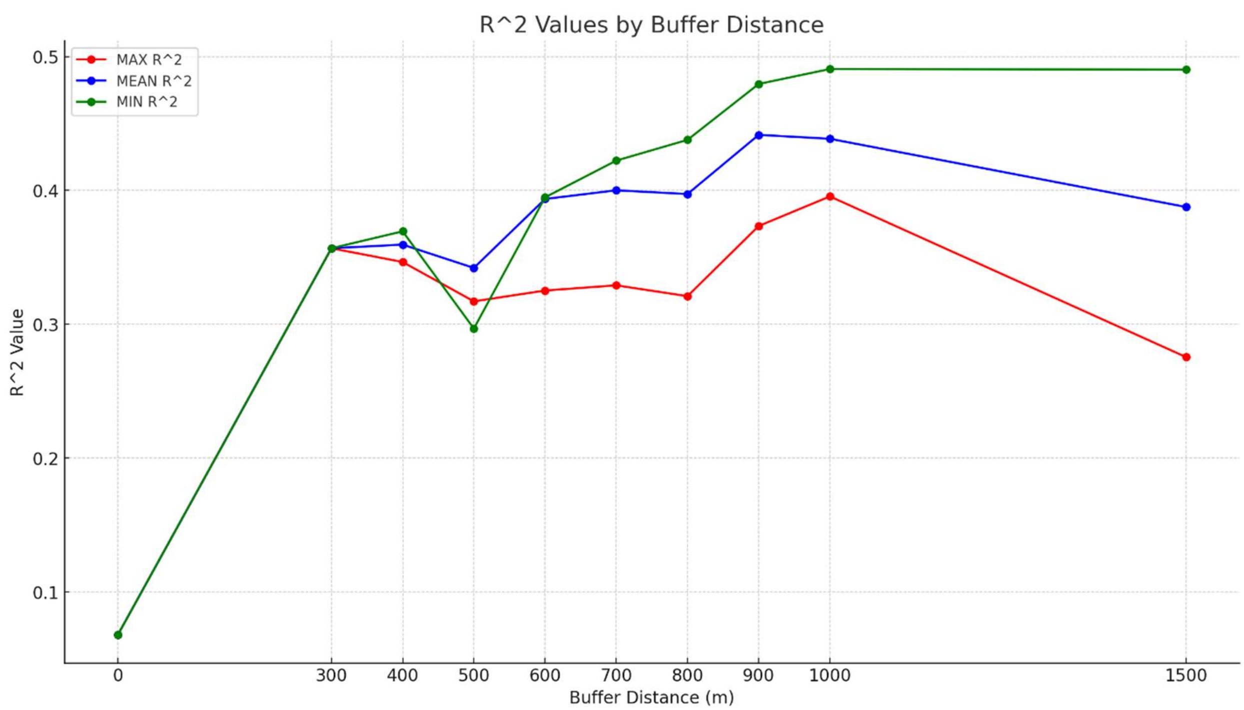

Figure 28: Milford Lake Zone C R^2 Values by Buffer Distance

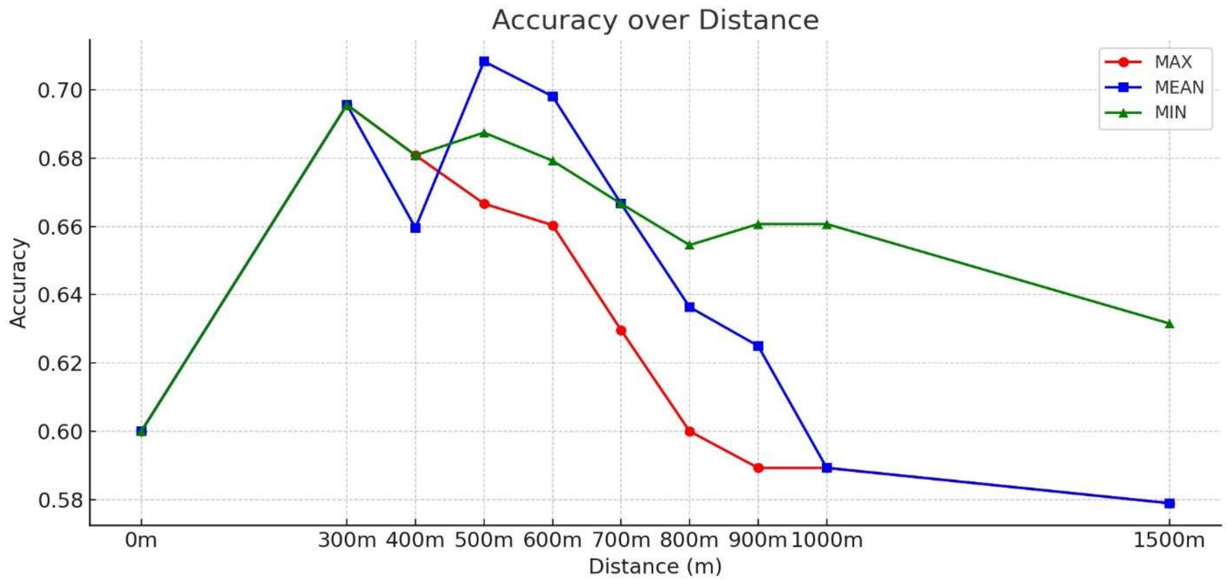

Figure 29: Milford Lake Zone C KDHE HAB Level Classification - Accuracy

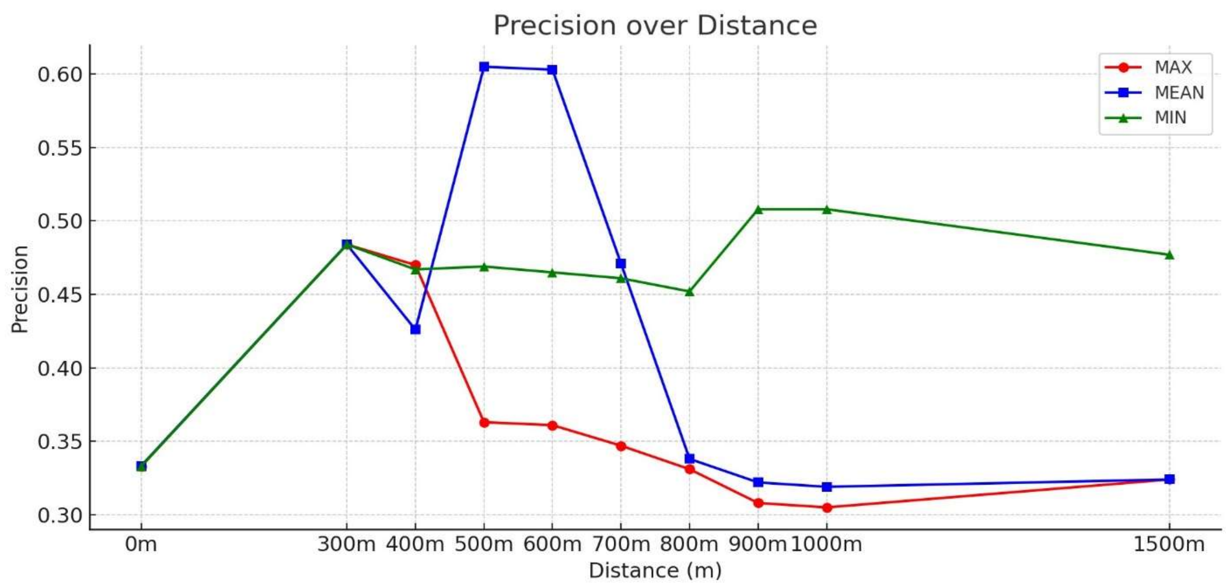

Figure 30: Milford Lake Zone C KDHE HAB Level Classification - Precision

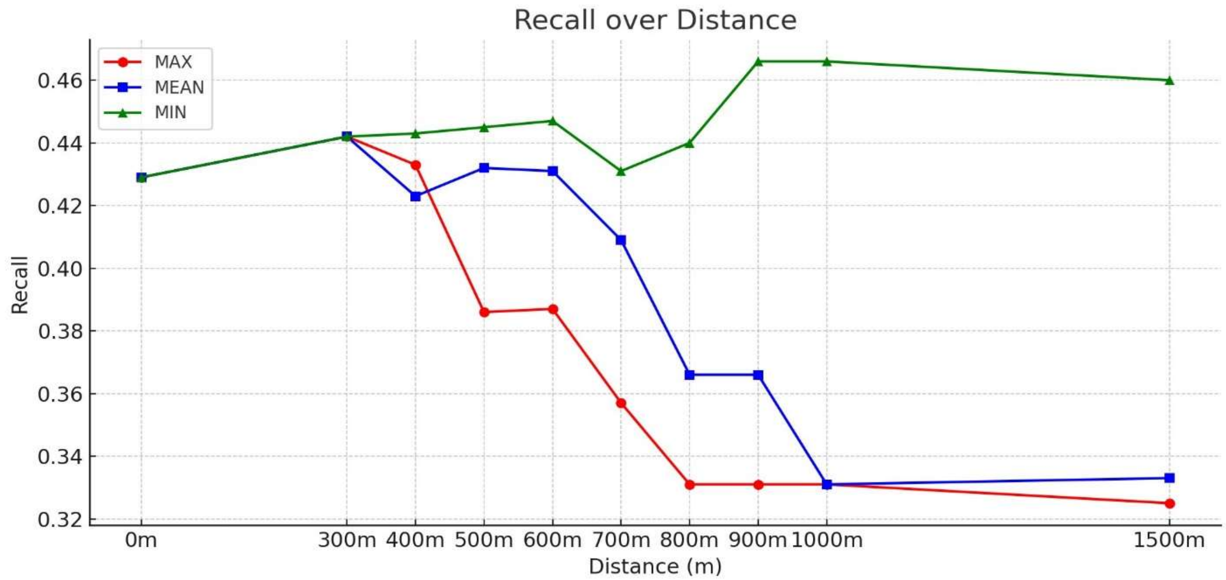

Figure 31: Milford Lake Zone C KDHE HAB Level Classification - Recall

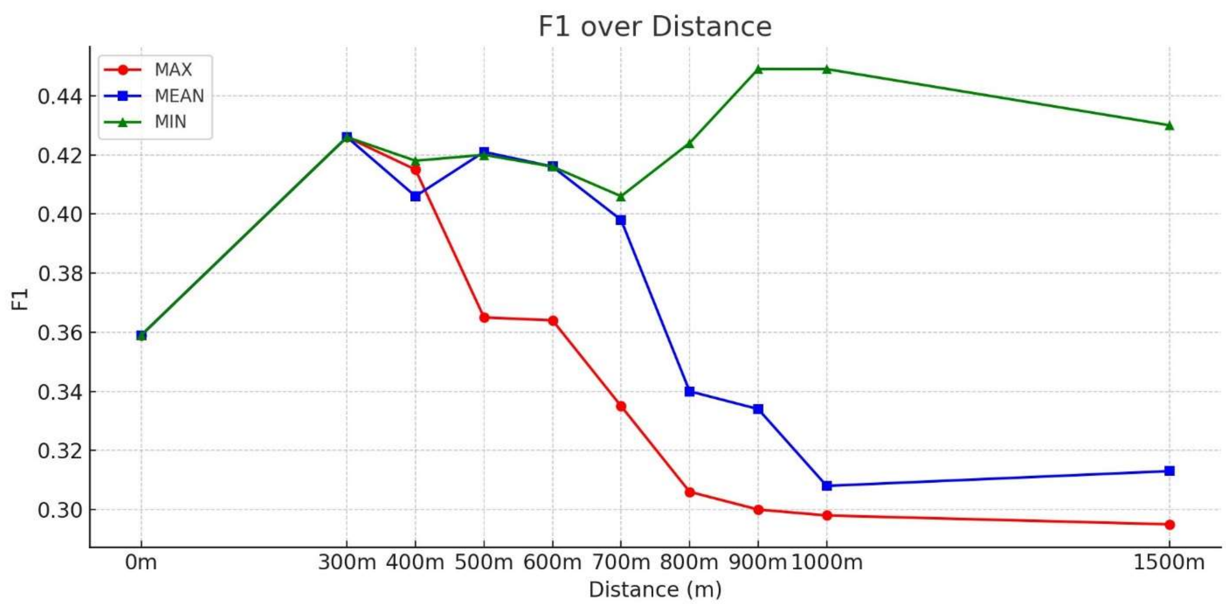

Figure 32: Milford Lake Zone C KDHE HAB Level Classification - F1 Score

Webster Lake

Initially, with the point analysis, the predictive capacity of the models was extremely limited, with an R-squared value of just 0.001 for MIN, MEAN, and MAX evaluations. This minuscule R-squared value indicates that the satellite data accounted for an almost negligible portion of the variance in KDHE measurements. Additionally, the consistently high p-values around 0.90 for these analyses underscored the lack of statistical significance in these relationships. As the buffer distances expanded, minor improvements in model performance were observed, though they remained generally low. At the 300m buffer, the R-squared values slightly increased to 0.023 for MAX analysis. However, p-values remained relatively high, continuing to show the models' statistical insignificance. Despite these slight increases in R-squared values, the substantial standard errors and biases indicated ongoing challenges in model accuracy and reliability. Progressing to intermediate buffer distances between 400m and 900m, the models still displayed very low effectiveness. The R-squared values in this range were marginally better but did not exceed 0.009, pointing to a continued weak predictive ability. The p-values across these distances also suggested that the models failed to achieve statistical significance, reinforcing the limited utility of the satellite data in accurately predicting ground-based measurements. The trend of diminishing model accuracy became even more evident at larger buffer distances, specifically at the 1000m and 1500m buffers. Here, the R-squared values, such as 0.016 for MAX analyses at 1000 meters, remained low, reflecting persistent inadequacies in the models' explanatory power. The high p-values and large biases at these extended distances further highlighted significant discrepancies in model accuracy, with RMSE values escalating, indicating increased prediction errors. See **Table 7**: Webster Lake Cell Concentration Analysis Results and **Figure 34**: Webster Lake R2 Values by Buffer Distance. It is important to note the presence of multiple data points where CyAN values are zero despite KDHE indicating bloom conditions. Again, this discrepancy highlights potential limitations or anomalies in detecting or reporting bloom presence under certain measurement or environmental conditions. See **Figure 33**: Webster Lake 800m Buffer Analysis Cell Concentration Results.

For the analysis of reclassified CyAN data into KDHE's harmful algal bloom categories at Webster Lake, the performance metrics—accuracy, precision, recall, and F1 score—were evaluated across various buffer distances for MAX, MEAN, and MIN statistics. Accuracy across all statistics began at approximately 0.643 for the point analysis. It peaked at 300m for MAX and MEAN at about 0.652, indicating an initial increase in accuracy at closer distances. However, for all statistics, there was a noticeable decline beyond 300m, with MIN experiencing the most significant drop, reaching as low as 0.459 at 1500m. This decline in accuracy at greater distances suggests a diminishing reliability of the satellite data as the buffer from the measurement point increases. Precision was highest for all statistics at 300m, with MAX achieving 0.329 and MIN peaking slightly lower at 0.321. Following this peak, there was a gradual decrease across all statistics, with MIN experiencing a steeper decline to 0.240 at 1500m. This reduction