WORKFLOW FOR CHLOROPHYLL CONCENTRATION MODELING USING MULTISPECTRAL SATELLITE DATA

Tomáš BRUNCLÍK

Institute of Environmental and Chemical Engineering, Faculty of Chemical Technology, University of Pardubice, Studentská 573, 53210, Pardubice, Czech Republic
Tomas.Brunclik@upce.cz

Abstract

This paper deals with work in progress to create usable workflow to assess chlorophyll-a concentration from remote sensing data to be used in lake and reservoir water quality monitoring to extend existing monitoring results. The workflow should be practicable with different software tools, and with various multispectral satellite data of suitable spatial resolution. The work is carried on with Landsat data, water quality data from reservoirs by Povodí Labe, a.s. and using Quantum GIS and GRASS GIS open source software. The workflow should use only the satellite data and commonly available geographic data.

Preliminary study showed very good correlation between raw Landsat data and chlorophyll-a concentration for good atmospheric conditions and larger reservoirs, but no agreement of model between various times and reservoirs. Workflows for three use cases are investigated. First use case leads to map of chlorophyll-a concentration for one reservoir and one time, second for multiple water bodies and one time. The third workflow scheme should use time-independent model of chlorophyll-a concentration. The first workflow scheme have been already verified. Details of the other two are being worked on.

Few scripts for preprocessing of Landsat data and for the DOS atmospheric correction were developed in the process of testing various steps of the workflow in GRASS GIS. These and possibly further scripts for other image-based atmospheric correction methods will be published under GPL license.

Keywords: Landsat, Chlorophyll-a, Reservoir, Modeling

1. INTRODUCTION

The subject of modeling chlorophyll-a concentration and water clarity in inland water using multispectral remote sensing data is widely covered in literature. Most articles focus on developing a model with good correlation between chlorophyll-a concentration and satellite data, usually based on chlorophyll-a concentration measurements in one relatively large lake, bay or reservoir (compared to average lakes and reservoirs in Czech Republic) and single sampling date, i.e. [1 2 3].

Proportionally less articles deal with regional scale studies, trying to extend the model over more water bodies and/or multiple sampling dates. An example might be study [4], dealing with modeling Secchi disk depth (SDD) parameter over approx. 12,000 lakes of Minnesota, or approach to retrieve chlorophyll-a concentration from airborne sensor data for 11 lakes in southern Finland [5].

This article deals with work in progress to create workflow for lake and reservoir chlorophyll-a modeling from satellite multispectral data usable in routine water quality monitoring. One of the objectives is also to verify if the methods used in above mentioned literature are applicable to relatively small reservoirs in the Czech Republic and to find optimal procedure. The work is carried on with Landsat data, water quality data from reservoirs by Povodí Labe, a.s. (east Bohemia region) and using Quantum GIS and GRASS GIS open source software, but the resulting workflow should be as software-independent as possible.

As the correlation used in creating a model is based on data from on-site sampling, the modeling of chlorophyll-a concentrations using remote sensing data could not replace the current monitoring methods, but could extend their results.

2. USE CASES

Results of conventional chlorophyll-a monitoring are concentrations on the sampling sites from various depths. One reservoir is sampled at 3 to 12 points depending on reservoir size .[6]

If there is suitable satellite image available, taken around the given date of conventional monitoring, there are three possibilities how use it to extend results of conventional chlorophyll-a monitoring:

1. It is possible to create detailed map of chlorophyll-a surface concentration for the reservoir monitored and date of samples collection. Currently, maps are constructed by the catchment area authorities based on the concentration measurements interpolation and knowledge of the water body constellation. The existing maps are constructed in 3D space and published on the web pages ,[6] but are spatially approximate and less detailed. Maps created using satellite data are more detailed and should be more accurate, although showing only surface concentrations. Map of this type was already created (Fig. 1) during the preliminary study mentioned in the next chapter [7]. It may be also possible to combine information from detailed surface maps and existing 3D maps to create more detailed 3D maps. This possibility should be yet investigated by experts in hydrology.

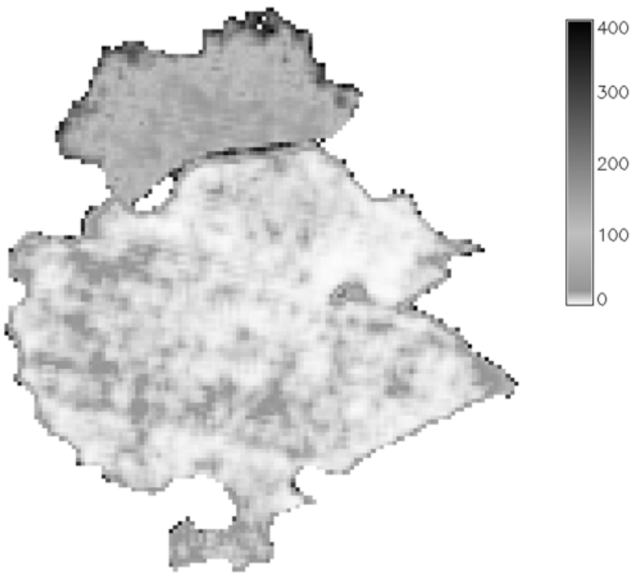

Fig. 1: Detail level of surface chlorophyll-a concentration map

2. It is also possible to create map mentioned in previous chapter for all water areas visible on the satellite image. As it involves applying model created on one or few reservoirs in the image to many different types of lakes, reservoirs and other water bodies, the results should be considered approximate, but they should give insight of the trophic state even of water bodies otherwise not monitored at all. This approach can be beneficial also for routinely monitored reservoirs, where it can add more information about changes in time. One reservoir or group of geographically near reservoirs in the area managed by Povodí Labe a.s. is currently typically monitored about every 50 days. As the monitoring takes place at different times for different reservoirs or groups, creating the map over the whole region of a satellite image should allow to add data more frequently for individual reservoirs or groups.
3. The last possibility is to create a time independent model of chlorophyll-a concentration for given set of reservoirs and satellite sensor. With such model, it would be possible to monitor reservoirs and lakes using satellite data independently on dates of conventional monitoring. Feasibility of this approach for the Czech Republic water bodies has to be yet investigated.

3. METHODOLOGY

3.1. Preliminary study

A preliminary study had been carried on to assess usability of Landsat Thematic Mapper (TM) data for chlorophyll-a concentration modeling [7]. This study, carried on with five on-site measurements of chlorophyll-a content from five reservoirs in eastern Bohemian part of Labe river basin, showed very good correlation between chlorophyll-a concentration and Landsat 5 data for the largest water body, reservoir Rozkoš. Every of the measurements contained results from 6 to 12 sampling points. Best correlation have been achieved with ratio of Landsat bands L1/L3 with the coefficient of determination R^2 = 0.93. The bands and band combinations found to have good correlation are from the range of Landsat TM bands 1-4 (Table 1), which corresponds with results commonly found by others [1 2 3 5].

Table 1: Best correlation of bands and band ratios with log of chlorophyll-a concentration, Rozkoš, 24-jul 2006

Landsat band or ratio	R^2
L1/L3	0.93
L3	0.82
L2/L3	0.82
L4	0.74

With other, smaller reservoirs the best coefficients of determination were ranging from 0.56 to 0.98, but they were for different bands or band combinations for every water body. R^2 for L1/L3 band combination had for other reservoirs and dates values 0.02, 0.12, 0.13, and 0.62. No single band or simple band combination with good coefficients of determination for all water bodies had been found.

3.2. The workflow schemes

Every use case mentioned in chapter 0 needs different workflow. Workflow developed in this work should be practicable with various software tools, and with various multispectral satellite data of suitable spatial resolution. The practical testing and development of methods is carried on in Quantum GIS software with GRASS GIS plug-in. The GRASS GIS software is not specialized remote sensing application, but it contains most basic tools needed in this field. The testing and development of methods is done with portability in mind, that is, whenever possible, the methods are carried on using basic tools like map calculator, unsupervised classifier etc. in parallel with specialized or more advanced tools which may be available only in particular software

The workflow schemes mentioned below are just basic steps, the merit of this work lies in details of individual steps and selection of methods with minimal data and software demands. The final results in case of success will be more detailed schemes with particular methods and algorithms, few of which will be mentioned in the next chapter.

The surface concentration map for one reservoir and one time is the simplest case. General workflow for this case, already tested in the preliminary study, consists of few steps:

1. Geometric correction of satellite data (optional, if needed).
2. Import of sampling point locations into GIS as map layer.
3. Reading values of all satellite layers on the sampling point locations into the sampling locations layer attribute table.

4. Examining correlation in a spreadsheet.
5. Using the best regression function to model the surface chlorophyll-a concentration in water.
    - To restrict the analysis to water areas, an existing map of water areas or satellite image classification can be used to create water-only image.

Similar workflow, with small variations, is used in other works [1 2 3]. While in the literature is often used atmospheric correction [1 3], in this work it is hypothesised that it is not necessary in this type of workflow. The step 1 is not needed with most Landsat images, since most of them are already geometrically corrected with satisfactory precision.

The second use case, constructing the water chlorophyll-a concentration map for the whole area of the satellite image can use similar workflow. If the satellite image is cloud-free and atmospheric conditions (especially haze) is uniform over the whole image, the workflow could be used without change. Unfortunately, absolutely cloud-free, uniform images are quite rare in the Czech Republic territory. If this is the case, the workflow would have additional steps:

1. Geometric correction of satellite data (optional, if needed).
2. Atmospheric correction or normalization (optional, if needed).
3. Masking or otherwise eliminating cloud areas (optional, if needed) and creating water-only image.
4. Import of sampling point locations into GIS as map layer.
5. Reading values of satellite bands on the sampling point locations into the sampling locations layer attribute table.
6. Examining correlation in a spreadsheet.
7. Using the best regression function to model the surface chlorophyll-a concentration in water.

The step 3 can be implemented by classification methods. In this step the water areas for step 7 could be also classified. Generally, various methods could be successfully used for relatively clear and uniform images. Problems to classify an image, not preprocessed by optional step 2, indicate that the image is not uniform in terms of atmospheric haze over the whole area, and step 2 is necessary.

The primary task of step 2 in this workflow is not to convert DN values to absolute reflectance, but to equalize DN values response to surface properties in different areas of the image to compensate spatial non-uniformity of atmospheric conditions, especially of atmospheric haze. This means all methods globally changing all pixels by the same offset or function would not work, while only relative (atmospheric normalization), but spatially sensitive method would be sufficient. The appropriate feasible method of atmospheric correction or normalization for this workflow is still searched for among several possibilities recommended in literature, i.e. [8 9 10 11].

Basic workflow for the last use case (time independent chlorophyll-a model) would be similar to the previous one, but the step 2 is crucial part, not optional one and step 6 would consist of examining data from several reservoirs and sampling times. It may also depend on the step of atmospheric correction or normalization, i.e. some atmospheric correction methods may need to eliminate clouds before the correction. As with previous use case the method of correction of atmospheric effects is still searched for. Unlike the previous use case, here would be of some use even methods assuming uniform atmospheric conditions in the study area.

3.3. Partial results

As this is work in progress, no final and detailed workflow has been created so far, but some steps are already semi-finished in more detail.

The first use case general workflow, which is simple and commonly used, had already been confirmed to work well for Labe reservoirs even without atmospheric correction. It had been proved in the preliminary study by constructing the map of chlorophyll-a concentrations and by good to very good levels of correlation between concentration and satellite data without the correction.

For the step of removing clouds and creating water-only image, several classification methods were tested, including the method of two step unsupervised classification recommended in [12]. Another quite successful method was supervised classification using GRASS built-in SMAP classifier .[13] Both methods worked well for most images, but for other failed – as mentioned earlier, problems with classification indicate inconsistency of atmospheric conditions in various parts of the satellite scene and should be solved by appropriate atmospheric correction. The two methods of classification will be used as alternatives in the final workflow.

In the part of finding correlation between chlorophyll concentrations and satellite bands was found, that best linear correlation for band digital number (DN) or simple band ratio indicated by square of Pearson coefficient R^2 is achieved with logarithm of chlorophyll-a concentration. Using logarithm of band DN, recommended in [1], does not improve the R^2. It was also tried, if preprocessing the satellite data by 3x3 pixels smoothing filter, as recommended in [1] , could improve the correlation. In the Table 2 are results for Landsat bands and reservoir Rozkoš.

Table 2: Influence of 3x3 smoothing on correlation.

Landsat band	L1	L2	L3	L4	L5	L7
R^2, raw DN	0,02	0,43	0,77	0,87	0,25	0,32
R^2, smooth 3x3	0,02	0,54	0,85	0,85	0,41	0,38

As can be seen, for the greatest reservoir Rozkoš there is some improvement of coefficient of determination in most cases, but for the smaller reservoirs the effect was mostly negative. In the final workflow there would be both variants and selection of the better one.

As mentioned earlier, model of chlorophyll-a concentration for reservoir Rozkoš agrees with those found in other works, while other, smaller reservoirs exhibit poor correlation in combinations of bands 1-3 and better correlations in combinations containing bands 5 and 7. Quick test with the 6 processed monitoring samplings from the preliminary study shows strong dependence of correlation of chlorophyll-a - band L3 on range of chlorophyll concentrations in the reservoir (Fig. 2). There was not so strong, but still significant, dependence on reservoir size (R^2=0,57).

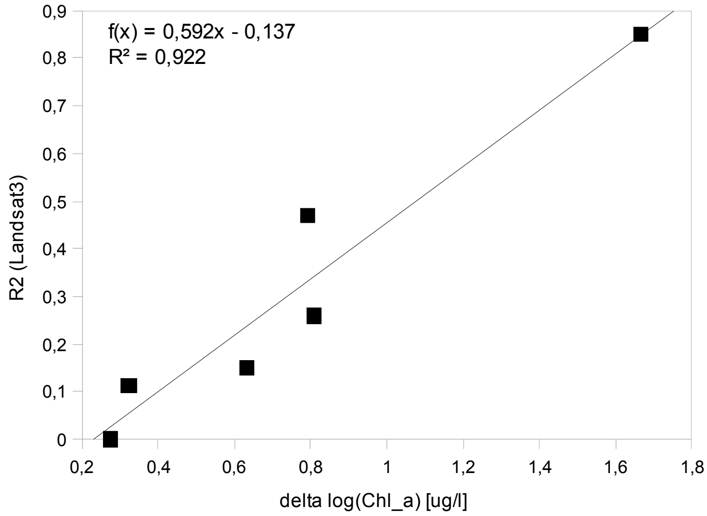

Fig. 2: Dependence of coefficient of determination of model on range of chlorophyll-a concentrations found in the reservoir.

These results are currently in the process of verification using data from more reservoirs and sampling times. If the dependence of correlation between various band combinations and reservoir size, depth or other parameter is confirmed there is possibility to incorporate such dependence into an universal model needed for the second and especially for the third use case workflow.

3.4. GRASS scripts

In the process of testing methods of the workflow in the GRASS GIS software, several GRASS scripts were developed as a by-product. Some of these could be generally useable and it is planned to release them under General Public License.

The script i.slcoff.fill fills gaps in the Landsat 7 ETM+ bands caused by Scan Line Corrector (SLC) failure [14] using interpolation. This script is in fact just modified r.fillnulls script which is part of GRASS GIS. The original script was not able to correct Landsat SLC-off bands.

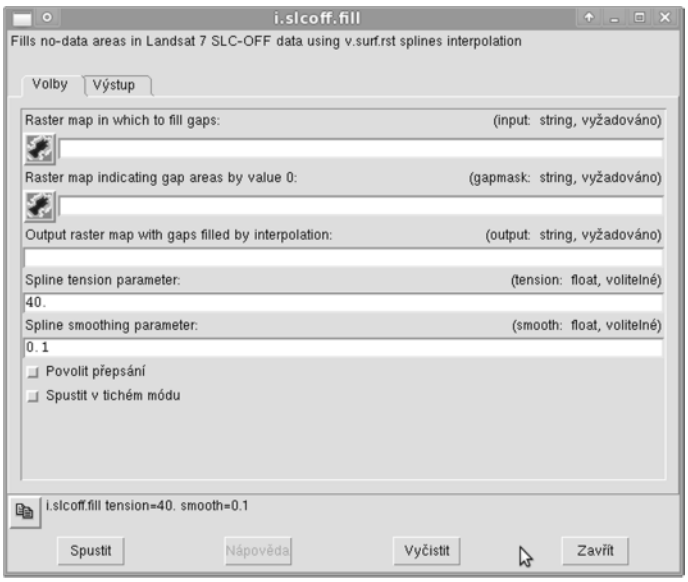

Fig. 3: User interface of the i.slcoff.fill script.

The script i.atcorr.dos is intended for quick atmospheric correction of all bands of a scene by well known dark object subtraction (DOS) method. It implements optional moving window averaging of the image before searching for the dark pixel – this approach should eliminate influence of image noise or dead pixels on the results.

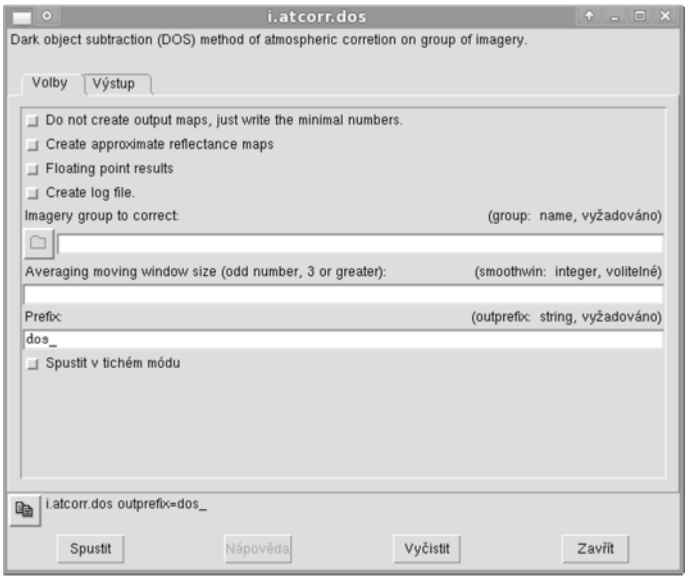

Fig. 4: User interface of the i.atcorr.dos script.

