УДК 004.94
UDK 004.94

М.ЛАУЕРМАН, Ф.БЕТЦ, Б.СЫФКА
M.LAUERMANN, F.BETZ, B CYFFKA
E.mail. ksucta@elcat.kg

ВЫДЕЛЕНИЕ СЕТИ СТОКОВ ИЗ ЦИФРОВЫХ МОДЕЛЕЙ РЕЛЬЕФА - ОЦЕНКА ПРИМЕНЕНИЯ ПРОГРАММ С ОТКРЫТЫМ ИСХОДНЫМ КОДОМ

CHANNEL NETWORK DERIVATION FROM DIGITAL ELEVATION MODELS – AN EVALUATION OF OPEN SOURCE APPROACHES

Гидрологиялык изилдөөлөрдө көп багыттар боюнча ASTERдин сүрөттөрүн жана SRTM рельефинин моделдерин кеңири пайдаланышат, мисалы, өзөндүн нугун аныктоодо булар стандарт процедура болуп саналат. Рельефтин моделинин тереңдигине кыртышта өскөн чөп же жөн эле ызы-чуу үндөр да суунун агымын аныктоодогу алгоритмдерге таасирин тийгизет. Бул маселени чечүүнүн эки жолу бар: көңдөйлөрдү толтуруу жана нугу жок рельефтин моделин түзүү же өзгөртүлбөгөн ЦММди пайдалануу. Бул изилдөөдө биз Кыргызстандагы Нарын дайрасынын өзөнүн мисал катары алабыз да ачык коддору SAGA жана GRASS болгон ГИСтин кеңири пайдаланылып жүргөн эки программасын колдонуп эки методду салыштырабыз. Санарипке салынган эталон өзөн нугун статистикалык салыштыруу көрсөткөндөй, GRASSта ишке ашырылган кыска жолдун методу эсептелген жана санарипке салынган нуктун бир топ эле азыраак аралыгын көрсөтүп турат. Андан тышкары, SRTMдан алынган рельефтин модели ASTERден тартылган сүрөттөргө салыштырганда эталонго жакынырак. Ошентип биз ири суу бассейндерин анализдөө үчүн SRTM рельефинин моделин айкалыштыруу менен GRASS ГИСти прайдаланууну сунуштайбыз.

Ачкыч сөздөр: *жерди анализдөө, агымдар тармагын моделдөө, ЦММ, ири масштабдагы суу бассейндери.*

Использование снимков ASTER грубого разрешения и модели рельефа SRTM является очень распространенным явлением для многих целей в гидрологических исследованиях, стандартной процедурой является, например, определение речных стоков. Тем не менее, углубления моделей рельефа в результате растительного покрова или просто шум влияют на алгоритмы определения стоков воды. Есть два распространенных способа справиться с этим: заполнение пустот и создание модели рельефа без углублений или применение неизмененной ЦММ. В данном исследовании мы берем водосбор реки Нарын в Кыргызстане в качестве примера и сравним эти два метода, используя двух широко используемых ГИС программ с открытыми исходными кодами - SAGA и GRASS. Статистическое сравнение оцифрованного эталонного стока показывает, что подход наиболее короткой пути, реализованный в GRASS, дает значительно меньшие расстояния вычисленного и оцифрованного стоков. Кроме того, модель рельефа из SRTM ближе к эталонному чем из снимка ASTER. Таким образом, мы предлагаем ГИС GRASS в сочетании с моделью рельефа SRTM для анализа крупных водосборных бассейнов.

Ключевые слова: *анализ местности, моделирование сети стоков, ЦММ, крупномасштабный водосбор.*

The use of the coarse resolution ASTER and SRTM elevation model is very common for many purpose in hydrological research, a standard procedure is for instance the extraction of river channels. However, depressions in the elevation models resulting from vegetation cover or just from noise in the data cause challenges for the flow routing algorithms used for the channel derivation. There are two common methods to handle this: filling the sinks and creating a depressionless elevation model or applying a least cost path approach routing through the unmodified DEM. In this study we take the catchment of the

Naryn River in Kyrgyzstan as an example and compare these two methods using the two widely used open source GIS systems SAGA and GRASS. The statistical comparison with a digitized reference stream shows that the least cost path approach implemented in GRASS GIS gives significant smaller distances of the computed to the digitized stream. Furthermore the derivation from the SRTM elevation model is closer to the reference as the one from ASTER. In summary, we suggest GRASS GIS in combination with the SRTM elevation model for the analysis of large scale watersheds.