in precision indicates a decrease in the ability of the data to predict true positive cases of harmful algal blooms as the distance from the sampling point increases. Recall remained relatively stable for MAX and MEAN with slight variations, peaking at 0.375 at 300m. MIN showed less stability, with a peak at 0.350 at 300m and a decrease to 0.304 at 1500m. The generally stable recall rates suggest that the sensitivity of the model to detect all relevant instances of blooms does not drastically diminish with distance, although MIN shows some reduction at the furthest range. F1 Score, balancing precision and recall, mirrored the trends seen in precision and recall. It was highest at 300m for all statistics, with MAX and MEAN scoring 0.349 and MIN slightly lower at 0.334. Thereafter, F1 scores decreased with increasing buffer distance, reflecting the combined impact of reducing precision and slightly variable recall rates. These trends indicate that while the satellite-derived data initially offers strong predictive capability at closer ranges, its effectiveness in accurately predicting harmful algal bloom categories decreases with distance. See **Figure 35**: Webster Lake KDHE HAB Level Classification - Accuracy, **Figure 36**: Webster Lake KDHE HAB Level Classification - Precision, **Figure 37**: Webster Lake KDHE HAB Level Classification - Recall, and **Figure 38**: Webster Lake KDHE HAB Level Classification - F1 Score.

Webster Lake Results

Table 7: Webster Lake Cell Concentration Analysis Results

Analysis Type	Analysis	Slope	Intercept	R-squared	P-value	Standard Error	PBIAS	RMSE	# of CyAN pixels	# of KDHE observations
Point Analysis	MIN vs KDHE_CC	0.03	551164	0.001	9.0E-01	0.242	5.2	1013110	14	14
Point Analysis	MAX vs KDHE_CC	0.03	551164	0.001	9.0E-01	0.242	5.2	1013110	14	14
Point Analysis	MEAN vs KDHE_CC	0.03	551164	0.001	9.0E-01	0.242	5.2	1013110	14	14
300m Buffer	MIN vs KDHE_CC	0.08	365542	0.011	6.3E-01	0.160	7.4	826666	36	23
300m Buffer	MAX vs KDHE_CC	0.13	419629	0.023	4.9E-01	0.184	-12.1	822752	36	23
300m Buffer	MEAN vs KDHE_CC	0.10	393694	0.016	5.7E-01	0.171	-2.0	824210	36	23
400m Buffer	MIN vs KDHE_CC	0.03	348284	0.002	8.1E-01	0.139	21.1	854313	54	27
400m Buffer	MAX vs KDHE_CC	0.07	418265	0.008	6.7E-01	0.160	1.7	846812	54	27
400m Buffer	MEAN vs KDHE_CC	0.05	382277	0.004	7.5E-01	0.149	12.1	849961	54	27
500m Buffer	MIN vs KDHE_CC	0.04	324415	0.004	7.6E-01	0.136	23.7	842494	73	28
500m Buffer	MAX vs KDHE_CC	0.08	408820	0.009	6.4E-01	0.158	0.5	833878	73	28
500m Buffer	MEAN vs KDHE_CC	0.06	367719	0.006	7.1E-01	0.146	12.3	836972	73	28
600m Buffer	MIN vs KDHE_CC	0.05	295788	0.006	7.1E-01	0.137	29.7	845679	84	28
600m Buffer	MAX vs KDHE_CC	0.07	409653	0.009	6.4E-01	0.158	0.3	833892	84	28
600m Buffer	MEAN vs KDHE_CC	0.06	358788	0.006	6.9E-01	0.145	14.2	837356	84	28
700m Buffer	MIN vs KDHE_CC	-0.01	296031	0.000	9.4E-01	0.135	34.1	850551	144	28
700m Buffer	MAX vs KDHE_CC	0.06	472485	0.005	7.2E-01	0.159	-12.9	837102	144	28
700m Buffer	MEAN vs KDHE_CC	0.02	377495	0.001	8.6E-01	0.144	12.9	838622	144	28
800m Buffer	MIN vs KDHE_CC	0.00	288309	0.000	9.7E-01	0.128	32.5	812916	176	31
800m Buffer	MAX vs KDHE_CC	0.07	462236	0.007	6.5E-01	0.152	-16.8	801634	176	31
800m Buffer	MEAN vs KDHE_CC	0.04	362348	0.003	7.9E-01	0.136	11.5	801790	176	31
900m Buffer	MIN vs KDHE_CC	0.00	285995	0.000	9.8E-01	0.127	33.0	813277	191	31
900m Buffer	MAX vs KDHE_CC	0.07	474146	0.006	6.7E-01	0.152	-19.3	802953	191	31
900m Buffer	MEAN vs KDHE_CC	0.03	367398	0.001	8.4E-01	0.136	11.1	802002	191	31
1000m Buffer	MIN vs KDHE_CC	0.00	240719	0.000	9.9E-01	0.119	38.4	775512	274	35
1000m Buffer	MAX vs KDHE_CC	0.11	439825	0.016	4.7E-01	0.151	-26.3	762211	274	35
1000m Buffer	MEAN vs KDHE_CC	0.04	322595	0.003	7.5E-01	0.130	13.7	761885	274	35
1500m Buffer	MIN vs KDHE_CC	0.00	188759	0.004	6.9E-01	0.008	90.5	9809171	608	37
1500m Buffer	MAX vs KDHE_CC	-0.01	516871	0.013	5.0E-01	0.012	74.2	9754207	608	37
1500m Buffer	MEAN vs KDHE_CC	-0.01	327083	0.009	5.8E-01	0.009	83.7	9784706	608	37

