European Commission

JRC TECHNICAL REPORTS

Modelling nutrient pollution in the Danube River Basin: a comparative study of SWAT, MONERIS and GREEN models

Anna Malagó[1], Markus Venohr[2], Andreas Gericke[2], Olga Vigiak[1], Faycal Bouraoui[1], Bruna Grizzetti[1], Adam Kovacs[3]

1. European Commission, Joint Research Centre (JRC), Ispra, Italy
2. Leibniz-Institute of Freshwater Ecology and Inland Fisheries, Berlin, Germany
3. International Commission for the Protection of the Danube River (ICPDR), Vienna, Austria

2015

The mouth of the Isar River (Olga Vigiak, September 2014)

EUR 27676 EN

Joint Research Centre

JRC Science Hub
https://ec.europa.eu/jrc

JRC99193

EUR 27676 EN

ISBN 978-92-79-54239-8 (PDF)
ISBN 978-92-79-54240-4 (print)

ISSN 1831-9424 (online)
ISSN 1018-5593 (print)

doi:10.2788/156278 (online)

Printed in Italy

How to cite: Anna Malagó, Markus Venohr, Andreas Gericke, Olga Vigiak, Faycal Bouraoui, Bruna Grizzetti, Adam Kovacs, 2015. Modelling nutrient pollution in the Danube River Basin: a comparative study of SWAT, MONERIS and GREEN models. EUR 27676 EN, doi:10.2788/156278.

Contents

Extended abstract

The Water Framework Directive requires the development and implementation of river basin management plans for improving the ecological status of freshwater bodies throughout Europe. The scientific community supports this process by developing decision-support tools for identifying the principal sources of water pollution. Integrated Basin Models provide a transparent framework to assess the dominant processes affecting nutrient emissions pathways and retention. Models, however, are imperfect representations of the real world, and are conditioned by structural uncertainty, implicit in the description of biophysical processes and data uncertainty, implicit in the environmental data upon which the models were developed. Hence, decision makers must plan management actions on the basis of the best available, however still incomplete, knowledge. The comparison of independent assessments may offer insights that are useful for decision-making, e.g. for identifying knowledge gaps, identifying data uncertainties, consolidating investigation results, and increasing stakeholders' acceptance.

The Danube River is the second largest and most international river of Europe. Its basin covers approximately 803,000 km^2 of Central and South-Eastern Europe and is shared by 19 countries. Within the context of fostering scientific collaboration in the Danube region and under the auspices of the International Commission for the Danube River Protection (ICPDR), three independent model-based assessments of nutrient pollution in the Danube River Basin were compared with the objective of reaching a shared appraisal of nutrient pressures and drivers in the Basin. The three models (SWAT, MONERIS and GREEN) are used world-wide to assess nutrient pollution. Considering that they differ in their structural complexity and data requirement, the comparison focused on model outputs. Annual water discharge (Flow, m^3/s) as well as annual loads of total nitrogen (TN, ton/y) and total phosphorus (TP, ton/y) were compared at the outlet of 18 ICPDR regions for the decade 2000-2009. For each region, mean annual values, correlation, standard deviation, and root mean square error of model simulations were analysed.

Results showed that water discharge was simulated well by all models. Lack of TN observations limits conclusions on the comparison of TN loads and only four regions had available observations (Delta, Pannonian Danube, Drava and Lower Danube). Models were generally in good agreement with each other, but differences were found for the Morava, Tisa, and Lower Danube tributaries of Romania. SWAT and MONERIS had comparable long mean annual TP loads, but differed for amplitude and phases; while GREEN generally overestimated TP loads.

In conclusion, good water discharge simulations across the basin confirmed that hydrology was correctly represented in all models. The nutrients comparison revealed for some assessment regions the need to intensify monitoring especially for TN concentrations at the region outlets, and TP in the Middle and Lower Danube. Despite differences in model approaches and input data, the three assessments were coherent, and all three models may be confidently used for river basin management of the region.

Modelling nutrient pollution in the Danube River Basin: a comparative study of SWAT, MONERIS and GREEN models

1. Introduction –Models inter-comparison

Challenges in reducing nutrient pollution

Most European rivers suffer significant nutrient pollution. The Water Framework Directive (WFD; European Commission, 2000) requires regional authorities to develop and implement river basin management plans to improve current conditions. River basin management plans are developed with the involvement of stakeholders to design appropriate remediation actions, based on the analysis of the socio-economic and biophysical river basin characteristics. The scientific community supports the process by providing information on nutrient fluxes and pathways within the river basin and in developing decision support tools such as Integrated Basin Models to help identifying cost-effective strategies to reduce water pollution.

Integrated Basin Models provide a transparent framework to address the dominant processes that affect nutrient sources and pathways, as well as their multiple interactions. They capture the current understanding of biophysical processes that govern the transport of nutrients from land to rivers and to the sea. At the same time, they allow assessing the potential impacts of envisaged management plans. Given the complexities of real world systems, Integrated Basin Models can only provide a partial representation of the biophysical processes, capturing the dominant trends as inferred from observation data. Observed nutrient data usually consist of concentration measured at monitoring stations in rivers. By definition they provide local information, representing at the same time an integral of the entire upstream catchment, and are assumed to be representative for the considered period (month, year). Additionally, data of the main environmental factors upon which models simulate the biophysical processes, such as land use/coverage, climate, geology, soil, and topography are of limited spatio-temporal resolution, extent and accuracy. Any Integrated Basin Model is thus suffering a structural uncertainty, i.e. the representation of the main processes and interactions, and a data uncertainty, i.e. implicit in the environmental data used to feed them or to calibrate/evaluate their outputs. Beyond that, as a result of differences in the process representations, included in the models and in the data used to create them, different Integrated Basin Models applied to the same basin may differ in the type and resolution of their outputs, and in the partition of nutrient sources and pathways.

Since both data and models are affected by uncertainty, decision makers have to plan management actions on the basis of incomplete knowledge of the systems. Comparison of results from different Integrated Basin Models however, may provide further insights to help decision-making, for example in pointing out knowledge gaps or consolidated assessments and increase acceptability by end-users (Grizzetti et al., 2015).

The Danube River Basin context

The Danube River Basin is the second largest river basin in Europe, covering approximately 803,000 km^2 of Central and South-Eastern Europe. In the year 2015, 19 countries are sharing the catchment, 14 of which are called 'Danube countries' (with catchment areas >2000 km^2). The biggest shares of the catchment belong to Romania (30%) followed by Hungary, Serbia, and Austria (around 10%) (Habersack et al., 2013; Fig. 1). Due to its vast area and its topography ranging from lowlands to mountains above 3,000 m a.s.l., the Danube River Basin exhibits a pronounced climatic variability. The western region is influenced by the Atlantic climate, whereas the eastern region is characterised by a continental climate leading to lower precipitation and typically colder winters. The mean annual precipitation for the whole Danube basin for the period 1980 to 2009 was 597 mm y^{-1}, ranging from 220 mm y^{-1} near the outlet of the river to 1510 mm y^{-1} in the Alps (Pagliero et al., 2014). The mean annual temperature for the period was 9.7°C, ranging from 0.8 to 13°C. The mean annual discharge at the outlet was estimated of about 6387 $m^3 s^{-1}$ (Pagliero et al., 2014). More than 83 million people live in the Danube River Basin. They have a significant impact on its environment settings, which results in significant pressures on water quality and quantity (i.e., organic emissions, nutrient emissions, hazardous substance emissions, and hydromorphological pressures). The Danube River Basin mainly consists of forest (35%), arable land (34%), and grassland (17%).

The Danube River can be divided into four main sections: the Upper, Middle, Lower Danube, and Danube Delta (Habersack et al., 2013). The Upper Danube basin reaches from the sources in the Black Forest Mountains to the Gate of Devın, near Bratislava. The Middle Danube basin covers a large area reaching from the Gate of Devın to the impressive gorges of the Danube at the Iron Gate (Iron Gate I and Iron Gate II), which divides the Southern Carpathian Mountains in the north and the Balkan Mountains in the south. The Middle Danube basin is confined by the Carpathians in the north and the east, and by the Karnic Alps, the Karawankas, the Julian Alps, and the Dinaric Mountains in the west and south. The Lower Danube basin covers the Romanian–Bulgarian Danube sub-basin downstream of Cazane Gorge and the sub-basins of the Siret and Prut River. It is confined by the Carpathians in the north, by the Bessarabian Upland Plateau in the east, and by the Dobrogea and Balkan Mountains in the south. Finally, the Danube Delta from the confluence of the Prut River (Ukraine) to the mouth into the Black Sea (Ukraine) covers an area of 5,640 km^2.

Fig. 1. The Danube Basin

The overarching legal instrument regulating co-operation in the matter of transboundary water management in the Danube River Basin is the Danube River Protection Convention. It was signed by 11 of the Danube Riparian States (Austria, Bulgaria, Croatia, the Czech Republic, Germany, Hungary, Moldova, Romania, Slovakia, Slovenia, and Ukraine) and the European Community, and came into force in 1998. The convention aims to ensure that surface waters and groundwater within the Danube River Basin are managed and used sustainably and equitably. The International Commission for the Protection of the Danube River (ICPDR) is the transnational body established to implement the Convention. The ICPDR adopts a river basin approach for water management, fostering close international cooperation between all the countries within the basin, and bringing together upstream and downstream stakeholders. In compliance with the WFD, the ICPDR coordinates the development of the Danube River Basin Management Plans. The first plan was published in 2009 and established Basin management for the period 2009-2015. The second plan has been developed through 2015, and will be adopted by February 2016.

The EU Strategy for the Danube Region (EUSDR; European Commission, 2010) is a macro-regional strategy developed by the European Commission together with the Danube Region countries. It provides a framework for policy integration and coherent development of the Danube Region. The strategy addresses four priority areas: connecting the region; protecting the environment; building prosperity; and strengthening the capacities.

Nutrient pollution has been estimated to be responsible for 65% of the Danube River length not achieving good standards required by the WFD (ICPDR, 2009). The independent research institute

IGB (Leibniz-Institute of Freshwater Ecology and Inland Fisheries) supports ICPDR by applying and improving the integrated basin model MONERIS to assess nutrient fluxes in the River Basin District (RBD). The JRC supports the EUSDR by addressing scientific needs and strengthening collaboration. In this context, the JRC has developed independent Integrated Basin Models and nutrient assessments.

Aim of the work

Within the context of fostering scientific collaboration in the Danube region, an inter-comparison of independent assessments was envisaged to identify opportunities and knowledge gaps in the region. The overall objective of the inter-comparison was to reach a shared and robust assessment of nutrient pressures and the drivers in the Danube River Basin to strengthen water management planning.

We compared assessments of water nutrient pollution in the Danube River Basin as estimated by three independent models. Accepting that the three models differed in structural complexity and data requirements, the comparison focused on model outputs of water balance, nitrogen and phosphorus fluxes. The analysis aimed at identifying regions where models agreement indicated congruent baseline assessments, and regions where models disagreement indicated uncertainty in the assessments, where more investigation would be warranted.

2. Methodology

Scope and approach of the inter-comparison

Model comparison exercises are not new in the scientific literature. For example, several Integrated Basin Models using the same data, whenever possible, have been compared in the project EUROHARP (Silgram et al., 2009). Unlike previous works, in this study the objective of the inter-comparison was not on the models, but on model outputs to gain insights useful from a management perspective. The comparison thus focused on the quality of the model results, including indications on expected model performances and pointing out the strengths and weaknesses in the region (Grizzetti et al. 2015; Thunis et al., 2011).

The spatial units selected for conducting the comparison were the 18 ICPDR main regions (Fig. 2) along the Danube River Basin itself, excluding the littoral regions for which no gauging station was available. The period for the comparison was the decade 2000-2009, which was considered long enough to allow for being relevant to management planning, representative of current conditions, and was covered in total or at least in part by all three models (Fig. 3). Model outputs considered in the comparison were annual water discharge (m^3/s), total nitrogen (TN, ton/y), and total phosphorus (TP, ton/y) loads at the outlet of each ICPDR region.

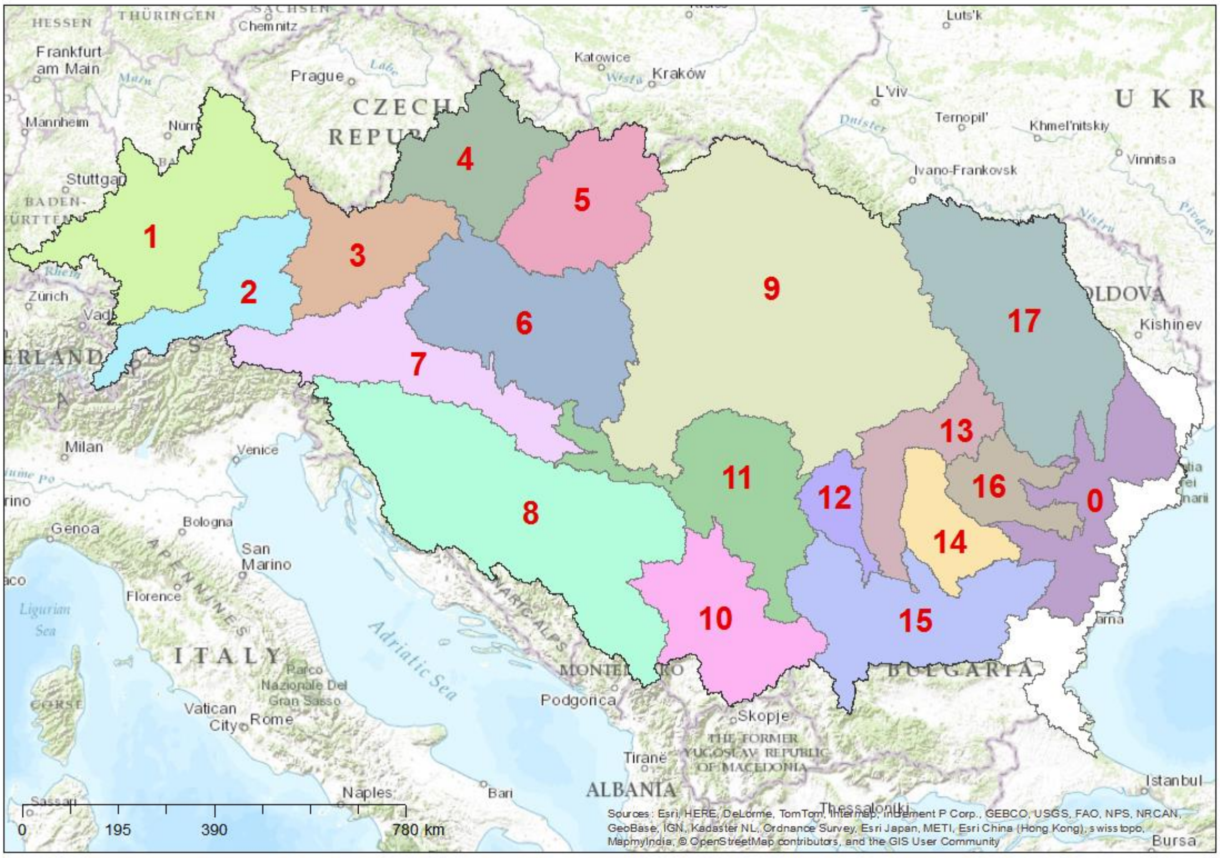

Fig. 2. Map of the 18 selected regions in the Danube Basin

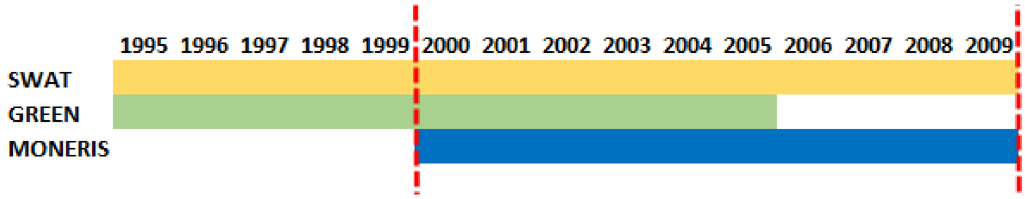

Fig. 3.Temporal scale of model simulations. The red lines define the period of interest for this study (2000-2009).