Keywords: *terrain analysis, channel network derivation, DEM, large scale watershed.*

1. Introduction

For many aims in environmental research hydrological derivations from digital elevation models (DEM) like the catchment area or the channel network are an important foundation. However, due to remote research areas or limitations in budget, the use of coarse resolution DEMs from the ASTER and the SRTM mission is very common (Metz et al., 2011). These data sets are available for free almost for the entire world (Hayakawa et al., 2008). When applying open source geographical information systems (GIS) for the analysis of these data sets, processing can be done for very low cost what is an asset for many applications in research and teaching. In this study, we compare GRASS GIS and SAGA GIS with respect to their capability for deriving channel networks from elevation data. These two GIS are among the most popular systems and contain elaborated methods for the processing of digital elevation data (Neteler and Mitasova, 2008; Hengl et al., 2009; Conrad et al., 2015).

Stream extraction from DEMs is based on the computation of flow accumulation. But the derivation of channels from ASTER and SRTM pose some challenges: These data sets include large depressions caused by vegetation or noise in the data what causes problems for flow routing algorithms (Arge et al., 2003; Metz et al., 2011). A widely applied method for handling depressions in elevation data is sink filling and creating depressionless DEMs (Planchon and Darboux, 2001; Wang and Liu, 2006). According to Conrad (2007) all SAGA modules for drainage network derivation use the deterministic 8 algorithm routing to the lowest neighboring grid cell. Thus, SAGA GIS is dependent on the depressionless DEMs and requires sink filling as pre-processing step. An alternative for handling depressions in digital elevation data is using a least cost path (LCP) approach with unmodified data. Ehlschlaeger (1989) was among the first to suggest this approach using an A^T search algorithm for flow routing through digital elevation data. Including further improving, this algorithm is still implemented in the hydrological tools of GRASS GIS (Jasiewicz and Metz, 2011; Metz et al., 2011).

In this study, we take the catchment of the Naryn River till the Toktogul Reservoir as an example and evaluate the sink filling approach in SAGA and the LCP approach in GRASS comparing the computed streams with a digitized reference stream. Furthermore, we compare the performance of ASTER and SRTM.

2. Material and Methods

2.1 Elevation Data and general Workflow

The elevation data used for this study is the SRTM-1 elevation model (Farr et al. 2007) and the ASTER global digital elevation model (ASTER GDEM; Tachikawa et al. 2011). These data sets have a spatial resolution of 1 arc second which results in grid cells of 24.26 m (ASTER) resp. 24.05 m (SRTM-1). For the entire Naryn Catchment with 52,130 km², this results in approx. 75 x 10^7 grid cells and a file size >1.5 GB. Beside the DEM input, mapped channel heads are required inputs.
These have been digitized from referenced topographic maps. Furthermore, the main stream of the Naryn River has been digitized from Google Earth to have a reference for evaluating the quality of the computed channels.

The workflow for the channel network derivation is differing between SAGA GIS and GRASS GIS, figure 1 gives an overview including the module names within the software. All steps have been performed for both elevation models, SRTM-1 and ASTER.

