Model of chlorophyll-a concentrations in inland water bodies based on Landsat data

TOMAS BRUNCLIK, KWASI ASARE BAFFOUR DANQUAH
Institute of Environmental and Chemical Engineering
University of Pardubice
Studentská 95, 532 10 Pardubice
CZECH REPUBLIC

Abstract: - This work seeks to create a model with the use of remote sensing as a tool in the measuring of chlorophyll-a and possible other water quality parameters in water bodies in the Czech Republic. The objective is to create a better correlation between in-situ measurement of chlorophyll-a and that of the data extracted from satellite images. This is to serve as a tool for the estimation and monitoring of chlorophyll-a distribution in water bodies in real time. The use of Quantum GIS and Grass GIS was adapted in this study to process the data collected from satellite(Landsat 7) and in-situ measurement made.

Key-Words: - Landsat, Remote sensing, chlorophyll-a, model, water quality, parameters

1 Introduction
Water bodies such as lakes and ponds do need to be carefully managed as their quality has a lot of effects on humans using it, aquatic organisms and on the environment in general. Making sure these water bodies have a standard quality is tedious and sometimes time consuming as well as costly work to do. Measuring of such quality parameters has to be done in-situ or samples taken to the laboratory for measurements. If one has a number of water bodies to deal with, one can imagine the work involved. Chlorophyll can be one of such parameters to measure with respect to water quality and specifically chlorophyll-a.

Chlorophyll-a (chlorophyll alpha) is a type of chlorophyll which is most common and predominant in all oxygen-evolving photosynthetic organisms such as higher plants, red and green algae. In this regard Remote Sensing could be one of the possible tools with which we can try to monitor some of these parameters. Remote sensing is defined as the technique of obtaining information about objects through the analysis of data collected by special instruments that are not in physical contact with the objects of investigation [1]. This tool could be effectively used in the measuring and estimating of some of these parameters [2] [3] [4] [5]. In this paper we will try to create a model that would help in the estimating of chlorophyll-a specifically in water bodies. A number of water bodies have been monitored and samples collected from them to analyze in helping create a suitable model to help monitor the chlorophyll levels in them. Its also interesting to know how other parameters such as Nitrogen, Carbon, Organic Carbon, Phosphate, Temperature and Water depth would have on the water bodies. A number of water bodies in the Pardubice and Hradec Kralove Region would be used in this research.

2 Study area
This work concentrates on water bodies near town Pardubice.

Fig. 1: Study area
Water bodies in this area are mainly fishponds and lakes created by sand mining. All the water bodies are relatively small, those sampled in this work range from 1.7 to 90 hectares, approximately. The fish ponds were established in middle ages \[^6\] and have inlet and outlet. Water bodies originating from sand mining are relatively new, the mining on some stopped few years ago, and do not have inlet or outlet. Most of water bodies of both types are used for both fishing and swimming.

3 Materials and methods

3.1 Water Sampling and Analysis

The sampling takes place in one day, at the satellite overpass, if possible. Samples have been taken from inflatable boat and stored in box with ice to ensure cold and dark conditions. Samples have been taken from the depth of few centimeters under the water level. Readings of temperature and Secchi disk depth (SDD) were recorded at the time of sampling. GPS unit Trimble Juno SB was used to record the sampling points location and the track of the boat.

Chlorophyll-a analysis was done in the laboratory according to ISO-10260 \[^7\], using microfibre glass filters and absorption into hot ethanol to concentrate the chlorophyll-a before spectrophotometric determination.

Total carbon (TC), total organic carbon (TOC) and total nitrogen (TN) analysis was done using TOC/TN analyzer directly from collected samples. Phosphate ions were analyzed spectrophotometrically.

3.2 Landsat data and GIS processing

The sampling started in autumn 2011, when Landsat 5 images were still available, but since the failure of Landsat 5 only Landsat 7 ETM+ data are used.

The study area is fully covered by both Landsat scenes of WRS-2 number 191/25 and 190/25, which gives average satellite revisit time of 8 days.

For this stage of the project, where we analyze data from individual sampling dates, atmospheric correction of the imagery is not necessary and was not used. Pre-processing of the data consists of loading the data into GIS software (GRASS, or Quantum GIS with GRASS plug-in), setting null value. Positional precision was checked and so far no adjustments were needed.

The first stage of processing of the measurements in GIS involves importing the measurement sites points as vector map layer and loading the values of digital numbers at the point locations into its attribute table. The problem of missing data strips in Landsat 7 imagery have been solved by moving the sampling point to the closets valid water pixel, or by removing that sampling point from the processing, if the move would be more than about 60 meters.

Smoothing of the Landsat bands was tested, but in most cases it had negative impact on the results. We presume it is connected with the small size of the water areas analyzed, as was indicated also in previous work on reservoirs \[^8\].