The evaluation was conducted using two different types of graphics: bar plots with 95% confidence intervals (CI) and Taylor diagrams (Taylor, 2001). Bar plots with confidence intervals were used to compare observed and simulated mean annual values of discharge, TN and TP for the period 2000-2009. The 95% confidence interval was computed by multiplying the standard error of the mean (i.e. the ratio of standard deviation to the square root of the sample size) by 1.96. The value of 1.96 is based on the fact that 95% of the area of a normal distribution is within 1.96 standard deviations of the mean.

Taylor diagrams allow a combined presentation of different statistical indicators, such as the Pearson's correlation coefficient R, the centred root-mean-square error RMS[1] between observations and models, and the standard deviation of observed and models samples. Fig. 4 is an example showing how to interpret it. The standard deviation and RMS error are on the correspondent arcs, whereas the correlation coefficient is on the circumference. The reference point (red dot) is always collocated on the x-axis, whereas the model point moves on the arcs of the quadrant. The position of a model (black dot) on the plot quantifies how closely the model outputs match the reference: the closer the black dot is to the red dot, the better the model performance.

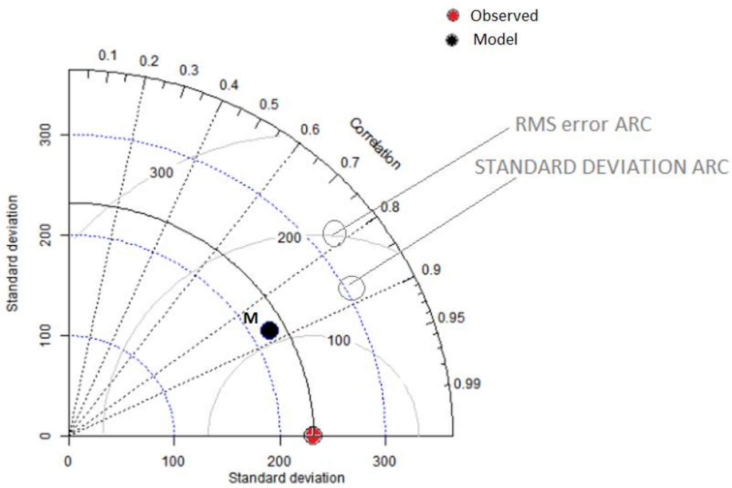

Fig. 4. Example of a Taylor diagram

Observations of discharges and nutrient loads were not available at all region outlets. Table 1 summarizes the monitoring stations for which observed data of at least 5 out of 10 years (2000-2009)

[1] In this document the RMS is the Centered Root Mean Square error calculated as described in Taylor (2001). It was calculated based on the assumption that the Centered Root Mean Square Error is a function of standard deviation of reference, standard deviation of test, and the correlation coefficient between the test and reference fields.

were available. Fig. 5 shows the location of the gauging stations. Mean annual discharge was computed from daily discharge data. Nutrient loads used as observations were estimated from nutrient concentration and daily flow, and calculated based on different methods: flow weighted concentrations method proposed by Moatar and Meybeck (2005) as well as the ICPDR method described in the ICPDR yearbooks (ICPDR, 2000-2009) and literature estimations (Gils, 2004).

Water discharge data near the outlet were available for most regions, except for Region 5 (Vah-Hron-Ipel), Region 8 (Sava), Region 9 (Tisa), Region 10 (Velika Morava), Region 14 (Arges-Vedea), and Region 16 (Buzau-Ialomita). Estimates of TN loads from gauging stations were available only for Region 0 (Delta), Region 6 (Pannonian Danube), Region 7 (Drava) and Region 15 (Middle Danube). In contrast, estimates of TP loads were available in most regions. In order to include more observations in the analysis we have considered the stations Dravaszabolcs in region 7 (Drava Basin) as the outlet of region because the drainage area of station Dravaszabolcs (about 37,500 km^2) is very close to the drainage area of the whole Drava river (about 39,700 km^2). Similarly, for Region 6 (Pannonian) and Region 8 (Sava) the station Hercegszanto and Sremska Mitrovica respectively were used as references for the correspondent outlet. For Region 17 (rivers Siret and Prut), we included the Buzau basin (5,240 km^2), a tributary of river Siret.

Where no observed data was available, the arithmetic means of estimates of the three models were used as reference solely for understanding similarities and differences between the models. Henceforth, we use the abbreviation AVG for the arithmetic means of the models. In the bar plots, the AVG is shown in grey, and observations in black.

Table 1. ICPDR regions code, names, drain area (km²), and gauging stations of reference used in the inter-comparison. The term "available" indicates that observed data for the variable were available for at least five years in the decade 2000-2009.

Regions	Area (km^2)	Name of Regions	Station Name	WATER DISCHARGE	TN LOADS	TP LOADS
0	802,032	DELTA	Reni Chilia	available	available	available
1	49,769	Upper Danube	Upstream Achleiten station (excluding the Passau Ingling)	available		
2	25,999	Inn	Passau Ingling	available		available
3	101,803	Austrian Danube	Wien-Nussdorf	available		available
4	26,628	Morava	Zahorska Ves	available		available
5	30,589	Vah-Hron-Ipel				
6	211,103	Pannonian Danube	Hercegszanto	available	available	available
7	39,679	Drava	Dravaszabolcs	available	available	available
8	100,102	Sava	Sremska Mitrovica	available		
9	149,567	Tisa				
10	37,702	Velika Morava				
11	582,414	Middle Danube	Pristol/Novo Selo harbour	available		available
12	10,333	Jiu	Zaval	available		
13	23,841	Olt	Izbiceni reservoir	available		
14	18,118	Arges-Vedea				
15	685,320	Lower Danube	Chiciu/Silistra	available	available	available
16	16,358	Buzau-Ialomita				
17	71,490	Siret-Prut-Buzau	Sendreni and Giurgiulesti	available		available

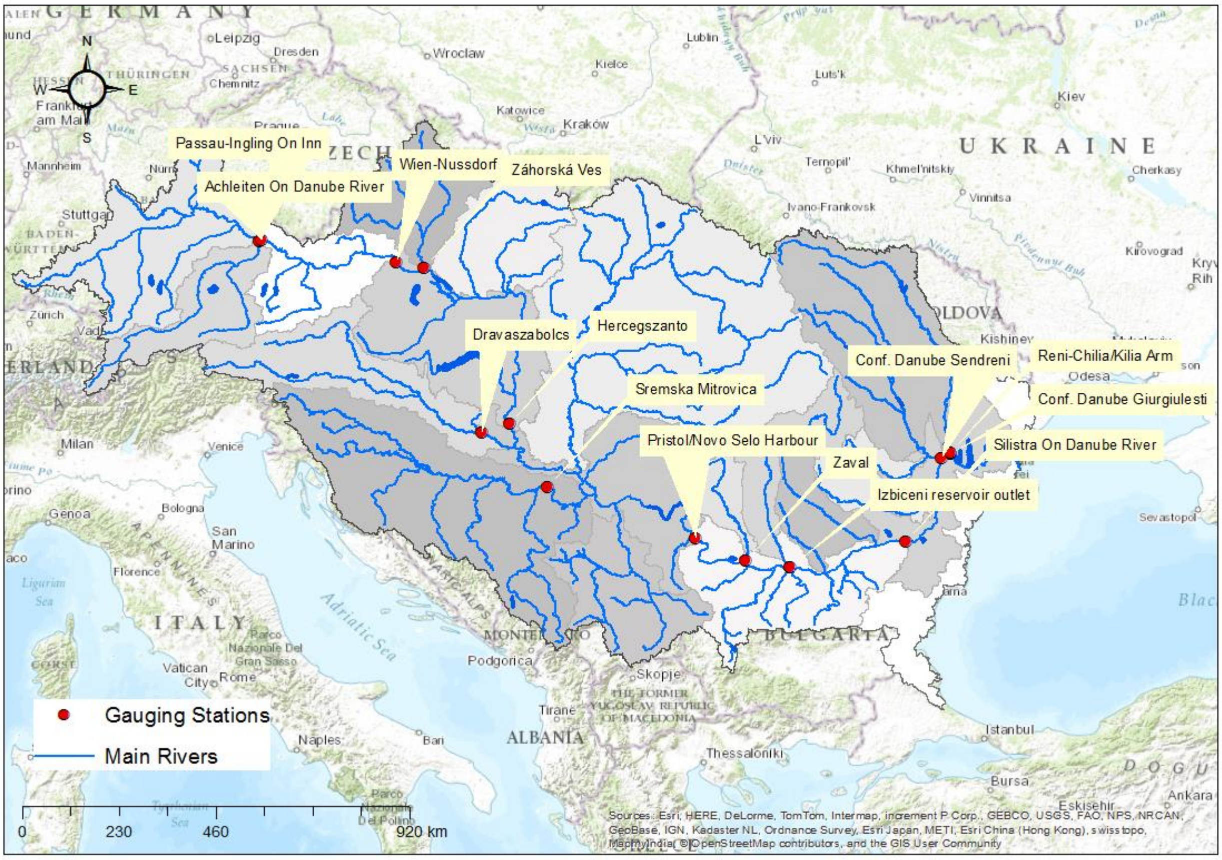

Fig. 5. Gauging stations used in this study.

Modelling tools

There is a large number of models used to assess nutrients loads in river basins. They vary in process descriptions, spatial and temporal scale and data requirements. In this study two conceptual models, GREEN (Grizzetti et al., 2008) and MONERIS (Venohr et al., 2011) were compared with the process-based model SWAT (Arnold et al., 1998). They are among the most commonly used models to assess diffuse and point source nutrient pollution. A comparison of the models for assessing remediation measures for nitrogen water pollution have been explored in Bouraoui and Grizzetti (2014). Here we provide a shorth description of each model, focusing on (i) processes representation; (ii) the spatial and temporal scale; and (iii) the main characteristics in the application to the Danube River Basin.

Importantly, in this study we considered the outcomes of the model GREEN in the application at the European scale presented in Grizzetti et al. (2012), where the model was calibrated by European macro regions and not specifically for the Danube basin, while SWAT and MONERIS models have been set-up and calibrated (partly) for the Danube River Basin. Further, it is important to note that the three models used different datasets during the calibration process.

SWAT

The Soil and Water Assessment Tool (SWAT, Arnold et al., 1998) is a physically based, semi-distributed, and long-term continuous watershed scale model that runs on a daily time step to predict the impact of climate, landuse, soil type, topographic characteristics, and land management practices on hydrology, sediment and nutrients in large ungauged watersheds. The basis of spatial resolution is the Hydrologic Response Unit (HRU) in which water, sediment and nutrient losses are calculated. The responses at HRU level are aggregated at the sub-basin level and routed to the sub-basin outlet through the channel network.

SWAT daily water balance considers actual evapotranspiration, canopy interception, plant transpiration, soil evaporation, surface runoff, and vertical water movement in the unsaturated soil zone to the aquifer. SWAT simulates nitrogen and phosphorous cycling and losses to the stream network in various forms (dissolved and particulate). Nutrient pathways include surface runoff, sediment, tile drainage, and aquifer. The model requires a cascade calibration approach (water-sediments-nutrients) of sensitive parameters, grouped on the basis of underpinning processes (Malago' et al., 2015; Vigiak et al., 2015a).

The SWAT model of the Danube River Basin was built using the 2012 version (SWAT2012). The Danube River Basin area of 833,908 km^2 was subdivided in 4,663 sub-basins (with an average area of 180 km^2) and 5,181 HRUs. The main input data are described in Pagliero et al. (2014). Improvements from Pagliero et al. (2014) version included revised climate and landuse inputs, and inclusion of current best management practices, such as tile drainage systems and riparian filtering (Malagò et al. forthcoming; Vigiak et al., 2015b). The calibration procedure of water discharge was improved by extending the calibration and validation gauging station network and period. Calibration and validation of nutrient concentrations and loads (nitrates, total nitrogen and total phosphorous) for the period 1995-2009 at annual and monthly time step is described in Malago' et al. (forthcoming). For the inter-comparison, SWAT outputs at the outlet of the 18 regions were aggregated at annual time step.

MONERIS

The model Modeling Nutrient Emissions into River Systems (MONERIS, Behrendt et al., 2000; Venohr et al., 2011) is a semi-empirical, semi-distributed, steady-state model for monthly, annual and long-term average nutrient emissions in river basins (Behrendt et al., 2000; Venohr et al., 2011). The model considers point emissions and several pathways and sources of nutrients, and takes into account the retention and transformation of nutrients in soils, groundwater, rivers and lakes for calculating nutrient loads. MONERIS aims at a moderate demand of input data, a short computing time, and offers an easy application to large river basins.

While the empirical approaches for the pathways included in MONERIS have been developed and calibrated independently, the in-stream retention and remobilisation is calibrated against observation data (i.e. nutrient loads) and the output of other models. Runoff is derived via a simplified runoff distribution procedure, considering precipitation, evaporation, water withdrawal/addition and it is validated against obsereved data. For each pathway, the flow components and nutrient concentrations are modelled. In MONERIS, groundwater discharge comprises the natural interflow and base flow (Venohr et al., 2011) and it is calculated as the residual of total flow and all other flow components.

MONERIS uses "analytical units" as smallest modelling unit based on topography, hydrological catchments and administrative units. In this study, the Danube River Basin was subdivided into 1,578 units with an average area of 510 km^2. It was also the first monthly application of MONERIS for the Danube River Basin (Venohr et al., 2015).

GREEN

GREEN is a conceptual statistical model that consists of a regression equation based on spatially referenced data (Grizzetti et al., 2008; 2012). The model estimates annual nitrogen and phosphorus loads in surface waters. It uses a routing structure to establish the emitting-receiving sub-basins relationship. It considers two different pathways of nutrient transfer from sources to the basin outlet: diffuse and point sources. GREEN does not simulate nutrient cycles but calculates the diffuse emissions (from land to the streamflow) as sum of all diffuse sources reduced by a basin reduction factor that takes into account the retention in the soil, aquifers and the nitrogen removed by plants. Finally, it estimates the loads in the river considering a river reduction factor (river retention). The model requires the calibration of only two parameters of retention. The water discharge in the rivers was calculated outside GREEN using the Budyko framework approach (Grizzetti et al., 2012). The basis of spatial resolution is the sub-basin, the same used in SWAT (4,663 sub-basins), with an average area of about 180 km^2.

Since the models were run for different periods, in this report results focused on the 2000-2009 decade (Fig. 3), however results for the more extended periods (1996-2005 and 1995-2009) are reported in **Appendix A**. In addition, in **Appendix B** the statistics represented in the Taylor diagram (with also the average of observations and model results in the period 2000-2009) were reported for each model and variable.

3. Results – Comparing assessments of nutrient water pollution in the Danube

Water Discharge

The comparison between observed and simulated long-term mean annual discharge (m^3/s) in the Danube River Basin is shown in Fig. 6. The map highlights the strong correlation between the models and the good agreement with the observations. Regional outputs are reported in the bar plots in Fig. 7 (a-r).

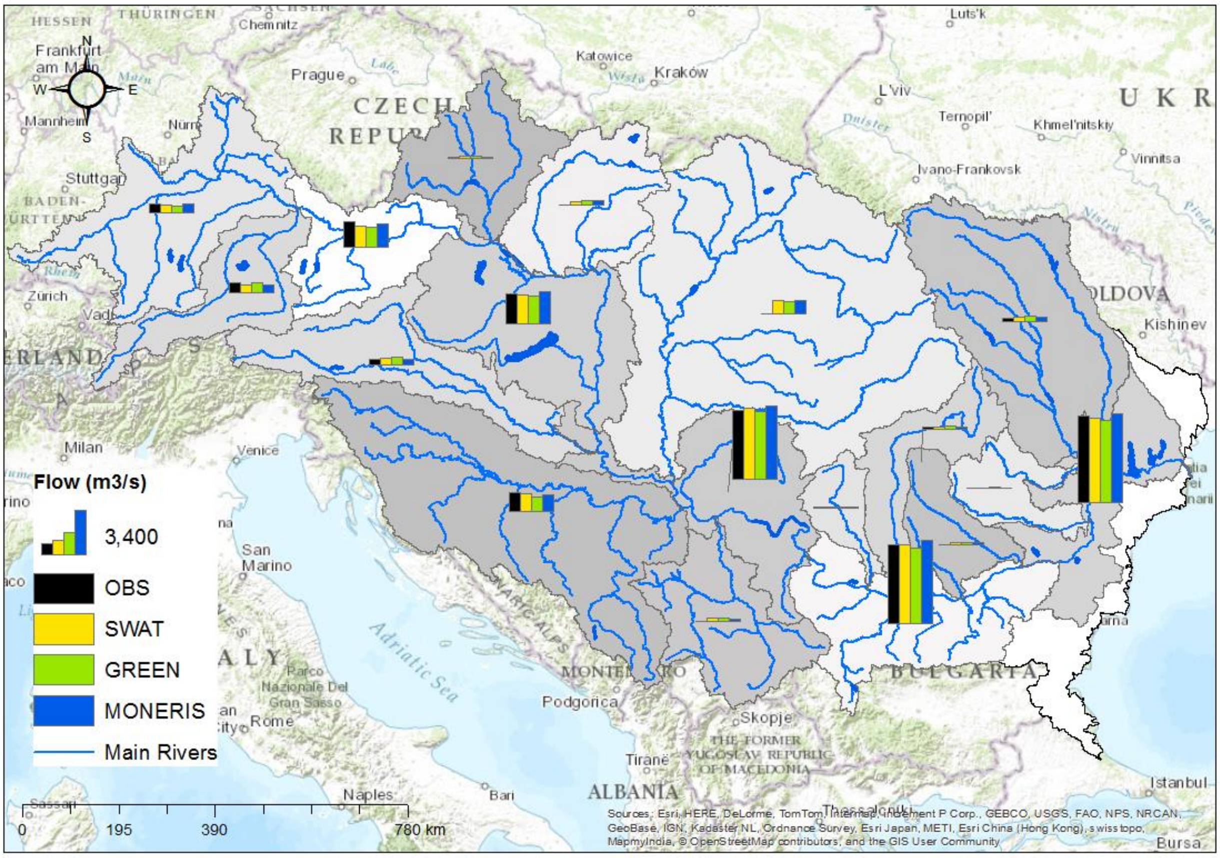

Fig. 6. Map of the mean annual water discharge (period 2000-2009) in the Danube Basin. The three model outputs are compared to available observations.

The Taylor diagrams, however, highlight some important differences and similarities between the models (Fig. 8 a-r). In *Region 0* (mouth of Danube at Reni Chilia (Fig. 8 a) MONERIS generally agreed best with the observations with the smallest RMS error (around 144 m^3/s), the highest correlation estimated equal 1, and a standard deviation of 1,103 m^3/s, very close to the standard deviation of observations (1,211 m^3/s) indicating similar annual variations in observed and modelled streamflow. SWAT had the same standard deviation of MONERIS (1,061 m^3/s), high correlation with observations (0.84), but higher RMS error (656 m^3/s). GREEN resulted in a slightly lower coefficient of correlation (around 0.81), a large RMS error of 1,102 m^3/s and the highest standard deviation. These results highlight the different methodologies used in the calibration process. While for MONERIS monitoring stations evenly distributed in the entire Danube River Basin were considered for the streamflow calibration, SWAT calibrated only the headwaters (10% for the whole) and used extrapolation techniques to extend the calibrated parameter to the whole river basin.

In *Region 1* (Upper Danube, Fig. 7 b) GREEN and SWAT slightly underestimated the long mean annual water discharge, whereas MONERIS agreed best with the observations. The Taylor diagram (Fig. 8 b) shows that GREEN and MONERIS had the correct standard deviation of 151 and 136 m^3/s respectively and very high correlation with the observations (0.78 and 0.92). GREEN, however, had a slightly higher RMS error (97 m^3/s) than MONERIS (56 m^3/s). SWAT RSM error was close to that of MONERIS (around 64 m^3/s), but SWAT overestimated the standard deviation. As a consequence, the amplitude of variations of annual discharges was simulated better by MONERIS and GREEN than SWAT.

Fig. 8 c-d-e confirm the strong correspondence between MONERIS and the observations in *Region 2* (Inn Basin), *3* (Austrian Danube) and *4* (Morava), followed by SWAT and then GREEN. In particular, in *Region 2* (Fig. 8 c) MONERIS and SWAT generally agreed well with observations, each with about the same RMS error (50 and 58 respectively). GREEN scored a higher RMS error (84 m^3/s) and lower correlation (around 0.68), but the standard deviation was similar to that of MONERIS (around 111 m^3/s) and observations (standard deviation of observations was around 93 m^3/s). In *Region 3*, the three models had very high correlation with the observations (Fig. 8 d). MONERIS simulated best the annual variations of discharge (standard deviation of 271 m^3/s, very close to the standard deviation of observations that was a 288 m^3/s) with the lowest RMR error. In *Region 4* MONERIS performed better than SWAT and GREEN, because it laid relatively close to the reference point of observations (Fig. 8 e). Similarly to MONERIS, SWAT had high correlation with observation (0.88), but the standard deviation was larger (around 37 m^3/s) than that of the observations (21 m^3/s). As well, GREEN standard deviation was larger than that of observations, but the correlation was the lowest among the models (around 0.56).

In *Region 5* (Vah-Hron-Ipel, Fig. 8 f) the three models had about the same correlation and RMS errors, but SWAT and MONERIS simulated the annual variability better than GREEN with respectively a standard deviation of 50 and 41 m^3/s compared to the reference value of 46 m^3/s.

In *Region 6* (Pannonia Danube, Fig. 8 g) SWAT water discharge agreed best with observations, in terms of both error and annual variability (with a standard deviation 373 m^3/s compared to observed value of 382 m^3/s) and a relatively small RMS error. MONERIS had about the same RMS error of SWAT, but higher standard deviation. In addition, the correlation coefficients of predictions with observations of SWAT and MONERIS were higher than GREEN.

In *Region 7* (Drava) GREEN overestimated the long mean annual water discharge, whereas SWAT and MONERIS were close to observations (Fig. 7 h). The Taylor diagram (Fig. 8 h) confirms that GREEN had the lowest correlation with observations and the highest RMS error (about 67 m^3/s) among the three models.

In *Region 8* (Sava, Fig. 8 i) SWAT and MONERIS had the same annual variability of the observations (the standard deviation was around 256 and 229 m^3/s compared to 204 m^3/s of observed value). GREEN had about the same RMS error of MONERIS but higher standard deviation, highlighting an overestimation of the amplitude of annual discharge variations.

Similar results were found in *Region 9* (Tysa, Fig. 8 j). SWAT and MONERIS agreed well with reference values: they had high correlation coefficients and same standard deviation (291 and 280 m^3/s respectively), very close to the observed value (273 m^3/s). GREEN had the highest correlation with the observations; however, the standard deviation was overestimated, highlighting the largest amplitude of annual discharge variations.

For *Region 10* (Velika Morava) SWAT and MONERIS described better the reference values (Fig. 8 k), but while SWAT slightly overestimated the amplitude of the annual discharge variations (standard deviation 72 m^3/s), MONERIS slightly underestimated it (54 m^3/s). GREEN had the highest standard deviation (102 m^3/s).

In *Region 11* (Middle Danube, Fig. 7 l and Fig. 8 l) MONERIS and SWAT agreed well with observations, but SWAT had a slightly higher RMS error than MONERIS (494 m^3/s of RMS error for SWAT compared to 208 m^3/s of MONERIS). GREEN overestimated the annual variations of water discharge, with higher RMS error around 878 m^3/s and a smaller correlation coefficient.

In *Region 12* (Jiu, Fig. 7 m and Fig. 8 m) MONERIS simulated better the long mean annual discharge than SWAT and GREEN, which slightly underestimated the observations. MONERIS and SWAT simulated correctly the variability of observations as measured by the standard deviation, even though SWAT had a higher RMS error. GREEN overestimated the standard deviation of observations with the highest RMS error.

In *Region 13* (Olt, Fig. 7 n and Fig. 8 n) MONERIS and SWAT simulated better than GREEN the long mean annual discharge. In particular, GREEN overestimated the water discharge and its annual variability, with a standard deviation (106 m^3/s) that was about twice the observed one (59 m^3/s).

In *Region 14* (Arges-Vedea) SWAT agreed best with the reference values showing a high correlation coefficient (0.97), low RMS error (13 m^3/s), and standard deviation very close to the reference value (57 m^3/s compared to the observed standard deviation of 55 m^3/s; Fig. 7 o and Fig. 8 o). MONERIS had about the same performances of SWAT, but had a slightly larger RMS error (20 m^3/s). GREEN's RMS error was about 30 m^3/s, and the standard deviation was higher (82 m^3/s) than that of observations. All three models had high correlation coefficients.

In *Region 15* (Lower Danube) the long mean annual values of the models were very close to observations (Fig. 7 p). MONERIS resulted in the lowest RMS error (around 150 m^3/s; Fig. 8 p) and the annual variability was close to observations. SWAT simulated correctly the amplitude of annual variations (the standard deviation was estimated around 934 m^3/s compared to the observed value of 985 m^3/s) but had a larger error than MONERIS (538 m^3/s). GREEN overestimated the observed annual variation, and its RMS error was the highest.

In *Region 16* (Buzau-Ialomita) GREEN largely overestimated the long mean annual discharge (Fig. 7 q). The Taylor diagram (Fig. 8 q) confirms the overestimation; the GREEN point falls outside the diagram due to the large standard deviation (75 m^3/s). SWAT and MONERIS results were similar, with low RMS error (around 14 and 19 m^3/s respectively) and similar standard deviation (28 and 31 m^3/s respectively).

In *Region 17* (Siret-Prut-Buzau), SWAT and MONERIS agreed well with the observations in terms of long mean annual discharge and performance statistics (Fig. 7 r and Fig. 8 r), with both models scoring small RMS error (in the range of 65-85 m^3/s) and simulating well the amplitude of mean annual variability, with high correlation coefficients. GREEN was not able to reproduce correctly the annual variability, resulting in the highest RMS error and the lowest correlation coefficient (0.63).

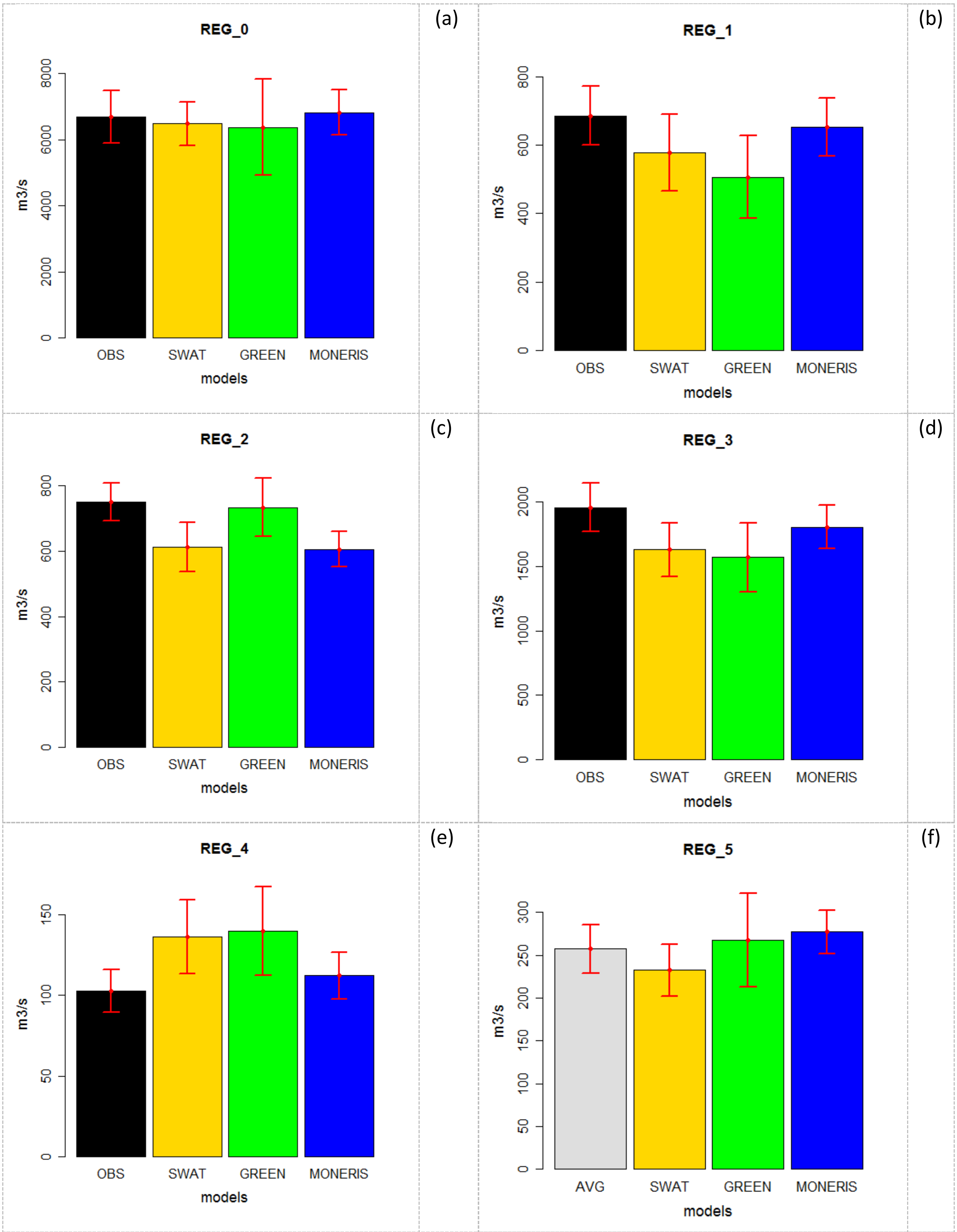

Fig. 7. a-f. Bar plots of long mean annual water discharges (period 2000-2009). The grey bar represents the average of the three model predictions as observations were not available. The error bars in red indicate the 95% confidence intervals.

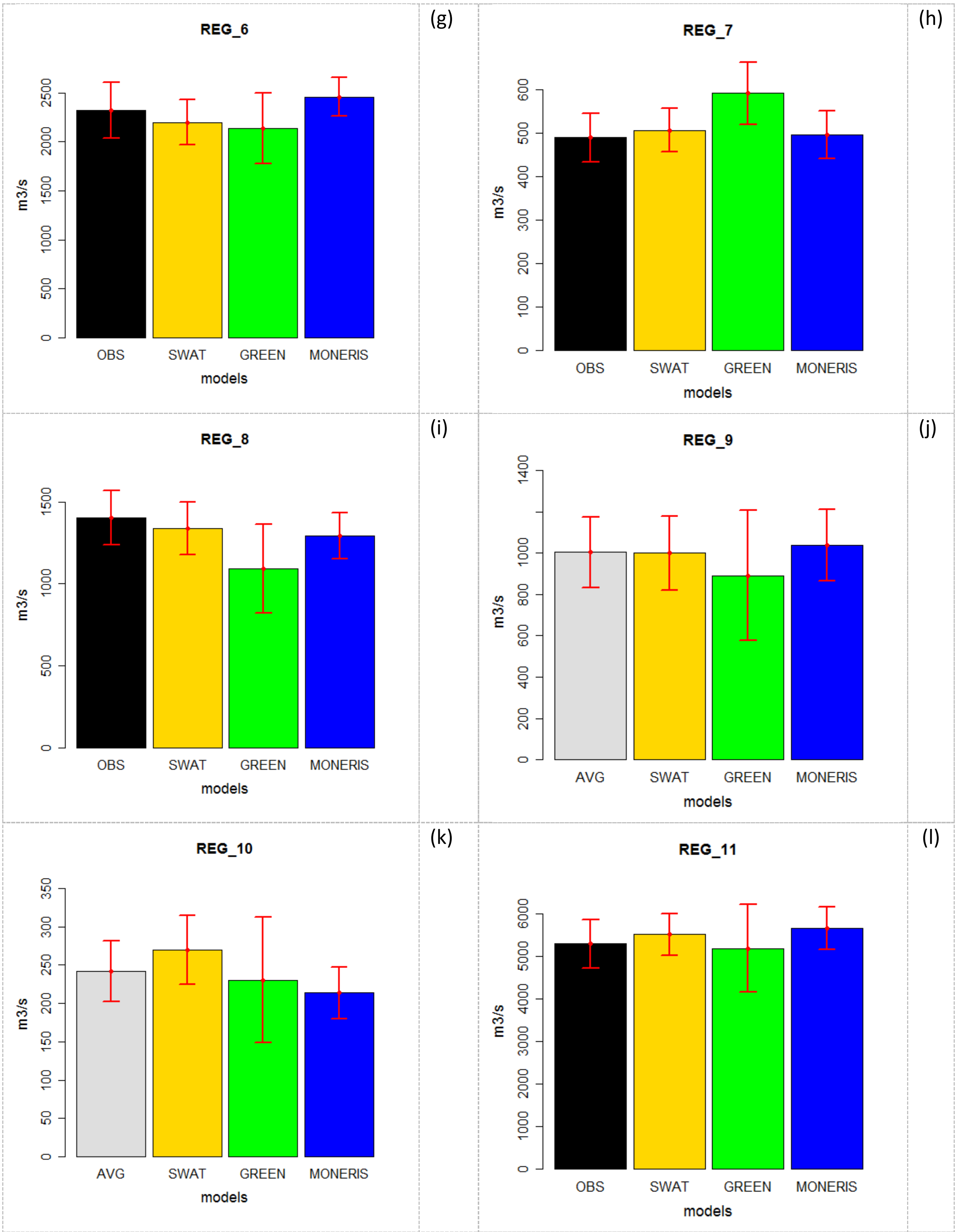

Fig.7. g-l. Bar plots of long mean annual water discharges (period 2000-2009). *The grey bar represents the average of the three model predictions as observations were not available. The error bars in red indicate the 95% confidence intervals.*

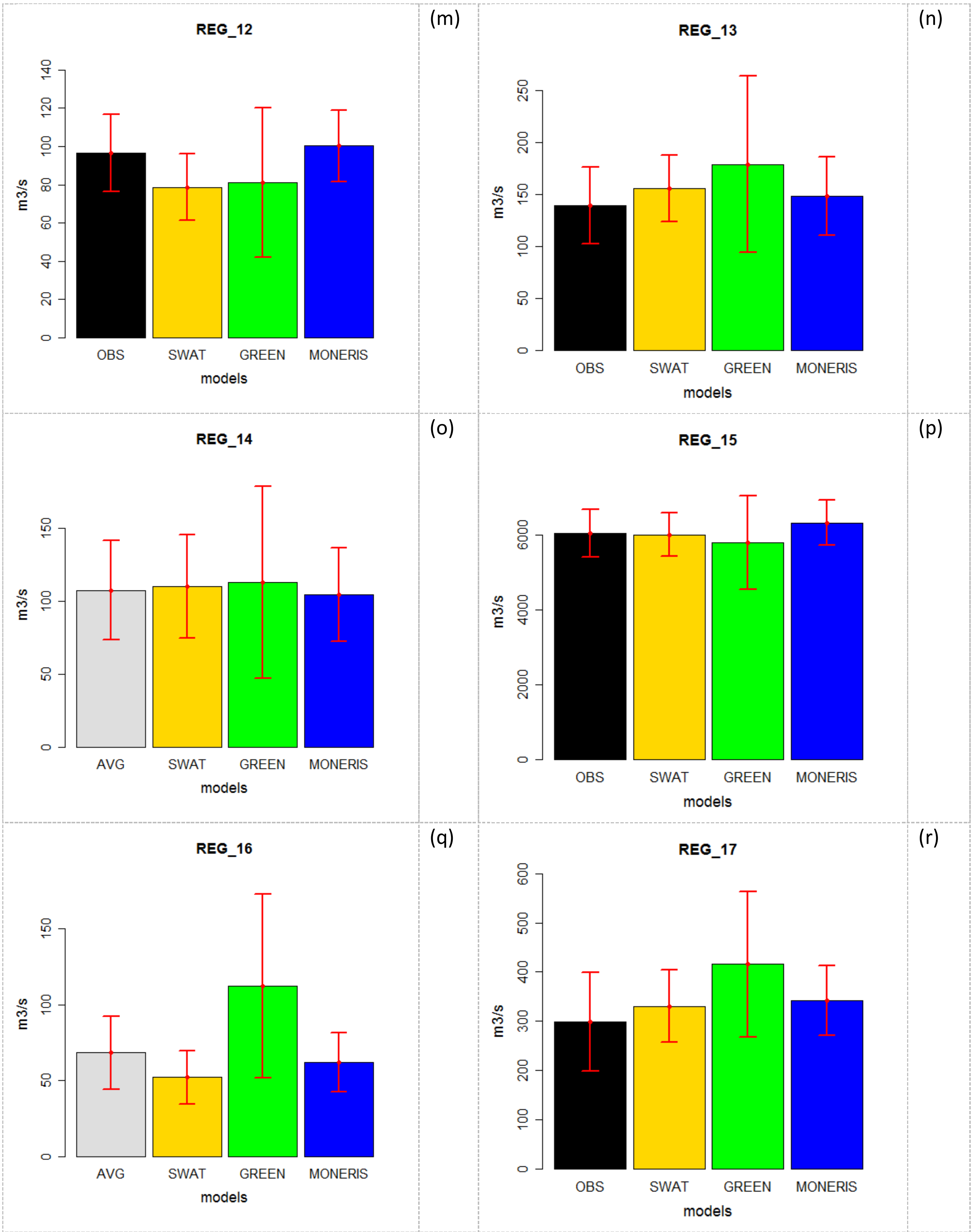

Fig.7. m-r. Bar plots of long mean annual water discharges (period 2000-2009). *The grey bar represents the average of the three model predictions as observations were not available. The error bars in red indicate the 95% confidence intervals.*

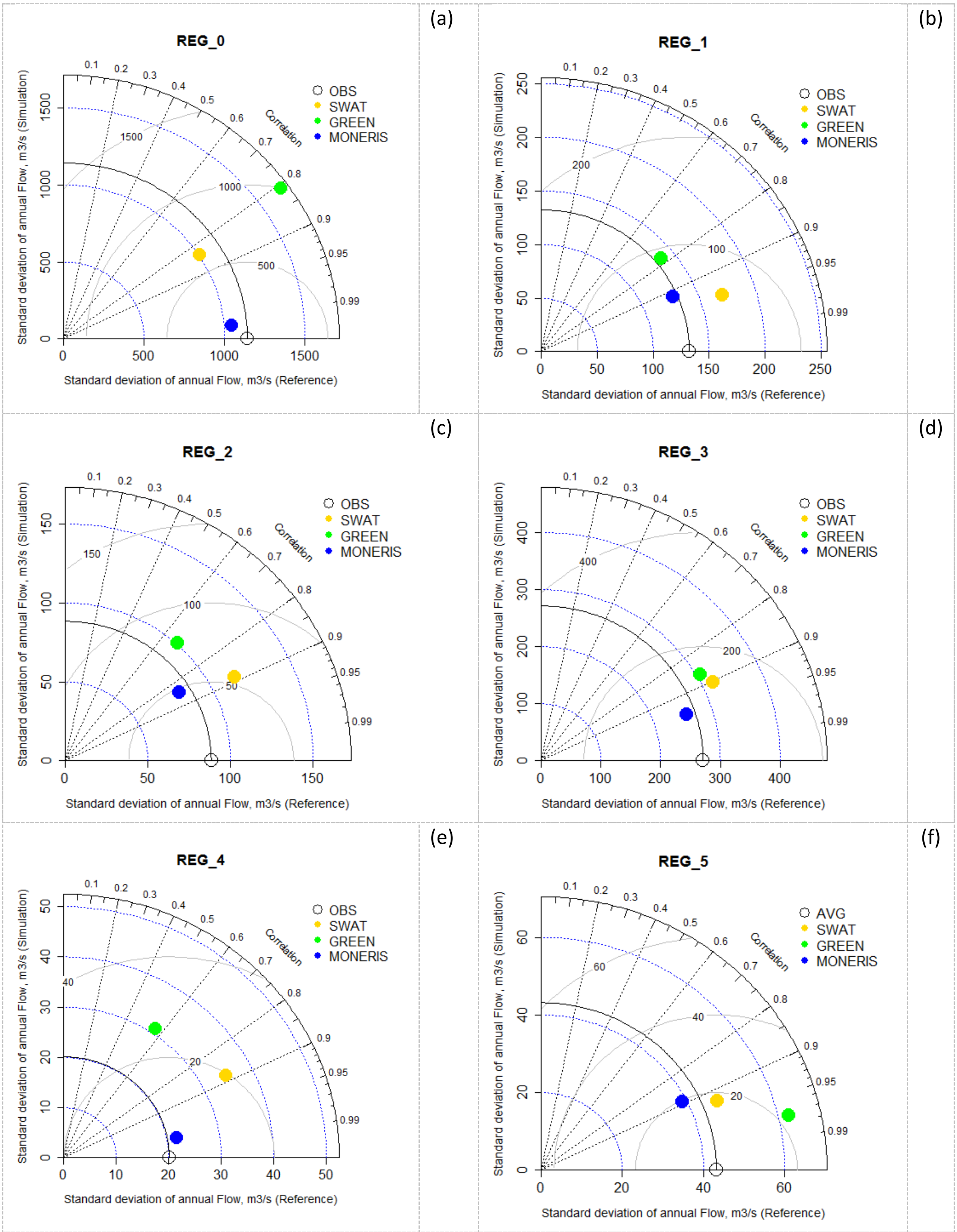

Fig. 8. a-f. Taylor diagram of annual water discharge (period 2000-2009). AVG represents the standard deviation of average of the three model predictions as observations were not available.

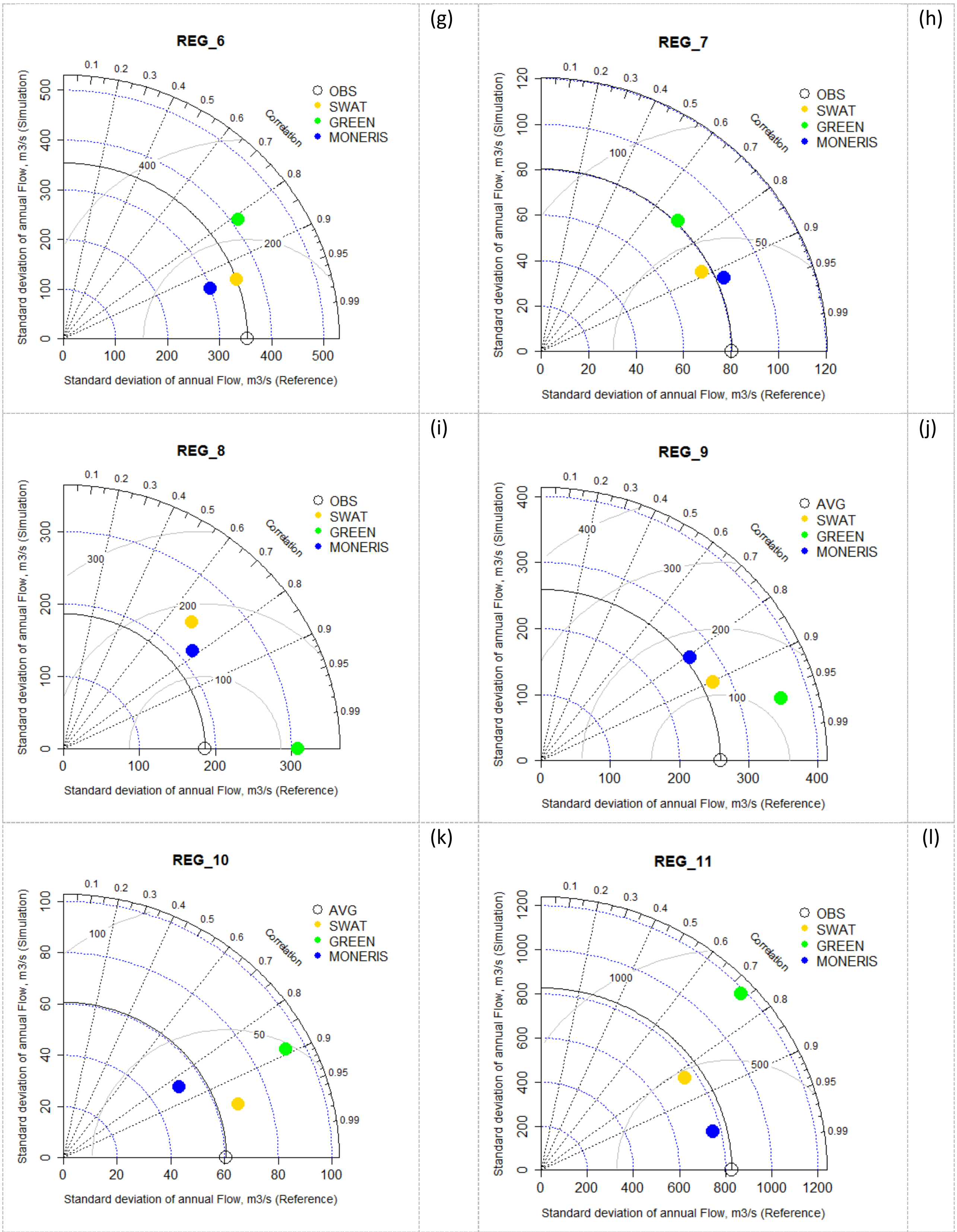

Fig.8. g-l. Taylor diagram of annual water discharge (period 2000-2009). AVG represents the standard deviation of average of the three model predictions as observations were not available.

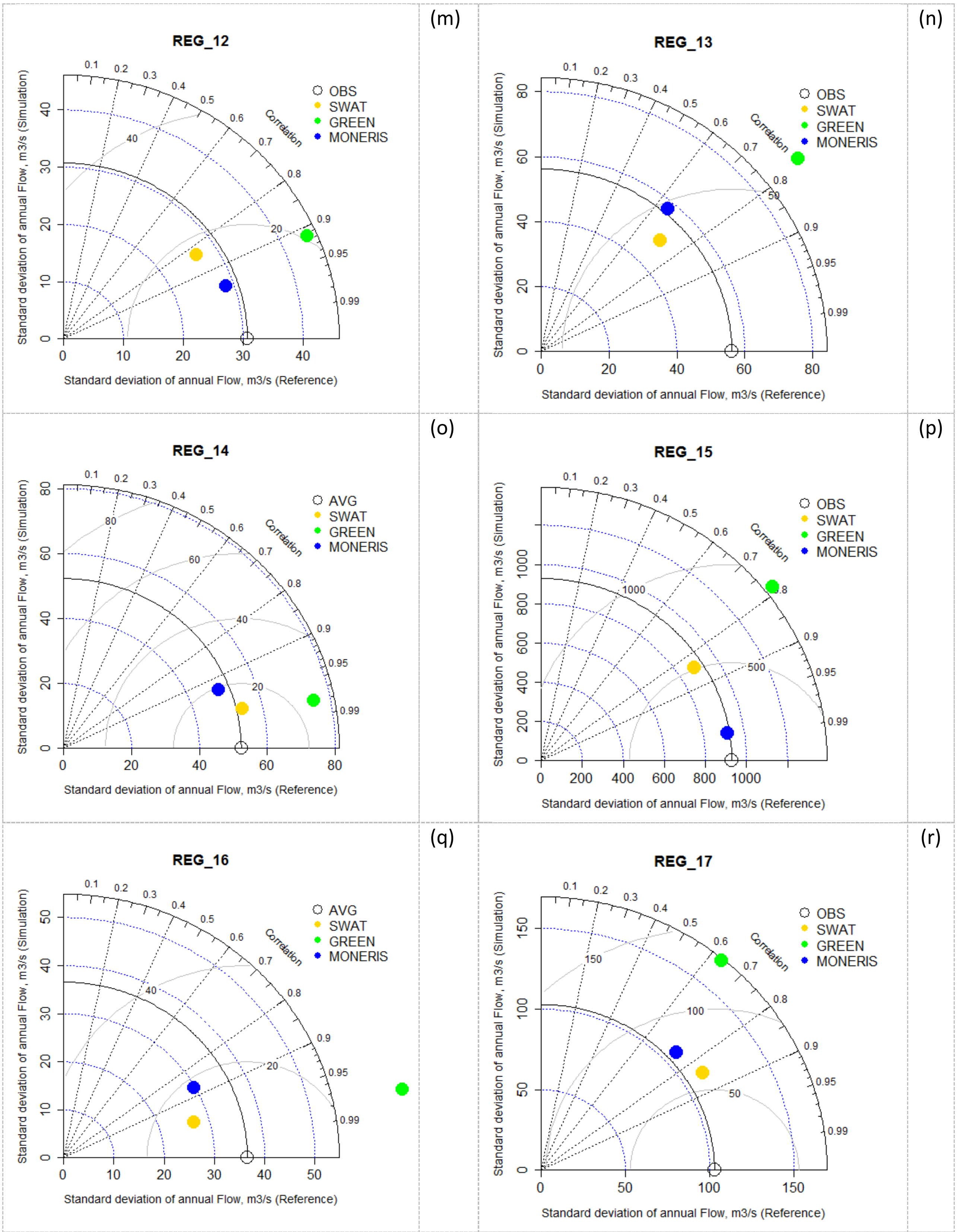

Fig.8. m-r. Taylor diagram of annual water discharge (period 2000-2009). AVG represents the standard deviation of average of the three model predictions as observations were not available.

Total Nitrogen

The map in Fig. 9 shows the long mean annual TN load in the 18 selected main regions of the Danube basin as estimated by the three models and from measured concentrations. Details per region are reported in Fig. 10 (a-r). The three models had comparable mean values in the upper Danube macro-region (*Regions 1-3*), along the Danube river (the Pannonian, Middle and Lower Danube; *Regions 6*, *11* and *15*), in the tributaries Vah-Hron-Ipel, Sava, and Velika Morava (*Region 5, 8, and 10*) and at the outlet at Reni-Chilia station (*Region 0)*. In the other 8 regions, differences between the models were more marked. Fig. 11 (a-r) summarizes the statistics of model simulations in each region.

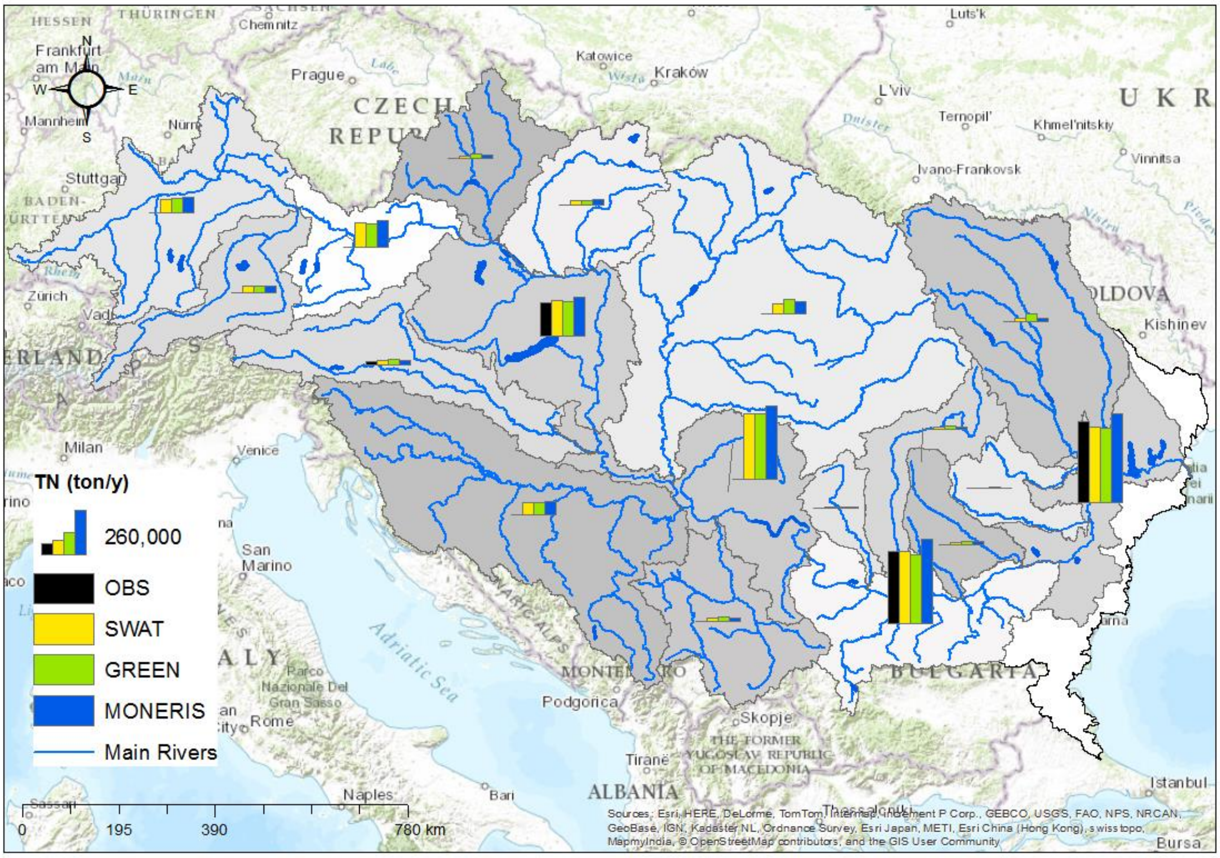

Fig. 9. Map of long mean annual total nitrogen (TN, ton/y) for the decade 2000-2009 in the Danube River Basin. The three model outputs are compared to available observations.

In *Region 0* (Delta), SWAT and GREEN slightly underestimated the long mean annual TN (453,210 ton/y for SWAT, 442,614 ton/y for GREEN and 484,290 ton/y for the observations), whereas MONERIS slightly exceeded (527,157 ton/y) the reference value (Fig. 10 a). All three models, however, had high correlation coefficients, simulating correctly the annual phase in observed loads, but none of the models simulated correctly the amplitude of the annual variations (Fig. 11 a).

While in *Regions 1* (Upper Danube, Fig. 10 b and Fig. 11 b) and *3* (Austrian Danube, Fig. 10 d and Fig. 11 d) all models were similar, in *Region 2 (*Inn, Fig. 10 c and Fig. 11 c) the standard deviation for SWAT (around 7,200 ton/y) was higher than for MONERIS and GREEN (5,306 and 5,057 ton/y respectively compared to the reference value of 5,275 ton/y).

In *Region 4* (Morava), GREEN simulated the highest TN annual load (26,325 ton/y), whereas SWAT and MONERIS simulated about 15,000 and 17,000 ton/y respectively (Fig. 10 e). GREEN and MONERIS standard deviations (4,042 and 4,016 ton/y) were very close to the reference value (3,794 ton/y), but they had different RMS error. SWAT slightly differed from MONERIS and GREEN, with the lowest correlation coefficient (0.78) and standard deviation (2,791 ton/y). The RMS error of SWAT and MONERIS were similar (2,362 and 2,282 ton/y) (Fig. 11 e).

In *Region 5* (Vah-Hron-Ipel), MONERIS predicted the largest mean TN load (28,340 ton/y) followed by GREEN (24,965 ton/y) and SWAT (23,908 ton/y; Fig. 10 f). Although all models were highly correlated with the observations (Fig. 10 f), MONERIS better represented the standard deviation (4,302 ton/y compared to the reference value of 4,071 ton/y). Despite similar RMS errors, SWAT slightly overestimated the standard deviation compared to MONERIS (5,221 ton/y). GREEN had the smallest RMS error (1,285 ton/y) and standard deviation (3,141 ton/y), the latter being slightly lower than the observed value.

In *Region 6* (Pannonian Danube), all models agreed well with the observations with similar mean annual loads (Fig. 10 g). GREEN, however, agreed best with the observations, with the lowest RMS error, the highest correlation coefficient, and similar standard deviation (Fig. 11 g). Instead, MONERIS had the largest long mean annual load and the lowest correlation coefficient (0.27) with the observations.

In *Region 7* (Drava), SWAT and MONERIS agreed best with the observed value of 25,271 ton/y (Fig. 10 h), but SWAT had a higher correlation (0.83) coefficient with the observation and a lower RMS error (Fig. 10 h). GREEN overestimated the long mean annual loads (34,521 ton/y) respect to the observations.

In *Region 8* (Sava), the three models had similar long-term annual TN load (Fig. 10 i), but different standard deviations. GREEN had the highest standard deviation (12,445 ton/y), followed by MONERIS (11,450 ton/y) and SWAT (9,157 ton/y; Fig. 11 i).

In *Region 9* (Tisa) GREEN simulated the highest annual TN load (85,802 ton/y). Its standard deviation (20,620 ton/y) also exceeded the SWAT and MONERIS values (Fig. 10 j). SWAT and MONERIS had similar RMS error but opposite estimation of annual variability. The standard deviation for SWAT (13,310 ton/y) was lower than the reference standard deviation and those of other models.

In *Region 10* (Velika Morava) GREEN long mean annual TN load (25,125 ton/y) was slightly larger than that of the other models (Fig. 10 k). GREEN standard deviation (6,417 ton/y, Fig. 11 k) was about twice those of SWAT and MONERIS. SWAT and MONERIS had very close standard deviations and high correlation with the reference value. However, SWAT simulated better TN reference loads, with the smallest RMS error.

In *Region 11* (Middle Danube), SWAT and GREEN had similar TN loads, whereas MONERIS slightly exceeded the others and the reference value (Fig. 10 l). This was confirmed also by the Taylor diagram (Fig. 11 l), in which the blue point (MONERIS model) is laid near the dashed arc of standard deviation correspondent to 60,000 ton/y, slightly far from the reference value (51,481 ton/y).

In *Regions 12, 13 and 14* (Jiu, Olt, and Arges-Vedea), GREEN simulated the highest mean annual TN loads, followed by MONERIS and SWAT (Fig. 10 m, n, o). All models had high correlation coefficients

(Fig. 11 m, n, o) but different standard deviations, with MONERIS being closest to the reference value.

In *Region 15* (Lower Danube), all three models agreed well with the observations (Fig. 10 p). GREEN had perfect correlation with the observed values, followed by MONERIS (0.87), and SWAT (0.71). However, all models clearly underestimated the annual variability as measured by the standard deviation of 56,210 ton/y for SWAT, 68,301 ton/y for GREEN and 79,105 ton/y for MONERIS, compared to the observed 138,315 ton/y (Fig. 10 p).

In *Region 16* (Buzau-Ialomita) and *17* (Siret-Prut-Buzau), *S*WAT and MONERIS agreed well with each other, whereas GREEN resulted in the highest mean annual TN load (12,862 ton/y in *Region 16* and 44,745 ton/y in *Region 17;* Fig. 10 q, r), standard deviation and coefficient of correlation (Fig. 11 q, r). SWAT and MONERIS produced comparable results. In *Region 16,* the average TN load was 3,065 ton/y according to SWAT and 4,825 ton/y according to MONERIS and their standard deviations were in the range of 1,100-1,350 ton/y. In *Region 17,* the average TN load was 21,021 ton/y according to SWAT and 20,054 ton/y according to MONERIS and their standard deviations were in the range of 5,900-6,000 ton/y.

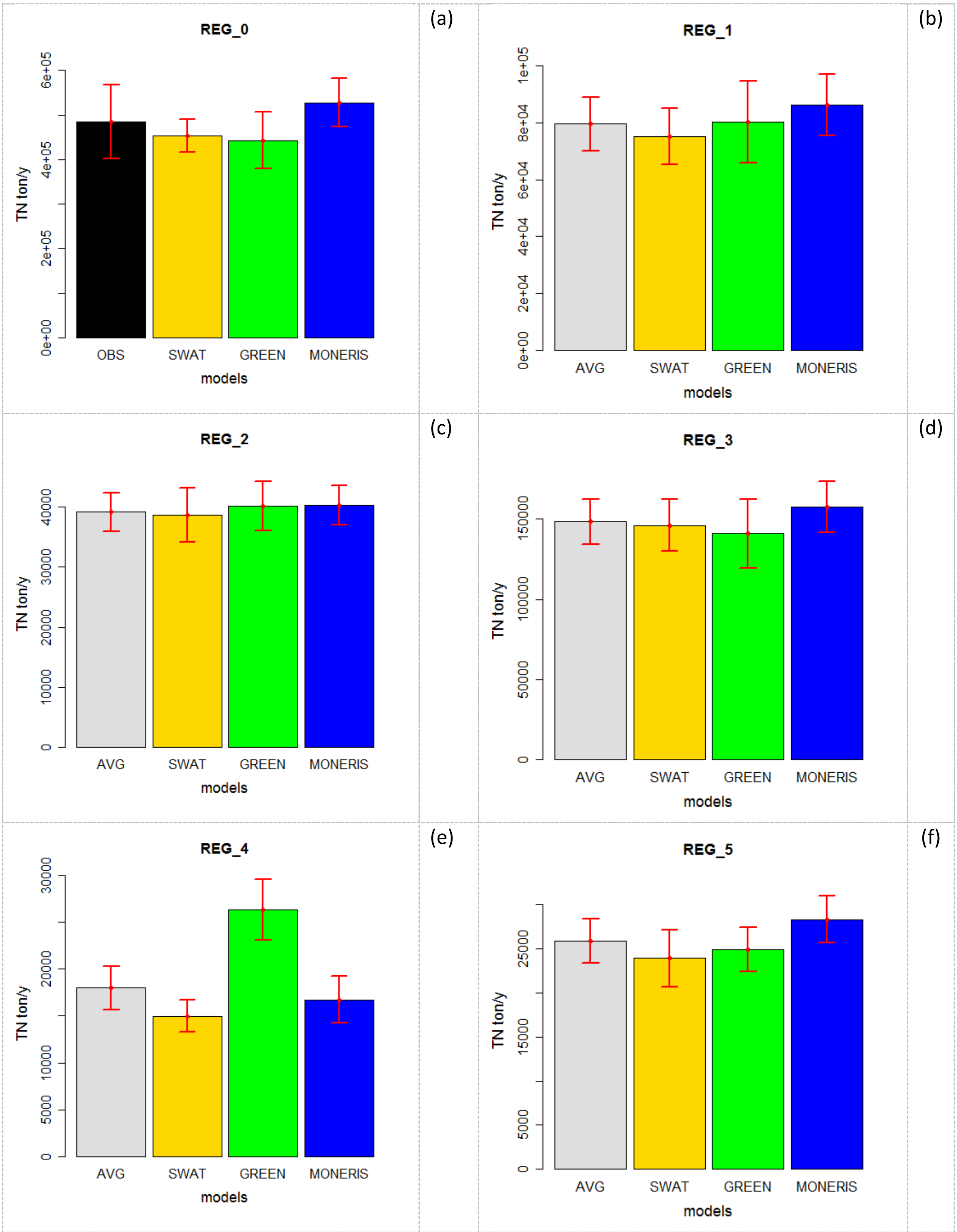

Fig. 10. a-f. Bar plots of long mean annual TN loads (period 2000-2009). *The grey bar represents the average of the three model predictions as observations were not available. The error bars in red indicate the 95% confidence intervals.*

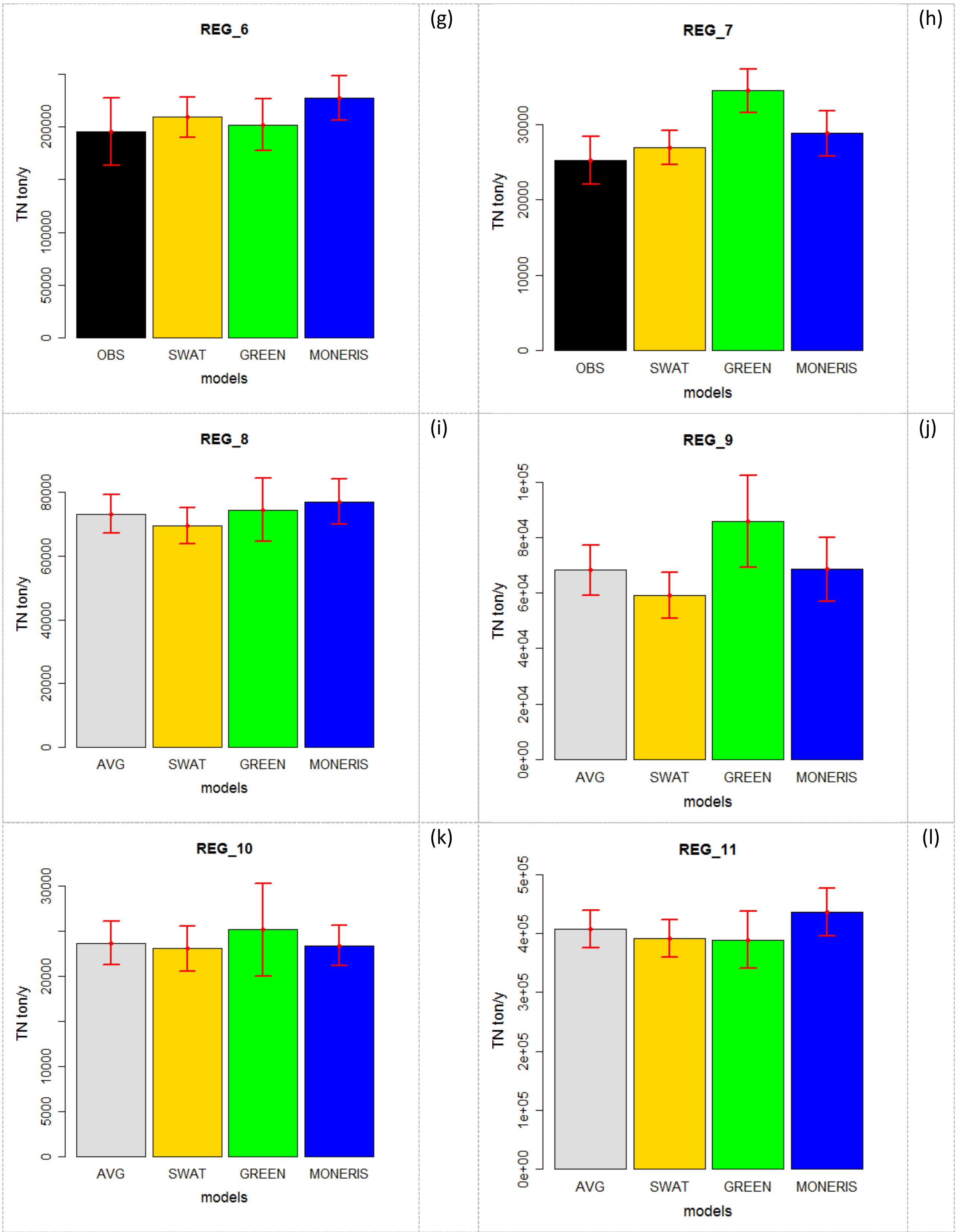

Fig.10. g-l. Bar plots of long mean annual TN loads (period 2000-2009). *The grey bar represents the average of the three model predictions as observations were not available. The error bars in red indicate the 95% confidence intervals.*

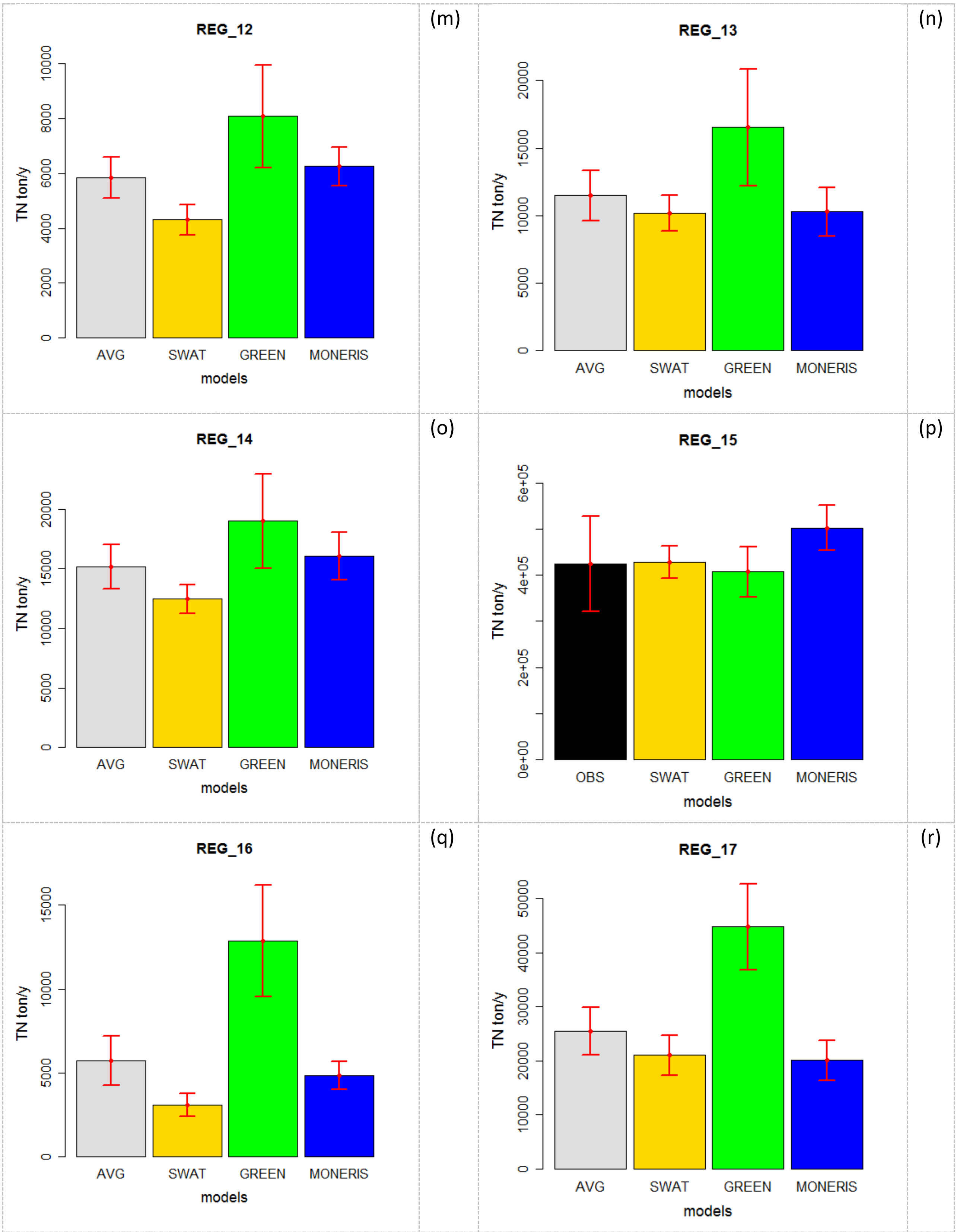

Fig.10. m-r. Bar plots of long mean annual TN loads (period 2000-2009). *The grey bar represents the average of the three model predictions as observations were not available. The error bars in red indicate the 95% confidence intervals.*

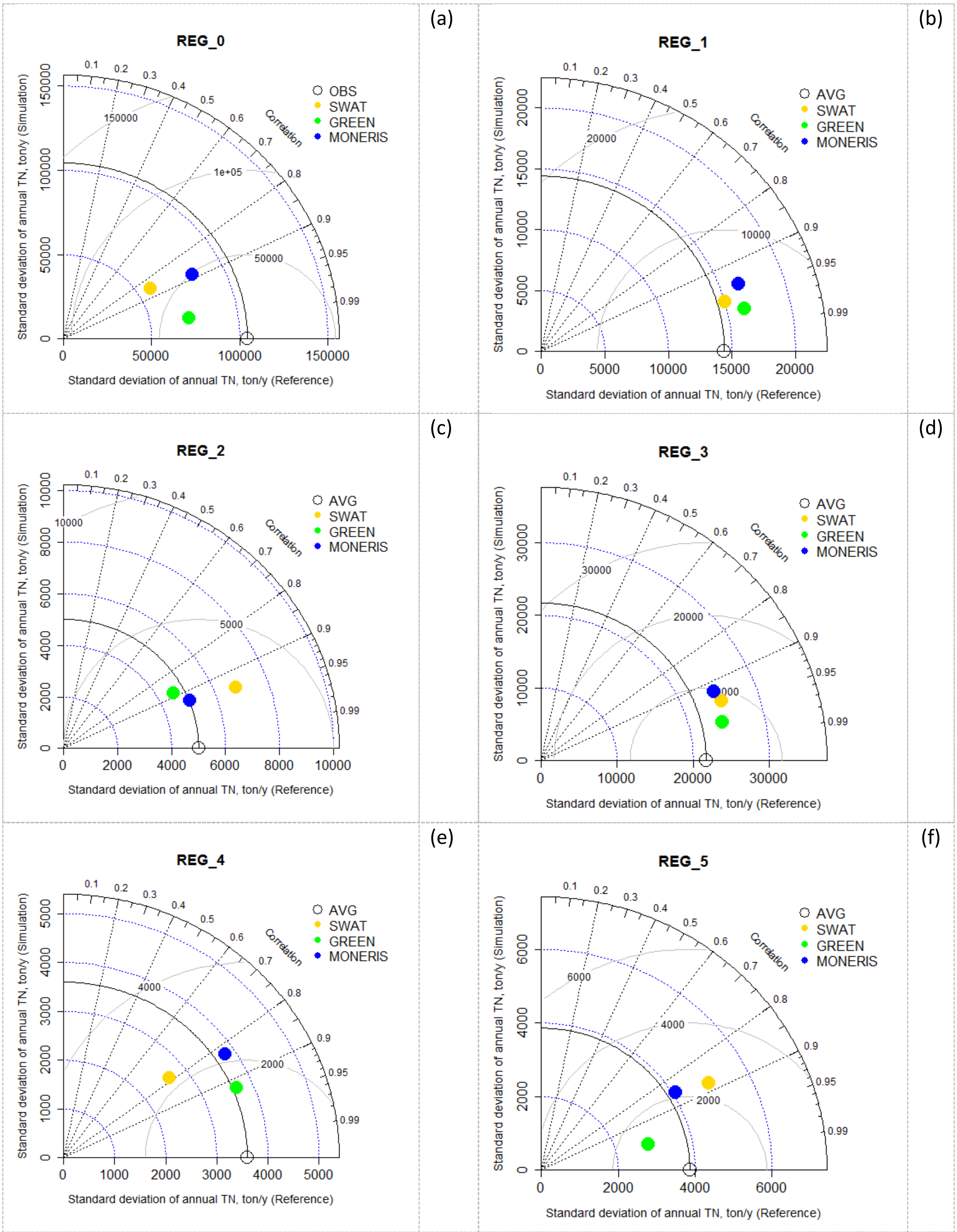

Fig. 11. a-f. Taylor diagram of annual TN loads (period 2000-2009). AVG represents the standard deviation of average of the three model predictions as observations were not available.

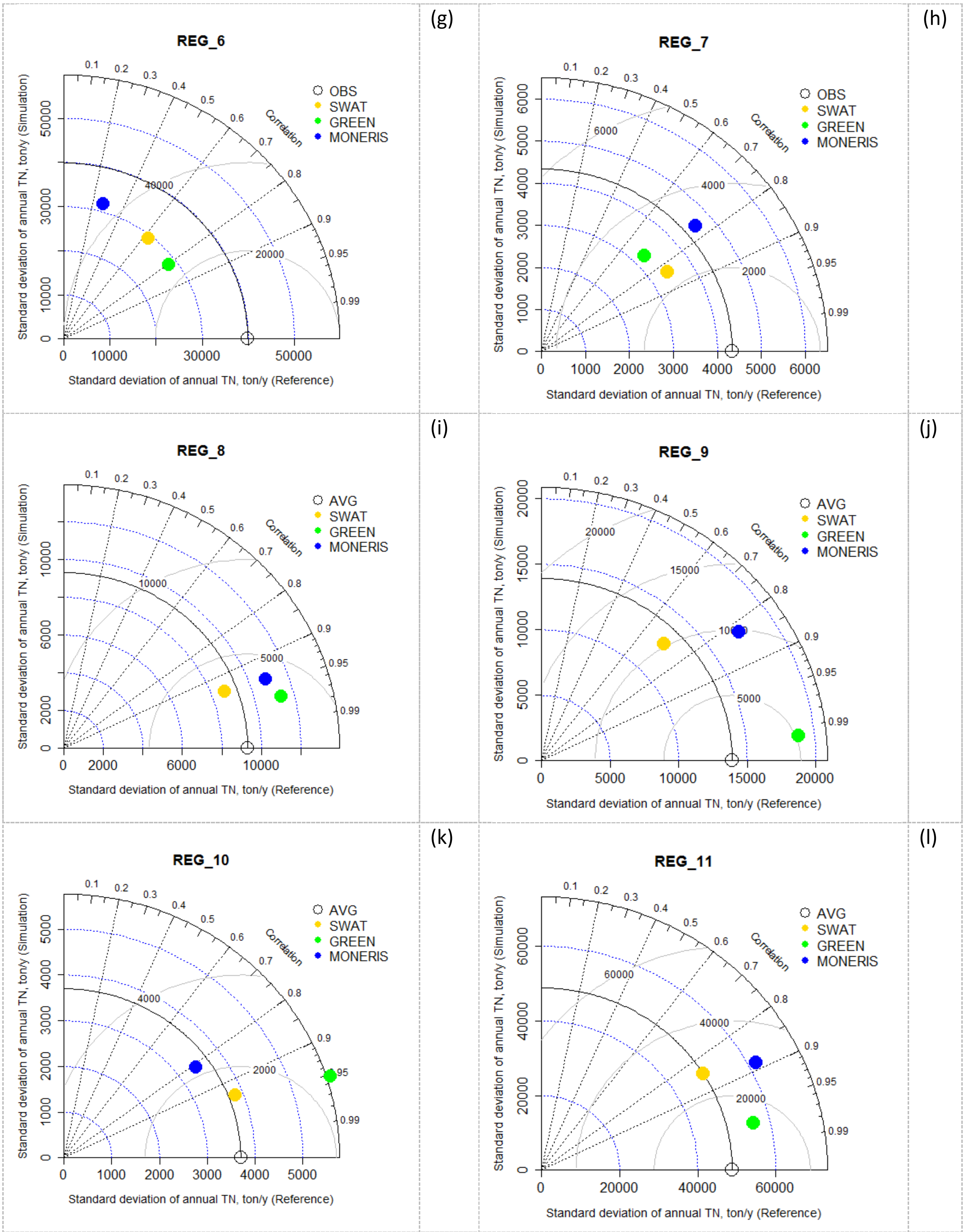

Fig.11. g-l. Taylor diagram of annual TN loads (period 2000-2009). AVG represents the standard deviation of average of the three model predictions as observations were not available.

