Training students in Python programming skills and WPS wrapping for geoprocessing tasks by using examples of less commonly applied thematic mapping methods

Christian Stern\textsuperscript{a} (corresponding author) and Gertrud Schaab\textsuperscript{a}

\texttt{christian.stern@h-ka.de, gertrud.schaab@h-ka.de}

\textsuperscript{a}Faculty of Information Management and Media, Karlsruhe University of Applied Sciences, Karlsruhe, Germany

\textbf{Abstract.} Python programming and the application of web services technology can be considered important skills for geomatics students when looking for employment later. A course on Web Processing Services (WPS) as part of the International Geomatics master programme at Karlsruhe University of Applied Sciences takes this into consideration while using the opportunity of making students aware of less-known thematic mapping methods. Students are introduced to automating a complex geoprocessing task and finally wrapping it into a WPS. It is achieved by following the phases of first conceptualizing the single steps before then using ArcGIS ModelBuilder and Python scripts in combination to build the processing chain. The alternate band map method serves as example in the hands-on exercise. Later students are asked to automate the generation of the geometries for the map face and/or map legend supporting other thematic mapping methods like the square density map or the absolute value raster map. The paper discusses the teaching approach and experiences made so far.

\textbf{Keywords:} service-oriented web cartography, geoprocessing web service, alternate band map, square density map, absolute value raster map

\section*{1 Introduction and Background}

Python is an object-oriented scripting language that is appreciated for being readable and thus easy to reuse and maintain. Its general-purpose kind renders it useful in almost all fields of science and industry. It is said to boost the developers’ productivity because they have to write less code and do not need to compile the code (Lutz, 2013). Web service technology supports interoperability by common standards and the accessibility to and sharing of distributed resources in the internet. Through the use of Application Programming Interfaces (API) functional components can be assembled to run complex geoprocessing tasks (Zhao et al., 2012). The standards for spatial content and associated services are developed by the Open Geospatial Consortium (OGC). Its non-propriety web mapping approach is based on open interfaces, encoding and schemas. The interface Web Processing Service (WPS) enables to describe a service that provides geoprocessing functionalities to be executed in a web environment (Rautenbach et al., 2012). In order to successfully teach students how to apply Python programming and/or how to make use of WPS for geoprocessing different approaches are followed. Institutions of higher education build on integrating e.g. problem-based learning (Deng and Di, 2008), game-based approaches (Priyaadharshini et al., 2019) or their own e-learning tools (Javed et al., 2013) in traditional classroom teaching.

Building on Service-Oriented Architecture (SOA) to support web services which automatically generate thematic maps is receiving increased attention for two reasons: the many geoportals with their potential to visualize the manifold data which is already stored in Spatial Data Infrastructures (SDI) (Cerba and Cepicky, 2012; Kraak et al., 2018; Zepner et al., 2020) and the new challenges encountered for analyzing big geospatial point data supported by intelligent visualization (Coetzee and Rautenbach, 2017; Gröbe and Burghardt, 2017; Pillay et al., 2019). Thematic cartography has been a focus of university training at Karlsruhe University of Applied Sciences (HKA) from early on (Schaab, 2014). Nowadays teaching students in the necessary programming and software development skills may be considered of higher importance, which however is not necessarily the easier task. In Jan 2014 HKA received the Esri Developer Center (EDC) label. Here we aim at server applications supporting correct and appealing visualization. Hence, the question arose how to motivate students to get a grip on programming while still conveying knowledge on
sophisticated thematic cartography. After an initial workshop with students in May 2015, a new course on WPS was introduced to the International Geomatics master programme (start: summer semester 2018). It aims at imparting knowledge on implementing cartographic web services, which deliver the required map face geometries and legend building blocks to be applied in the design of thematic maps.

