remote sensing

MDPI

Article

Implementation of a Satellite Based Inland Water Algal Bloom Alerting System Using Analysis Ready Data

Tim J. Malthus [1,*], Eric Lehmann [2], Xavier Ho [3], Elizabeth Botha [4] and Janet Anstee [4]

1 Coasts Program, CSIRO Oceans and Atmosphere, Brisbane, QLD 4102, Australia
2 CSIRO Data61, Canberra, ACT 2601, Australia; eric.lehmann@csiro.au
3 CSIRO IMT, now CSIRO Data61, Melbourne, VIC 3008, Australia; xavier.ho@csiro.au
4 Coasts Program, CSIRO Oceans and Atmosphere, Canberra, ACT 2601, Australia; elizabeth.botha@csiro.au (E.B.); janet.anstee@csiro.au (J.A.)
* Correspondence: tim.malthus@csiro.au; Tel.: +61-7-3833-5583

Received: 20 November 2019; Accepted: 9 December 2019; Published: 10 December 2019

Abstract: Water managers need tools to assist in the management of ever increasing algal bloom problems over wide spatial areas to complement sparse and declining in situ monitoring networks. Optical methods employing satellite data offer rapid and widespread coverage for early detection of bloom events. The advent of the Analysis Ready Data (ARD) and Open Data Cube concepts offer the means to lower the technical challenges confronting managers, allowing them to adopt satellite tools. Exploiting Landsat ARD integrated into the Digital Earth Australia data cube, we developed a prototype algal bloom alerting tool for the state of New South Wales, Australia. A visualization portal allows managers to gain insights into bloom status across the state as a whole and to further investigate spatial patterns in bloom alerts at an individual water body basis. To complement this we also proposed an algal alert system for trial based on chlorophyll and TSM levels which requires further testing. The system was able to retrieve the status of 444 water bodies across the state and outputs from the visualization system are presented. Time series of image acquisitions during an intense bloom in one lake are used to demonstrate the potential of the system. We discuss the implications for further development and operationalisation including the potential for augmentation with alternative algorithms and incorporation of other sensor ARD data to improve both temporal and spectral resolutions.

Keywords: Landsat Analysis Ready Data; algal blooms; total suspended matter; turbidity; cyanobacteria; remote sensing; water quality; Open Datacubes; Earth observation

1. Introduction

Along with water quantity, water quality forms a critical component of global fresh water security and ecosystem health. Indeed, access to safe freshwater is such a significant challenge for much of the world's population that the UN has prioritized access to clean water and sanitation as Goal 6 of its Sustainable Development Goals (SDGs) [1]. Poor water quality is a problem that is faced by high, middle and low-income countries alike [2]; for example, Australia's inland water quality is ranked among the worst of the advanced economies and it is getting poorer [3,4].

Existing data related to water quality is scarce and declining, of questionable accuracy and of poor geographic and temporal coverage [5]. Thus, more consistent and more accurate information on inland water quality over wider areas is required. Only then can current conditions be assessed and changes investigated in response to other impacts such as changes in land use, fires, flooding and climate change [6,7].

Cyanobacterial (blue-green algal) blooms are a particular increasing concern in inland waters across the globe. Some species produce potent toxins that pose a major hazard to human health, livestock, wildlife and the aquatic environment [2]. Blooms result in significant economic impact to affected communities (e.g., [8,9]) in Australia, algal blooms result in >AU$250 million costs to affected communities per annum [10].

Traditional field monitoring for algal blooms involves sampling followed by species identification and cell counting [11,12]. However, whilst reliable, these methods take time, can be limited in spatial extent and risk under-sampling of episodic processes like algal blooms. Furthermore, with downward pressures on budgets that resource in situ sampling where repeated travel to remote areas and costly laboratory analyses are required, countries of both emerging and advanced economies need solutions to overcome issues related to the vast areas and number of water bodies required to be monitored.

Additional tools are therefore required to assist in more rapid and widespread monitoring at temporal resolutions that allow for early detection of blooms, thus providing early warning for management intervention. As one solution, satellites and other forms of optical remote sensing may provide an appealing complement to water managers for algal bloom and related water quality monitoring over multiple spatial scales. Recently launched satellite sensors (Landsat 8, Sentinel 2A, 2B, Sentinel 3A) offer the potential of free, medium resolution, wide scale and frequent monitoring of water quality in inland water bodies of a range of sizes in support of the development of bloom alerts for water managers (e.g., [13–16]).

Whilst data availability is no longer a limiting factor, for many agencies charged with oversight of water quality issues there can be significant challenges and barriers to the adoption of remotely sensed methods. These may include:

- Data deluge: the increasingly vast volumes of satellite data being offered, often free of charge from a variety of sensors.
- The need to develop sophisticated platforms and systems to handle the volumes of data and its complexities in preparation, handling, storage and analysis.
- The need to yield results with low latency: being able to rapidly process the frequently acquired data required to deliver a warning system with timely alerts, and for mitigation needs to be assessed.
- The need to effectively communicate information succinctly: being able to filter the vast quantity of data available to deliver relevant and timely information to water managers.

The above factors lead to hesitation in adoption of remote sensing technologies by water management agencies where the technical barriers are perceived to be too high. From the perspective of developing an algal bloom alerting system, key technical considerations cover a number of additional key aspects:

- Computational requirements, i.e., the characteristics of an appropriate implementation platform; these include the computational complexity of the method(s) used to generate the desired outputs, the desired spatial and temporal extents, software coding language and computing infrastructure.
- Data infrastructure requirements, including a) satellite sensor data suitability (with respect to spectral, spatial and temporal resolutions), availability and access, including to future data streams; b) data volume, storage and access; c) data pre-processing, i.e., the potentially complex process involving multiple steps such as orthorectification, cloud masking, reflectance calibration, terrain illumination correction, identification of corrupted pixels and compositing into mosaics. An algal bloom monitoring system typically needs fully-pre-processed/normalised surface reflectance data. Data latency is also relevant here.
- Desired outputs, i.e., the form of the resulting outputs (e.g., spatial maps, temporal animations and specific warnings for point locations), the users to be targeted, communication medium and update frequency.

The recent paired development of "analysis ready data" (ARD, http://ceos.org/ard/) and "open data cube" (ODC, https://www.opendatacube.org) concepts and initiatives represents a potential technological solution to agencies seeking to overcome the challenges above. ODC concepts exploit technological developments associated with new computing and data infrastructures; this international initiative represents a new paradigm in the management, analysis and distribution of Earth Observation (EO) data and products [17]. It enables the realisation of the value and impact of EO data by providing an open and freely accessible exploitation architecture to foster the breadth and depth of applications for societal benefits. At the heart of such systems are stacks of spatially aligned time series of ARD where consistent and endorsed pre-processing is addressed to reduce the data preparation burden on users and to allow for the rapid development of new applications [18]. Central to the ethos of these approaches is access to EO data and analysis platforms that are traceable, transparent, interoperable and open source. The open source paradigm encourages free access, code reuse, and a rapid expansion of user capabilities. Examples include Digital Earth Australia (DEA) [17,19,20], the USGS ARD initiative [18], Digital Earth Africa and the Swiss Data Cube [21].