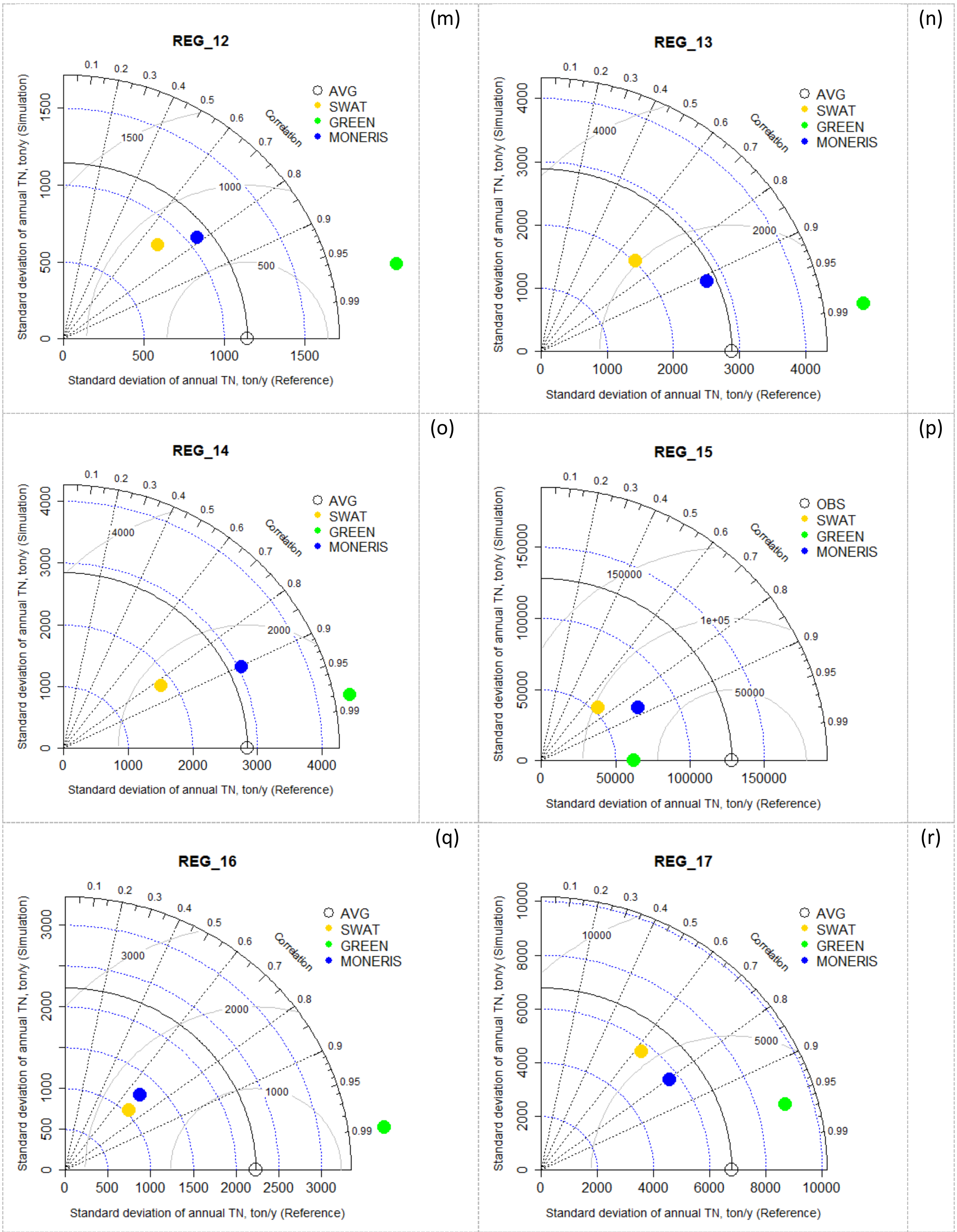

Fig.11. m-r. Taylor diagram of annual TN loads (period 2000-2009). AVG represents the standard deviation of average of the three model predictions as observations were not available.

Total Phosphorous

The map in Fig. 12 shows the spatial distribution of long mean annual loads of total phosphorous (TP, ton/y) estimated in the Danube River Basin with the three models and from measured concentrations. It highlights the overall agreement between the models and the observations, especially in the Lower Danube and Delta Regions. Conversely, in the tributaries some differences were more noticeable. The bar plots in Fig. 13 a-r and the Taylor diagrams in Fig. 14 a-r help analysing differences and similarities in detail.

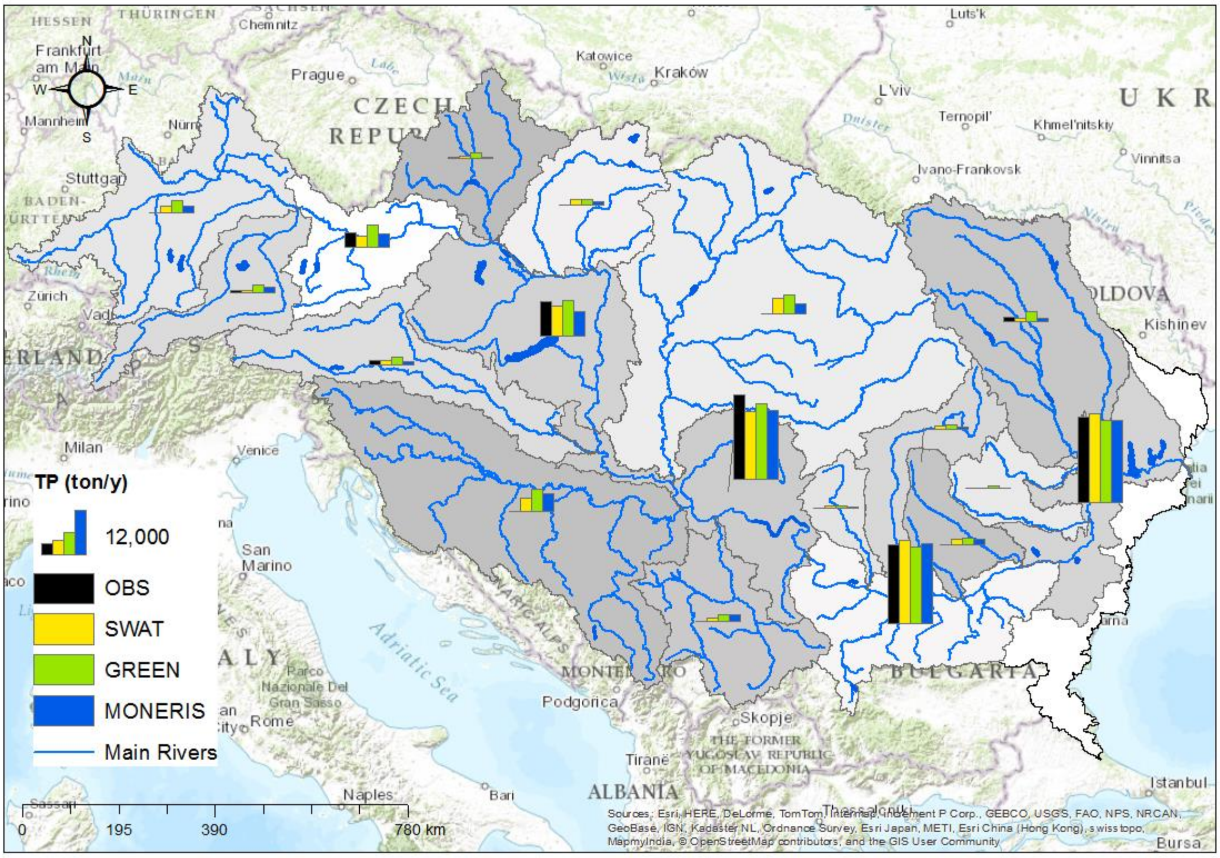

Fig. 12. Maps of the long mean annual total phosphorous comparisons (period 2000-2009) in each region. The three model outputs are compared to available observations.

In *Region 0* (Delta) the long mean annual loads of the models were very close to observations (Fig. 13 a) with similar coefficients of correlation and RMS errors (Fig. 14 b). However, the observations had a standard deviation markedly higher (around 20,489 ton/y) than those predicted by the models.

In *Region 1* (Upper Danube), GREEN estimated the highest annual TP load (Fig. 13 b). SWAT and MONERIS estimations were similar in mean value and standard deviations (226 and 255 ton/y respectively; Fig. 14 b).

In *Region 2* (Inn), SWAT agreed best with the observed annual TP load (738 ton/y compared to observed 655 ton/y; Fig. 13 c), but SWAT standard deviation (106 ton/y) was lower than the observed value (307 ton/y). MONERIS and GREEN overestimated the TP loads, with standard deviations of 132 and 391 ton/y respectively. GREEN was the least correlated to observations (coefficient of 0.30) and had the highest RMS error.

In *Region 3* (Austrian Danube), MONERIS and SWAT gave similar annual TP loads (3,020 and 3,638 ton/y; Fig. 13 d), standard deviation (around 400 ton/y; Fig. 14 d), RMS error (around 2,500 ton/y) and coefficient of correlation (0.41 and 0.55 respectively). GREEN resulted in higher mean load (6,296 ton/y) and lower correlation (coefficient of 0.13) with observations.

In *Region 4* (Morava), SWAT and MONERIS predicted loads comparable with the observations (509 and 412 ton/y respectively), unlike GREEN (1,662 ton/y, Fig. 13 e). According to Fig. 14 e, MONERIS agreed best with the observations with its small RMS error (76 ton/y), its standard deviation of 92 ton/y being close to the observed value (108 ton/y), and a high correlation coefficient (0.72). Albeit SWAT mean TP load agreed better than GREEN to observations, its coefficient of correlation was lower (0.23) and the RMS error (167 ton/y) was higher than for GREEN (coefficient of correlation of 0.6; RMS error of 130 ton/y).

In *Region 5* (Vah-Hron-Ipel), MONERIS simulated the lowest annual TP load (632 ton/y) (Fig. 13 f). MONERIS and GREEN had similar standard deviations (119 ton/y and 115 ton/y respectively), whereas SWAT standard deviation was higher (225 ton/y), with a higher RMS error too (Fig. 14 f). However, all three models had high correlation with the observations (coefficients of correlation greater than 0.7).

In *Region 6* (Pannonian Danube), the GREEN mean load (about 9,900 ton/y) agreed best with the observed value of 9,539 ton/y (Fig. 13 g). MONERIS underestimated long mean annual TP load (6,680 ton/y) and the amplitude of annual variations was lower than for observations, with standard deviation of 864 ton/y compared to the observed value of 2,515 ton/y. SWAT long mean annual TP load (8,388 ton/y) was closer to the observations than MONERIS, and SWAT simulated better than MONERIS the amplitude of annual variations (standard deviation of 1,311 ton/y) (Fig. 14 g). However, a significant decrease in observed loads could be detected after the year 2002, indicating that conditions at this station might have changed during the decade 2000-2009.

In *Region 7* (Drava) SWAT long mean annual TP of 1,289 ton/y was in very good agreement with the observations (1,474 ton/y), whereas GREEN (2,134 ton/y) and MONERIS (1,106 ton/y) overestimated and underestimated it respectively (Fig. 13 h). However, none of the model could simulate correctly the amplitude of annual variations and the annual phase (Fig. 14 h).

In *Region 8* (Sava) the three model markedly differed from each other (Fig. 13 i and Fig. 14 i). GREEN estimated a long mean annual TP load of 6,052 ton/y (standard deviation of 489 ton/y); MONERIS estimated 5,051 ton/y (standard deviation of 766 ton/y), whereas SWAT estimated 3,732 ton/y (standard deviation of 334 ton/y). It is difficult to quantify model performances because there were no sufficient observations available at the Sava outlet. Thus, there is high uncertainty for TP load estimation that warrants more research effort in the future.

In *Region 9* (Tisa) long mean annual TP loads of the three models were very different (Fig. 13 j). All models had comparable standard deviations and high correlation with the reference value, although GREEN agreed best with the reference (Fig. 14 j).

In *Region 10* (Velika Morava) MONERIS and GREEN predicted similar mean long annual TP loads (1,739 and 1,797 ton/y respectively) and standard deviation (210 and 195 ton/y respectively), whereas SWAT estimated lower loads (955 ton/y) and standard deviation (131 ton/y; Fig. 13 k and Fig. 14 k).

In *Region 11* (Middle Danube) all models slightly underestimated the observed long mean annual TP load (Fig. 13 l). Models had similar RMS errors (around 11,000 ton/y) and standard deviations (ranging in 2,000-3,000 ton/y) (Fig. 14 l). However, MONERIS scored the lowest correlation coefficient (0.17). Observations had a standard deviation markedly higher (around 12,380 ton/y) than those predicted by the models.

In *Regions 12, 13*, *14* (Jiu, Olt, Arges-Vedea) and *16* (Buzau-Ialomita) GREEN had the highest long mean annual TP load, followed by SWAT and MONERIS (Fig. 13 m,n,o,q). Also, GREEN standard deviations were the highest (Fig. 14 m,n,o,q). In *Region 12* MONERIS agreed better with the reference value, whereas in *Regions 13*, *14* and *16* models diverged from the reference in different ways (i.e. higher or lower coefficient of correlations) and no observation was available.

In *Region 15* (Lower Danube) all models had similar long mean predicted TP loads and in agreement with the observed value (Fig. 13 p). As in *Region 11*, models underestimated the standard deviation of the observations (Fig. 14 p).

In *Region 17* (Siret-Prut-Buzau) long mean annual TP loads of SWAT (1,129 ton/y) and MONERIS (1,027 ton/y) agreed well with the observation (1,351, ton/y, Fig. 13 q), albeit they slightly underestimated it. Conversely, GREEN overestimated TP load at about 3,000 ton/y. As in *Regions 0*, *11* and *15* all models markedly underestimated the observed standard deviation (Fig. 14 q).

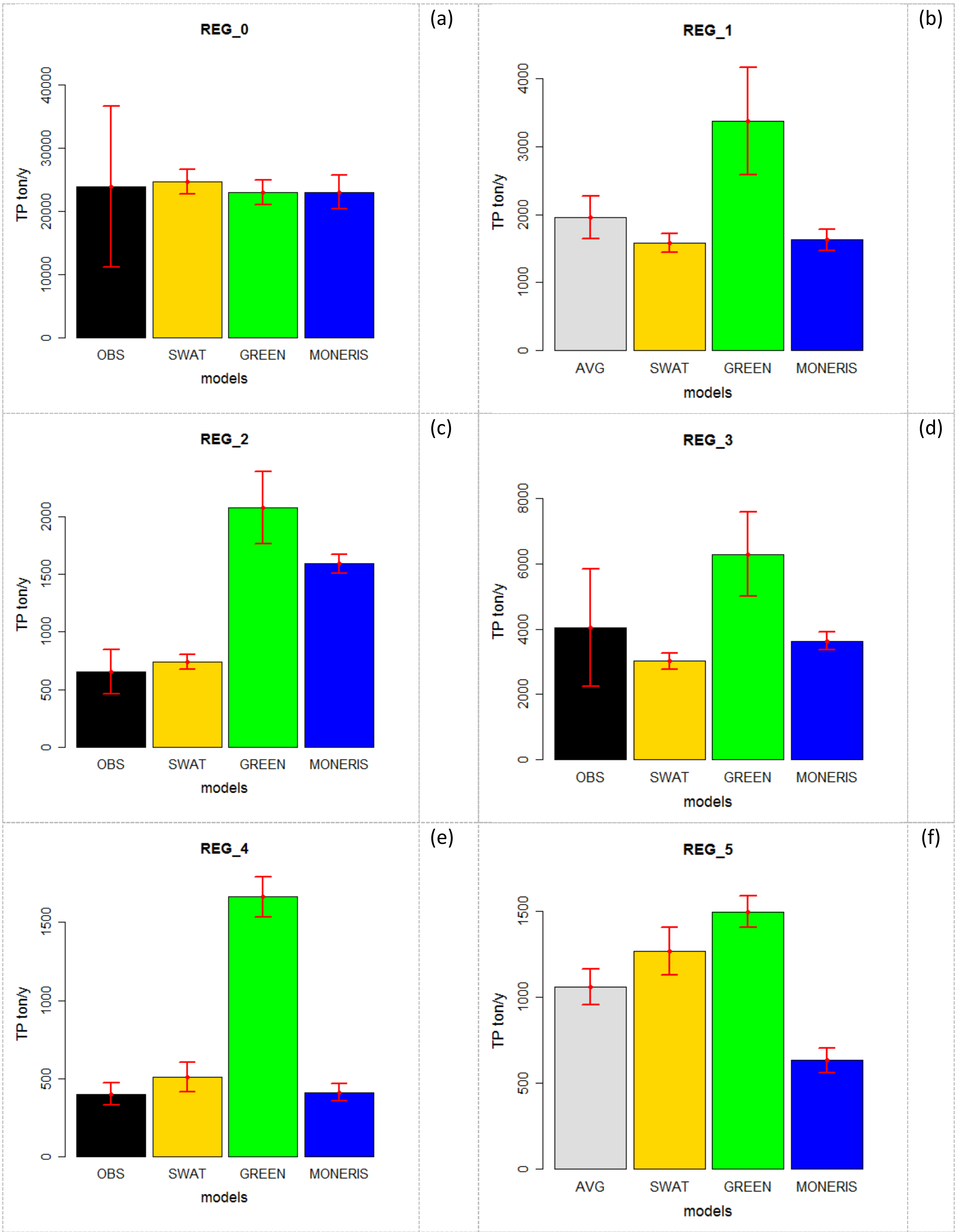

Fig. 13. a-f. Bar plots of long mean annual TP loads (period 2000-2009). *The grey bar represents the average of the three model predictions as observations were not available. The error bars in red indicate the 95% confidence intervals.*