4 Results

4.1 Chlorophyll-a correlation

The analysis of correlation of chlorophyll-a (Table 1) was done to bring out the bands with the best correlation as well as band combinations. In all 17 samples were analyzed on these 2 dates, 9 samples on the Sept-15 2011 and 8 on the May-29 2012. Correlation was done for bands and only the ones with correlation above 0.50 are shown in the table above. The combinations of bands did have some considerable improvement in certain band combinations as compared to the individual bands. L3 can be seen to 0.58 and L1 is 0.002374 but when combined there is a positive effect bringing about a better correlation of 0.84 also L1/L2 and L2/L1 gave the highest correlation of 0.86 and 0.87 respectively. The log of chlorophyll measured was also processed to see if it has any impact on the level of correlation it brought a positive correlation in the combination of bands L2/L7 as well as L7/L3. Correlation for chlorophyll for the next date 29.05.12 had L6 with the highest with a value of 0.85 and L2/L1 with highest in the combination of bands with a value of 0.79.Log of chlorophyll brought no change in the value of L6 which the highest measured without.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Date</th>
<th>L2</th>
<th>L3</th>
<th>L7</th>
<th>L1/L2</th>
<th>L1/L3</th>
<th>L1/L7</th>
<th>L2/L7</th>
<th>L3/L5</th>
<th>L3/L7</th>
<th>L4/L7</th>
<th>L2/L1</th>
<th>L3/L1</th>
<th>L7/L1</th>
<th>L7/L2</th>
<th>L7/L3</th>
<th>L7/L4</th>
</tr>
</thead>
<tbody>
<tr>
<td>R² chl</td>
<td>15/09/11</td>
<td>0.59</td>
<td>0.52</td>
<td>0.86</td>
<td>0.81</td>
<td>0.54</td>
<td>0.69</td>
<td>0.73</td>
<td>0.53</td>
<td>0.87</td>
<td>0.84</td>
<td>0.60</td>
<td>0.63</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R² log chl</td>
<td>15/09/11</td>
<td>0.51</td>
<td>0.79</td>
<td>0.87</td>
<td>0.76</td>
<td>0.64</td>
<td>0.77</td>
<td>0.61</td>
<td>0.78</td>
<td>0.63</td>
<td>0.87</td>
<td>0.78</td>
<td>0.60</td>
<td>0.71</td>
<td>0.71</td>
<td>0.55</td>
<td></td>
</tr>
<tr>
<td>R² chl</td>
<td>29/05/12</td>
<td>0.68</td>
<td>0.65</td>
<td>0.85</td>
<td>0.79</td>
<td>0.65</td>
<td>0.78</td>
<td>0.79</td>
<td>0.72</td>
<td>0.76</td>
<td>0.71</td>
<td>0.78</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R² log chl</td>
<td>29/05/12</td>
<td>0.52</td>
<td>0.65</td>
<td>0.85</td>
<td>0.69</td>
<td>0.7</td>
<td>0.83</td>
<td>0.69</td>
<td>0.74</td>
<td>0.82</td>
<td>0.79</td>
<td>0.83</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 1: Correlation of chlorophyll-a
Logarithm but it brought a positive change in the combination of bands L1/L6 with a value of 0.83. In all these situations the Pearson coefficient squared was used as the tool for correlation.

4.2 Correlation of other parameters

The parameters in the above Table 2 were analyzed based on extracted digital numbers from the software used for processing data Landsat Thematic Mapper data (Grass GIS and Quantum GIS). It is to investigate if there is some consistency in the correlation of the parameters vis-à-vis certain bands and some band combinations. Band L6 which is thermal infrared (TIR) band L6 0.85. This impacted on the high correlation values that measure as its shows a positive correlation between temperature $T$ [°C] and the other parameters. In the combination of bands L1/L6 and its reverse combination L6/L1 had good correlation in respect of Total Chlorine (TC), Total Organic Carbon (TOC), Total Nitrogen (TN), Temperature (T) with values ranging between 0.57 to 0.93. The combination of these elements affects the growth algae in these water bodies which also directly affect the level of chlorophyll in them.

4.3 Model of chlorophyll-a

The model created was based on correlation between chlorophyll-a and the band combination L1/L2 and measurements from Sept-15. By optimizing parameters $a$ and $b$ of linear model

$$\log Chl-a = a \times L1/L2 + b \quad (1)$$

using solver in spreadsheet.

Resulting model has the parameters:

$$a = -9.060040$$

$$b = 14.343895$$

The model predicts the values of chlorophyll in the pixels corresponding to sampling points as it can be seen in Table 3, with RMSE = 11.29.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>PO$_4$ $^3-$ [mg/l]</td>
<td>$R^2$</td>
<td>0.76</td>
<td>0.79</td>
<td>0.72</td>
<td>0.77</td>
<td>0.89</td>
<td>0.73</td>
<td>0.56</td>
<td>0.82</td>
<td>0.73</td>
<td>0.55</td>
<td>0.68</td>
<td>0.53</td>
<td>0.69</td>
<td>0.39</td>
<td>0.71</td>
<td>0.54</td>
<td>0.8</td>
<td>0.54</td>
<td>0.86</td>
<td>0.81</td>
<td></td>
</tr>
<tr>
<td>$R^2 \log$</td>
<td>0.66</td>
<td>0.64</td>
<td>0.65</td>
<td>0.53</td>
<td>0.89</td>
<td>0.59</td>
<td>0.87</td>
<td>0.33</td>
<td>0.73</td>
<td>0.78</td>
<td>0.38</td>
<td>0.76</td>
<td>0.61</td>
<td>0.80</td>
<td>0.43</td>
<td>0.88</td>
<td>0.71</td>
<td>0.58</td>
<td>0.6</td>
<td>0.69</td>
<td>0.62</td>
<td></td>
</tr>
<tr>
<td>TC [mg/l]</td>
<td>$R^2$</td>
<td>0.96</td>
<td>0.57</td>
<td>0.93</td>
<td>0.57</td>
<td>0.61</td>
<td>0.58</td>
<td>0.93</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TOC [mg/l]</td>
<td>$R^2$</td>
<td>0.96</td>
<td>0.59</td>
<td>0.5</td>
<td>0.92</td>
<td>0.59</td>
<td>0.54</td>
<td>0.62</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TN [mg/l]</td>
<td>$R^2$</td>
<td>0.95</td>
<td>0.63</td>
<td>0.52</td>
<td>0.9</td>
<td>0.63</td>
<td>0.57</td>
<td>0.64</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$R^2 \log$</td>
<td>0.61</td>
<td>0.09</td>
<td>0.54</td>
<td>0.61</td>
<td>0.78</td>
<td>0.55</td>
<td>0.66</td>
<td>0.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$T$ [°C]</td>
<td>$R^2$</td>
<td>0.5</td>
<td>0.61</td>
<td>0.93</td>
<td>0.55</td>
<td>0.61</td>
<td>0.78</td>
<td>0.55</td>
<td>0.66</td>
<td>0.5</td>
<td>0.69</td>
<td>0.78</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$R^2 \log$</td>
<td>0.61</td>
<td>0.09</td>
<td>0.54</td>
<td>0.61</td>
<td>0.78</td>
<td>0.55</td>
<td>0.66</td>
<td>0.5</td>
<td>0.69</td>
<td>0.78</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SDD [cm]</td>
<td>$R^2$</td>
<td>0.53</td>
<td>0.55</td>
<td>0.89</td>
<td>0.6</td>
<td>0.58</td>
<td>0.77</td>
<td>0.61</td>
<td>0.61</td>
<td>0.59</td>
<td>0.78</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$R^2 \log$</td>
<td>0.65</td>
<td>0.88</td>
<td>0.94</td>
<td>0.59</td>
<td>0.76</td>
<td>0.72</td>
<td>0.85</td>
<td>0.82</td>
<td>0.75</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2: Other parameters correlation

The parameters in the above Table 2 were analyzed based on extracted digital numbers from the software used for processing data Landsat Thematic Mapper data (Grass GIS and Quantum GIS). It is to investigate if there is some consistency in the correlation of the parameters vis-à-vis certain bands and some band combinations. Band L6 which is thermal infrared (TIR) band L6 0.85. This impacted on the high correlation values that measure as its shows a positive correlation between temperature $T$ [°C] and the other parameters. In the combination of bands L1/L6 and its reverse combination L6/L1 had good correlation in respect of Total Chlorine (TC), Total Organic Carbon (TOC), Total Nitrogen (TN), Temperature (T) with values ranging between 0.57 to 0.93. The combination of these elements affects the growth algae in these water bodies which also directly affect the level of chlorophyll in them.

Based on this model a map of chlorophyll-a concentrations would be created using the model on the values of channels L1 and L2 in GIS. Result of such model can be seen on Fig. 2.

<table>
<thead>
<tr>
<th>chl-model</th>
<th>chlorophyll</th>
</tr>
</thead>
<tbody>
<tr>
<td>7,316841</td>
<td>5,698</td>
</tr>
<tr>
<td>4,398930</td>
<td>6,438</td>
</tr>
<tr>
<td>5,657294</td>
<td>2,96</td>
</tr>
<tr>
<td>185,5857024</td>
<td>145,04</td>
</tr>
<tr>
<td>20,857186</td>
<td>11,84</td>
</tr>
</tbody>
</table>

Table 3: Results of model
5 Conclusion

Employment of Landsat remote sensing data allows to create a map of chlorophyll concentrations based on model created with the help of results of measurement. The map created covers also water bodies which were not covered by the measurements. This way the model estimation allows to extend the monitoring on more water bodies. The other water parameters also had good correlation with respect to the Landsat data. But its possibly due to inter-correlation between chlorophyll-a and the other parameters as for example temperature, TN or PO₄³⁻ influences algae growth. It was also that there was strong correlation between thermal band and chlorophyll-a as well all the water parameters measured which possibly could be attributed to inter-correlation of the parameters and temperature.

6 Acknowledgements

This work would not be possible without support from University of Pardubice. We would like to thank also for Landsat data by National Aeronautics and Space Administration, made freely available through United States Geological Survey web service Earth Explorer

References: