MODELLING SPATIAL DISTRIBUTION OF GLOBAL TOTAL COLUMN OZONE IN QGIS AND GRASS GIS ENVIRONMENT

E. Tcherkezova, P. Kaleyna, Pl. Mukhtarov

National Institute of Geophysics, Geodesy and Geography, str. Acad. G. Bonchev, bl 3, Sofia1113, Bulgaria, e-mail: et@geophys.bas.bg, pkaleyna@geophys.bas.bg, pmuhtarov@geophys.bas.bg

Abstract. Total Column Ozone (TCO) has an impact on the Earth's atmosphere. In this article the measurements are used for modelling the spatial distribution of the TCO in the atmosphere to observe the values of TCO and the state of the ozone layer as a whole. Spatial interpolation allows an extrapolation of large TCO point data sets to a larger area of interest and is an important method for such purpose. This article focuses on modelling spatial distribution of global TCO using two spatial interpolation methods applying the free open source software Quantum GIS (QGIS) and GRASS GIS. The priority is given to elaborate the workflow showing how to apply Quantum GIS and GRASS GIS and how to implement some spatial interpolation methods making a contribution to the use of geographic information systems (GIS) in modelling spatial distribution of the TCO in Earth's atmosphere. The main goal is, therefore, to demonstrate different means for integration of raw TCO data into GIS and modelling its spatial distribution using free and open source GIS software. The results show that QGIS and GRASS GIS are appropriate tools for modelling spatial distribution of TCO. There are two main advantages of this use: no-cost for software and GIS application that allow also more complex spatial analysis of TCO in GIS environment using additional geographic data and analytical tools for spatial analysis.

Key words: Total Column Ozone (TCO), spatial interpolation, Free and Open Source Software (FOSS), Quantum GIS (QGIS), GRASS GIS.

Introduction

The ozone is one of the so-called "small components" of the Earth's atmosphere (Schwartz et al. 2005). It represents allotropic form of oxygen gas with a triatomic molecule containing three oxygen atoms. In the atmosphere it is presented through very small concentrations, in total, the ozone makes up only 0.6 parts per million of the atmosphere. Despite these low concentrations, the ozone plays a vital role in the existence and development of the Earth's biosphere, because it has the ability to absorb this amount of solar ultraviolet radiation that penetrates the Earth's stratosphere and prevents the ultraviolet radiation to reach dangerous quantity for living creatures. Absorption of solar energy, on the other hand, has an impact on temperature conditions of the stratosphere and its dynamics.

The part of the ultraviolet radiation of the Sun, which is absorbed by the ozone, depends on TCO (Total Column Ozone): a measurement of the total amount of atmospheric ozone in a given column. TCO is measured in Dobson Units (DU). One Dobson unit refers to a layer of gas that would be 10 μm thick under standard temperature and pressure, sometimes referred to as a 'milli-atmo-centimeter' (Schwartz et al. 2005). A baseline value of 220 DU is chosen as the starting point for an ozone hole since total ozone values of less than 220 Dobson Units were not found in the historic observations over Antarctica prior to 1979. Also, from direct measurements over Antarctica, a column ozone level of less than 220 Dobson units is a result of the ozone loss from chlorine and bromine compounds (Ozone Hole Watch 2007). The Dobson unit is named after Gordon Dobson, who was a researcher at the University of Oxford. In the 1920s, he built the first instrument to measure total ozone from the ground, now called the Dobson ozone spectrophotometer. The modern measurements of TCO are made through satellite equipment (Ozone Monitoring Instrument (OMI) Data User's Guide, 2008).

In this article TCO data is used from the Ozone Monitoring Instrument (OMI). OMI is a nadir-viewing near-UV/Visible CCD spectrometer aboard NASA's Earth Observing System's (EOS) Aura satellite. OMI measurements cover a spectral region of 264-504 nm (nanometers) with a spectral resolution between 0.42 nm and 0.63 nm and a nominal ground footprint of 13 x 24 km2 at nadir. The OMI instrument observes Earth's back scattered radiation with a wide-field telescope feeding two imaging grating spectrometers (Ozone Monitoring Instrument (OMI) Data User's Guide, 2008).

Essentially complete global coverage of TCO is achieved in one day (OMI Ozone). Those data is available as large data sets in tabular form (e. g. in delimited text format) which contain geographic coordinates (latitude and longitude) and TCO-value at each measured sample point. This format allows modelling spatial distribution of TCO using different methods of interpolation and software.

The main purpose of this work is to perform spatial interpolation methods for modeling global TCO using Free and Open Source Software (FOSS). The priority is given to elaboration of workflow showing in detail how to apply QGIS and GRASS GIS making a contribution to the use of geographic information systems (GIS) in modelling spatial distribution of global TCO. It includes processing precise tasks like data input into QGIS and GRASS GIS, spatial interpolation of TCO data, and 2D and 3D visualization.

Data

TCO data from OMI is available on the website http://ozoneaq.gsfc.nasa.gov/ in tabular form. OMI daily global data set is reproduced by using equal angle grid 1 degree × 1 degree cover the whole globe. The data for portion of the ground in many cases is missing due to technical reasons, which demands, when we need an overall picture of the distribution of TCO, to resort to interpolation. In this work is shown the global distribution of TCO on 21.01.2009 during Stratospheric sudden warming. Stratospheric sudden warming (SSW) is one of the most prominent phenomena in the middle atmosphere. The selected day coincides with the maximum of SSW (Jin et al., 2012). This is a large-scale stratospheric anomaly that manifests itself in a sudden increase in stratospheric temperature in the North polar region with dozens of degrees and change the direction of the stratospheric zonal wind. During such anomalies is observed an uneven distribution of total ozone at high latitudes, which is evident in the figures in this article.

Additionally, we have used the free geographical data sets for land borders and for country boundaries on land and offshore for maps composition from the public domain map dataset 'Natural Earth' (Source: http://www.naturalearthdata.com/).

Software

Quantum GIS and GRASS GIS are free open source desktop GIS software projects. QGIS project is established as a project on SourceForge in June 2002 (QGIS Development Team 2004-2013). It is developed as separate projects for GNU/Linux, Windows, Mac OSX, and Androids. The current stable version of Quantum GIS on Windows is 2.0.1.

During the period 1982-1995 GRASS GIS (Geographic Resources Analysis Support System) was developed as software by the U.S. Army Construction Engineering Research Laboratories (CERL) to support land management in military installations. In 1999 the new development team of GRASS GIS has adopted the GNU GPL license (Blackwell Publishing Ltd. 2004).

Currently GRASS GIS is open source project that is permanent developed for GNU/Linux, Windows, Mac OSX and sponsored by numerous sites worldwide (e. g. Centre for Scientific and Technological Research, in Trento, Italy, http://grass.itc.it). The current stable version of GRASS GIS on Windows is 6.4.x. GRASS GIS is also integrated as a tool in QGIS. GRASS GIS is a hybrid geographic information system for vector, raster, image analysis, modelling and 2D and 3D-visualization.

QGIS and GRASS GIS were released under the GNU General Public License (GPL). This license means "...that users have the freedom to distribute copies of free software (and charge for this service if they wish), that users receive source code or can get, that users can change the software or use pieces of it in new free programs; and that users know they can do these things." (Free Software Foundation, Inc. 1991).

QGIS and GRASS GIS provide a powerful collection of tools for the management, spatial data analysis and visualization. Therefore GRASS GIS and QGIS are used for different applications around the world by academic, governmental agencies and commercial institutions.

In this work we have used the capabilities of those software to demonstrate how raw TCO data can be imported into QGIS and GRASS GIS and how can be modelled in order to create spatial surfaces of global TCO.

Methods

According to Burrough (1986) the values of properties at unsampled locations on the set of observed values at known locations can be estimated through spatial interpolation. Interpolation is therefore, process of estimation unknown values that are located between surrounding known values across an area and is used to create a surface of models a sampled phenomenon. Spatial interpolation procedures are reviewed by different authors, e. g. Burrough (1986), Burrough and McDonnell (2000), Mitasova and Hofierka (1993), Mitasova et al. (1995), Mitas and Mitasova (1999), Cressie (1993). They can be grouped as follow:

- local neighborhood approach (Inverse Distance Weighted interpolation - IDW, Natural Neighbor interpolation, interpolation based on a Triangulated Irregular Network (TIN),
- geostatistics approach (e.g. Kriging),
- variational approach to interpolation and approximation (Thin Plate Spline - TPS function, Regularized Spline with Tension (RST), and other forms of smoothness semi-norm) (Mitas and Mitasova 1999).

According to Chiles and Delfiner (1999) the quality of any analysis depends on interpolation of observed data and is subject to a degree of uncertainty. Different interpolation techniques that are performed on the same data sets can therefore generate different spatial predictions at same locations.

Assessment of error (uncertainty) or comparison of some interpolation methods were presented by different authors and applications, e. g. MacEachren and Davidson (1987), Mitas and Mitasova (1999), Siska and Hung (2005), Collins and Bolstad (1996), Hůnová et al. (2012) and others. According to MacEachren and Davidson (1987) data measurement accuracy, data density, data distribution, and spatial variability had impact on interpolation accuracy.

Siska and Hung (2005) concluded that IDW, Kriging, Thiessen polygons and TIN interpolations performed almost on the same level, but TIN appears to be a leading method in predicting the unknown values on a more uniform, less varied data set (flat surface).

The results of Collins and Bolstad (1996) outline that certain a priori data characteristics (in their research they are temperature range, temperature variance and temperature correlation with elevation), spatial scale, relative spatial density and distribution of sampling locations impact on interpolation and affect the choice of spatial interpolation technique.

The results of Hůnová et al. (2012) confirm the importance of assessment of performance of different interpolation methods.

Generally, can be concluded that the spatial interpolation accuracy is crucial in accurate estimation of different type of sampled data such meteorological, soil, terrain, bathymetry and others and should be took into account in spatial modelling and analysis.

Spatial interpolation procedures that were used in this article are the Inverse Distance Weighted interpolation (IDW) and Regionalized Spline Interpolation with Tension (RST) without focus on estimating the interpolation accuracy, but on methodological approach for spatial modeling in GIS environment.

The IDW interpolation is a local and exact deterministic interpolation technique. The value in an unsampled location is computed as weighted average considering the values of the sample points and the distance between then and the estimated cell. According to Burrough (1986) weights are inversely proportional to a power of distance. That means that points closer to the evaluated cell are more weighted than sample points that are further away. The result is a surface that does not pass through the data points. IDW produces local maxima and minima at the sample points (Mitas and Mitasova 1999).

Spline methods belong to the variational approach for interpolation and approximation (Mitas and Mitasova 1999). They are based on the assumption that the sum of deviations from the sample points should be minimized in order to pass exactly or closely as possible to them and smoothing surface at the same time (Mitasova et al. 1995; Mitas and Mitasova 1999).

The Regionalized Spline Interpolation with Tension (RST) method used in this article incorporates the properties of Thin Plate Spline (TPS) with tension and function with regular second- and possibly higher-order derivatives (Mitas and Mitasova 1988, 1999). The mathematical functions of TPS with tension and RST are presented and discussed in detail by Mitas and Mitasova (1988), Mitasova and Mitas (1993), Mitas and Mitasova (1999). According to Mitasova et al. (1995) and Hofierka et al. (2002) the RST method enables smoothing of noisy data. This interpolation method is implemented in GRASS GIS (Neteler and Mitasova 2004).

Interpolation methods are integrated at most GIS software packages and are used in different applications. For example, interpolation methods are often used in climatology (e.g. Hutchinson 1995; Hofierka et al. 2002; Hancock and Hutchinson 2006; Luo et al. 2008), geomorphology and geomorphometry (e.g. Mitasova and Mitas 1993; Mitasova and Hofierka 1993; Mitasova et al. 2004, Mitasova et al. 2005 and others), and others.

Results

The spatial analysis is the most important functionality of geographic information systems (GIS). Including other GIS capabilities such spatial data processing, storage in geographical database and visualization, GIS is a powerful tool that is widely used currently as technology and method in different applications. The choice of methodological approach in each GIS application depends on the specific goal. Therefore the selection or the elaboration of appropriate work processes has an important role in each GIS project (Tcherkezova 2004).

In this article the following methodological approaches for processing and visualization of global TCO data in QGIS and GRASS GIS environment were demonstrated:

- TCO data processing and visualization using QGIS,
- TCO data processing and visualization using GRASS GIS, and
- TCO data processing and visualization using GRASS GIS as a QGIS tool.

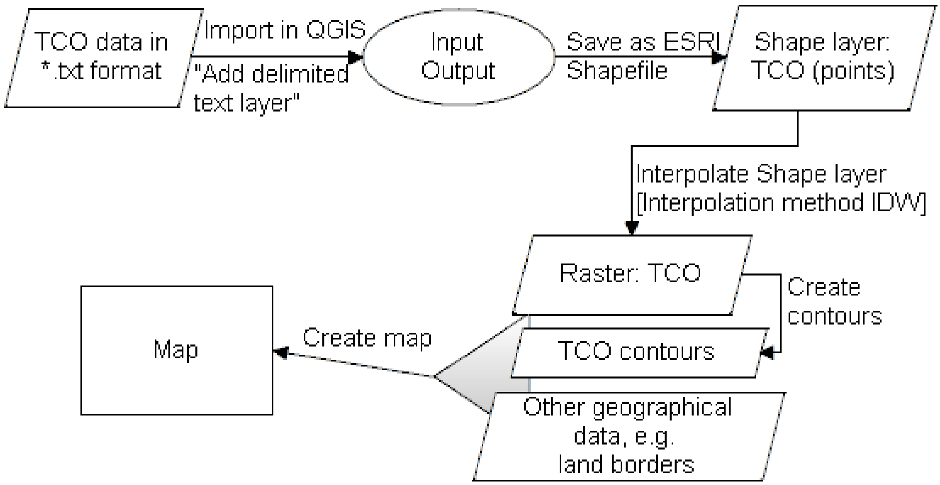

Fig. 1. Flowchart of raw global TCO data processing and spatial modelling in QGIS environment

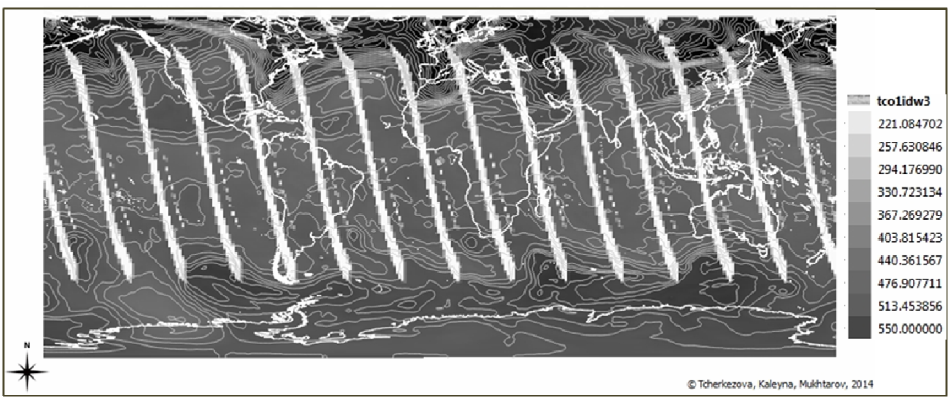

Fig. 2. Global TCO spatial distribution based on IDW interpolation method overlapped with some main TCO contours and land boundaries

Figure 1 shows in detail how to import raw TCO daily data with ‘no data’ values as delimited text into QGIS and how to generate surface of global spatial distribution using IDW method that is implemented in QGIS. The first step was to open the raw data in Notepad++ a free and source code editor for Windows (http://www.notepad-plus-plus.org/) in order to proof the data format and if necessary to convert it in "delimited text format".

The result is presented in Figure 2 and outlines that the ‘no data’ values are not interpolated as zero values.

In the Figure 3 were presented three possible work steps for integration of the raw TCO data into GRASS GIS. In the first case the attribute table was transformed in DBF format and imported using the GRASS GIS command db.in.org (File → Import database table → Multiple import formats using OGR). It includes three fields: longitude, latitude and TCO. The imported table was used to create a vector point layer in GRASS GIS environment (v.in.db , Vector → Generate points → Generate from database). The created vector point layer was used then to interpolate surfaces that models the global TCO spatial distribution using IDW and RST interpolation methods (Raster - Interpolate Surfaces - IDW from vector points whereby the TCO values in the attribute table of the vector layer should be selected as Z value field) and command v.surf.rst (Raster - Interpolate Surfaces - Regularized Spline Tension whereby the TCO values in the attribute table of the vector layer should be also selected as Z value field).

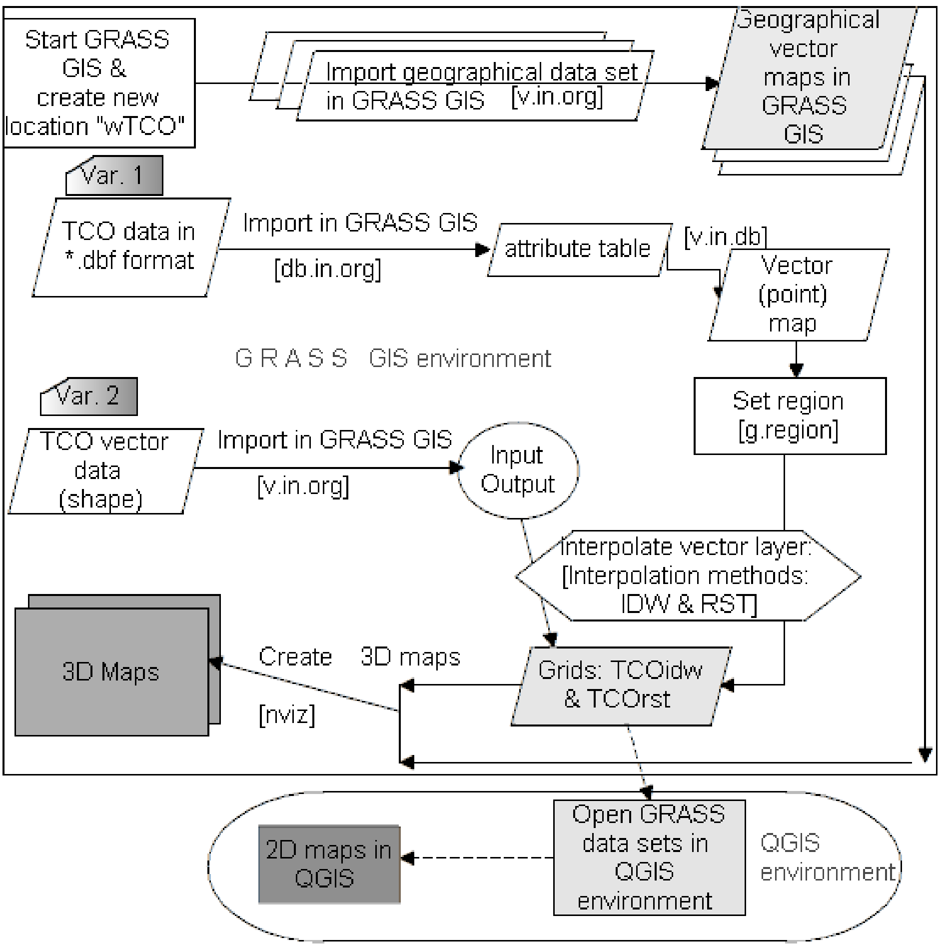

Fig. 3. Flowchart of raw global TCO data processing and spatial modeling in GRASS GIS environment and visualization

The layers that are stored in GRASS GIS geo-database can be opened, analysed and visualized also in QGIS environment (Figure 3 in the ellipse). This is another way to manipulate the geodata combining QGIS and GRASS GIS.

The result of IDW interpolation and created contours was presented as 3D map using the NVIZ tool for visualization and animation of GRASS data (Figure 4). In the interpolation IDW and RST procedures in GRASS GIS environment the 'no data' values were ignored.

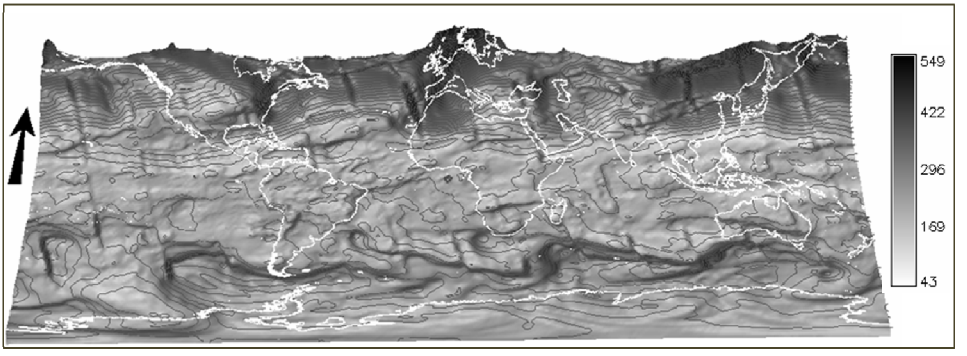

Fig. 4. Spatial model of global TCO which is generated using IDW method in GRASS GIS environment

The three dimensional visualization helps for better understanding the spatial distribution of the Total Column Ozone. We have visualized the interpolated IDW surface of TCO spatial distribution using 'invert grey color' legend (Figure 4).

The figure below shows the result of spatial modelling of Total Column Ozone using Regularized Spline Tension method ignoring 'no data' values (Figure 5). In order to highlight better the differences by close contrast values, the legend was colored in grey using histogram equalization. The raster image was overlapped with land boundaries and created TCO contours.

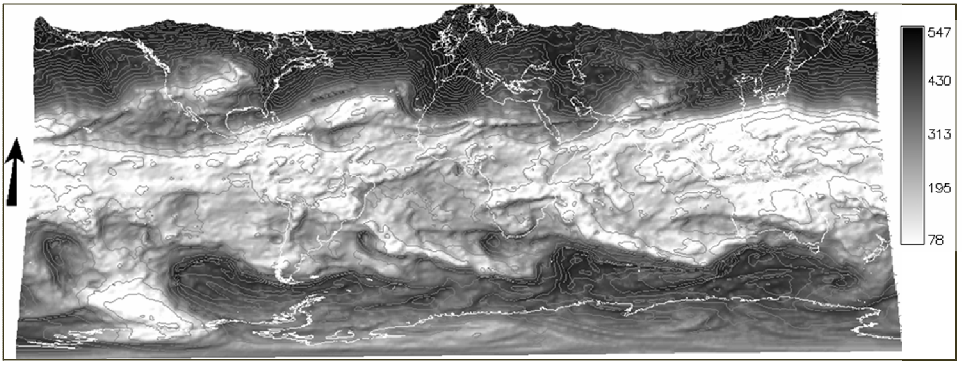

Fig 5. Spatial model of global TCO which is generated using RST method in GRASS GIS environment

Although the results of IDW and RST are visually similar, but there are differences which can be presented trough visual comparison of IDW and RST contours (Figure 6) or through analysis of errors.

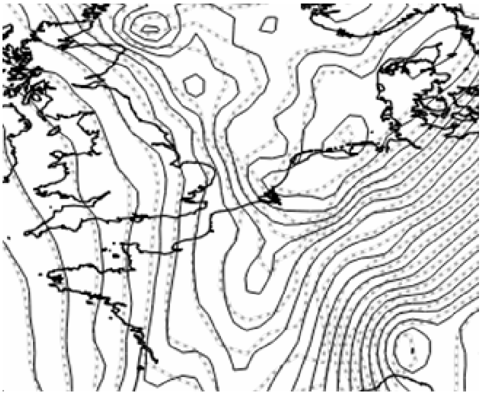

Fig. 6. Differences between IDW and RST interpolation results (black lines: contours of RST interpolation; grey point line style: contours of IDW interpolation)

The results present that the spatial distribution of global TCO can be modelled using different methodological approaches. The import of raw TCO data in table format in QGIS is more user-friendly as their import in GRASS GIS, but the GRASS GIS offers more interpolation methods than QGIS and an easy of access 3D visualization. Additionally, the functionality of GRASS GIS for creation of an own GRASS geodatabase in PERMANENT and users map sets allows using the stored data in QGIS. The stored data in the PERMANENT map set can't be changed or damaged from additional users (Dassau et al. 2004-2005).

For this reasons the mostly appropriate methodological approach includes the following steps:

- Import of raw TCO data in table format in QGIS.
- Save the imported data as vector data (points).
- Import of TCO point vector layer in GRASS GIS.
- Interpolate surface using different interpolation methods in GRASS GIS environment.
- 3D visualization of the results or mapping TCO in QGIS environment.

The results demonstrate that the QGIS and GRASS GIS are appropriate for the article's purposes and offer a realistic alternative to commercial GIS and to other software. The advantages of this use are the following:

- no-cost for software and
- GIS application that allows also more complex spatial analysis of TCO in GIS environment using additional geographic data and analytical tools for spatial analysis.

Conclusions

In this article were demonstrated different methodological approaches for integration, manipulation and interpolation of raw global TCO data using free and open source QGIS and GRASS GIS.

The results show how the global TCO can be modelled in GIS environment. The elaborated flowcharts present in detail the workflow for import, processing and visualization applying QGIS, GRASS GIS and GRASS GIS as a QGIS tool.

The presented methodological approaches can be adopted for modelling of TCO and other meteorological phenomena such as air pollution at regional scale.

This work also link interdisciplinary expertise to help investigation and spatial modelling of TCO at more advanced level which results can be used directly in other studies or for more complex spatial analysis and to help the understanding variation of TCO. Additionally, QGIS and GRASS GIS effort high degree of customisation and users can develop scripts for other interpolation methods.

Acknowledgment. In this work we have used only Free and Open Source Software (QGIS, GRASS GIS and Notepad++) and free available data (OMI Ozone data and Natural Earth land data sets) those links are presented in the references.

References

Burrough, P.A. (1986). Principles of Geographical Information Systems for Land Resources Assessment. Clarendon Press. ISBN 0198545924. 194 p.

Burrough, P.A. and McDonnell, R. (2000). Principles of Geographical Information Systems. Oxford University Press. 333p. Chiles, J.P. and P. Delfiner, 1999. Geostatististics and Modelling Spatial Uncertainty. 1st Edition, Wiley-Interscience, USA. ISBN-13: 978-0471083153. 720p.

Collins, F.C. and Bolstad, P.V. (1996). A Comparison of Spatial Interpolation Techniques in Temperature Estimation. In Proceedings of the Third International Conference/Workshop on Integrating GIS and Environmental Modeling. National Center for Geographic Information Analysis (NCGIA): Santa Fe, NM, Santa Barbara, CA: January 21–25. Available:

Cressie, Noel. A.C. (1993) Statistics for Spatial Dat. John Wiley & Sons, Inc., New York, [revised edition], 900p.

Dassau, O., Holl, S., Neteler, M., and Redslob, M. (2004-2005). Eine Einfürung in den praktischen Umgang mit dem Freien Geographischen Informationssystem GRASS 6.0. Version 1.2. GDF Hannover bR. 157.

Hancock, P.A. and Hutchinson, M.F. (2006). Spatial interpolation of large climate data sets using bivariate thin plate smoothing splines. Environmental Modelling and Software, 21(12). 1684-1694.

Hofierka, J., Parajka, J., Mitasova, H., and Mitas, L.(2002). Multivariate Interpolation of Precipitation Using Regularized Spline with Tension. - Transactions in GIS, 6(2): 135-150.

Hůnová I., Horálek, J., Schreiberová, M., and Zapletal, M. (2012). Ambient Ozone Exposure in Czech Forests: A GIS-Based Approach to Spatial Distribution Assessment. - The ScientificWorld Journal ,Volume 2012, Article ID 123760, 10 pages. Doi:10.1100/2012/123760.

Hutchinson, M.F. (1995). Interpolating Mean Rainfall Using Thin Plate Smoothing Splines. International Journal of Geographical Information Systems, 9(4). 385-403.

Jin, H. , Miyoshi, Y., Pancheva, D., Mukhtarov, P., Fujiwara, H. and Shinagawa, H. (2012) Response of migrating tides to the stratospheric sudden warming in 2009 and their effects on the ionosphere studied by a whole atmosphere-ionosphere model GAIA with COSMIC and

TIMED/SABER observations, Journal of Geophysical Research, VOL. 117, A10323, doi:10.1029/2012JA017650

Luo, W., Taylor, M.C., Parker, S.R. (2008). A comparison of spatial interpolation methods to estimate continuous wind speed surfaces using irregularly distributed data from England and Wales. Int. J. Climatol., 28. 947-959. Doi:101002/joc.1583.

MacEachren, A.M., J.V. Davidson (1987). Sampling and Isometric Mapping of Continuous Geographic Surfaces. - The American Cartographer, Vol. 14, No. 4, 299-320.

Mitas, L., Mitasova, H. (1988). General variational approach to the interpolation problem. - Comput. Math. Application, Vol. 16, No. 12, Pergamon Press, 983-992.

Mitas, L., Mitasova, H. (1999). Spatial Interpolation. In Longley, P., Goodchild, M.F., Maguire, D.J., Rhind, D.W. (Eds.), Geographical Information Systems: Principles, Techniques, Management and Applications, GeoInformation International, Wiley, 481-492.

Mitasova, H. and Hofierka, J. 1993. Interpolation by Regularized Spline with Tension: II. Application to Terrain Modelling and Surface Geometry Analysis. - Mathematical Geology, Vol. 25, No. 6. 657-669.

Mitasova, H., Drake, T.G., Harmon, R.S. and Bernstein, D. (2004). Quantifying rapid changes in coastal topography using modern mapping techniques and GIS. Env. and Engineering Geoscience, vol. 10. 1-11.

Mitasova, H., Mitas, L. (1993). Interpolation by regularised spline with tension: I. Theory and implementation. Mathematical Geology 25: 641–655.

Mitasova, H., Mitas, L., Brown, W.M., Gerdes, D.P., Kosinovsky. I., Baker. T. (1995) Modelling spatially and temporally distributed phenomena: new methods and tools for GRASS GIS. –

Mitasova, H., Overton, M., Harmon, R.S. (2005). Geospatial analysis of a coastal sand dune field evolution: Jockey’s Ridge, North Carolina. Geomorphology 72.204–221. Available: www.sciencedirect.com. Last visit: 9th Dec. 2013.

Neteler, M. and Mitasova, H. (2004). Open Source GIS: A GRASS GIS Approach, 2nd Ed. Kluwer International Series in Engineering and Computer Science 773, Springer/Kluwer Academic Press,NY/Boston. ISBN: 1-4020-8064-6. 424p.

Schwartz, S. E. and Warneck, P. (1995). Units for use in atmospheric chemistry. Pure Appl. Chem. 67 (8-9): 1377–1406. doi:10.1351/pac199567081377.

Siska, P. P., Hung, I. K. (2005). Assessment of kriging accuracy in the GIS environment. Available: http://gis.esri.com/library/userconf/proc01/professional/papers/pap280/p280.htm. Last visit: 17th Jan 2014.

Tcherkezova, E. 2004. (2004). GIS-gestützte geomorphologische Untersuchungen an Beispielen aus dem nördlichen Teil der Bucht von Burgas (Schwarzmeerküste, Bulgarien). – Berliner Geographische Studien. Band 52. D 83. Berlin. 191 Seiten mit CD-Rom. ISBN 3 7983 1945 6. 191.

Data sources

Natural Earth: http://www.naturalearthdata.com/. Last visit: 15th Aug 2013.
OMI Ozone: http://ozoneaq.gsfc.nasa.gov/qa/faq_omi.md#g_1. Last visit: 25th Apr 2013

Моделиране пространственото разпределение на глобалното съдържание на озон чрез QGIS и GRASS GIS

Е. Черкезова, П. Калейна, Пл. Мухтаров

Резюме. Тоталното съдържание на озон (TCO) оказва влияние върху атмосферата на Земята. В тази статия измерванията на TCO са използвани за моделиране на пространственото разпределение на TCO в атмосферата, с цел да се наблюдават стойностите на TCO и състоянието на озоновия слой като цяло. Пространствена интерполация позволява екстраполация на големи бази точкови данни за TCO за голями площи и е важен метод за тази цел. Тази статия се фокусира върху моделиране на глобалните данни за TCO, използвайки два метода за пространствена интерполация чрез безплатния отворен софтуера Quantum GIS (QGIS) и GRASS GIS. Основния приоритет на статията е да се представят фигури-таблици, показващи как да се прилагат Quantum GIS и GRASS GIS, как да се прилагат някои методи за пространствена интерполация при използването на географските информационни системи (GIS) в моделирането на пространственото разпределение на TCO в земната атмосфера. Основната цел следователно е, да се покажат различни средства за интегриране на необработените данни за TCO в GIS и моделиране на неговото пространствено разпределение, използвайки безплатения софтуер с отворен код GIS. Резултатите показват, че QGIS и GRASS GIS са подходящи инструменти за моделиране на пространствено разпределение на TCO. Има две основни предимства при използването на GIS: безплатен софтуер и GIS приложение, което позволява още по-сложен пространствен анализ на TCO в GIS среда, с помощта на допълнителни инструменти за пространствен анализ на географски и аналитични данни.