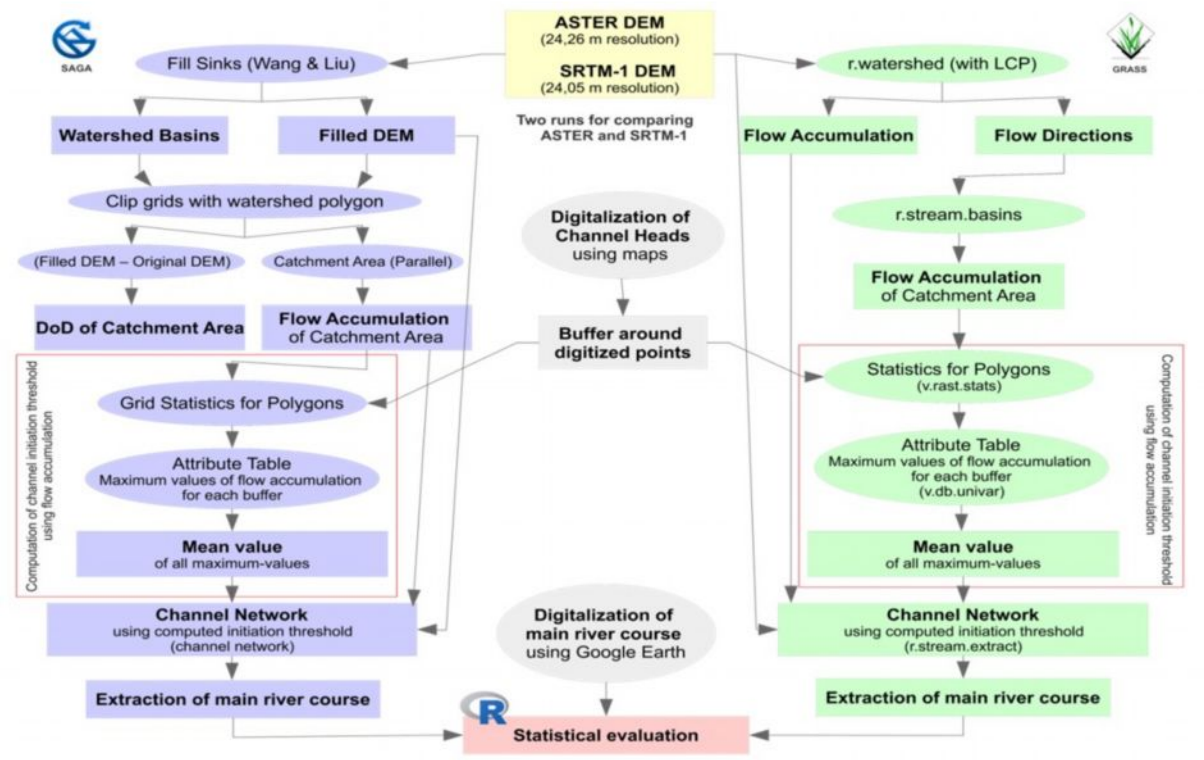

Figure 1. Workflow for computing the channel networks in SAGA and GRASS as well as for statistical evaluation

As SAGA depends on a depressionless DEM, the workflow is starting with the sink filling procedure. We use the algorithm of Wang and Liu (2006) as it offers a reasonable computing efficiency – an important issue regarding the file size. The respective SAGA module produces beside the "corrected" elevation model also already the catchment boundaries ("watershed basins"), this allows to clip the raster file and reduce the file size for enhancing the processing speed in the subsequent steps. Next, the flow accumulation was computed using a combination of the multiple flow direction (Freeman 1991) and D8 (O'Callaghan and Mark 1984) flow routing methods. The threshold for changing from multiple flow direction to single flow was kept at the default of 500 grid cells. Now, the river channels could be extracted. The required initiation threshold of the flow accumulation was created using a 50 m buffer around the digitized channel heads, extracting the maximum of each buffer and selecting the mean of all buffer maxima as initiation threshold (see also Bhowmik et al. 2015 for details about this method).

For GRASS GIS, the process starts with the computation of the flow accumulation and flow direction using directly the unmodified DEM. For the flow routing, also a combination of a multiple and single flow direction was applied using the standard settings of GRASS (Holmgren 1994). The flow direction grid is the input for delineating the catchment. For extracting the channel network we used again a threshold of the flow accumulation as input with the same method described above (just the GIS module names differ). For further statistical evaluation, the main river channel – the Naryn River – has been extracted from the computed channel network.

2.2 Statistical Evaluation

For the statistical evaluation, we compared the distances between the digitized reference channel and the computed river channels. The distances have been calculated using the spatstat package within R (Baddeley and Turner, 2005; R Core Team 2015). This tool splits the vector lines of the river channels into segments and generates the distances between the respective segments. For ASTER, 37'952 segments, for SRTM-1 37'942 segments have been used for comparison. The
Mann-Whitney test was applied to test for significance of the difference between the distances.

3. Results and Discussion

3.1 Statistical Evaluation of Channel Derivation

For the ASTER GDEM, the median of the distances between the computed channel and the reference stream is 149.31 m for SAGA GIS and 94.23 m for GRASS. For SRTM-1 this relative distance is also visible, while the overall distances are much smaller. Here, SAGA comes to a median distance of 79.94 m and GRASS to 30.80 m. The boxplots of the distances for all segments are shown in figure 2.

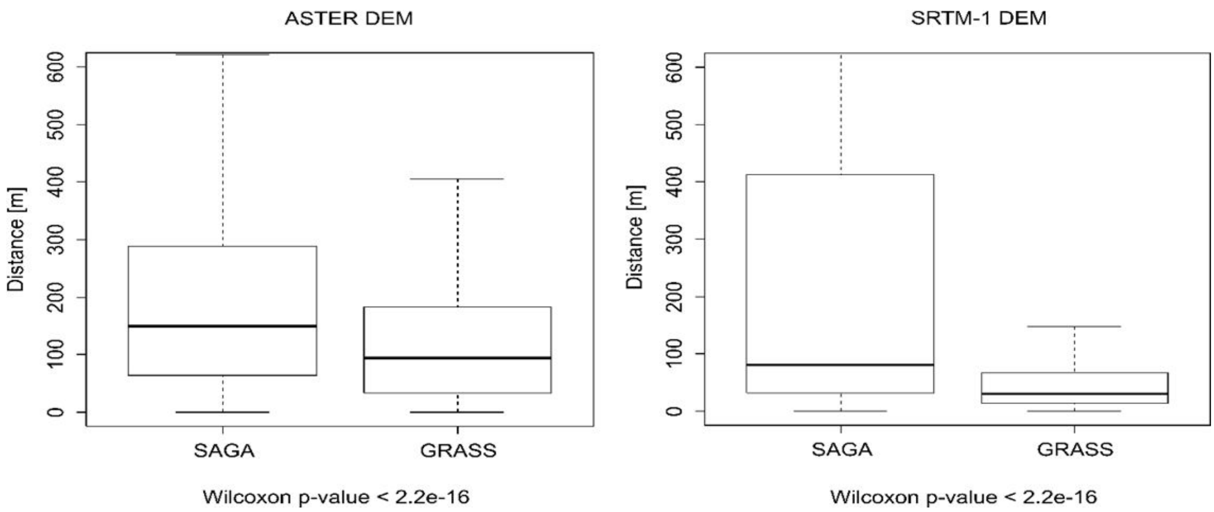

Figure 2. Distances of the stream line computed with GRASS and SAGA; the left figure shows the results for the ASTER GDEM, the right one for SRTM-1

The differences are highly significant even under the 99 % confidence interval. Thus, the LCP algorithm performing flow routing through unmodified DEMs is able to generate channel networks significantly closer to reality than algorithms depending on elevation models with removed sinks. This corresponds well with results from Metz et al. (2011) who also tested the performance of the GRASS LCP algorithm against the sink filling procedure. Figure 3 shows the differences between the computed stream lines exemplified for one river bend.

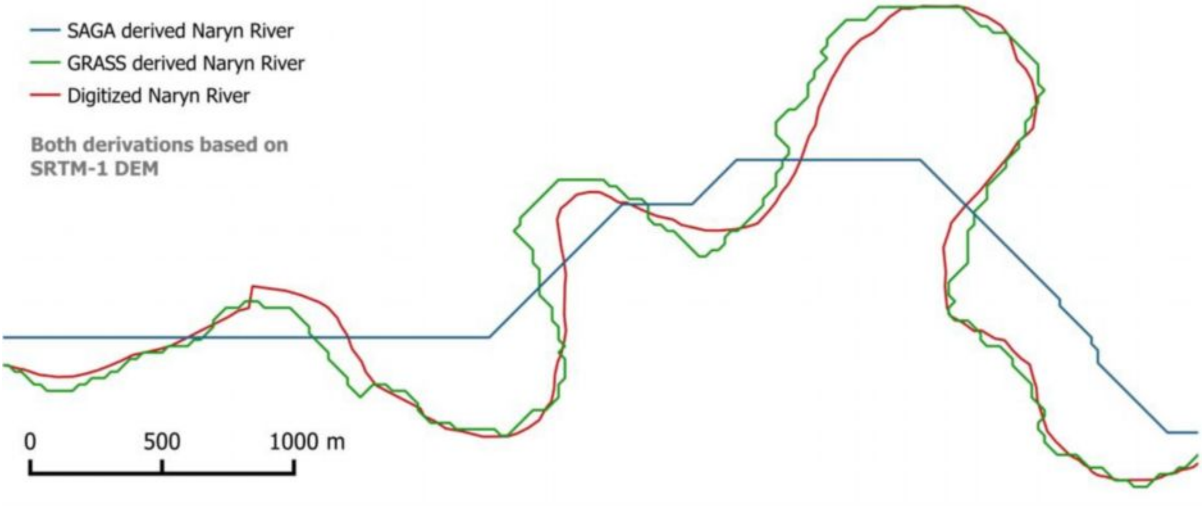

Figure 3. Comparison of computed river channels with the digitized one

As the sink filling procedure created almost flat areas, the resultant river channel shows relatively straight lines and follows the bend only rudimental. The LCP stream line shows a slight noise but follows the river bend in accurate manner. From geomorphological point of view, the straight SAGA stream lines cannot be interpreted as a realistic representation of this river section while the LCP stream line is representing the river form in sufficient way.

A further point of interest was the performance of the ASTER GDEM in comparison with the SRTM-1 elevation model. Figure 4 shows the distances of stream lines derived from the ASTER GDEM resp. the SRTM-1 to the reference stream line.

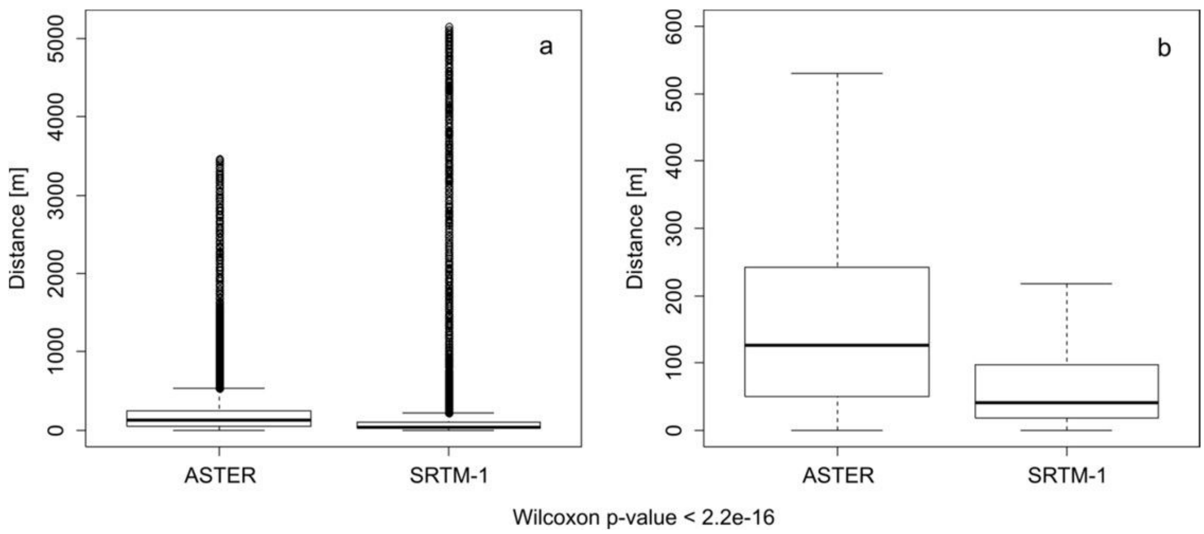

Figure 4. Distances channels derived from ASTER GDEM and SRTM-1, both computed with the LCP algorithm

The channel derivation from the SRTM-1 elevation model shows significant smaller distances to the reference line than the one computed based on ASTER. The explanation is a higher noise of the ASTER GDEM compared to SRTM-1. The relatively worse quality of the ASTER GDEM for hydrological research is a known issue. For instance, Jarihani et al. (2015) advice in a recent study that the use of the SRTM elevation data for hydrodynamic modeling as ASTER showed bigger elevation errors when compared with GPS control points. A potential reason of this fact is the different generation method of the DEMs. While ASTER is a photogrammetric product from overlapping satellite imagery, SRTM is based on radar interferometry and thus independent from atmospheric disturbance. Figure 5 gives two profile lines of the Naryn River, one based on ASTER, one based on SRTM-1. Both lines show some artefacts, but it is very obvious that the SRTM profile contains much less noise than the ASTER derived.

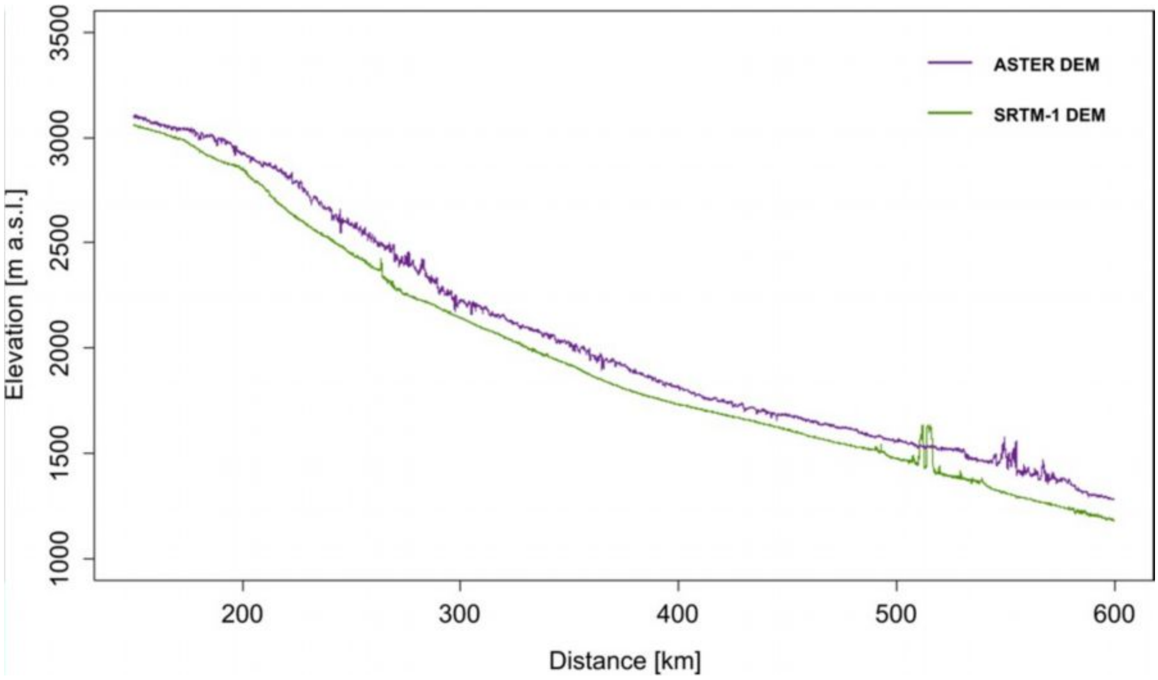

Figure 5. Profile lines of the Naryn River derived from ASTER and SRTM-1

3.2 General performance of the GIS Systems

In the previous section we demonstrated that the SRTM-1 elevation model in combination with the LCP algorithm implemented in GRASS GIS was able to generate the channels closest to the reference. Beside the performance of the flow routing and channel extraction algorithms, there are further differences between the two evaluated GIS systems. Probably most important is the fact that SAGA is handling the entire data set in memory, thus there is a limit of the maximum file size for processing (Hengl et al. 2009). Contrary, GRASS GIS offers the opportunity to handle the data from hard disk offering the ability to process even very large files on regular personal computers (Metz et al. 2011). Functionalities regarding further raster or vector processing are similar in GRASS and SAGA (Neteler and Mitasova, 2008; Conrad et al., 2015). However, with the recently established r.stream toolkit GRASS has a very comfortable tool for stream network analysis. Further issue necessary to be mentioned is the lack in contingent documentation of the modules of SAGA GIS, for instance the explication of parameter units or further information about algorithms is often missing (Hengl et al., 2009). While SAGA has the more user friendly interface, GRASS tends to be the better choice when going beyond the default settings of modules as there is a detailed documentation for all modules, often accompanied by scientific references.

4. Conclusion

In this study we evaluated the derivation of river channels from digital elevation data using a least cost path algorithm in GRASS GIS and a "classical" algorithm working with a depressionless DEM implemented in SAGA GIS. We performed this evaluation with two freely available DEMs,ASTER GDEM and SRTM-1. The results indicate that the LCP algorithm in GRASS GIS in combination with the SRTM-1 elevation model resulted in the most realistic river channels when compared to a digitized reference stream. Thus, considering this fact in combination with the convenient tools for further stream network analysis and the opportunity for handling very large data sets, we recommend GRASS GIS in combination with the SRTM-1 elevation model as open source variant for morphometric investigation of large scale river systems.

References

1. Arge L, Chase J.S., Halpin P, Toma L, Vitter J.S., Urban D, Wickremesinghe R (2003): Efficient flow computation on massive grid terrain datasets, *Geoinformatica*, 7(4): 283-313.
2. Baddeley A, Turner R (2005): spatstat: An R package for analyzing spatial point patterns, *Journal of statistical software,* 12(6): 1-42.
3. Bhowmik AK, Metz M, Schäfer RB (2015): An automated, objective and open source tool for stream threshold selection and upstream riparian corridor delineation, *Environmental Modelling and Software,* 63: 240-250.
4. Conrad O (2007): SAGA-Entwurf, Funktionsumfang und Anwendung eines Systems für automatisierte geowissenschaftliche Analysen, *PhD Thesis, University of Goettingen.* [in German]
5. Conrad O, Bechtel B, Dietrich H, Fischer E, Gerlitz L, Wehberg J, Wichmann V, Böhner J (2015): System for automated geoscientific analysis (SAGA) v.2.1.4, *Geoscientific Model Development*, 8: 1991-2007.
6. Ehlschlaeger C (1989): Using the A^T search algorithm to develop hydrologic models from digital elevation data, *Proceedings of International Geographical Information Symposium, 18-19 March 1989, Baltimore, USA, pp.275-289.*
7. Farr TG, Rosen PA, Caro E, Crippen R, Duren R, Hensley S, Kobrick M, Paller M, Rodriguez E, Roth L, Seal D, Shaffer S, Shimada J, Umland J, Werner M, Oskin M, Burbank D, Alsdorf D (2007): The shuttle radar topography mission, *Reviews of Geophysics*, 45: RG2004.
8. Freeman GT (1991): Calculating catchment area with divergent flow based on a regular grid, *Computers and Geosciences*, 17: 413-422.
9. Hayakawa YS, Oguchi T, Lin Z (2008): Comparison of new and existing global digital elevation models: ASTER G-DEM and SRTM-3, *Geophysical Research Letters,* 35: L17404.
10. Hengl T, Grohmann C, Bivand R Conrad O, Lobo A (2009): SAGA vs. GRASS: a comparative analysis of the two open source desktop GIS for the automated analysis of elevation data. *Proceedings of Geomorphometry, 31. August-2.September 2009, Zurich, Switzerland, pp.22- 27.*
11. Holmgren P (1994): Multiple flow direction algorithms for runoff modelling in grid based elevation models: an empirical evaluation, *Hydrological Processes*, 8(4): 327-334.
12. Jarihani AA, Callow JN, McVicar TR, Van Niel TG, Larsen JR (2015): Satellite

derived digital elevation model (DEM) selection, preparation and correction for hydrodynamic modelling in large, low-gradient and data-sparse catchments, *Journal of Hydrology*, 524:489-506.

13. Jasiewicz J, Metz M (2011): A new GRASS toolkit for hortonian analysis of drainage networks, *Computers and Geosciences* 37(8): 1162-1173.

14. Metz M, Mitasova H, Harmon RS (2009): Fast stream extraction from large, radar-based elevation models with variable level of detail, *Proceedings of Geomorphometry, 31. August-*
2.September 2009, Zurich, Switzerland, pp.237-242.

15. Metz M, Mitasova H, Harmon RS (2011): Efficient extraction of drainage networks from massive, radar-based elevation models with least cost path search, *Hydrology and Earth System Science,* 15: 667-678.

16. Neteler M, Metasova H (2008): *Open source GIS. A GRASS GIS approach*, Heidelberg: Springer.

17. O'Callaghan J, Mark D (1984): The extraction of drainage networks from digital elevation data, *Computer Vision, Graphics and Image Processing*, 28: 323-344.

18. R Core Team (2015): R: A language and environment for statistical computing. Vienna, Austria. url: http://R-project.org.

19. Tachikawa T, Hato M, Kaku M, Iwasaki A (2011): Characteristics of the ASTER GDEM version 2, *IEEE International,* 2011: 3657-3660.