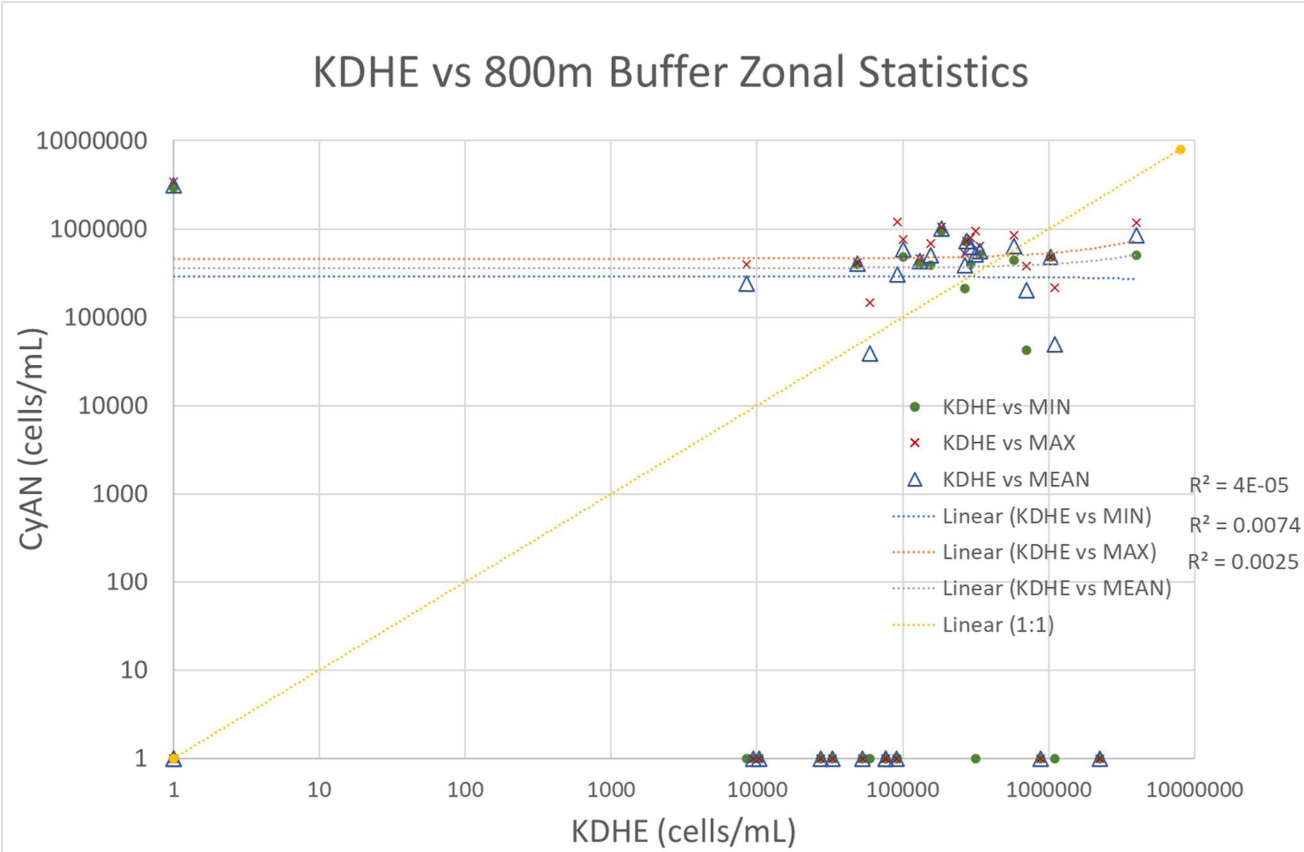

Figure 33: Webster Lake 800m Buffer Analysis Cell Concentration Results

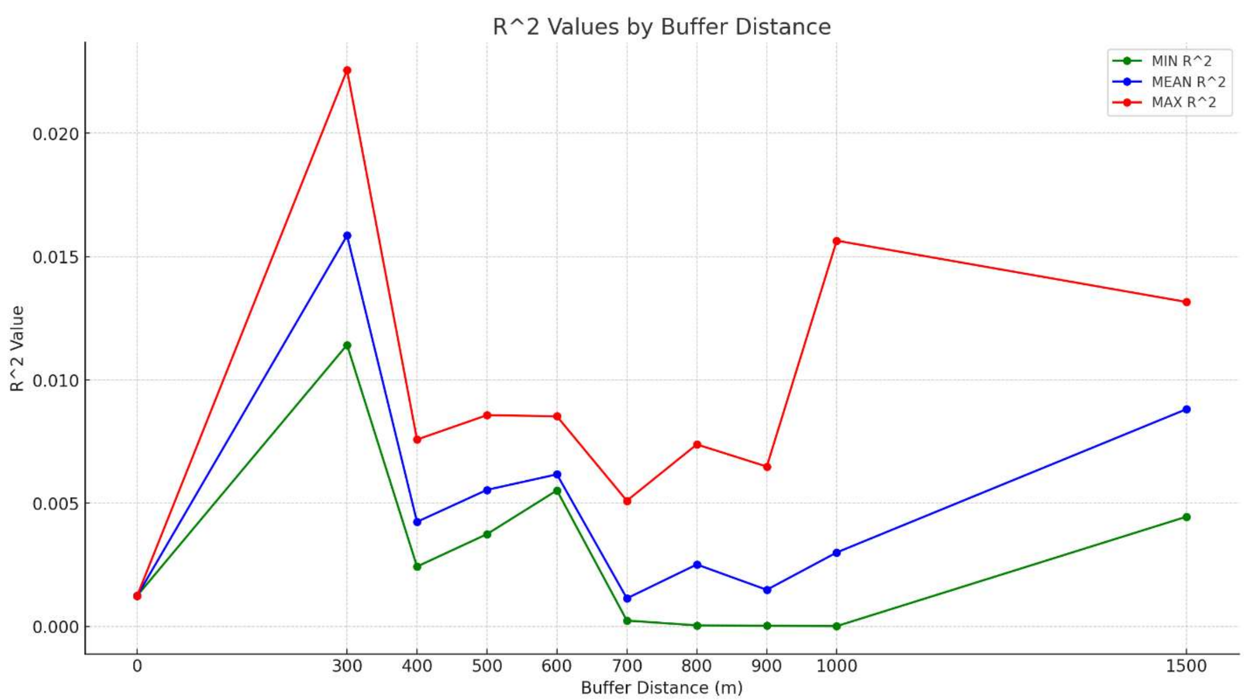

Figure 34: Webster Lake R^2 Values by Buffer Distance

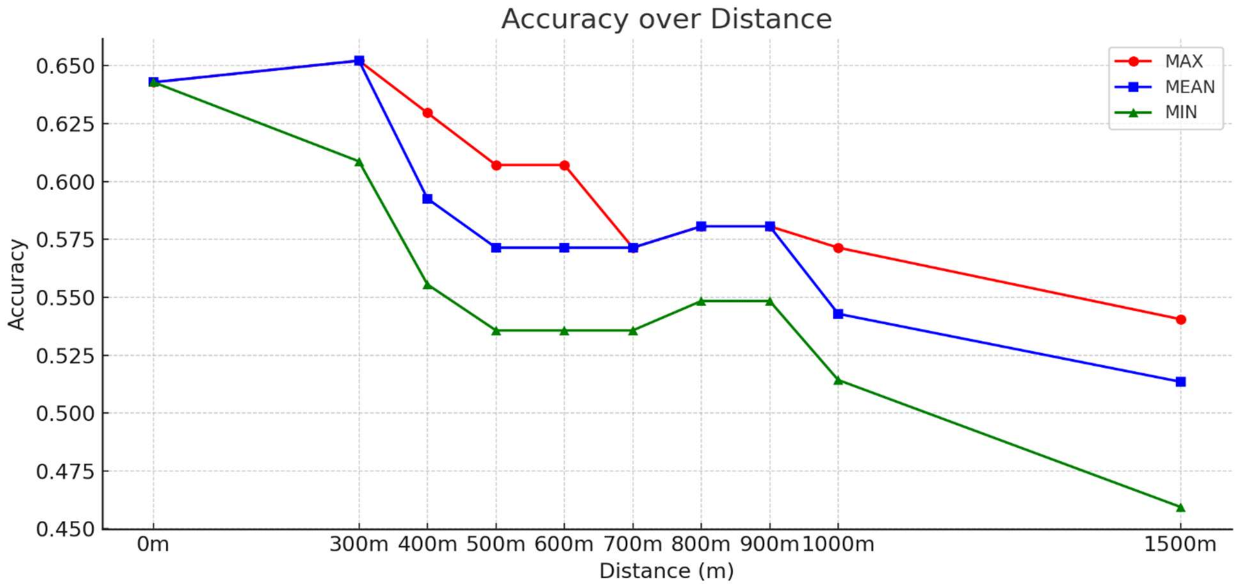

Figure 35: Webster Lake KDHE HAB Level Classification - Accuracy

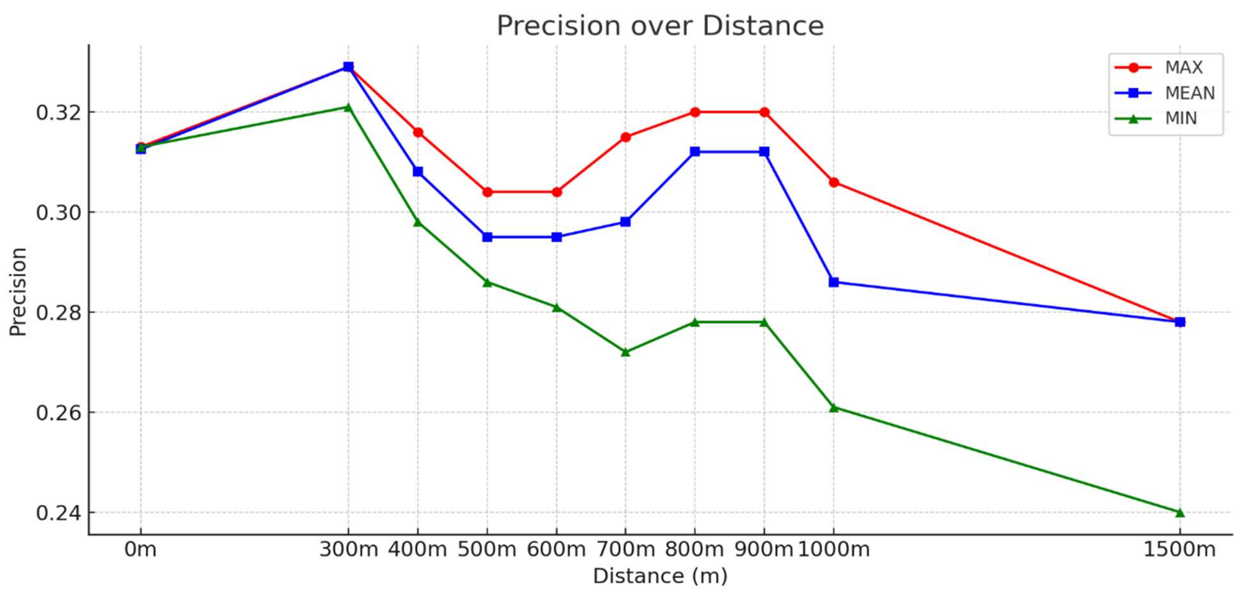

Figure 36: Webster Lake KDHE HAB Level Classification - Precision

Figure 37: Webster Lake KDHE HAB Level Classification - Recall

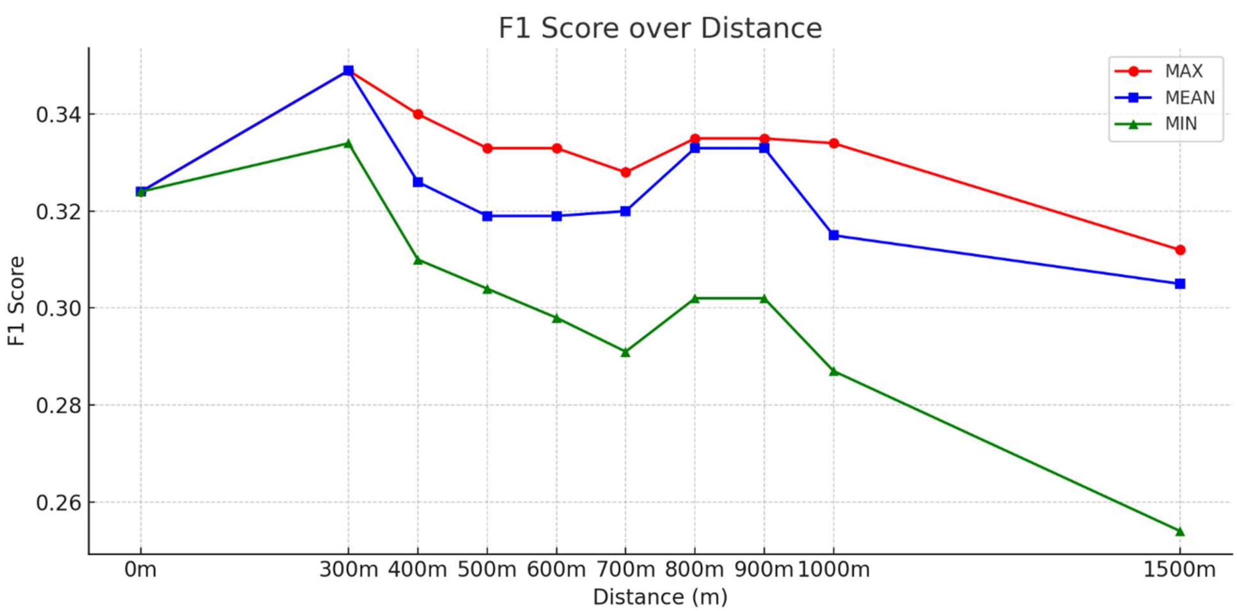

Figure 38: Webster Lake KDHE HAB Level Classification - F1 Score

Discussion

The core objective of this study was to conduct a comparative analysis of cyanobacterial concentration data, collected through two methodologies: the in-situ sampling by the Kansas Department of Health and Environment (KDHE) and satellite-derived data from the Cyanobacteria Assessment Network (CyAN) data portal. The study also examined the agreement between the KDHE HAB warning levels and the CyAN values classified within the same warning level cell concentration ranges.

This study was focused on the correlation of KDHE in-situ sampling and CyAN satellite derived vales at the waterbody scale but this was constrained by the number of data points made available by KDHE. This was further limited by the number of days the satellite was able to get a reading whether that be due to cloud cover or scheduled down time. A sample size of 30 was selected as it is a common statistical benchmark and it allowed for a feasible balance between comprehensive coverage of the various sites and the practical constraints of the study. While a sample size of this magnitude allows for preliminary analysis, it may not be sufficient to generalize the findings across all waterbodies or to capture the full variability inherent in cyanobacterial concentrations. The distribution of these 30 samples across multiple testing sites in each waterbody introduces further complexity to the analysis. Given that the samples were split between at least two locations per waterbody, the effective sample size for each specific site becomes even smaller. This subdivision reduces the statistical power of the analysis at each individual site and increases the uncertainty associated with the findings. Moreover, it challenges the ability to detect nuanced environmental differences or subtle effects of various factors influencing cyanobacterial concentrations at specific locations.

The analytical approach involved linear regression to establish a correlation between these datasets, complemented by Root Mean Square Error (RMSE) calculations to evaluate the predictive accuracy of satellite data across various spatial resolutions. While correlation coefficients provided a measure of the linear relationship between datasets, additional metrics such as accuracy, precision, recall, and F1 Score were calculated to assess the performance of satellite data in classifying HAB levels. These metrics revealed variability but generally had peaks at closer buffer distances, suggesting that increasing the buffer distance increased the noise of the signal.

The significant variation in correlation coefficients between the KDHE in-situ and CyAN satellite-derived data across different waterbodies highlights the challenges of accounting for diverse lake geometries, chemistries, and ecologies when predicting cyanobacteria cell concentrations. For example, at Marion Reservoir, a notable correlation coefficient ($R^2 = 0.581$) was observed at a 300m buffer distance. However, this value decreased significantly at a 1500m buffer ($R^2 = 0.0598$). In contrast, Milford Lake showed varying degrees of correlation across different zones; the strongest being an R^2 value of 0.214 within a 300m buffer in Zone A. Webster Lake, however, consistently displayed lower correlation coefficients, with R^2 values not exceeding 0.020. This suggests potential limitations of satellite data in certain environmental contexts, particularly in smaller lakes like Webster Lake. Moreover, regardless of the waterbody or lake zone, RMSE values exceeding 800,000 across all buffer distances and types of zonal statistics highlight a major limitation in using CyAN data to predict cell concentrations.

A significant observational discrepancy arises from the fact that KDHE often records readings of cyanobacterial blooms when CyAN does not. This phenomenon could be attributed to the accumulation of pond scum, which KDHE can detect through direct visual and physical sampling methods. In contrast, CyAN's satellite-derived data may not capture such localized and surface-level phenomena effectively. The CyAN technology relies on reflectance data that may not adequately represent the actual conditions when cyanobacteria are concentrated in clumps rather than uniformly distributed across the water surface.

The discrepancies observed could also be attributed to the inherent differences between satellite remote sensing and in-situ sampling methodologies. Each method comes with its own set of accuracy levels purely based on sampling characteristics, contributing to the variance observed. When taking samples at waterbodies with more than one sampling site KDHE will prioritize the sample with the greenest visual appearance and send it to a lab to be counted (KDHE HAB Response Plan, 2022). This method highlights the potential hazards but is likely not represented evenly across a 300m x 300m area which is the size of a CyAN pixel.

A notable factor in the data discrepancy is temporal variability. The consistently higher KDHE values could be indicative of differences in the timing of data collection, which introduces a temporal bias when comparing the datasets. This aspect is particularly crucial given the dynamic nature of cyanobacterial movement within the water column between the morning hours when KDHE samples and solar noon when CyAN samples. A temporal weighting factor increasing CyAN values was not added for this study but should be considered in future comparisons.

Spatial factors also significantly influence the data concordance. The observed lower R^2 values at increased distances highlight the challenge of noise in satellite data aggregation. This issue is compounded by adjacency effects from nearby landforms. The variation in results across different locations could also be partially attributed to environmental dynamics, such as water flow and circulation patterns, which play a pivotal role in the distribution of cyanobacteria blooms. Furthermore, the volumetric geometries of the waterbodies significantly impact bloom distribution, as evidenced by the widely varied metrics across locations. Local pollution sources as well as watershed management practices could also emerge as critical factors influencing the differences observed across the waterbodies.

The discrepancies between in situ and satellite data highlight the opportunity of adopting a multifaceted approach to HAB management. Currently, KDHE protocols don't integrate CyAN extensively but rather use it as supplementary data. Integration of the two data sources could enhance the monitoring and reporting for HAB occurrences in Kansas.

The findings from this study open several avenues for future research that could enhance our understanding and management of HABs. One potential area for further investigation involves the application of a "nearest value" analytical approach. By utilizing the nearest CyAN data points relative in-situ testing locations, this method could significantly increase the amount of usable data. Moreover, the acquisition of more comprehensive datasets over time presents an opportunity to conduct seasonal analyses. Splitting the data by season would allow the examination of the temporal variability of cyanobacterial concentrations and the effectiveness of satellite and in-situ methodologies in different climatic conditions. Such seasonal differentiation could be crucial in understanding the cyclic patterns of HABs and tailoring management strategies to be more effective during peak bloom periods. A second opportunity would be incorporating advanced machine learning algorithms to improve data classification and prediction accuracy. The use of sophisticated models could enhance our ability to distinguish between cyanobacteria and other reflective substances in satellite data, a challenge highlighted by the variable RMSE values observed across different buffer distances. A third opportunity is in the enhancements in satellite data processing, particularly concerning spatial and temporal resolution, which could also lead to more accurate and reliable monitoring of HABs. As technology advances, the integration of higher-resolution satellite images and more frequent sampling could provide unprecedented insights into the dynamics of cyanobacterial blooms, improving both the detection and management of these environmental hazards.

Conclusion

This study's comparative analysis of KDHE in-situ and CyAN satellite-derived data has provided valuable insights into the challenges and potentials of using different methodologies for monitoring harmful algal blooms (HABs). Despite the complexities and some inherent discrepancies in data correlation, the satellite data from CyAN, available freely, proves to be remarkably valuable. For a freely available resource, CyAN data demonstrates substantial utility in providing initial assessments of cyanobacterial presence, albeit with some limitations in precision.

The use of government-provided data like that from CyAN exemplifies the benefits of leveraging available resources to enhance public health and environmental management services. By utilizing such data, governmental agencies can improve their monitoring capabilities without significant investments in additional resources. It is evident that when government data is effectively integrated into service delivery, the enhancement in service quality is substantial, particularly in environmental monitoring and management.

Currently, CyAN data is most effective as a binary indicator for the presence or absence of cyanobacteria. This capability, while basic, is crucial for initial screening and quick decision-making processes in public health and safety contexts (Mishra et al, 2021). It allows for prompt responses and preventive measures, reducing potential risks associated with cyanobacterial blooms. However, when it comes to predicting bloom intensity, the results from this study suggest that CyAN data does not provide accurate predictions at the cell concentration level. Instead, better results are achieved when the CyAN values are classified into the KDHE classifications of watch, warning, and hazard.

Overall, the appropriateness of the CyAN product as a management tool for reporting bloom intensity at the water's edge is varied widely across waterbodies and even within lake zones and should be considered on a case by case basis. This variability highlights the challenge of using a single remote sensing product to uniformly monitor diverse aquatic environments.

Moving forward, the integration of more refined analytical techniques and the continuous improvement of data processing methods will be essential. Such advancements, like the release of Version 5 of the CyAN satellite product not only bolster the accuracy and reliability of satellite data but also enhance its applicability to more nuanced cyanobacteria phenomena. As this field evolves, the continued exploration of synergies between various data sources and methodologies will undoubtedly play a pivotal role in advancing our understanding and management of environmental threats like HABs.

References

"Cyanobacteria Assessment Network (CyAN)." EPA, Environmental Protection Agency, 12 July 2023, www.epa.gov/water-research/cyanobacteria-assessment-network-cyan.

Dawson, R.M. "The toxicology of Microcystins." Toxicon, vol. 36, no. 7, July 1998, pp. 953–962, https://doi.org/10.1016/s0041-0101(97)00102-5.

Dzialowski, Andrew R., et al. "Effects of non-algal turbidity on cyanobacterial biomass in seven turbid Kansas reservoirs." *Lake and Reservoir Management*, vol. 27, no. 1, 15 Mar. 2011, pp. 6–14, https://doi.org/10.1080/07438141.2011.551027.

Elliott, J. Alex. "The seasonal sensitivity of cyanobacteria and other phytoplankton to changes in flushing rate and water temperature." Global Change Biology, vol. 16, no. 2, 4 Jan. 2010, pp. 864–876, https://doi.org/10.1111/j.1365-2486.2009.01998.x.

Handler, Amalia M., et al. "Identifying lakes at risk of toxic cyanobacterial blooms using satellite imagery and field surveys across the United States." *Science of The Total Environment*, vol. 869, Apr. 2023, p. 161784, https://doi.org/10.1016/j.scitotenv.2023.161784.

"Health Effects Support Document for the Cyanobacterial Toxin Microcystins", United States Environmental Protection Agency, June 2015, www.epa.gov/sites/default/files/2017-06/documents/microcystins-support-report-2015.pdf.

"History of Marion Reservoir." U.S. Army Corps of Engineers, Tulsa District, U.S. Army Corps of Engineers, n.d. https://www.swt.usace.army.mil/Locations/Tulsa-District-Lakes/Kansas/Marion-Reservoir/History/. Accessed 21 Oct. 2023.

Hu, Yufei, et al. "A review of neurotoxicity of Microcystins." Environmental Science and Pollution Research, vol. 23, no. 8, 9 Feb. 2016, pp. 7211–7219, https://doi.org/10.1007/s11356-016-6073-y.

Jean Lindsey. "Webster, Kansas History of Webster." n.d. Solomon Valley Highway 24 Heritage Alliance, www.hwy24.org/uploads/2/6/1/8/26189167/webster.pdf. Accessed 12 May 2024.

Kansas Department of Health and Environment. (n.d.). TMDL Vision Process. Retrieved November 13, 2023, from https://www.kdhe.ks.gov/1304/TMDLVisionProcess

Kansas Department of Wildlife & Parks. "Webster." Webster / Locations / State Parks / KDWP - KDWP, ksoutdoors.com/State-Parks/Locations/Webster. Accessed 13 May 2024.

Kasinak, Jo-Marie E., et al. "Benchtop fluorometry of phycocyanin as a rapid approach for estimating cyanobacterial biovolume." Journal of Plankton Research, vol. 37, no. 1, 29 Oct. 2014, pp. 248–257, https://doi.org/10.1093/plankt/fbu096.

"KDHE Agency HAB Response Plan – 2022." KDHE, 2022, www.kdhe.ks.gov/DocumentCenter/View/6266/2022-HAB-Response-Plan-Complete-PDF?bidId=.

List of All Impaired & Potentially Impaired Waters, 2022, www.kdhe.ks.gov/DocumentCenter/View/22777/2022-303d-List-PDF?bidId=.

Lunetta, Ross S., et al. "Evaluation of cyanobacteria cell count detection derived from Meris imagery across the Eastern USA." Remote Sensing of Environment, vol. 157, Feb. 2015, pp. 24–34, https://doi.org/10.1016/j.rse.2014.06.008.

Marion Lake Fact Sheet. n.d. Kansas Department of Health and Environment, https://www.kdhe.ks.gov/DocumentCenter/View/14796/Marion-Lake-Fact-Sheet-PDF. Accessed 21 Oct. 2023.

"Milford Reservoir Watershed Restoration and Protection Strategy." Kansas Wraps, 21 Aug. 2019, kswraps.org/wp-content/uploads/2020/10/Milford-WRAPS-Plan-Update-2018-Final.pdf.

Mishra, Sachidananda, et al. "Evaluation of a satellite-based cyanobacteria bloom detection algorithm using field-measured microcystin data." Science of The Total Environment, vol. 774, June 2021, p. 145462, https://doi.org/10.1016/j.scitotenv.2021.145462.

NASA Applied Sciences. (n.d.). Projects: Cyan – Cyanobacteria Assessment Network. Retrieved November 13, 2023, from https://appliedsciences.nasa.gov/what-we-do/projects/cyan-cyanobacteria-assessment-network

"NASA Ocean Color." NASA, NASA, oceancolor.gsfc.nasa.gov/about/projects/cyan/. Accessed 13 Apr. 2024.

"NASA Ocean Color." NASA, NASA, oceancolor.gsfc.nasa.gov/data/reprocessing/projects/cyan/version/5/. Accessed 14 Apr. 2024.

Stumpf, R. P., Wynne, T. T., Baker, D. B., & Fahnenstiel, G. L. (2012). Interannual variability of cyanobacterial blooms in Lake Erie. PLoS ONE, 7(8). https://doi.org/10.1371/journal.pone.0042444

U.S. Army Corps of Engineers. (n.d.). Learn about the lake: Milford Lake. Retrieved November 13, 2023, from https://www.nwk.usace.army.mil/Locations/District-Lakes/Milford-Lake/Learn-About-the-Lake/

U.S. Army Corps of Engineers. (n.d.). [PDF file]. Retrieved November 13, 2023, from https://usace.contentdm.oclc.org/utils/getfile/collection/p266001coll1/id/9595

"Webster Reservoir, Kansas." Recreation.Gov, www.recreation.gov/camping/gateways/3127. Accessed 13 May 2024.

Whitman, Peter, et al. “A validation of satellite derived cyanobacteria detections with state reported events and recreation advisories across U.S. lakes.” Harmful Algae, vol. 115, June 2022, p. 102191, https://doi.org/10.1016/j.hal.2022.102191.

Wynne, Timothy T., et al. “Characterizing a cyanobacterial bloom in western Lake Erie using satellite imagery and Meteorological Data.” Limnology and Oceanography, vol. 55, no. 5, 18 Aug. 2010, pp. 2025–2036, https://doi.org/10.4319/lo.2010.55.5.2025.