4. CONCLUSION

Three types of workflows are investigated. The workflow leading to map of chlorophyll-a concentration for single lake and time was confirmed to be feasible for reservoirs in the Czech Republic. Details of other workflows are being worked on, so far two map classification techniques are confirmed to be suitable and satellite image smoothing filter influence tested. Influence of water body parameters on the chlorophyll-a concentration model is being examined, with promising preliminary results. These have to be studied further on a broader data set. As a by-product of testing of the workflow steps in GRASS GIS, two possibly generally useable scripts for this software were created.

5. ACKNOWLEDGEMENTS

This work was financially supported by the Ministry of Education, Youth and Sports of the Czech Republic by the project MSM 0021627502. This work would not be possible without water monitoring data kindly provided by Povodí Labe, a.s. and Landsat Thematic Mapper data by National Aeronautics and Space Administration, made freely available through United States Geological Survey web service EarthExplorer.

6. REFERENCES

1. Luoheng Han, Jordan K..(2005) Estimating and mapping chlorophyll-a concentration in Pensacola Bay, Florida using Landsat ETM+ data. International Journal of Remote Sensing. 26 23, 5245-5254.

2. Duan H, Zhang Y, Zhang B, Song K, Wang Z. (2006) Assessment of Chlorophyll-a Concentration and Trophic State for Lake Chagan Using Landsat TM and Field Spectral Data. Environ Monit Assess. 129 1-3, 295-308.

3. Matthews MW, Bernard S, Winter K. (2010) Remote sensing of cyanobacteria-dominant algal blooms and water quality parameters in Zeekoevlei, a small hypertrophic lake, using MERIS. *Remote Sensing of Environment*. 114 9, 2070-2087.

4. Olmanson LG, Bauer ME, Brezonik PL. (2002) Use of Landsat imagery to develop a water quality atlas of Minnesota's 10,000 lakes. In: Pecora 15/Land Satellite Information IV/ISPRS Commission I/FIEOS 2002 Conference Proceedings.Vol XXXIV. Denver, CO, USA: ISPRS, 10-15 November 2002, pp. 8. Available at: http://www.isprs.org/proceedings/XXXIV/part1/paper/00102.pdf.

5. Pulliainen J, Kallio K, Eloheimo K, et al. (2001) A semi-operative approach to lake water quality retrieval from remote sensing data. The Science of The Total Environment. 268 1-3, 79-93.

6. Povodí Labe. Water quality in reservoirs 4.2. http://www.pla.cz/portal/jvn/en/index.htm [Accessed October 31, 2010].

7. Šimek J. (2010) Využití družicových multispektrálních snímků k hodnocení obsahu chlorofylu ve vodních nádržích. University of Pardubice, Pardubice.

8. Song C, Woodcock CE, Seto KC, Lenney MP, Macomber SA. (2001) Classification and Change Detection Using Landsat TM Data: When and How to Correct Atmospheric Effects? Remote Sensing of Environment. 75 2, 230-244.

9. Guanter L, et al. (2010) Atmospheric correction of ENVISAT/MERIS data over inland waters: Validation for European lakes. Remote Sensing of Environment. 114 3, 467-480.

10. Canty MJ, Nielsen AA, Schmidt M. (2004) Automatic radiometric normalization of multitemporal satellite imagery. Remote Sensing of Environment. 91 3-4, 441–451.

11. Yuan D, Elvidge C. (1996) Comparison of relative radiometric normalization techniques. *ISPRS Journal of Photogrammetry and Remote Sensing*. 51 3, 117-126.

12. Olmanson LG, Kloiber SM, Bauer ME, Brezonik PL. (2001) Image processing protocol for regional assessments of lake water quality. Water Resources Centre and Remote Sensing Laboratory, University of Minnesota.

13. Image processing in GRASS GIS. http://grass.fbk.eu/gdp/html_grass64/imageryintro.html [Accessed November 1, 2010].

14. Williams DL, Goward S, Arvidson T. (2006) Landsat: yesterday, today, and tomorrow. Photogrammetric engineering and remote sensing. 72 10, 1171.