The purpose of this paper is to outline the development of a prototype bloom alerting system built for the first time exploiting Landsat ARD made available in the open data cube concept. The system was developed for the Australian state of New South Wales (NSW) and involved two major components: 1) the application of an algorithm in a processor to allow rapid conversion of Landsat satellite data to suitable water quality data and 2) development of a visualisation for NSW water managers system providing timely alerts that allow for a rapid overview of both the state-wide extent of the problem as well as a focused view on individual water bodies.

2. Materials and Methods

2.1. Digital Earth Australia

In contrast to other continents, the sinuous and spatially complex size of freshwater bodies in Australia necessitates the use of medium to high spatial resolution sensors. In an analysis of 'observable' water bodies using a vectorised layer of Australian water bodies, Hestir et al. [22] showed that few natural freshwater ecosystems in Australia are detectable at MODIS (500 m) resolution. However, resolvability increases to as high as ~75% of water bodies at the Landsat (30 m) pixel scale.

We therefore implemented our system on Digital Earth Australia (DEA, formally known as the Australian Geoscience Data Cube) an open source data cube initially built using the Landsat archive [17,19]. The data cube is delivered on the National Computing Infrastructure (NCI, https://nci.org.au) at the Australian National University, a petascale High Performance Computing (HPC) facility. It consists of 300,000+ Landsat archive images (representing 10^{13} potential "measurements") for the entire Australian continent extending back to 1986, encompassing the sequence of scenes captured by Landsats 5 TM, 7 ETM+ and 8 OLI [17]. Ingestion of Landsat scenes into the DEA involves the following steps to ensure consistent processing (see [17] for further details):

1) Spatial alignment, where geometric corrections are applied such that pixels are stacked as time series of observations;
2) Radiometric correction involving the Nadir BRDF Adjusted Reflectance (NBAR) conversion of raw digital numbers (DNs) to consistent and comparable measurements of normalised surface reflectance [23];
3) Quality assessment, where all observations are retained but quality flags are applied at the dataset and pixel levels to allow for fitness for purpose determinations. Flags applied include cloud, cloud shadow, instrument saturation, [24], ACCA [25] and Fmask [26] cloud assessments, tests for sensor saturation, and zero values and sea/ocean flags;
4) Spatial partitioning, including tiling and NetCDF packaging for delivery, where the NetCDF format supports the efficient creation, access and sharing of array-oriented scientific data.

2.2. Algal Bloom Alerting Implementation

This implementation used Digital Earth Australia version 2, as outlined in [17]. An API provided a set of both high and lower level Python functions which can be used to efficiently search, and read DEA data and to apply tailored processing such as algorithms for specific applications. The final workflow describing the overall algal bloom alerting implementation for NSW water bodies is presented in Figure 1 and was implemented in the Python language on the NCI's Raijin supercomputing facility. The next sections describe the key components of this implementation. The various processes shown in Figure 1 are executed in parallel for each waterbody to leverage the computational power of modern multi-core infrastructures such as the NCI.

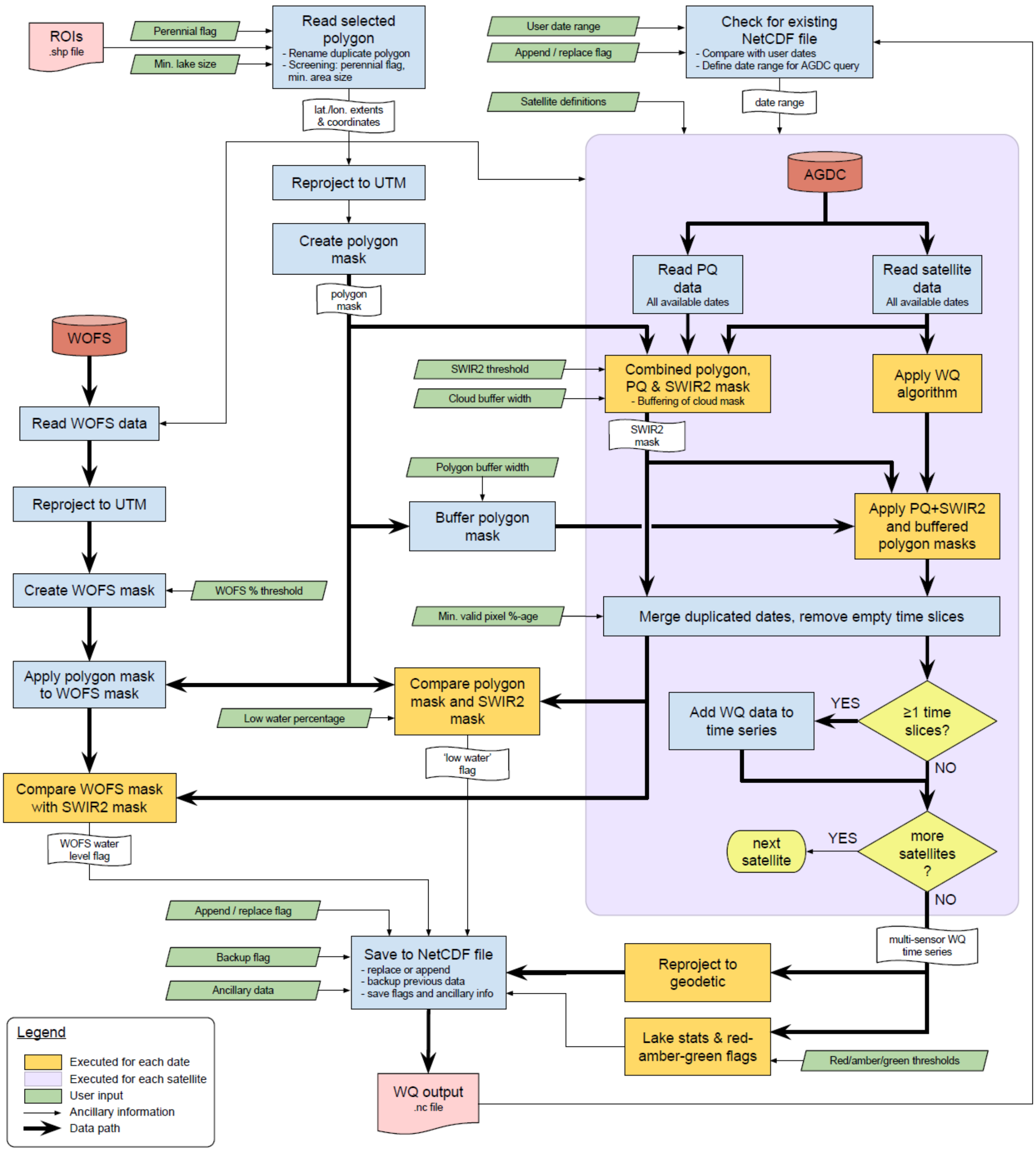


Figure 1. Detail of the implementation workflow to apply the algal alert system for NSW water bodies and showing the detail of various sub-components of the algorithm. The main processing task is highlighted in purple.