Appendix A - KDHE Data

Waterbody Name	Number_of_Samples	First_Sample_Date	Most_Recent_Sample_Date
AGRA LAKE	2	7/6/2021	8/2/2021
ALTAMONT CITY LAKE NORTH	4	7/12/2021	9/27/2021
ALTAMONT CITY MAIN LAKE	4	7/12/2021	9/27/2021
AMESBURY LAKE	2	10/9/2023	10/23/2023
ANTHONY CITY LAKE	12	8/26/2010	6/25/2018
ANTIOCH PARK LAKES	12	8/6/2012	9/8/2014
ATCHISON CITY LAKE 9	1	10/2/2023	10/2/2023
ATCHISON CO. STATE FISHING LAKE	7	9/17/2018	8/21/2023
ATCHISON COUNTY PARK LAKE	21	9/19/2011	10/14/2019
ATWOOD TOWNSHIP LAKE	4	8/20/2018	7/26/2021
AUGUSTA CITY LAKE	9	8/8/2011	6/24/2019
AUGUSTA SANTA FE LAKE	25	8/6/2010	9/11/2023
BANNER CREEK LAKE	2	6/27/2012	6/27/2012
BELLEVILLE CITY LAKE	1	7/28/2014	7/28/2014
BIG ELEVEN LAKE	20	6/10/2019	10/30/2023
BIG HILL LAKE	23	7/12/2011	9/7/2021
BROWN CO STATE FISHING LAKE	25	7/9/2012	9/7/2021
BUHLER CITY LAKE	4	7/15/2013	9/13/2021
CAMP HYDE LAKE	2	7/7/2011	7/7/2011
CARBONDALE WEST LAKE	4	7/23/2018	10/22/2018
CEDAR BLUFF LAKE	4	6/11/2012	6/25/2012
CEDAR BROOK POND	2	6/30/2014	7/14/2014
CEDAR CREST LAKE	1	10/31/2011	10/31/2011
CEDAR LAKE	5	8/5/2014	6/26/2023
CENTRAL PARK LAKE	51	8/18/2010	10/21/2019
CENTRALIA LAKE	43	7/26/2010	10/8/2012
CHENEY LAKE	60	7/28/2011	7/31/2023
CHISHOLM CREEK PARK LAKE	32	7/16/2012	10/26/2015
CLARION WOODS PARK LAKE	4	6/25/2018	9/17/2019
CLINTON LAKE	13	8/22/2011	6/13/2022
COLBY CITY POND	2	8/22/2017	7/29/2019
COLWICH CITY LAKE	14	7/15/2013	9/26/2022
CONCANNON STATE FISHING LAKE	2	7/6/2015	7/20/2015
COUNCIL GROVE CITY LAKE	4	6/23/2014	9/17/2018
COUNCIL GROVE LAKE	4	6/26/2023	8/29/2023

COWLEY CO STATE FISHING LAKE	2	10/9/2017	10/30/2017
CRYSTAL LAKE	10	8/10/2015	10/31/2022
DEANNA ROSE FARMSTEAD FISHING POND	7	7/18/2012	9/15/2014
DILLON PARK LAKE #1	4	8/29/2011	10/10/2011
DORNER PARK LAKE	1	7/10/2023	7/10/2023
DOUGLAS CO STATE LAKE	1	9/6/2017	9/6/2017
EISENHOWER PARK POND IN MARQUETTE	2	10/18/2021	6/20/2022
EL DORADO LAKE	3	8/22/2022	8/22/2022
ELK CITY LAKE	1	9/6/2017	9/6/2017
ELKHORN LAKE	5	8/27/2019	10/17/2022
ELKS LAKE	1	8/19/2014	8/19/2014
ELLIS CITY LAKE	2	8/29/2023	9/11/2023
FALL RIVER LAKE	5	9/26/2011	6/18/2012
FELDHAUSEN, JOHNSON & MITCHELL POND	4	7/25/2011	8/5/2011
FLINT HILLS N.W.R.	1	7/19/2011	7/19/2011
FORD COUNTY LAKE	21	7/25/2016	10/31/2022
FORT SCOTT CITY LAKE	1	10/18/2021	10/18/2021
FOSSIL LAKE	6	8/21/2023	10/30/2023
FRAIZER LAKE	8	7/17/2018	10/30/2023
FRISCO LAKE	1	7/16/2012	7/16/2012
GAGE PARK LAKE	4	8/20/2012	9/30/2019
GARNETT LAKE	2	6/13/2022	7/11/2022
GATHERING POND	56	6/24/2019	10/31/2022
GEARY CO STATE FISHING LAKE	1	7/19/2021	7/19/2021
HAIN SFL	4	9/27/2021	9/12/2022
HARGIS LAKE	4	7/26/2021	8/16/2021
HARVEY CO. EAST LAKE	49	8/9/2011	4/1/2024
HARVEY CO. WEST PARK LAKE	26	8/29/2011	8/12/2013
HARVEY COUNTY CAMP HAWK LAKE	7	8/29/2011	10/7/2019
HERINGTON CITY LAKE	27	6/27/2011	6/6/2022
HERINGTON RESERVOIR	9	7/11/2012	6/6/2022
HIAWATHA CITY LAKE	29	8/27/2012	10/14/2019
HIDDEN WOODS PARK POND	1	7/17/2023	7/17/2023
HILLSDALE LAKE STATION 1	12	8/1/2011	6/7/2021
HODGEMAN CO. SFL/W.A.	22	9/11/2017	10/4/2021
HORSE THIEF LAKE	7	7/25/2011	10/23/2023
JEFFREY ENERGY CENTER NORTH LAKE	2	8/21/2023	9/5/2023

JERRY DISHMAN LAKE	1	8/3/2015	8/3/2015
JERRY IVEY POND	10	10/29/2018	4/3/2023
JEWELL CO. STATE FISHING LAKE	8	7/14/2014	6/29/2015
KANOPOLIS LAKE	7	6/25/2019	6/27/2023
KANSAS JAYCEES CEREBRAL PALSY LAKE	5	8/23/2010	9/7/2010
KFG HATCHERY AND PONDS	2	8/17/2015	8/24/2015
KILL CREEK PARK LAKE	2	8/20/2012	8/20/2012
KINGSTON LAKE	5	8/29/2022	8/21/2023
KIRWIN LAKE	14	6/12/2017	6/18/2018
LA CYGNE LAKE	6	7/11/2016	6/22/2022
LAKE AFTON	91	8/2/2011	10/30/2023
LAKE COLDWATER	2	8/15/2011	8/15/2011
LAKE JEANETTE	7	9/9/2019	10/25/2021
LAKE JIVARO	3	9/23/2011	4/13/2012
LAKE MEADE STATE PARK	53	7/27/2010	10/24/2016
LAKE OF THE FOREST	1	9/7/2011	9/7/2011
LAKE SCOTT STATE PARK	20	7/16/2012	9/12/2022
LAKE SHAWNEE	51	10/8/2013	10/16/2023
LAKE TANKO (CHERRYVALE CITY LAKE	1	7/25/2022	7/25/2022
LAKE WARNOCK	16	8/2/2011	10/2/2023
LAKEWOOD PARK LAKE	10	8/6/2018	10/28/2019
LEAVENWORTH CO. STATE FISHING LK	2	8/30/2016	7/31/2017
LEBO CITY PARK LAKE	6	6/25/2012	10/7/2019
LENEXA CITY LAKE	1	7/18/2022	7/18/2022
LINN VALLEY LAKE	7	8/20/2018	7/19/2021
LOGAN CITY LAKE	22	6/21/2011	6/30/2014
LOUISBURG OLD LAKE	6	7/18/2022	8/7/2023
LOVEWELL LAKE	121	7/12/2010	10/30/2023
MARAIS DES CYGNES W.A.	16	8/27/2018	9/9/2019
MARION CO. STATE FISHING LAKE	61	6/27/2011	6/5/2023
MARION RESERVOIR	265	6/27/2011	7/11/2023
MARYS LAKE	5	7/16/2018	10/15/2018
MARYSVILLE CITY/GOLF COURSE LAKE	1	7/26/2011	7/26/2011
MCKINLEY LAKE	5	8/16/2010	11/15/2010
MCLAUGHIN LAKE	3	6/13/2016	8/8/2016
MCPHERSON CO STATE FISHING LAKE	14	7/30/2012	10/19/2015
MEADOWBROOK PARK LAKE	3	8/20/2019	10/14/2019

MELVERN LAKE	55	8/20/2018	5/23/2023
MELVERN OUTLET POND	62	8/3/2015	9/18/2023
MELVERN OUTLET SWIM POND	33	8/2/2017	9/5/2023
MEMORIAL PARK LAKE	46	8/16/2010	10/30/2017
MILFORD LAKE ZONE A	229	7/18/2011	6/12/2023
MILFORD LAKE ZONE B	136	7/18/2011	6/12/2023
MILFORD LAKE ZONE C	229	7/18/2011	6/12/2023
MISSION LAKE	15	8/6/2012	8/2/2021
MIZE LAKE	1	8/20/2019	8/20/2019
NEMAHA CO. SFL/W.A.	6	8/3/2015	10/26/2015
NEOSHO CO STATE FISHING LAKE	7	8/19/2019	10/25/2021
NEW YATES CENTER LAKE	1	7/8/2019	7/8/2019
NORTH PARK LAKE	1	7/17/2023	7/17/2023
NORTH POINT LAKE	1	5/18/2018	5/18/2018
NORTON LAKE	44	7/7/2014	5/30/2023
OLPE JONES PARK LAKE	1	6/27/2012	6/27/2012
OVERBROOK CITY KIDS POND	3	10/1/2018	6/24/2019
OVERBROOK CITY LAKE	26	8/1/2016	10/17/2022
PARSONS LAKE	22	7/26/2021	9/11/2023
PERRY LAKE ZONE A	17	7/12/2011	8/27/2018
PERRY LAKE ZONE B	14	7/12/2011	8/27/2018
PERRY LAKE ZONE C	27	7/12/2011	7/6/2021
PERRY LAKE ZONE D	15	7/12/2011	7/6/2021
PLAINVILLE TOWNSHIP LAKE	5	7/13/2015	12/29/2015
POMONA LAKE	87	8/21/2018	6/19/2023
PONY CREEK LAKE	4	7/26/2021	10/4/2021
RIGGS PARK LAKE	16	7/25/2011	7/11/2022
RIVER PARK POND	14	8/26/2011	8/29/2023
RIVERWALK LANDING POND	3	8/13/2021	10/18/2021
ROCK GARDEN POND	6	7/19/2018	9/18/2023
ROOKS CO. STATE FISHING LAKE	5	7/9/2018	10/31/2022
ROSES LAKE	10	8/7/2013	7/25/2022
SABETHA CITY LAKE	5	6/6/2011	8/31/2015
SAMS POND SYRACUSE	8	6/8/2017	10/2/2017
SCHOOL STREET POND	3	7/17/2023	8/14/2023
SOUTH PARK LAKE	58	5/30/2012	10/23/2023
STONE LAKE	3	9/8/2014	10/27/2014
STROWBRIDGE RESERVOIR	15	6/19/2019	7/25/2023
SUMMERCREST LAKE	3	8/15/2022	8/7/2023
TALLWOOD DRIVE LAKE	1	2/20/2014	2/20/2014
TOMAHAWK PARK LAKES	2	9/3/2013	10/15/2018

TORONTO LAKE	4	5/21/2012	7/28/2014
TUTTLE CREEK LAKE	3	9/23/2013	9/23/2013
VETERANS LAKE ARK CITY	3	8/23/2017	10/2/2017
VINEYARD PARK POND	1	10/31/2016	10/31/2016
WABAUNSEE CO LAKE	22	7/17/2018	10/1/2018
WACONDA LAKE	3	7/13/2011	7/13/2011
WEBSTER LAKE	164	6/30/2015	7/10/2023
WELLINGTON LAKE	6	8/30/2011	8/14/2023
WELLINGTON NEW CITY LAKE	6	8/30/2011	8/14/2023
WEST CAMPUS LAKE KU	1	10/1/2018	10/1/2018
WILDERNESS LAKE	1	7/16/2019	7/16/2019
WINDMILL LAKE	1	9/29/2014	9/29/2014
WINFIELD CITY LAKE	9	5/21/2012	10/9/2023
WOLF POND	3	6/26/2017	7/24/2017
WOODLAND PARK POND MOUNT HOPE	2	6/19/2023	8/16/2023
WOODSON CO. STATE FISHING LAKE	1	6/4/2018	6/4/2018
WYANDOTTE CO STATE FISHING LAKE	7	5/17/2020	4/2/2024

Appendix B - Marion Confusion Matrices

BufferedOutputTable_1000m.csv - MIN_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MIN_LVL 3	0	0	0	0
MIN_LVL 2	0	6	1	6
MIN_LVL 1	0	4	0	4
MIN_LVL 0	0	2	0	12
BufferedOutputTable_1000m.csv - MAX_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MAX_LVL 3	0	0	0	0
MAX_LVL 2	0	10	1	2
MAX_LVL 1	0	5	0	3
MAX_LVL 0	0	3	1	10
BufferedOutputTable_1000m.csv - MEAN_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MEAN_LVL 3	0	0	0	0
MEAN_LVL 2	0	8	1	4
MEAN_LVL 1	0	5	0	3
MEAN_LVL 0	0	2	1	11
BufferedOutputTable_1500m.csv - MIN_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MIN_LVL 3	0	0	0	1
MIN_LVL 2	0	5	1	8
MIN_LVL 1	0	3	0	5
MIN_LVL 0	0	2	0	12
BufferedOutputTable_1500m.csv - MAX_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MAX_LVL 3	0	0	0	1
MAX_LVL 2	0	11	0	3
MAX_LVL 1	0	5	0	3
MAX_LVL 0	0	4	1	9
BufferedOutputTable_1500m.csv - MEAN_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MEAN_LVL 3	0	0	0	1
MEAN_LVL 2	0	8	2	4

MEAN_LVL 1	0	5	0	3
MEAN_LVL 0	0	2	1	11
BufferedOutputTable_300m.csv - MIN_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MIN_LVL 3	0	0	0	0
MIN_LVL 2	0	7	1	2
MIN_LVL 1	0	2	0	1
MIN_LVL 0	0	3	0	7
BufferedOutputTable_300m.csv - MAX_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MAX_LVL 3	0	0	0	0
MAX_LVL 2	0	8	0	2
MAX_LVL 1	0	2	0	1
MAX_LVL 0	0	3	0	7
BufferedOutputTable_300m.csv - MEAN_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MEAN_LVL 3	0	0	0	0
MEAN_LVL 2	0	8	0	2
MEAN_LVL 1	0	2	0	1
MEAN_LVL 0	0	3	0	7
BufferedOutputTable_400m.csv - MIN_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MIN_LVL 3	0	0	0	0
MIN_LVL 2	0	6	2	3
MIN_LVL 1	0	3	0	1
MIN_LVL 0	0	3	0	9
BufferedOutputTable_400m.csv - MAX_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MAX_LVL 3	0	0	0	0
MAX_LVL 2	0	8	0	3
MAX_LVL 1	0	3	0	1
MAX_LVL 0	0	3	0	9
BufferedOutputTable_400m.csv - MEAN_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MEAN_LVL 3	0	0	0	0
MEAN_LVL 2	0	7	1	3
MEAN_LVL 1	0	3	0	1

MEAN_LVL 0	0	3	0	9
BufferedOutputTable_500m.csv - MIN_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MIN_LVL 3	0	0	0	0
MIN_LVL 2	0	6	2	3
MIN_LVL 1	0	3	0	2
MIN_LVL 0	0	3	0	9
BufferedOutputTable_500m.csv - MAX_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MAX_LVL 3	0	0	0	0
MAX_LVL 2	0	8	0	3
MAX_LVL 1	0	3	0	2
MAX_LVL 0	0	3	0	9
BufferedOutputTable_500m.csv - MEAN_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MEAN_LVL 3	0	0	0	0
MEAN_LVL 2	0	7	1	3
MEAN_LVL 1	0	3	0	2
MEAN_LVL 0	0	3	0	9
BufferedOutputTable_600m.csv - MIN_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MIN_LVL 3	0	0	0	0
MIN_LVL 2	0	5	2	4
MIN_LVL 1	0	3	0	2
MIN_LVL 0	0	2	0	10
BufferedOutputTable_600m.csv - MAX_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MAX_LVL 3	0	0	0	0
MAX_LVL 2	0	8	0	3
MAX_LVL 1	0	3	0	2
MAX_LVL 0	0	3	0	9
BufferedOutputTable_600m.csv - MEAN_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MEAN_LVL 3	0	0	0	0
MEAN_LVL 2	0	7	1	3
MEAN_LVL 1	0	3	0	2
MEAN_LVL 0	0	3	0	9

BufferedOutputTable_700m.csv - MIN_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MIN_LVL 3	0	0	0	0
MIN_LVL 2	0	5	1	5
MIN_LVL 1	0	3	0	2
MIN_LVL 0	0	2	0	10
BufferedOutputTable_700m.csv - MAX_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MAX_LVL 3	0	0	0	0
MAX_LVL 2	0	8	1	2
MAX_LVL 1	0	4	0	1
MAX_LVL 0	0	3	1	8
BufferedOutputTable_700m.csv - MEAN_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MEAN_LVL 3	0	0	0	0
MEAN_LVL 2	0	7	1	3
MEAN_LVL 1	0	4	0	1
MEAN_LVL 0	0	3	0	9
BufferedOutputTable_800m.csv - MIN_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MIN_LVL 3	0	0	0	0
MIN_LVL 2	0	6	1	5
MIN_LVL 1	0	4	0	2
MIN_LVL 0	0	2	0	11
BufferedOutputTable_800m.csv - MAX_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MAX_LVL 3	0	0	0	0
MAX_LVL 2	0	9	1	2
MAX_LVL 1	0	5	0	1
MAX_LVL 0	0	3	1	9
BufferedOutputTable_800m.csv - MEAN_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MEAN_LVL 3	0	0	0	0
MEAN_LVL 2	0	8	1	3
MEAN_LVL 1	0	5	0	1
MEAN_LVL 0	0	2	1	10
BufferedOutputTable_900m.csv - MIN_LVL				

	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MIN_LVL 3	0	0	0	0
MIN_LVL 2	0	6	1	5
MIN_LVL 1	0	4	0	2
MIN_LVL 0	0	2	0	11
BufferedOutputTable_900m.csv - MAX_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MAX_LVL 3	0	0	0	0
MAX_LVL 2	0	9	1	2
MAX_LVL 1	0	5	0	1
MAX_LVL 0	0	3	1	9
BufferedOutputTable_900m.csv - MEAN_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MEAN_LVL 3	0	0	0	0
MEAN_LVL 2	0	8	1	3
MEAN_LVL 1	0	5	0	1
MEAN_LVL 0	0	2	1	10
OutputTable.csv - MIN_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MIN_LVL 3	0	0	0	0
MIN_LVL 2	0	6	0	2
MIN_LVL 1	0	0	0	1
MIN_LVL 0	0	2	0	3
OutputTable.csv - MAX_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MAX_LVL 3	0	0	0	0
MAX_LVL 2	0	6	0	2
MAX_LVL 1	0	0	0	1
MAX_LVL 0	0	2	0	3
OutputTable.csv - MEAN_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MEAN_LVL 3	0	0	0	0
MEAN_LVL 2	0	6	0	2
MEAN_LVL 1	0	0	0	1
MEAN_LVL 0	0	2	0	3

Appendix C - Milford Lake Zone A

Confusion Matrices

BufferedOutputTable_1000m.csv - MIN_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MIN_LVL 3	0	0	0	0
MIN_LVL 2	0	1	2	8
MIN_LVL 1	0	0	1	9
MIN_LVL 0	0	3	2	23
BufferedOutputTable_1000m.csv - MAX_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MAX_LVL 3	0	0	0	0
MAX_LVL 2	0	4	3	4
MAX_LVL 1	0	2	1	7
MAX_LVL 0	0	7	3	18
BufferedOutputTable_1000m.csv - MEAN_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MEAN_LVL 3	0	0	0	0
MEAN_LVL 2	0	3	1	7
MEAN_LVL 1	0	1	1	8
MEAN_LVL 0	0	4	4	20
BufferedOutputTable_1500m.csv - MIN_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MIN_LVL 3	0	0	0	0
MIN_LVL 2	0	1	2	8
MIN_LVL 1	0	0	1	9
MIN_LVL 0	0	3	2	23
BufferedOutputTable_1500m.csv - MAX_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MAX_LVL 3	0	0	0	0
MAX_LVL 2	0	4	3	4
MAX_LVL 1	0	2	2	6
MAX_LVL 0	0	9	1	18
BufferedOutputTable_1500m.csv - MEAN_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0

MEAN_LVL 3	0	0	0	0
MEAN_LVL 2	0	3	1	7
MEAN_LVL 1	0	1	1	8
MEAN_LVL 0	0	4	5	19
BufferedOutputTable_300m.csv - MIN_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MIN_LVL 3	0	0	0	0
MIN_LVL 2	0	1	1	3
MIN_LVL 1	0	0	0	5
MIN_LVL 0	0	4	3	10
BufferedOutputTable_300m.csv - MAX_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MAX_LVL 3	0	0	0	0
MAX_LVL 2	0	1	1	3
MAX_LVL 1	0	0	1	4
MAX_LVL 0	0	4	3	10
BufferedOutputTable_300m.csv - MEAN_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MEAN_LVL 3	0	0	0	0
MEAN_LVL 2	0	1	1	3
MEAN_LVL 1	0	0	1	4
MEAN_LVL 0	0	4	3	10
BufferedOutputTable_400m.csv - MIN_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MIN_LVL 3	0	0	0	0
MIN_LVL 2	0	2	3	6
MIN_LVL 1	0	0	0	8
MIN_LVL 0	0	4	4	17
BufferedOutputTable_400m.csv - MAX_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MAX_LVL 3	0	0	0	0
MAX_LVL 2	0	2	3	6
MAX_LVL 1	0	0	1	7
MAX_LVL 0	0	4	4	17
BufferedOutputTable_400m.csv - MEAN_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MEAN_LVL 3	0	0	0	0

MEAN_LVL 2	0	2	3	6
MEAN_LVL 1	0	0	1	7
MEAN_LVL 0	0	4	4	17
BufferedOutputTable_500m.csv - MIN_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MIN_LVL 3	0	0	0	0
MIN_LVL 2	0	2	1	8
MIN_LVL 1	0	0	1	9
MIN_LVL 0	0	3	3	21
BufferedOutputTable_500m.csv - MAX_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MAX_LVL 3	0	0	0	0
MAX_LVL 2	0	4	1	6
MAX_LVL 1	0	0	2	8
MAX_LVL 0	0	5	4	18
BufferedOutputTable_500m.csv - MEAN_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MEAN_LVL 3	0	0	0	0
MEAN_LVL 2	0	3	1	7
MEAN_LVL 1	0	0	2	8
MEAN_LVL 0	0	3	6	18
BufferedOutputTable_600m.csv - MIN_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MIN_LVL 3	0	0	0	0
MIN_LVL 2	0	2	1	8
MIN_LVL 1	0	0	1	9
MIN_LVL 0	0	3	3	22
BufferedOutputTable_600m.csv - MAX_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MAX_LVL 3	0	0	0	0
MAX_LVL 2	0	4	1	6
MAX_LVL 1	0	1	1	8
MAX_LVL 0	0	5	4	19
BufferedOutputTable_600m.csv - MEAN_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MEAN_LVL 3	0	0	0	0
MEAN_LVL 2	0	3	1	7

MEAN_LVL 1	0	1	1	8
MEAN_LVL 0	0	4	4	20
BufferedOutputTable_700m.csv - MIN_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MIN_LVL 3	0	0	0	0
MIN_LVL 2	0	2	1	8
MIN_LVL 1	0	0	1	9
MIN_LVL 0	0	3	3	22
BufferedOutputTable_700m.csv - MAX_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MAX_LVL 3	0	0	0	0
MAX_LVL 2	0	4	2	5
MAX_LVL 1	0	2	0	8
MAX_LVL 0	0	6	3	19
BufferedOutputTable_700m.csv - MEAN_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MEAN_LVL 3	0	0	0	0
MEAN_LVL 2	0	3	1	7
MEAN_LVL 1	0	1	1	8
MEAN_LVL 0	0	4	4	20
BufferedOutputTable_800m.csv - MIN_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MIN_LVL 3	0	0	0	0
MIN_LVL 2	0	2	1	8
MIN_LVL 1	0	0	1	9
MIN_LVL 0	0	3	3	22
BufferedOutputTable_800m.csv - MAX_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MAX_LVL 3	0	0	0	0
MAX_LVL 2	0	4	2	5
MAX_LVL 1	0	2	0	8
MAX_LVL 0	0	7	2	19
BufferedOutputTable_800m.csv - MEAN_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MEAN_LVL 3	0	0	0	0
MEAN_LVL 2	0	3	1	7
MEAN_LVL 1	0	1	1	8

MEAN_LVL 0	0	4	4	20
BufferedOutputTable_900m.csv - MIN_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MIN_LVL 3	0	0	0	0
MIN_LVL 2	0	2	1	8
MIN_LVL 1	0	0	1	9
MIN_LVL 0	0	3	3	22
BufferedOutputTable_900m.csv - MAX_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MAX_LVL 3	0	0	0	0
MAX_LVL 2	0	4	3	4
MAX_LVL 1	0	2	0	8
MAX_LVL 0	0	7	3	18
BufferedOutputTable_900m.csv - MEAN_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MEAN_LVL 3	0	0	0	0
MEAN_LVL 2	0	3	1	7
MEAN_LVL 1	0	1	1	8
MEAN_LVL 0	0	4	4	20
OutputTable.csv - MIN_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MIN_LVL 3	0	0	0	0
MIN_LVL 2	0	0	0	1
MIN_LVL 1	0	0	0	5
MIN_LVL 0	0	1	2	7
OutputTable.csv - MAX_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MAX_LVL 3	0	0	0	0
MAX_LVL 2	0	0	0	1
MAX_LVL 1	0	0	0	5
MAX_LVL 0	0	1	2	7
OutputTable.csv - MEAN_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MEAN_LVL 3	0	0	0	0
MEAN_LVL 2	0	0	0	1
MEAN_LVL 1	0	0	0	5
MEAN_LVL 0	0	1	2	7

Appendix D - Milford Lake Zone B

Confusion Matrices

BufferedOutputTable_1000m.csv - MIN_LVL				

	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MIN_LVL 3	0	0	0	0
MIN_LVL 2	0	9	1	1
MIN_LVL 1	0	8	3	5
MIN_LVL 0	0	2	0	8
BufferedOutputTable_1000m.csv - MAX_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MAX_LVL 3	0	0	0	0
MAX_LVL 2	0	10	0	1
MAX_LVL 1	0	14	1	1
MAX_LVL 0	0	5	1	4
BufferedOutputTable_1000m.csv - MEAN_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MEAN_LVL 3	0	0	0	0
MEAN_LVL 2	0	10	0	1
MEAN_LVL 1	0	11	3	2
MEAN_LVL 0	0	4	1	5
BufferedOutputTable_1500m.csv - MIN_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MIN_LVL 3	0	0	0	0
MIN_LVL 2	0	8	1	2
MIN_LVL 1	0	8	3	6
MIN_LVL 0	0	1	0	10
BufferedOutputTable_1500m.csv - MAX_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MAX_LVL 3	0	0	0	0
MAX_LVL 2	0	10	0	1
MAX_LVL 1	0	16	0	1
MAX_LVL 0	0	6	0	5
BufferedOutputTable_1500m.csv - MEAN_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MEAN_LVL 3	0	0	0	0
MEAN_LVL 2	0	10	0	1
MEAN_LVL 1	0	12	4	1
MEAN_LVL 0	0	4	1	6
BufferedOutputTable_300m.csv - MIN_LVL				

	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MIN_LVL 3	0	0	0	0
MIN_LVL 2	0	10	0	0
MIN_LVL 1	0	10	3	2
MIN_LVL 0	0	2	2	5
BufferedOutputTable_300m.csv - MAX_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MAX_LVL 3	0	0	0	0
MAX_LVL 2	0	10	0	0
MAX_LVL 1	0	10	3	2
MAX_LVL 0	0	2	2	5
BufferedOutputTable_300m.csv - MEAN_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MEAN_LVL 3	0	0	0	0
MEAN_LVL 2	0	10	0	0
MEAN_LVL 1	0	10	3	2
MEAN_LVL 0	0	2	2	5
BufferedOutputTable_400m.csv - MIN_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MIN_LVL 3	0	0	0	0
MIN_LVL 2	0	10	0	0
MIN_LVL 1	0	10	3	2
MIN_LVL 0	0	2	1	6
BufferedOutputTable_400m.csv - MAX_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MAX_LVL 3	0	0	0	0
MAX_LVL 2	0	10	0	0
MAX_LVL 1	0	11	2	2
MAX_LVL 0	0	3	2	4
BufferedOutputTable_400m.csv - MEAN_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MEAN_LVL 3	0	0	0	0
MEAN_LVL 2	0	10	0	0
MEAN_LVL 1	0	10	3	2
MEAN_LVL 0	0	3	1	5
BufferedOutputTable_500m.csv - MIN_LVL				

	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MIN_LVL 3	0	0	0	0
MIN_LVL 2	0	10	0	0
MIN_LVL 1	0	10	3	2
MIN_LVL 0	0	2	1	6
BufferedOutputTable_500m.csv - MAX_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MAX_LVL 3	0	0	0	0
MAX_LVL 2	0	10	0	0
MAX_LVL 1	0	11	2	2
MAX_LVL 0	0	3	2	4
BufferedOutputTable_500m.csv - MEAN_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MEAN_LVL 3	0	0	0	0
MEAN_LVL 2	0	10	0	0
MEAN_LVL 1	0	10	3	2
MEAN_LVL 0	0	3	1	5
BufferedOutputTable_600m.csv - MIN_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MIN_LVL 3	0	0	0	0
MIN_LVL 2	0	9	1	0
MIN_LVL 1	0	10	1	4
MIN_LVL 0	0	2	1	7
BufferedOutputTable_600m.csv - MAX_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MAX_LVL 3	0	0	0	0
MAX_LVL 2	0	10	0	0
MAX_LVL 1	0	13	1	1
MAX_LVL 0	0	4	1	5
BufferedOutputTable_600m.csv - MEAN_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MEAN_LVL 3	0	0	0	0
MEAN_LVL 2	0	10	0	0
MEAN_LVL 1	0	10	3	2
MEAN_LVL 0	0	3	1	6
BufferedOutputTable_700m.csv - MIN_LVL				

	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MIN_LVL 3	0	0	0	0
MIN_LVL 2	0	9	1	0
MIN_LVL 1	0	10	2	4
MIN_LVL 0	0	2	0	8
BufferedOutputTable_700m.csv - MAX_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MAX_LVL 3	0	0	0	0
MAX_LVL 2	0	10	0	0
MAX_LVL 1	0	14	1	1
MAX_LVL 0	0	5	1	4
BufferedOutputTable_700m.csv - MEAN_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MEAN_LVL 3	0	0	0	0
MEAN_LVL 2	0	10	0	0
MEAN_LVL 1	0	11	3	2
MEAN_LVL 0	0	4	1	5
BufferedOutputTable_800m.csv - MIN_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MIN_LVL 3	0	0	0	0
MIN_LVL 2	0	9	1	0
MIN_LVL 1	0	9	3	4
MIN_LVL 0	0	2	0	8
BufferedOutputTable_800m.csv - MAX_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MAX_LVL 3	0	0	0	0
MAX_LVL 2	0	10	0	0
MAX_LVL 1	0	14	1	1
MAX_LVL 0	0	5	1	4
BufferedOutputTable_800m.csv - MEAN_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MEAN_LVL 3	0	0	0	0
MEAN_LVL 2	0	10	0	0
MEAN_LVL 1	0	11	3	2
MEAN_LVL 0	0	4	1	5
BufferedOutputTable_900m.csv - MIN_LVL				

	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MIN_LVL 3	0	0	0	0
MIN_LVL 2	0	9	1	0
MIN_LVL 1	0	9	3	4
MIN_LVL 0	0	2	0	8
BufferedOutputTable_900m.csv - MAX_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MAX_LVL 3	0	0	0	0
MAX_LVL 2	0	10	0	0
MAX_LVL 1	0	14	1	1
MAX_LVL 0	0	5	1	4
BufferedOutputTable_900m.csv - MEAN_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MEAN_LVL 3	0	0	0	0
MEAN_LVL 2	0	10	0	0
MEAN_LVL 1	0	11	3	2
MEAN_LVL 0	0	4	1	5
OutputTable.csv - MIN_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MIN_LVL 3	0	0	0	0
MIN_LVL 2	0	3	0	0
MIN_LVL 1	0	3	1	1
MIN_LVL 0	0	0	1	3
OutputTable.csv - MAX_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MAX_LVL 3	0	0	0	0
MAX_LVL 2	0	3	0	0
MAX_LVL 1	0	3	1	1
MAX_LVL 0	0	0	1	3
OutputTable.csv - MEAN_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MEAN_LVL 3	0	0	0	0
MEAN_LVL 2	0	3	0	0
MEAN_LVL 1	0	3	1	1
MEAN_LVL 0	0	0	1	3

Appendix E - Milford Lake Zone Confusion Matrices

BufferedOutputTable_1000m.csv - MIN_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MIN_LVL 3	0	1	0	0
MIN_LVL 2	0	22	1	6
MIN_LVL 1	0	6	3	3
MIN_LVL 0	0	2	0	12
BufferedOutputTable_1000m.csv - MAX_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MAX_LVL 3	0	1	0	0
MAX_LVL 2	0	28	1	0
MAX_LVL 1	0	9	0	3
MAX_LVL 0	0	9	0	5
BufferedOutputTable_1000m.csv - MEAN_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MEAN_LVL 3	0	1	0	0
MEAN_LVL 2	0	28	1	0
MEAN_LVL 1	0	9	0	3
MEAN_LVL 0	0	5	4	5
BufferedOutputTable_1500m.csv - MIN_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MIN_LVL 3	0	1	0	0
MIN_LVL 2	0	19	2	8
MIN_LVL 1	0	2	3	7
MIN_LVL 0	0	1	0	14
BufferedOutputTable_1500m.csv - MAX_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MAX_LVL 3	0	1	0	0
MAX_LVL 2	0	28	1	0
MAX_LVL 1	0	10	0	2
MAX_LVL 0	0	9	1	5
BufferedOutputTable_1500m.csv - MEAN_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MEAN_LVL 3	0	1	0	0
MEAN_LVL 2	0	27	2	0
MEAN_LVL 1	0	9	0	3
MEAN_LVL 0	0	6	3	6
BufferedOutputTable_300m.csv - MIN_LVL				

	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MIN_LVL 3	0	0	0	0
MIN_LVL 2	0	22	1	1
MIN_LVL 1	0	7	1	2
MIN_LVL 0	0	3	0	9
BufferedOutputTable_300m.csv - MAX_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MAX_LVL 3	0	0	0	0
MAX_LVL 2	0	22	1	1
MAX_LVL 1	0	7	1	2
MAX_LVL 0	0	3	0	9
BufferedOutputTable_300m.csv - MEAN_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MEAN_LVL 3	0	0	0	0
MEAN_LVL 2	0	22	1	1
MEAN_LVL 1	0	7	1	2
MEAN_LVL 0	0	3	0	9
BufferedOutputTable_400m.csv - MIN_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MIN_LVL 3	0	0	0	0
MIN_LVL 2	0	21	1	3
MIN_LVL 1	0	7	1	2
MIN_LVL 0	0	2	0	10
BufferedOutputTable_400m.csv - MAX_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MAX_LVL 3	0	0	0	0
MAX_LVL 2	0	22	1	2
MAX_LVL 1	0	7	1	2
MAX_LVL 0	0	3	0	9
BufferedOutputTable_400m.csv - MEAN_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MEAN_LVL 3	0	0	0	0
MEAN_LVL 2	0	21	2	2
MEAN_LVL 1	0	7	1	2
MEAN_LVL 0	0	3	0	9
BufferedOutputTable_500m.csv - MIN_LVL				

	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MIN_LVL 3	0	0	0	0
MIN_LVL 2	0	22	1	3
MIN_LVL 1	0	7	1	2
MIN_LVL 0	0	2	0	10
BufferedOutputTable_500m.csv - MAX_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MAX_LVL 3	0	0	0	0
MAX_LVL 2	0	25	1	0
MAX_LVL 1	0	8	0	2
MAX_LVL 0	0	4	1	7
BufferedOutputTable_500m.csv - MEAN_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MEAN_LVL 3	0	0	0	0
MEAN_LVL 2	0	25	0	1
MEAN_LVL 1	0	7	1	2
MEAN_LVL 0	0	4	0	8
BufferedOutputTable_600m.csv - MIN_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MIN_LVL 3	0	1	0	0
MIN_LVL 2	0	25	1	3
MIN_LVL 1	0	8	1	2
MIN_LVL 0	0	2	0	10
BufferedOutputTable_600m.csv - MAX_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MAX_LVL 3	0	1	0	0
MAX_LVL 2	0	28	1	0
MAX_LVL 1	0	9	0	2
MAX_LVL 0	0	4	1	7
BufferedOutputTable_600m.csv - MEAN_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MEAN_LVL 3	0	1	0	0
MEAN_LVL 2	0	28	0	1
MEAN_LVL 1	0	8	1	2
MEAN_LVL 0	0	4	0	8
BufferedOutputTable_700m.csv - MIN_LVL				