Using SOA and applying web services in cartography started with extending the symbolizing and labelling capabilities of Styled Layer Descriptor (SLD) and Symbology Encoding (SE) for the OGC Web Services (i.e. WMS, WCS and WFS) (e.g. Dietze and Zipf, 2007; Sae-Tang and Ertz, 2007; Iosifescu-Enescu et al., 2010). By wrapping existing or new geoprocessing functionality in a WPS and chaining the tasks, the limitations of the standards can be overcome and more sophisticated and diverse thematic map products can be generated (e.g. Cerba and Cepicky, 2012; Rautenbach et al., 2012; Zeng et al., 2013; Krimbacher, 2014; Moncrieff et al., 2016). For an automated, parallel processing of distributed big geospatial data, the software design pattern approach can be used to produce a series of thematic map designs which employ only those cartographic representation methods which are appropriate depending on the input data characteristics (Coetzee and Rautenbach, 2017).

For mapping areal distributions of quantitative area-related characteristics, the choropleth, dasymetric and chorodot map methods are widely used. All apply the relative method based on standardized data but use different reference units. They typically visualize census data on very limited basemaps, require an equal-area map projection and the consideration of an appropriate data classification (e.g. Slocum et al., 2010). But what about other areal cartographic methods which allow the representation of proportions (sections) making up a topic, of spatial variation within enumeration units or offering absolute and relative information at the same time?

Common to areal diagrams is that the entire map face is subdivided according to the percentages of the characteristics of the depicted variable (Kraak and Ormeling, 2010). The alternate band map (Poidevin, 1989) uses equal-sized stripes, distinctly partitioned into slices (bands) per enumeration unit, which form distinctive categories but need to represent and add to 100% (i.e. the method visualizes sections). However, an alternate band map does not provide information on the geographical distribution or spatial pattern within the reference units. In contrast, this is the case for the absolute value raster map (Quadratraster-Flächenkarto-

gramm, translating to ‘square raster areal diagram’, by M. Bürgener, 1957). It allows for an appropriate depicting of minorities and majorities in their true location, while visually revealing the proportions as stated in the statistics. Not being an aerial diagram map, but using the absolute method while offering relative information at the same time, for this the dot density map (Tyner, 2010) is known but not necessarily the square density map (Dichtequadratmethode by F. Scharner, 1958). The latter works with grid cell areas instead of dots, is visually pleasing and can benefit from an increased visual impact. In comparison to the standard classified choropleth maps, there is no loss of information because it does without classes. The inventor of the method points to the necessity of certain legend elements for adding understanding to the map method.

Having elaborated on the setting and background of the topic, in the following section 2 the teaching approach is described, also in comparison to others. Section 3 takes the example of one particular thematic map method to demonstrate the prototypic solving of the geoprocessing task as a web service. In section 4 we present and discuss the other methods which have been tackled by students so far and conclude on our teaching approach.

2 Teaching Methods

The field of information and communication technology (ICT) poses ever and fast evolving developments. At HKA we want to prepare geomatics students to be capable of taking on the resulting chances and challenges. It requires the students to be trained in the most modern methods, techniques and technologies. At the same time students are required not to lose grip with the old but still valid, i.e. well-established, methodologies which have proven to be beneficial in data analysis and visualization. In times of “Big Data” an increasing amount of statistical data, too, is readily available from nearly everywhere and at any time. It asks for experts who know how to facilitate the breaking down and packaging of data into understandable information for everyone.

Opinions and approaches vary of how to train the future experts in the application of Python and/or geoprocessing services. Priyaadharshini et al. (2019) regard traditional classroom teaching in programming as boring and monotonous due to the lack of interaction. They state that learners expect new teaching techniques, digital working tasks and continuous motivation-pushing.
Deng and Di report that GeoBrain helped to improve the geospatial science and engineering (Deng and Di, 2008) offers online resources for educational purposes inresulted not only in an extension offering the GRASS GIS and to allow for transforming the data into information teaching effectiveness in traditional classroom teaching asNASA's geospatial data by scientists and general users and to allow for transforming the data into information and knowledge by means of web services (Di, 2004). It resulted not only in an extension offering the GRASS GIS functionalities as web services (Qu et al., 2012), but offers online resources for educational purposes in geospatial science and engineering (Deng and Di, 2008). Deng and Di report that GeoBrain helped to improve the teaching effectiveness in traditional classroom teaching as well as in problem-based learning regarding WPS use in a cloud-based environment. Our approach for teaching Python and WPS within a single course aims at a versatile problem-based learning environment.

It all started after discussions with master students on curriculum development, in which the students expressed the wish to learn Python. Now the programming language is being taught in the second semester of the International Geomatics master study programme as part of a lecture on Web Processing Services (WPS) covering 2 SWS (2 contact ‘hours’ of 45 min per week, i.e. 30 in total). As Tab. 1 shows, the course starts with an emphasis on Python, but gradually evolves towards a focus on WPS. At the beginning a short introduction to the Python programming language is provided. It is meant to harmonize the level of knowledge within the usually rather heterogeneous group of students: students from an abroad mix with students who followed the geodata-related Python programming course within their bachelor studies on Geo-Information Management at HKA. Basics such as data types and structures and a comparison between Python versions 2.x and 3.x are covered. It is followed by an introduction to the ArcPy module of the ArcGIS Platform, which allows access to its geoprocessing resources. First small geoprocessing scripting tools for ArcGIS are programmed by the students to ensure practical experience.

To provide for more advanced and later-on required Python/ArcPy programming skills, concepts like cursors and methods for directly manipulating features and geometries are introduced and tested. It leads to a first assignment where each student has to choose a different task from a given list. The tasks are of varying level of difficulty to also challenge those students who have already a high affinity to programming. The common focus of the tasks is the programmatically direct manipulating of geometries while creating a well-documented scripting tool for the ArcGIS Platform. After that, about a third into the semester, the focus starts to change from mere teaching of Python/ArcPy towards introducing WPS and the concept of the geoprocessing web. The necessary background is achieved by jointly studying the OGC Web Services Common Standard (Whiteside and Greenwood, 2010), the OpenGIS Web Processing Service Standard (Schut, 2007) and the paper by Zhao et al. (2012). On the cartography side, a short overview or repetition of in particular areal representation methods for thematic maps is included, which also covers the example of the alternate band map being of relevance later. The last component, for a change again as a small hands-on session, is the setting up a simple WPS on the ArcGIS Platform using an Esri tutorial.

Next the alternate band map method is used as an example for learning how to go about creating a WPS. Over a couple of teaching units, this WPS is first jointly conceptualized, then implemented via coding and finally published. The units are based on an elaborate hands-on exercise worked on together with all students in the lab. Details regarding the steps involved are presented in section 3. Approximately the last third of the semester is filled by project group work on a specific task of again conceptualizing, programming and publishing a WPS rooted in thematic cartography. The group members (2 to 3 students) are responsible for subdividing the work and assigning the tasks as well as for organizing a successful project management. Similar to the student assignments, also here the students need to make their coding available as well as to prepare a documentation. In parallel, each student has to look for and study a piece of literature on the application of WPS, WPS chaining or other advancements of the geoprocessing web. Reporting about it by means of a lightning talk ensures that every student participant gains at least an idea of the current state-of-the-art regarding WPS. A demonstration of how to chain two WPS helps to illustrate web service orchestration and concludes the outlook.

The results of student assignment, project work and lightening talk form the basis for assessing the students’ performance. Feedback is provided and decisions made of who of the students is ready to sit the written exam. It concludes the course and allows to mark each student’s performance individually.
Data and Software Availability

The data, software and computational infrastructure used in the teaching course can be summarized by referring to the ArcGIS platform, which is available to members of the university. Python 2.7 (https://www.python.org) and the ArcPy library of ArcGIS Desktop 10.x are used to set up geoprocessing chains as ArcGIS ModelBuilder tools. Wrapped as a WPS the geoprocessing chain is published to a standalone ArcGIS Server at Karlsruhe University of Applied Sciences, which is available within the university network and runs on a virtual machine with Windows Server 2012 R2. It is planned to switch to ArcGIS Pro and the Python 3.x environment soon. The spatial data, as originally compiled for the Python workshop, were taken from Esri’s Data and Maps 2012 (accessible to Esri users). The statistical data stems from the Federal Statistical Office of Germany (https://www-genesis.destatis.de; for the data license by-2-0 see https://www.govdata.de/dl-de/by-2-0). The teaching material and coding scripts are not shared, because it would defeat the teaching aim and require the consent of the students.