2.3. Main Processing Component

The main processing task (purple box in Figure 1) reads the satellite data for a given region (waterbody) of interest, in a given time window for the selected sensor. Pixel quality (PQ) data are used to identify pixels affected by cloud, shadow and saturation issues. Multiple same-date and empty time slices are also detected and removed/merged. A waterbody polygon vector layer for NSW waterbodies is applied to the time series to mask land pixels, and the desired water quality algorithm is applied. An additional SWIR2-band (threshold) filter is used to further mask out pixels affected by residual water vapour, sunglint and floating vegetation [16,27]. The resulting unified time series of water quality data are then thresholded for green, amber and red alert status and are written or appended as new data to a NetCDF file.

Water Observations from Space (WOFS) [20] data, a time series of water extent for the Australian continent, were used in combination with SWIR2 data to identify low-water conditions.

A separate NetCDF file is generated for each waterbody. Together with the time series data, ancillary data are also saved including time series of whole-of-waterbody algal bloom alert status flags (green, amber or red), which Landsat sensor generated the data in the time series, and low-water flags and statistics (mean, median, minima and maxima) for the water quality values. This allows efficient storing and querying of waterbody data.

A set of required and optimal user-defined parameters (green rhomboids in Figure 1) allows for customisation of the process and for flexibility in code execution. These allow, for example, the setting of the data time series windows, control of the buffering for lake, cloud and cloud shadow edges, definition of the desired water quality algorithm and thresholds for algal alert status and for low-water status definitions.

2.4. The Water Quality Algorithm

In terms of water quality algorithms, the few broad VIS-NIR bands available on the Landsat series of sensors limits detection to variables that have a broad spectral response [28]. To demonstrate functionality of the approach, we have in the first instance implemented a two band total suspended matter algorithm (TSM) [29]:

$$TSM_{index} = \frac{(green + red)}{2} \tag{1}$$

where *green* and *red* correspond to bands 2 and 3 for Landsats 5 and 7 and bands 3 and 4 for Landsat 8. This algorithm was calibrated for Australian inland waters by [16]:

$$TSM\left(mg\ L^{-1}\right) = 3983\ TSM_{index}{}^{1.6246} \tag{2}$$

for Landsat 5 TM and Landsat 7 ETM+, and

$$TSM\left(mg\ L^{-1}\right) = 3957\ TSM_{index}{}^{1.6436} \tag{3}$$

for Landsat 8 OLI, and has been shown to reliably generate time series for Australian waters, particularly with the improved radiometric resolution of Landsat 8. We are confident that determination of turbidity is in most instances related to occurrences of phytoplankton growth (discussed further below). The algal bloom alerting framework (Figure 1) allows for the modular implementation of the water quality algorithm. This means that alternate and new algorithms may be easily defined and applied within the framework.

2.5. Translation to Algal Alert Modes

Current Australian cyanobacterial alert levels for recreational waters build on WHO guidelines [11] and are based on cell counts and biovolumes [12]. Whilst biovolume, representing the integration of numerical abundance and community size structure, is the appropriate, legislated indicator of phytoplankton and cyanobacterial biomass for many countries, its measurement is time consuming, not without its limitations [30] and not suited to estimation using satellite data. The specific references to chlorophyll in the WHO guidelines allow for a starting point for the development of a staged algal alerting system suited to earth observation approaches. Malthus et al. [31] proposed an alert system for trial on the basis of chlorophyll concentration ranges given in Table 1. Conversion of these to TSM concentration ranges (Table 1) was made following analysis of the TSM–CHL relationship from 192 in situ sample measurements made in 13 eastern Australian waterbodies (Log(Chl) = 0.424 Log(TSM) + 0.650, R^2 = 0.58). Whilst a useful start, the establishment of agreed boundaries for algal alerting based on pigment and turbidity/sediment concentrations requires further research and the consensus of water managers.

Table 1. Proposed algal alert level guidelines based on chlorophyll and total suspended matter concentrations adopted for this study.

Alert Level	Chlorophyll	TSM
Green surveillance mode	< 20 ug Chl l^{-1}	< 20 mg m^{-3}
Amber alert mode	>20 – 50 ug Chl l^{-1}	>20–70 mg m^{-3}
Red action mode	>50 ug Chl l^{-1}	>70 mg m^{-3}

2.6. Visualisation of the Output

The main output of the algal bloom alerting implementation in Figure 1 is a multi-sensor time series of water quality (TSM) maps and alert flags for NSW waterbodies saved in NetCDF format. This output dataset is then used as input to a visualisation module to display the data at multiple scales. Users can rapidly browse, search for water bodies in NSW and see the most recent satellite capture and algal status.

The data were served by Hyrax (https://www.opendap.org/software/hyrax-data-server), an open source package that provides partial, plaintext and structured endpoints for NetCDF files in a means that saves bandwidth and increases client performance. Hyrax is able to generate plaintext or JSON (Javascript Object Notation) files for the purpose of visualisation by querying a waterbody's algal status and statistics in a single browser request, thereby significantly speeding up data access. For the visualisation, the open source packages Leaflet (http://leafletjs.com/), React (https://facebook.github.io/react/) and D3 (https://d3js.org/) were used to map the data, build the interface and chart algal trends, respectively.

Key features of the interface design allow for the rapid appraisal of alert status at a state-wide overview level as a gauge on overall assessment of which lakes are stable, at risk, or improving. Using the thresholds in Table 1, the chart is coloured using the green, amber and red alert levels for intuitive interpretation. Search and zoom functions then allow to explore spatial patterns and temporal trends at the individual waterbody level. Algal trend at a glance is provided by putting recent measurements in a 24-month context in a separate interactive area-line chart.

3. Results

A total of 1947 polygons were contained within the waterbody vector layer for NSW. We used a conservative minimum threshold area of 0.5 ha, representing 8 Landsat pixels at 25 m resolution, that ensured only pure water pixels were considered for further processing; this resulted in a total of 912 perennial lakes above this threshold. Following application of the bloom alerting workflow, a total of 444 NetCDF files were created and saved to disk. No valid data were found for the remaining 468 lakes in the selected time window due to SWIR2 data and pixel quality data indicating missing data, insufficient water present, contaminated by sunglint, shallow enough to be dominated by aquatic macrophytes and other wetland vegetation or affected by cloud cover.