	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MIN_LVL 3	0	1	0	0
MIN_LVL 2	0	25	1	3
MIN_LVL 1	0	8	1	2
MIN_LVL 0	0	3	0	10
BufferedOutputTable_700m.csv - MAX_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MAX_LVL 3	0	1	0	0
MAX_LVL 2	0	28	1	0
MAX_LVL 1	0	9	0	2
MAX_LVL 0	0	6	1	6
BufferedOutputTable_700m.csv - MEAN_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MEAN_LVL 3	0	1	0	0
MEAN_LVL 2	0	27	1	1
MEAN_LVL 1	0	8	1	2
MEAN_LVL 0	0	5	0	8
BufferedOutputTable_800m.csv - MIN_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MIN_LVL 3	0	1	0	0
MIN_LVL 2	0	23	1	5
MIN_LVL 1	0	7	2	2
MIN_LVL 0	0	2	1	11
BufferedOutputTable_800m.csv - MAX_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MAX_LVL 3	0	1	0	0
MAX_LVL 2	0	28	1	0
MAX_LVL 1	0	9	0	2
MAX_LVL 0	0	8	1	5
BufferedOutputTable_800m.csv - MEAN_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MEAN_LVL 3	0	1	0	0
MEAN_LVL 2	0	28	0	1
MEAN_LVL 1	0	9	0	2
MEAN_LVL 0	0	5	2	7
BufferedOutputTable_900m.csv - MIN_LVL				

	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MIN_LVL 3	0	1	0	0
MIN_LVL 2	0	22	1	6
MIN_LVL 1	0	6	3	3
MIN_LVL 0	0	2	0	12
BufferedOutputTable_900m.csv - MAX_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MAX_LVL 3	0	1	0	0
MAX_LVL 2	0	28	1	0
MAX_LVL 1	0	9	0	3
MAX_LVL 0	0	8	1	5
BufferedOutputTable_900m.csv - MEAN_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MEAN_LVL 3	0	1	0	0
MEAN_LVL 2	0	28	0	1
MEAN_LVL 1	0	9	0	3
MEAN_LVL 0	0	5	2	7
OutputTable.csv - MIN_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MIN_LVL 3	0	0	0	0
MIN_LVL 2	0	7	0	0
MIN_LVL 1	0	5	0	1
MIN_LVL 0	0	2	0	5
OutputTable.csv - MAX_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MAX_LVL 3	0	0	0	0
MAX_LVL 2	0	7	0	0
MAX_LVL 1	0	5	0	1
MAX_LVL 0	0	2	0	5
OutputTable.csv - MEAN_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MEAN_LVL 3	0	0	0	0
MEAN_LVL 2	0	7	0	0
MEAN_LVL 1	0	5	0	1
MEAN_LVL 0	0	2	0	5

Appendix F - Webster Confusion Matrices

BufferedOutputTable_1000m.csv - MIN_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MIN_LVL 3	0	0	0	0
MIN_LVL 2	0	6	1	6
MIN_LVL 1	0	4	0	4
MIN_LVL 0	0	2	0	12
BufferedOutputTable_1000m.csv - MAX_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MAX_LVL 3	0	0	0	0
MAX_LVL 2	0	10	1	2
MAX_LVL 1	0	5	0	3
MAX_LVL 0	0	3	1	10
BufferedOutputTable_1000m.csv - MEAN_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MEAN_LVL 3	0	0	0	0
MEAN_LVL 2	0	8	1	4
MEAN_LVL 1	0	5	0	3
MEAN_LVL 0	0	2	1	11
BufferedOutputTable_1500m.csv - MIN_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MIN_LVL 3	0	0	0	1
MIN_LVL 2	0	5	1	8
MIN_LVL 1	0	3	0	5
MIN_LVL 0	0	2	0	12
BufferedOutputTable_1500m.csv - MAX_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MAX_LVL 3	0	0	0	1
MAX_LVL 2	0	11	0	3
MAX_LVL 1	0	5	0	3
MAX_LVL 0	0	4	1	9
BufferedOutputTable_1500m.csv - MEAN_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MEAN_LVL 3	0	0	0	1
MEAN_LVL 2	0	8	2	4
MEAN_LVL 1	0	5	0	3
MEAN_LVL 0	0	2	1	11
BufferedOutputTable_300m.csv - MIN_LVL				

	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MIN_LVL 3	0	0	0	0
MIN_LVL 2	0	7	1	2
MIN_LVL 1	0	2	0	1
MIN_LVL 0	0	3	0	7
BufferedOutputTable_300m.csv - MAX_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MAX_LVL 3	0	0	0	0
MAX_LVL 2	0	8	0	2
MAX_LVL 1	0	2	0	1
MAX_LVL 0	0	3	0	7
BufferedOutputTable_300m.csv - MEAN_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MEAN_LVL 3	0	0	0	0
MEAN_LVL 2	0	8	0	2
MEAN_LVL 1	0	2	0	1
MEAN_LVL 0	0	3	0	7
BufferedOutputTable_400m.csv - MIN_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MIN_LVL 3	0	0	0	0
MIN_LVL 2	0	6	2	3
MIN_LVL 1	0	3	0	1
MIN_LVL 0	0	3	0	9
BufferedOutputTable_400m.csv - MAX_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MAX_LVL 3	0	0	0	0
MAX_LVL 2	0	8	0	3
MAX_LVL 1	0	3	0	1
MAX_LVL 0	0	3	0	9
BufferedOutputTable_400m.csv - MEAN_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MEAN_LVL 3	0	0	0	0
MEAN_LVL 2	0	7	1	3
MEAN_LVL 1	0	3	0	1
MEAN_LVL 0	0	3	0	9
BufferedOutputTable_500m.csv - MIN_LVL				

	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MIN_LVL 3	0	0	0	0
MIN_LVL 2	0	6	2	3
MIN_LVL 1	0	3	0	2
MIN_LVL 0	0	3	0	9
BufferedOutputTable_500m.csv - MAX_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MAX_LVL 3	0	0	0	0
MAX_LVL 2	0	8	0	3
MAX_LVL 1	0	3	0	2
MAX_LVL 0	0	3	0	9
BufferedOutputTable_500m.csv - MEAN_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MEAN_LVL 3	0	0	0	0
MEAN_LVL 2	0	7	1	3
MEAN_LVL 1	0	3	0	2
MEAN_LVL 0	0	3	0	9
BufferedOutputTable_600m.csv - MIN_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MIN_LVL 3	0	0	0	0
MIN_LVL 2	0	5	2	4
MIN_LVL 1	0	3	0	2
MIN_LVL 0	0	2	0	10
BufferedOutputTable_600m.csv - MAX_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MAX_LVL 3	0	0	0	0
MAX_LVL 2	0	8	0	3
MAX_LVL 1	0	3	0	2
MAX_LVL 0	0	3	0	9
BufferedOutputTable_600m.csv - MEAN_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MEAN_LVL 3	0	0	0	0
MEAN_LVL 2	0	7	1	3
MEAN_LVL 1	0	3	0	2
MEAN_LVL 0	0	3	0	9
BufferedOutputTable_700m.csv - MIN_LVL				

	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MIN_LVL 3	0	0	0	0
MIN_LVL 2	0	5	1	5
MIN_LVL 1	0	3	0	2
MIN_LVL 0	0	2	0	10
BufferedOutputTable_700m.csv - MAX_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MAX_LVL 3	0	0	0	0
MAX_LVL 2	0	8	1	2
MAX_LVL 1	0	4	0	1
MAX_LVL 0	0	3	1	8
BufferedOutputTable_700m.csv - MEAN_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MEAN_LVL 3	0	0	0	0
MEAN_LVL 2	0	7	1	3
MEAN_LVL 1	0	4	0	1
MEAN_LVL 0	0	3	0	9
BufferedOutputTable_800m.csv - MIN_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MIN_LVL 3	0	0	0	0
MIN_LVL 2	0	6	1	5
MIN_LVL 1	0	4	0	2
MIN_LVL 0	0	2	0	11
BufferedOutputTable_800m.csv - MAX_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MAX_LVL 3	0	0	0	0
MAX_LVL 2	0	9	1	2
MAX_LVL 1	0	5	0	1
MAX_LVL 0	0	3	1	9
BufferedOutputTable_800m.csv - MEAN_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MEAN_LVL 3	0	0	0	0
MEAN_LVL 2	0	8	1	3
MEAN_LVL 1	0	5	0	1
MEAN_LVL 0	0	2	1	10
BufferedOutputTable_900m.csv - MIN_LVL				

	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MIN_LVL 3	0	0	0	0
MIN_LVL 2	0	6	1	5
MIN_LVL 1	0	4	0	2
MIN_LVL 0	0	2	0	11
BufferedOutputTable_900m.csv - MAX_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MAX_LVL 3	0	0	0	0
MAX_LVL 2	0	9	1	2
MAX_LVL 1	0	5	0	1
MAX_LVL 0	0	3	1	9
BufferedOutputTable_900m.csv - MEAN_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MEAN_LVL 3	0	0	0	0
MEAN_LVL 2	0	8	1	3
MEAN_LVL 1	0	5	0	1
MEAN_LVL 0	0	2	1	10
OutputTable.csv - MIN_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MIN_LVL 3	0	0	0	0
MIN_LVL 2	0	6	0	2
MIN_LVL 1	0	0	0	1
MIN_LVL 0	0	2	0	3
OutputTable.csv - MAX_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MAX_LVL 3	0	0	0	0
MAX_LVL 2	0	6	0	2
MAX_LVL 1	0	0	0	1
MAX_LVL 0	0	2	0	3
OutputTable.csv - MEAN_LVL				
	KDHE_CC_LVL 3	KDHE_CC_LVL 2	KDHE_CC_LVL 1	KDHE_CC_LVL 0
MEAN_LVL 3	0	0	0	0
MEAN_LVL 2	0	6	0	2
MEAN_LVL 1	0	0	0	1
MEAN_LVL 0	0	2	0	3

Glossary of Terms

- **Cyanobacteria**: Often referred to as blue-green algae, these are photosynthetic bacteria that can form harmful algal blooms in nutrient-rich waters.
- **CyAN (Cyanobacteria Assessment Network)**: A satellite remote sensing project involving multiple U.S. agencies, aimed at monitoring harmful algal blooms in freshwater systems across the United States.
- **HABs (Harmful Algal Blooms)**: Overgrowths of algae in water that can produce harmful toxins affecting the health of the environment, animals, and humans.
- **KDHE (Kansas Department of Health and Environment)**: The state agency responsible for monitoring environmental and public health throughout Kansas.
- **In-situ**: Refers to measurements or observations made at the site where phenomena are occurring without disturbing the system.
- **Raster:** Refers to a grid of pixels used in digital imaging and geographic data systems. Each pixel, or cell, in a raster image contains a value that represents information like temperature, elevation, or cell concentration.
- **Remote Sensing**: The process of detecting and monitoring the physical characteristics of an area by measuring its reflected and emitted radiation from a distance (typically from satellite or aircraft).
- **Cell Concentration**: The number of cells per unit volume in a sample, used here to quantify cyanobacteria in water samples.
- **Eutrophic**: Waters rich in nutrients that can support a dense growth of algae and other aquatic plants, often resulting in oxygen depletion.

- **Python**: A high-level programming language used here for data analysis and comparison between satellite data and ground measurements.
- **Raster Data**: A type of digital image captured by satellite imaging systems, represented in rows and columns of pixels.
- **Spatial Resolution**: The dimension of the cell size representing the area covered on the ground, expressed in distance units per pixel.
- **Zonal Statistics**: Statistical analysis of raster data values within a specific zone defined by a dataset.
- **Buffer Distance**: The radius around a point location used for sampling spatial data, relevant in the context of comparing satellite data with in-situ sampling points.