3. The Alternate Band Map Method as a WPS

The alternate band map method was chosen for use as the hands-on teaching example: it is rarely known, is

<table>
<thead>
<tr>
<th>Teaching block</th>
<th>Teaching unit</th>
<th>Learning aim (student efforts)</th>
<th>Level of difficulty</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Harmonizing students’ knowledge level regarding Python/ArcPy</td>
<td>Introduction to Python</td>
<td>Learning/repeating the basics of Python, including differences between version 2.x and 3.x</td>
<td>Bachelor level</td>
</tr>
<tr>
<td></td>
<td>Introduction to ArcPy</td>
<td>Learning/repeating the use of the ArcPy module for interaction with the ArcGIS platform</td>
<td></td>
</tr>
<tr>
<td>2. Providing for the necessary background and programming skills</td>
<td>Manipulating geometries with Python/ArcPy</td>
<td>Gaining knowledge and experience of working with and manipulating geometries</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Creating a scripting tool in ArcGIS</td>
<td>Application of the previously learned (individual student assignments)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Introduction to the alternate band map method</td>
<td>Learning the cartographic background</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Introduction to WPS</td>
<td>Learning about the OGC standards, the geoprocessing web and how to make use of documentations and literature</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Setting up a simple WPS on the ArcGIS Platform</td>
<td>Gaining knowledge and experience of using WPS</td>
<td>Master level</td>
</tr>
<tr>
<td>3. Using the alternate band map method as an example for creating a WPS</td>
<td>Conceptualizing a WPS for the alternate band map method</td>
<td>Learning how to develop a more complex WPS from scratch</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Setting up a geoprocessing workflow for the alternate band map method by means of the ArcGIS ModelBuilder and scripting tools</td>
<td>Gaining knowledge and experience in coding a complex geoprocessing workflow and its publishing as a WPS (elaborate lab exercise)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Publishing the geoprocessing workflow of the alternate band map method as WPS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Independent working, mainly on a WPS service which wraps a thematic mapping method</td>
<td>Chaining of WPS (meant as an outlook)</td>
<td>Independent knowledge gain by studying literature (lightening talks)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Project work: prototyping other cartographic methods as WPS</td>
<td>Independent application of the previously learned content, in groups (student group projects)</td>
<td></td>
</tr>
</tbody>
</table>
easy to understand and implement by students while being not too trivial, and provides a convincing output. An example of an alternate band map is provided in Fig. 3 to the left. It is based on statistical data of school graduation types attained for the German states. The data requires to be standardized, i.e. expressing relative values, which sum up to 100% per enumeration unit. In the process of creating the alternate bands, the map maker has to familiarize with the data first (e.g. checking out minimum and maximum values). Next, decisions are required on the direction and width of the stripes which determine the partitioning of the map face. The decision for the stripes’ orientation is a compromise between the overall orientation of the enumeration units and the fact that stripes being placed in cardinal direction are visually most appealing (Arnberger, 1997). The stripe’s width is determined by the number of slices (bands) and their minimum values to be visualized. Here the rule applies that each stripe must be displayed in its entire width at least once per enumeration unit (ibid.).

The conceptual phase of setting-up the WPS usually starts with a blank map of enumeration units (e.g. regions of France) on a sheet of paper and some tabular statistical data accommodating the cartographic representation method already. Having been introduced to the map method before, the students are asked to sketch the alternate band map method for some of the areal units. They should take notes about the requirements and rules applied for producing such a map. In the next step, the students are asked to think about automating the process. What input is needed? What are the required parameters? Are there also optional parameters? These are example questions which describe the determining of the WPS’s input. Next the processing of the spatial data needs to be conceptualized, which always turns out to be the most difficult part for the students. Ideas of how to break down the task need to be developed, which requires to think outside the box. Finally, the output needs to be described which again appears as an easy task after the demanding interim steps have been solved.

Figure 1: Conceptual model for geoprocessing the spatial geometries for an alternate band map based on GIS functionality. Roman numerals link the illustrations of the processed geometries to the workflow. The arrangement of the workflow reflects the ArcGIS ModelBuilder diagram in Fig. 2.
Fig. 1 visualizes the final concept including sketches of the processed geometries. Required input parameters to the Alternate Bands WPS are the input geometries (polygons) of the enumeration units, at least two attribute fields with statistical data (allowing for up to three additional fields), the scale of the resulting map and the desired stripe width in millimetres. The angle defining the orientation of the stripes is regarded optional and calculated automatically based on the spatial input data if not specified. All this covers the initial step (I) of the processing chain. For calculating the ‘optimal angle’ of the stripes, the average main orientation of all polygons is determined via the Minimum Bounding Geometry functionality and a value of 90 (degrees) subtracted to obtain the orthogonal direction. As such it remains for now a simple solution only with further optimization potential. Additionally, the input geometries are “repaired” in case that they do not follow Esri’s feature geometry specifications. In the second step (II) the input feature class is “described” in order to determine the rotation centre, which serves as input to the next action (step III). Here the geometries are rotated counterclockwise based on the ‘optimal angle’. This approach helps to simplify the upcoming calculations, especially when generating the slices (bands) within the stripes. The ‘describe’ functionality is run once more, this time for deriving the minimum and maximum coordinates of the bounding box extent. These variables are needed for the creation of the stripes in the subsequent processing step (IV). The stripes can be easily achieved by creating a fishnet by means of the respective ArcGIS tool. The most complicated step is reached with action V. Here, the stripes per polygon feature need to be sliced (subdivided) according to the relative values stored in the attribute table and specified as input parameters (via the field names). In an iterative process, per polygon feature the stripes are split into smaller slices (bands) and then intersected with the polygon feature’s geometry. In the final processing step (VI), the resulting geometries are also repaired in case that the newly created polygons do not follow the specifications and then rotated based on the ‘optimal angle’. Rotation is done clockwise to finally generate the output feature class ensuring its correct spatial positioning (VII). Colouring of the bands is not part of the processing chain.

Fig. 2 (see Appendix) reveals that the hands-on exercise for a complex geoprocessing task performed on the ArcGIS Platform requires a splitting of the processing chain in tasks. I.e. the students are handed over an incomplete model based on a ModelBuilder tool and incomplete Python scripts. For each task, learning aims are defined and the target what to achieve is set. Due to the combined use of the ModelBuilder and Python coding, the students need to learn how to branch in a model, how to create scripting tools, and how to call tools from within Python code. They are asked to make extensive use of the ArcGIS Online Help and to use other resources, too. Starting with smaller tasks, in step IV the students are asked to code a complete script. After documenting the resulting model tool by means of metadata and after a thorough testing in ArcGIS locally, the process is published as a WPS to ArcGIS Server. The Alternate Band WPS service can then be used by the students for designing their own maps (see e.g. Fig. 3, left).

4. Discussion and Conclusions

Fig. 3 shows examples of student work on thematic cartography WPS during the initial workshop or semester courses. The alternate band map method (Fig. 3, left) provides a visually appealing result, although two disadvantages remain: The method does not provide information on the actual location and spatial distribution within the reference units. Care has to be taken on how to place the stripes, because it depends on the reference units’ sizes and shapes as well as their overall arrangement if the stripes and their slices are equally often represented (Amberger, 1997). The WPS could be further improved: more tricky parts are still waiting for a solution like the automated determination of a threshold for the minimum proportion to be depicted, of the ‘optimal angle’ for the stripes by considering the shape of the enumeration units more closely, of the optimal stripe width depending on the size of the enumeration units, the map scale and the statistical data, and the automatic creation of a meaningful legend layer. Still also missing is a proper documentation and the sharing of the service, which will only make sense when we will have moved on to using a different hands-on exercise during the master course on WPS.

In case of the square density map (Fig. 3, centre) one wonders why this method is not more commonly used. By replacing each dot symbol of a grid-based dot density map with a square, dot densities are transformed into square densities with the advantage of representing one density dot by two sides of a density square. It increases the visual impact by more black ink and thus facilitates the distinction of even small differences (Scharner, 1958). As the student example of the WPS output reveals, also here it becomes evident that automatically creating a legend (or at least the required geometries) as part of the overall WPS (see e.g. Zeng et al., 2013) is as important as creating the features for the map face. It will help to avoid misleading legends and thus misinterpretation of maps by the map users. The need for automating also the creation of a legend is
highlighted by Chua and Vande Moere (2017) and Jenny et al. (2010).

The absolute value raster map presented in Fig. 3 to the right shows that the entire map face is subdivided by a grid of equal-sized squares. It leads to the same absolute area value per grid cell throughout but different proportions per grid cell because their numbers vary per enumeration unit. The optimal size of the grid cells depends on the map scale, the average reference unit size and the demand of depicting a given scattered distribution (Bürgener, 1957). This mapping of the geographical distribution within the enumeration units was not realized by the students. A solution could be to integrate a dasymetric modelling approach (see Mileu and Queirós, 2018 for an overview on GIS-based dasymetric mapping tools), which the students failed to realize. For none of the three thematic mapping methods, which we tackled together with our students so far, we are aware of an automated solution.

Students’ programming assignment work can only very sometimes be considered ready for sharing as final products. However, working on these prototypes can act as an eye-opener to a wider pool of cartographic visualization options as compared to what the students are aware of. They also learn how to get around the limitations of only applying standard GIS software functionality but experience the power of automating geoprocessing tasks. By programming in Python, a scripting language which is easy to understand, is portable between operating systems and open-source and provides many popular third-party libraries and applications built on top of it (Lutz, 2013), the students get the chance of becoming part of its big and active user and developer community. Being confronted with OGC specifications and gaining experience in setting-up web services, the concept of Software as a Service (SaaS) becomes all of a sudden approachable. In our course on WPS, the automation of processes plus making them available/sharing them with others is linked to the fast evolving field of service-oriented web cartography by generating map face geometries and legend layers. They can serve as input in straight-forward browser-based visualization tools of open data portals (Kraak et al., 2018) but also in graphically more enhanced map visualizations derived by means of Desktop Mapping (DTM) or by focusing on dynamic interactive features (as in traditional web mapping). In combination, all this adds to the motivation of the students.

Instead of mixing game-based learning and traditional classroom teaching (Priyaadharshini et al., 2019), for teaching we follow a problem-based learning approach to keep the students connected to the subject and interested. Interlinked with teaching how to set-up and use WPS as a state-of-the-art technology in service-oriented architectures, we teach them programming with Python by means of a complex hands-on exercise on a challenging cartographic task that helps to visualize geographical distributions within the enumeration units.
statistical data. It works, although Javed et al. (2013) point to the limitation for hands-on exercises on WPS due to the generally complex use cases. Similarly, the educational material offered by the GeoBrain project (Deng and Di, 2008) focuses on using WPS in a cloud-based environment and not on how to set-up and program WPS. Also our teaching falls short, this in regard to fully introducing WPS chaining.

The approach of teaching Python and WPS within a single course integrates theoretical knowledge (taught by the lecturer or presented by students) of established methods (thematic cartography) and new techniques (web services), hands-on programming exercises (geo-processing web services) as well as independent individual and group work (accommodating different knowledge/experience levels). This versatility in teaching means, in topics, techniques and methods covered, plus asking students to elaborate solutions for overlooked, i.e. not yet solved tasks, is considered to drive the students’ motivation. The course works closely with what the ArcGIS Platform offers, while a different module of the International Geomatics master programme is dedicated to open source software tools in the GIS field. In future we plan to make use of ArcGIS Notebooks (based on Jupyter Notebooks) for accommodating the hands-on exercises. Most challenging to our students proved to be the conceptualization phase of independently finding solutions for the automatization of a given complex problem as specified in the student project, which requires in-depth study of examples and related publications.

References


Figure 2: The ArcGIS ModelBuilder diagram for creating the spatial geometries for an alternate band map as used in teaching Python programming as part of setting-up a WPS. Roman numerals link to the descriptive workflow in Fig. 1.