In relation to observation frequency, for the currently operating Landsat 7 and 8 platforms, the revisit rate is 16 days; with both platforms operating in opposition, NSW—wide Landsat—based algal alert products can thus be expected to be updated at best once every 8 days. However, for water bodies lying in L7/L8 overlap regions, coverage can be obtained under certain situations within a 24-hour period. A surprisingly large number of water bodies in NSW fall into this category; indeed for some lakes lying within the overlap between adjacent Landsat paths, Lymburner et al. [16] were able to retrieve >500 valid observations of TSM between 1987 and 2014.

An example of the rapid state-wide overview from the visualisation module is shown in Figure 2. This shows broad lake-level flags indicating overall current alert status for individual waterbodies, as well as a search and lake-listing feature allowing the ability to view data at the level of individual waterbody. The interface also allows for the selection of both image and vector map background datasets to add context. A list of all returned waterbodies is provided on the right hand side of the overview interface, which, together with the zoom and search functions, allows for easy investigation of details for an individual water body. An example of visualisation at the lake level is shown in Figure 3. Colour coded for alert status at the pixel level, the spatial pattern of bloom formation can be explored to determine lake areas at greatest risk. The data for one period in time are set in context over a 24-month period for the water body extracted from the Landsat data using an area-line chart (right-hand side panel, Figure 3). The area shows the range of the data, and the line which runs through shows the median point. Dragging along the chart changes the selected date of the displayed algal alert colour map.

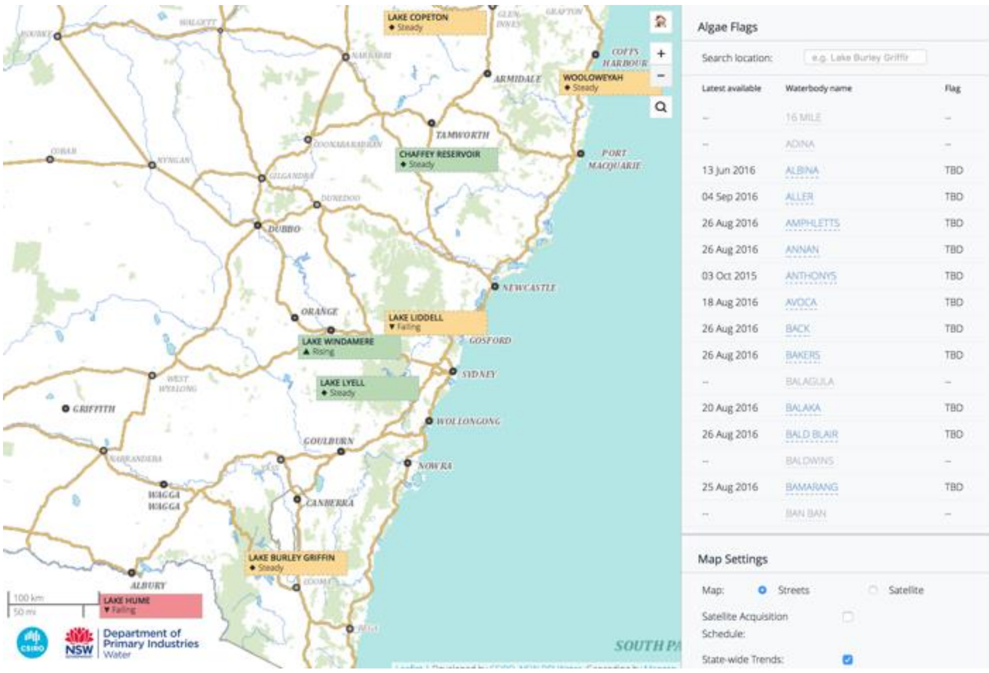


Figure 2. Example of the visualisation tool showing the statewide overview of bloom conditions.

To illustrate the power of the visualisation system, a time series of returned Landsat 7 and 8 alert images for Lake Hume, covering the summer to autumn period mid-January to early July 2016, are presented in Figure 4. This time series covers an intense period of bloom formation in Lake Hume [32]. The reservoir sits in an overlap region of Landsat orbits which means images may be acquired within 24 hours, thus increasing the frequency of coverage for this lake. Striping is evident in the Landsat 7 data owing to the SLC issue associated with this sensor (https://landsat.usgs.gov/slc-products-background). Other gaps in the data are due to the presence of clouds and varying water extent associated with changes in reservoir level. Turbidity in this deep reservoir is primarily driven by changes in phytoplankton concentrations [32]. The time series shows the development of an algal bloom from late

February through March and April. This bloom provided the 'seed' to stimulate a bloom in the River Murray downstream of the reservoir which affected some ~1600 km of river for three months to June 2016 [32,33]. A full time series of turbidities from January 2015 for this lake is depicted in Figure 3 (right hand side). The increase in status due to the bloom is indicated in rising median turbidity in this deep lake from March 2016 onwards.

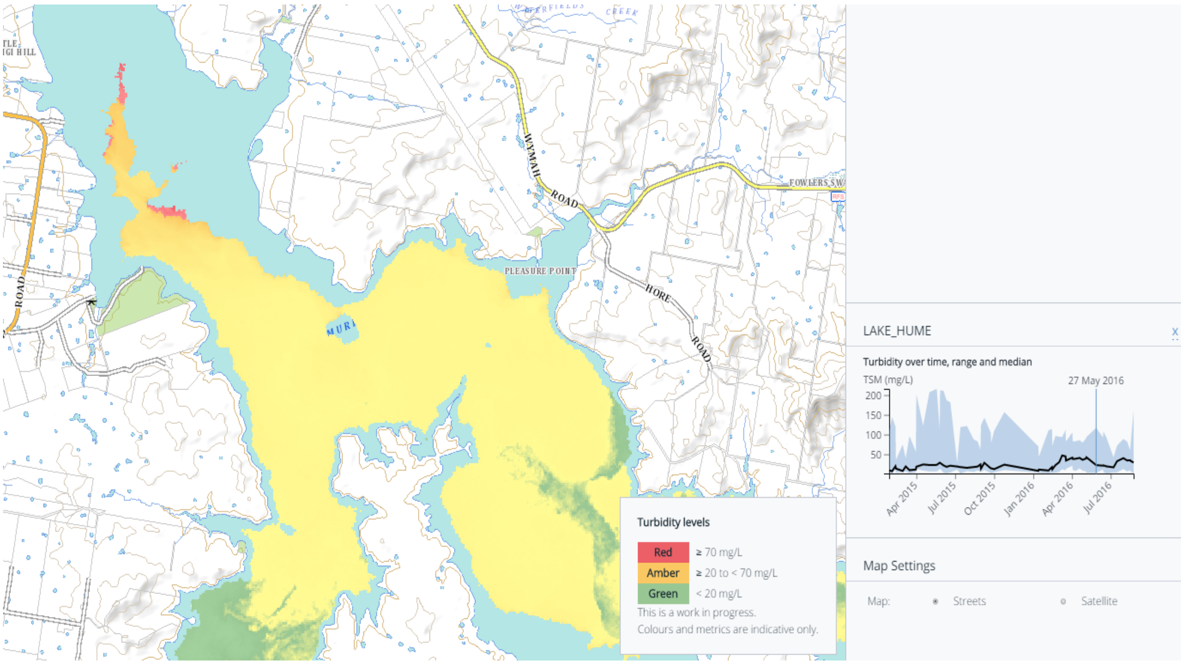
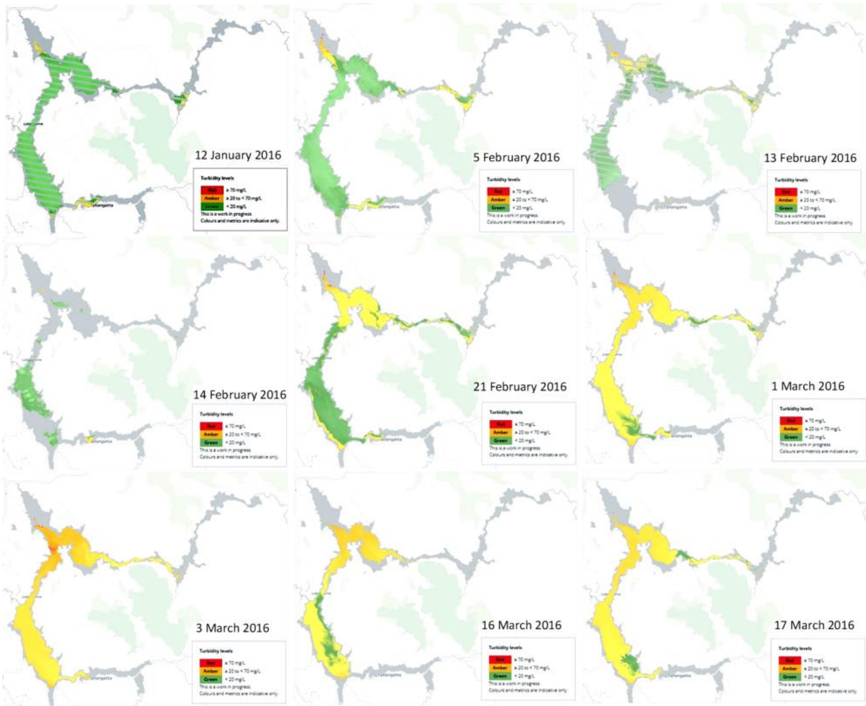
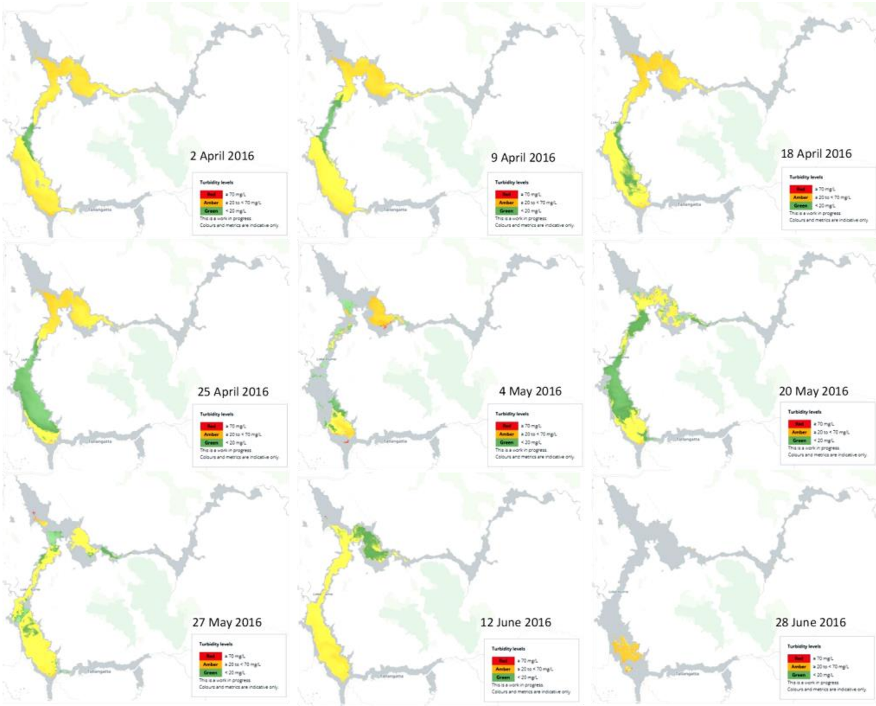


Figure 3. Example of the visualisation tool showing spatial detail for part of Lake Hume, NSW, on May 27, 2016.



(**a**)

Figure 4. *Cont.*



(**b**)

Figure 4. Time series of Landsat 7 and 8 turbidity images for Lake Hume using the algal bloom processing and visualisation system, covering the summer to autumn period mid-January to early July 2016. Turbidity levels are coloured on the basis of green, amber and red algal alert status, depicted bottom right in each image. (**a**) Image time series from 17 January 2016 to 17 March 2016; (**b**) image time series 2 April 2016 to 28 June 2016.

4. Discussion

For Australian waters, we have, in the first instance, produced a turbidity-based algal alerting system drawing on Landsat ARD held within the DEA open datacube. The challenge in the project was to develop a software framework and product delivery system that achieves automated rapid turnaround of satellite data streams to green-amber-red algal reports. To achieve this, the main processing code has been specifically optimised and main computations parallelised. Via a visualisation interface, the results can be reviewed by water managers for NSW at state level to provide a rapid state-wide overview of overall algal alert status or at the scale of the individual water body to allow the determination of spatial bloom dynamics. Current data can be displayed with historical data to allow the up-to-date situation to be put into a longer-term context. Together with the visualisation module interface, the Python code written to implement the workflow depicted in Figure 1 represents a self-contained, fully functional and customisable prototype of an algal alert monitoring system using satellite data streams, leveraging the flexibility and computational power of the ARD database and the NCI's high-performance computing facilities. The system was also designed such that it is capable of accommodating new sensors and data streams, and to be easily updated with alternative semi-empirical algorithms. However, sophisticated systems are required to handle the frequent image acquisition and rapid processing required to deliver a warning system with low latency; for many agencies, this can be a significant barrier to the adoption of remotely sensed methods.

As an indication of performance and efficiency on the HPC facility, using all available Landsat 7 and 8 ARD time slices over NSW over an 18-month period (January 1, 2015 and July 1, 2016), the shapefile of 1947 water bodies for NSW and operating on one node (16 CPUs) on the NCI Raijin supercomputer the multi-threaded workflow was executed in a wall time of 23.5 minutes (using about 23GB of memory), resulting in a total of about 6.27 CPU-hours, of which 35% of the time is spent querying, accessing and loading data from the data cube. Using the base rate for commercial access to NCI facilities and an 'express priority' multiplier this equates to a ~AU$1.50 cost to process the entire dataset and less than AU$0.30 to perform an update on new time slices to append to the database.

In selecting the DEA ODC as the tool upon which to develop the bloom alerting service, the key benefits and considerations influencing the decision were: 1) continuity of data archiving of high quality Landsat ARD (as surface reflectance) for the foreseeable future; 2) future integration of Copernicus Sentinel 2 ARD data, allowing for a high-frequency multi-sensor algal alert system; 3) the availability of an API allowing easy enquiry and access to speed up initial product development.

During the development of this work package, a number of issues and further improvements were identified which are further elaborated here. The cloud/shadow mask provided as part of the Landsat ARD's Pixel Quality layer at times missed significant regions of affected pixels; buffering around the edges of the mask was required to mitigate these effects, but some issues still remained.

A SWIR2 threshold (currently set at 1%) filter was implemented to filter out all pixels potentially affected by issues such as shallow water, shore adjacency, floating vegetation and sunglint. However, it was common for this filtering process to discard a significant proportion of the available time slices in a waterbody's time series. Further research should determine whether a more moderate SWIR2 threshold could be used without significantly affecting the false alarm rate of the algorithm.

At the heart of remote sensing approaches to the problem of bloom detection lies the estimation of turbidity or pigment concentrations. The scale of Australian water bodies demands medium resolution data obtained by either Landsat or Sentinel 2 data to obtain a sufficient coverage of water bodies which are typically small in size in comparison to those on other continents [22]. The use of Landsat ARD as the sensor-type for detection in this prototype limits the applications in both choice of algorithms available and its overpass frequency. In terms of temporal acquisitions, the Lake Hume example (Figure 4) illustrates the frequency with which Landsat acquisitions can be obtained under typical Australian atmospheric conditions. However, at an 8-day repeat cycle (for combined Landsat 7 and 8 acquisitions) temporal frequency may fail to provide adequate coverage for algal bloom alerting, particularly in more cloudy regions of the world [22,27]. Temporal resolution could be significantly improved through the integration of Landsat and Sentinel 2 ARD products.

The few broad VIS-NIR bands available on the Landsat series of sensors limits detection to variables that have a broad spectral response such as suspended sediment, light attenuation, transparency and turbidity using semi-empirical approaches [28]. To effect this demonstration, we developed the alerting system based on an algorithm accurate for the detection of TSM in Australian waters [16]; as stated, limitations in Landsat band numbers, positions and widths limits detection of spectral responses directly related to phytoplankton-driven spectral responses. However, our approach is further supported by an empirical TSM-CHL relationship, developed on data from 13 east Australian water bodies covering a large gradient in hydro-climatic conditions. This approach is most valid when algal biomass is the primary determinant of TSM, often the case in eutrophic conditions. The TSM-CHL relationship is stronger when our analysis is restricted to deeper water bodies (8 water bodies, R^2 = 0.60); the relationship is poorer for shallow lakes (5 water bodies, R^2 = 0.41) where contributions from non-algal sources such as resuspension may be greater. Similar, strong TSM-CHL relationships (R^2 = 0.78) have been observed in northern American lakes [34] indicating the approach may have wider applicability.

To be detected accurately, algal pigments such as chlorophyll-a and cyanobacterial pigments require spectral bands located appropriately to define their absorption features; the broad nature of the Landsat red band and differences in spectral response between the sensors on Landsats 5, 7 and 8 [16], implies that the ability of Landsat to discriminate pigment concentrations may be low [13], particularly at low concentrations [28]. Landsat-based detections of cyanobacterial blooms in previous studies (e.g., [35]) are likely to be the result of correlation between phycocyanin and other variables such as turbidity [22,27]. More sophisticated approaches will determine the presence of pigments in the remotely sensed data but will require sensors with sufficient spectral resolution [28]. Radiative transfer-based inversion has been shown to robustly retrieve chlorophyll and other water quality components simultaneously using Landsat ARD and other broad band satellite data corrected to surface reflectance [32,36–38], and although more work to code, could be implemented on our workflow.

Nevertheless, we have here relied on a turbidity algorithm in the first instance to demonstrate the potential of the approach. The algorithm has been shown to reliably generate time series for Australian waters, particularly with the improved radiometric resolution of Landsat 8 [16] and may be used as a proxy for the density of cyanobacteria blooms, particularly in deep waters; the Lake Hume time series during the course of a long lasting cyanobacterial bloom (Figure 4) demonstrates this ability well. We are confident that determination of turbidity is in most instances related to occurrences of phytoplankton growth in the NSW region. The system is designed to be used by water managers who will be aware of antecedent conditions such as rainfall, flooding and wind that might have contributed to high turbidities caused by sediments. As such, it may inform managers of the need to investigate in the field whether or not the bloom in a lake is associated with potentially toxic cyanobacterial species. Similarly, the sustained increase of potential bloom levels (e.g., sequence of amber to red alerts) provided by the system may be used as a trigger for the acquisition of fine-resolution/low-latency satellite data (e.g., WorldView-3, Planet Labs Dove) for more vigilant RS monitoring over specific areas.

With resources for in situ monitoring scarce and declining [16], the paucity of in situ observations of water quality parameters including TSM and chlorophyll in Australian inland waters limits our ability to robustly validate the algal alert system developed using independently measured observations. ‘Match-up’ validation data are required coincident with satellite overpasses that are consistent enough to be scaled up to the spatial scales of satellite pixels; this is a global challenge [39]. Nevertheless, our approach is based on a validated algorithms and qualitative evaluations (Lake Hume and other water bodies not presented here) highlight the value of the approach. As Australian efforts to monitor inland water quality from satellite data develops, greater effort will be required to synchronise in situ measurements with satellite overpasses to obtain sufficient data for validation of satellite-estimated water quality products. Emphasis on the use of continuously monitoring in situ equipment will be required to achieve this.

To be effective for water management, near real time satellite driven algal alerting systems such as that implemented here require rapid update, a process driven both by the acquisition frequency of the satellite data and the latency with which the latest satellite imagery can be processed to ARD form and ingested into a data cube. Latency itself can be driven more by the availability of the ancillary data (e.g., atmospheric data provided non-commercially by other institutions) required for satellite pre-processing than of the availability of the satellite data itself.

The greater adoption of optical approaches for algal detection would be greatly facilitated by the adaptation of WHO driven guidelines to include clearly stated pigment and turbidity levels. There are, however, clear challenges to the direct conversion of WHO guidelines based on cell numbers and biovolumes to those based on pigments and turbidity but we have here proposed green/amber/red alerting thresholds (Table 1) to enable this using a logical approach based on analysis of in situ data; these require further validation for wider application. Different sets of thresholds could ideally be used for waterbodies belonging to either the ‘deep/clear’ or ‘shallow/turbid’ water body classes. These could be driven by known long-term temporal variability of the waterbody’s turbidity levels, or morphometric parameters such as the ratio of water volume (lake capacity) to surface area, or the shoreline development index. The calibration of such water-type-dependent thresholds will require further research effort.

5. Conclusions

Reliable and consistent water quality mapping requires consistent satellite surface reflectance data and delivery of such can be a considerable barrier to entry for many organisations and agencies interested in implementing satellite data as part of their routine monitoring processes. This paper has demonstrated the significant benefit that ARD, organised within a data cube, has brought to the development of a valuable water quality alerting tool. The consistent pre-processing, quality flags and speed and efficiency of data access allows for a system that can be both rapidly implemented and rapidly updated for new acquisitions. Most importantly such an implementation allows for the generation of more consistent and more accurate information on inland water quality over wider geographical areas such that those responses can be assessed in response to other impacts such as changes in land use, fires, flooding and climate change and the immediate perceived threats mitigated through management.

Our system, based on optical remote sensing, offers a method to objectively assess inland water quality over multiple spatial scales. Our tools allow for rapid visualisation at a regional scale followed by in depth analysis at the water body scale. Analysis of the patterns at the water body scale exploits the rich spatial detail available in the satellite data, highlighting areas at particular risk for mitigation management; this both complements and enhances the limited spatial detail available from more traditional, and more costly, point-based in situ sampling methods.

The approach developed here offers an objective tool to further build on, with areas of further development including integration with other high spatial resolution sensor data to improve both temporal and spectral resolutions and improved algorithms.

Author Contributions: T.J.M. led the project and wrote the paper. E.L. coded the DEA implementation and wrote the report upon which the paper is based. X.H. developed and coded the visualisation interface. E.B. and J.A. contributed to Landsat analysis and algorithm development.

Funding: This project was supported by funding from the New South Wales Office of the Premier. Further support for development of the visualisation interface was also provided by internal CSIRO eResearch Collaboration Project ERRFP-394.

Acknowledgments: Highly constructive discussions and feedback from John Brayan, Mustak Shaikh and Lee Bowling of the NSW Department of Industry, Lands and Water Division, were highly valuable in the creation of the final alerting system. Advice from Geoscience Australia staff on implementation on the data cube is also gratefully acknowledged.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. United Nations Development Program (UNDP). UNDP Support to the Implementation of Sustainable Development goal 6: Sustainable Management of Water and Sanitation. United Nations Development Program. Available online: http://www.undp.org/content/undp/en/home/librarypage/sustainable-development-goals/undp-support-to-the-implementation-of-the-2030-agenda/ (accessed on 30 October 2019).
2. Damania, R.; Desbureaux, S.; Rodella, A.; Russ, J.; Zaveri, E. *Quality Unknown: The Invisible Water Crisis*; World Bank: Washington, DC, USA, 2019. [CrossRef]
3. Dekker, A.G.; Hestir, E. *Evaluating the Feasibility of Systematic Inland Water Quality Monitoring with Satellite Remote Sensing*; CSIRO: Canberra, Australia, 2012.
4. Emerson, J.; Hsu, A.; Levy, M.; de Sherbinin, A.; Mara, V.; Esty, D.; Jiteh, M. *Environmental Performance Index and Pilot Trend Environmental Performance Index*; Yale Center for Environmental Law and Policy: New Haven, CT, USA, 2012.
5. Srebotnjak, T.; Carr, G.; de Sherbinin, A.; Rickwood, C.A. Global Water Quality Index and hot-deck imputation of missing data. *Ecol. Indic.* **2012**, *17*, 108–119. [CrossRef]
6. Salmaso, N.; Mosello, R. Limnological research in the deep southern subalpine lakes: Synthesis, directions and perspectives. *Adv. Oceanogr. Limnol.* **2010**, *1*, 29–66. [CrossRef]
7. Vorosmarty, C.J.; McIntyre, P.B.; Gessner, M.O.; Dudgeon, D.; Prusevich, A.; Green, P.; Glidden, S.; Bunn, S.E.; Sullivan, C.A.; Reidy Liermann, C.; et al. Global threats to human water security and river biodiversity. *Nature* **2010**, *467*, 555–561. [CrossRef] [PubMed]
8. Hoagland, P.; Scatasta, S. The economic effects of harmful algal blooms. In *Ecology of Harmful Algae*; Graneéli, E., Turner, J.T., Eds.; Springer-Verlag Berlin Heidelberg: Heidelberg, Germany, 2006; Volume 189, pp. 391–402.

9. Willis, C.; Papathanasopoulou, E.; Russel, D.; Artioli, Y. Harmful algal blooms—The impacts on cultural ecosystem services and human well-being in a case study setting, Cornwall, UK. *Mar. Policy* **2018**, *97*, 232–238. [CrossRef]
10. Davis, J.; Koop, K. Eutrophication in Australian rivers, reservoirs and estuaries—A southern hemisphere perspective on the science and its implications. *Hydrobiologia* **2006**, *559*, 23–76. [CrossRef]
11. Chorus, I.; Bartram, J. *Toxic Cyanobacteria in Water: A Guide to Their Public Health Consequences, Monitoring and Management*; Published on behalf of UNESCO, WHO and UNEP; E. & F.N. Spon: London, UK, 1999.
12. Australian Government, National Health and Medical Research Council (NHMRC). Guidelines for Managing Risks in Recreational Water. 2008. Available online: https://www.nhmrc.gov.au/_files_nhmrc/publications/attachments/eh38.pdf (accessed on 30 October 2019).
13. Matthews, M.W. A current review of empirical procedures of remote sensing in inland and near-coastal transitional waters. *Int. J. Remote Sens.* **2011**, *32*, 6855–6899. [CrossRef]
14. Oyama, Y.; Matsushita, B.; Fukushima, T. Distinguishing surface cyanobacterial blooms and aquatic macrophytes using Landsat/TM and ETM + shortwave infrared bands. *Remote Sens. Environ.* **2015**, *157*, 35–47. [CrossRef]
15. Stumpf, R.P.; Davis, T.W.; Wynne, T.T.; Graham, J.L.; Loftin, K.A.; Johengen, T.H.; Gossiaux, D.; Palladino, D.; Burtner, A. Review: Challenges for mapping cyanotoxin patterns from remote sensing of cyanobacteria. *Harmful Algae* **2016**, *54*, 160–173. [CrossRef]
16. Lymburner, L.; Botha, E.; Hestir, E.; Anstee, J.; Sagar, S.; Dekker, A.; Malthus, T. Landsat 8: Providing continuity and increased precision for measuring multi-decadal time series of total suspended matter. *Remote Sens. Environ.* **2016**, *185*, 108–118. [CrossRef]
17. Lewis, A.; Oliver, S.; Lymburner, L.; Evans, B.; Wyborn, L.; Mueller, N.; Raevksi, G.; Hooke, J.; Woodcock, R.; Sixsmith, J.; et al. The Australian Geoscience Data Cube — Foundations and lessons learned. *Remote Sens. Environ.* **2017**, *202*, 276–292. [CrossRef]
18. Dwyer, J.L.; Roy, D.P.; Sauer, B.; Jenkerson, C.; Zhang, H.K.; Lymburner, L. Analysis ready data: enabling analysis of the Landsat archive. *Remote Sens.* **2018**, *10*, 1363.
19. Lewis, A.; Lymburner, L.; Purss, M.B.J.; Brooke, B.; Evans, B.; Ip, A.; Dekker, A.G.; Irons, J.R.; Minchin, S.; Mueller, N.; et al. Rapid, high-resolution detection of environmental change over continental scales from satellite data — The earth observation data cube. *Int. J. Digital Earth* **2015**, *9*, 106–111. [CrossRef]
20. Mueller, N.; Lewis, A.; Roberts, D.; Ring, S.; Melrose, R.; Sixsmith, J.; Lymburner, L.; McIntyre, A.; Tan, P.; Curnow, S.; et al. Water observations from space: Mapping surface water from 25 years of Landsat imagery across Australia. *Remote Sens. Environ.* **2016**, *174*, 341–352. [CrossRef]
21. Giuliani, G.; Chatenoux, B.; De Bono, A.; Rodila, D.; Richard, J.-P.; Allenbach, K.; Dao, H.; Peduzzi, P. Building an earth observations data cube: Lessons learned from the Swiss Data Cube (SDC) on generating analysis ready data (ARD). *Big Earth Data* **2017**, *1*, 1–18. [CrossRef]
22. Hestir, E.L.; Brando, V.E.; Bresciani, M.; Giardino, C.; Matta, E.; Villa, P.; Dekker, A.G. Measuring freshwater aquatic ecosystems: The need for a hyperspectral global mapping satellite mission. *Remote Sens. Environ.* **2015**, *167*, 181–195. [CrossRef]
23. Li, F.; Jupp, D.L.B.; Thankappan, M.; Lymburner, L.; Mueller, N.; Lewis, A.; Held, A.A. Physics-based atmospheric and BRDF correction for Landsat data over mountainous terrain. *Remote Sens. Environ.* **2012**, *124*, 756–770. [CrossRef]
24. Sixsmith, J.; Oliver, S.; Lymburner, L. A hybrid approach to automated Landsat pixel quality. In Proceedings of the IEEE International Geoscience and Remote Sensing Symposium – IGARSS, Melbourne, Australia, 21–26 July 2013; pp. 4146–4149.
25. Irish, R.R.; Barker, J.L.; Goward, S.N.; Arvidson, T. Characterization of the Landsat-7 ETM+ automated cloud-cover assessment (ACCA) algorithm. *Photogramm. Eng. Remote Sens.* **2006**, *72*, 1179–1188. [CrossRef]
26. Zhu, Z.; Woodcock, C.E. Object-based cloud and cloud shadow detection in Landsat imagery. *Remote Sens. Environ.* **2012**, *118*, 83–94. [CrossRef]
27. Kutser, T. Passive optical remote sensing of cyanobacteria and other intense phytoplankton blooms in coastal and inland waters. *Int. J. Remote Sens.* **2009**, *30*, 4401–4425. [CrossRef]

28. Dekker, A.G.; Pinnel, N.; Feasibility Study for an Aquatic Ecosystem Earth Observing System. Commonwealth Scientific and Industrial Research Organisation (CSIRO), Australia. 2018. Available online: http://ceos.org/document_management/Publications/Feasibility-Study-for-an-Aquatic-Ecosystem-EOS-v.2-hi-res_05April2018.pdf (accessed on 30 October 2019).
29. Dekker, A.G.; Vos, R.J.; Peters, S.W.M. Analytical algorithms for lake water TSM estimation for retrospective analyses of TM and SPOT sensor data. *Int. J. Remote Sens.* **2002**, *23*, 15–35. [CrossRef]
30. Kasprzak, P.; Padisak, J.; Koschel, R.; Krienitz, L.; Gervais, F. Chlorophyll concentration across a trophic gradient of lakes: An estimator of phytoplankton biomass. *Limnologica* **2008**, *38*, 327–338. [CrossRef]
31. Malthus, T.J.; Botha, E.; Anstee, J. *Early Warning System for Harmful Algal Blooms: Report on Work Package 1—Rapid Bloom Identification: Near Surface Sensing*; CSIRO Land and Water Flagship: Canberra, Australia, 2015; Available online: http://hdl.handle.net/102.100.100/92466?index=1 (accessed on 30 October 2019).
32. Joehnk, K.D.; Anstee, J.; Ford, P.; Botha, H.; Sherman, B. *Lake Hume Blue-Green Algal Risk Minimisation—Unpublished Report Prepared for the Murray Darling Basin Authority*; CSIRO Land and Water: Canberra, Australia, 2018.
33. Baldwin, D.S.; Gigney, H.; Wilson, J.; Watson, G.; Boulding, A.N. Drivers of water quality in a large water storage reservoir during a period of extreme drawdown. *Water Res.* **2008**, *42*, 4711–4724. [CrossRef] [PubMed]
34. Brezonik, P.L.; Bouchard, R.W.; Finlay, J.C.; Griffin, C.G.; Olmanson, L.G.; Anderson, J.P.; Arnold, W.A.; Hozalski, R. Color, chlorophyll a, and suspended solids effects on Secchi depth in lakes: Implications for trophic state assessment. *Ecol. Appl.* **2019**, *29*, e01871. [CrossRef] [PubMed]
35. Vincent, R.K.; Qin, X.; Mckay, R.M.L.; Miner, J.; Czajkowski, K.; Savino, J.; Bridgeman, T. Phycocyanin detection from LANDSAT TM data for mapping cyanobacterial blooms in Lake Erie. *Remote Sens. Environ.* **2004**, *89*, 381–392. [CrossRef]
36. Campbell, G.; Phinn, S.R.; Dekker, A.G.; Brando, V.E. Remote sensing of water quality in an Australian tropical freshwater impoundment using matrix inversion and MERIS images. *Remote Sens. Environ.* **2011**, *115*, 2402–2414. [CrossRef]
37. Mishra, S.; Mishra, D.R.; Lee, Z. Bio-optical inversion in highly turbid and cyanobacteria-dominated waters. *IEEE Trans. Geosci. Remote Sens.* **2014**, *52*, 375–388. [CrossRef]
38. Watanabe, F.; Alcaântara, E.; Imai, N.; Rodrigues, T.; Bernardo, N. Estimation of chlorophyll-a concentration from optimizing a semi-analytical algorithm in productive inland waters. *Remote Sens.* **2018**, *10*, 227. [CrossRef]
39. Odermatt, D.; Giardino, C.; Heege, T. Chlorophyll retrieval with MERIS Case-2-Regional in perialpine lakes. *Remote Sens. Environ.* **2010**, *114*, 607–617. [CrossRef]