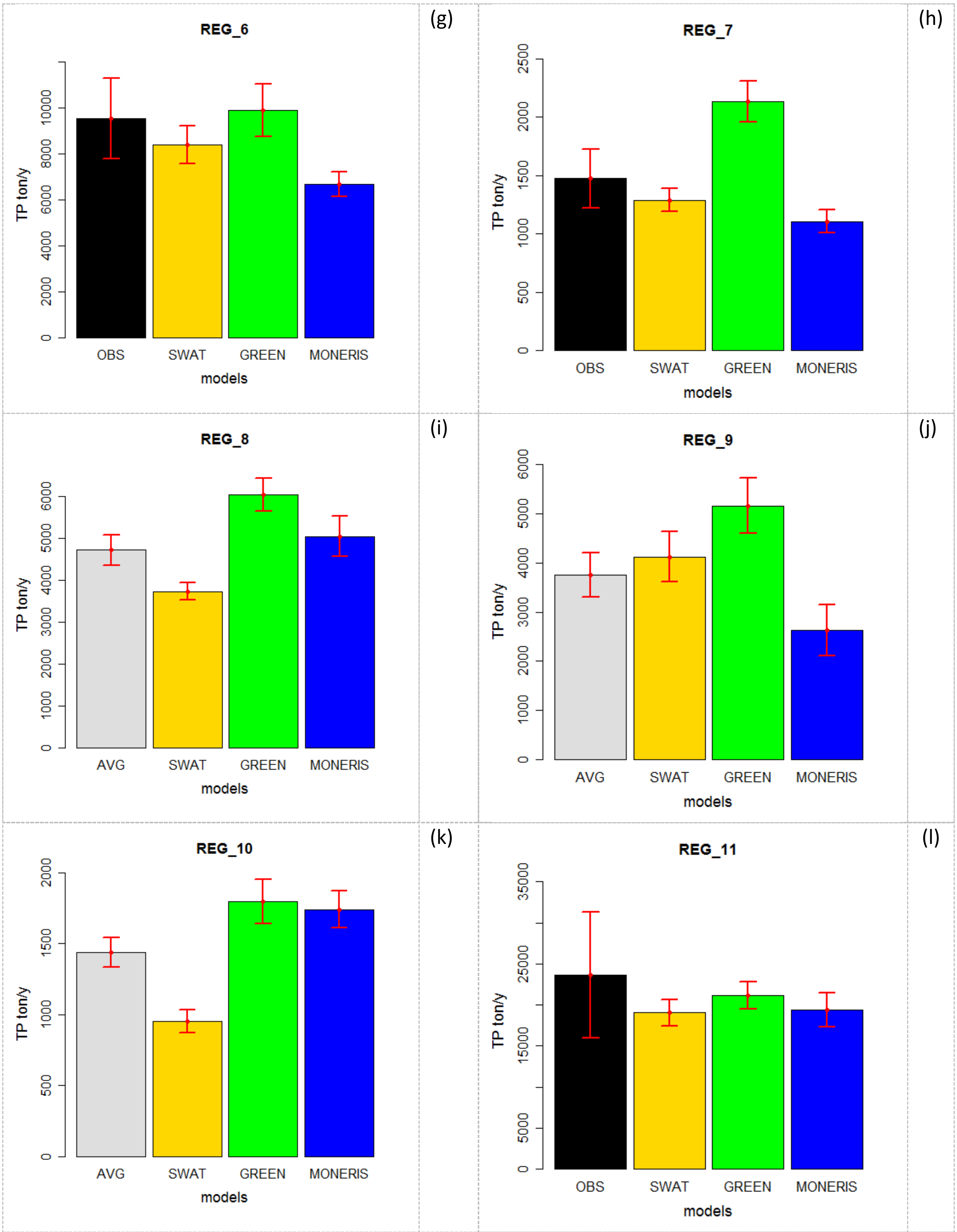

Fig.13. g-l. Bar plots of long mean annual TP loads (period 2000-2009). *The grey bar represents the average of the three model predictions as observations were not available. The error bars in red indicate the 95% confidence intervals.*

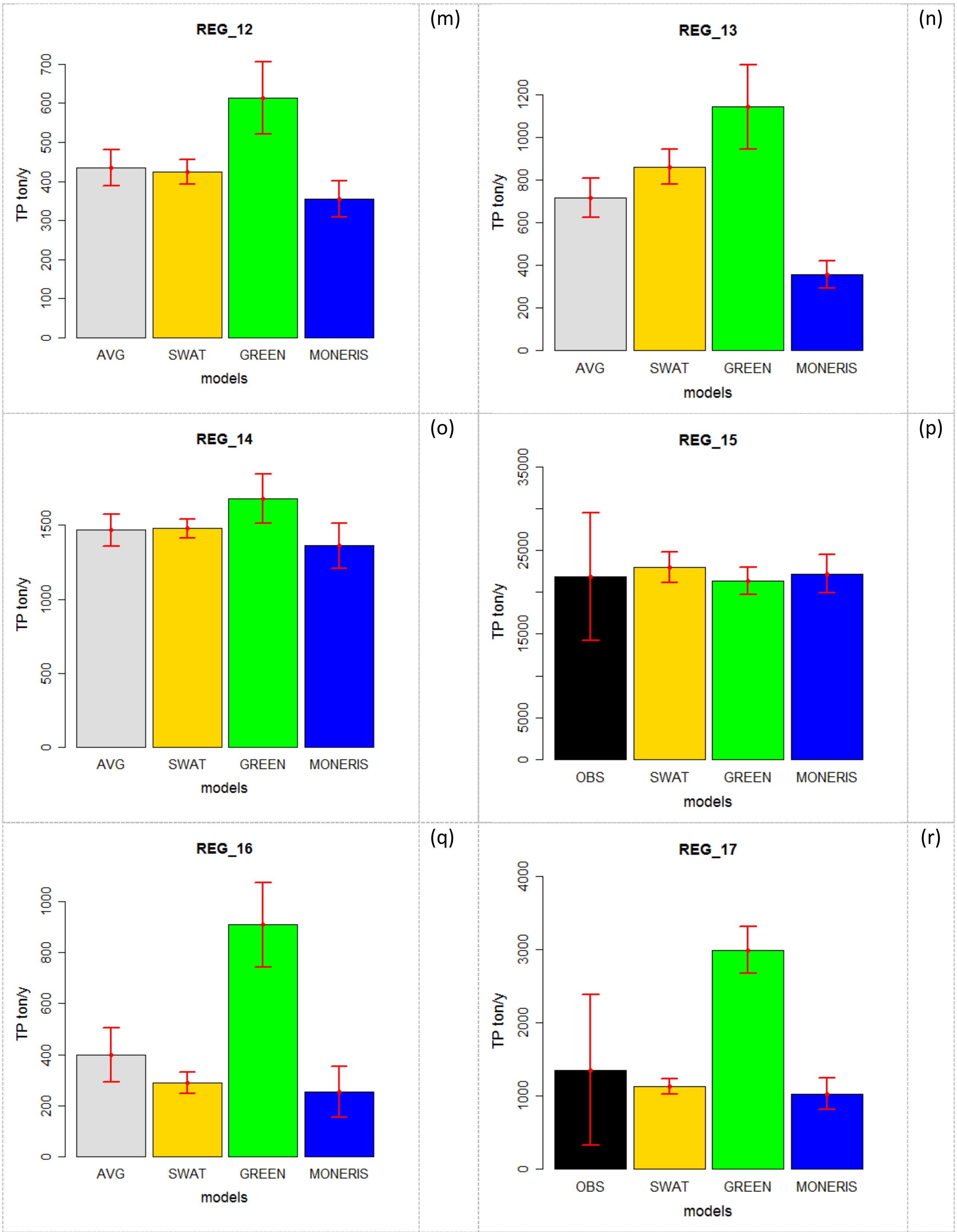

Fig. 13. m-r. Bar plots of long mean annual TP loads (period 2000-2009). *The grey bar represents the average of the three model predictions as observations were not available. The error bars in red indicate the 95% confidence intervals.*

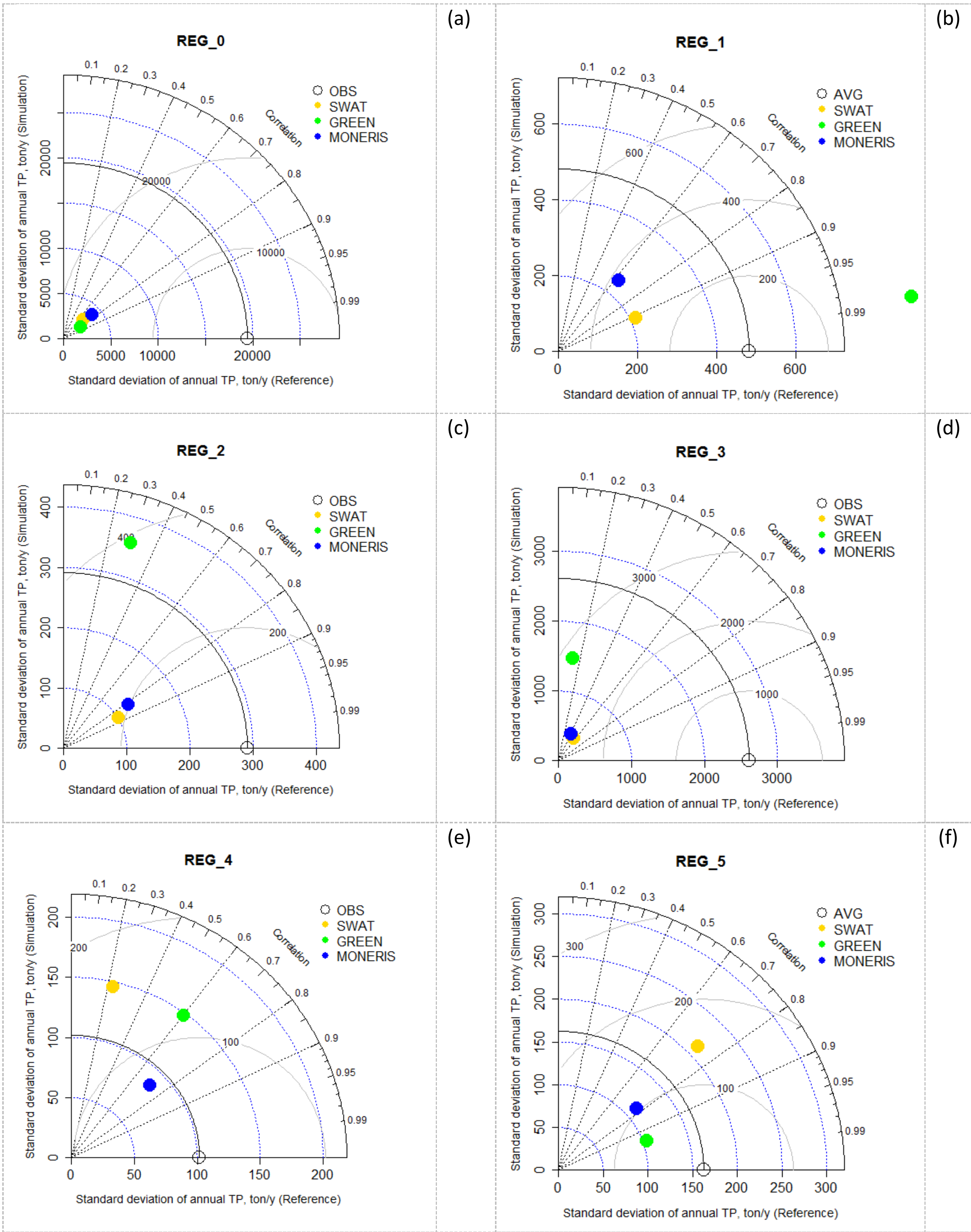

Fig. 14. a-f. Taylor diagram of annual TP loads (period 2000-2009). AVG represents the standard deviation of average of the three model predictions as observations were not available.

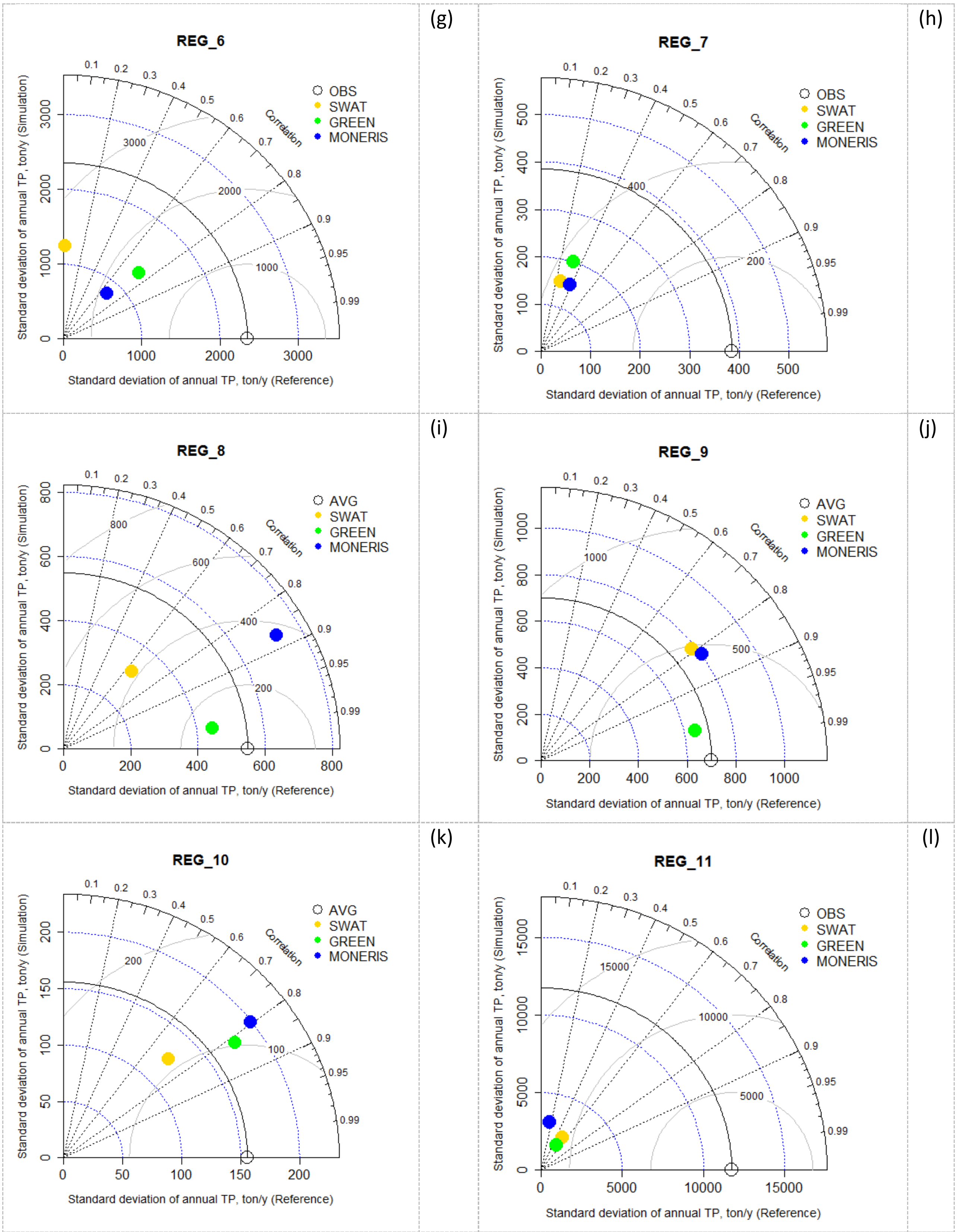

Fig.14. g-l. Taylor diagram of annual TP loads (period 2000-2009). AVG represents the standard deviation of average of the three model predictions as observations were not available.

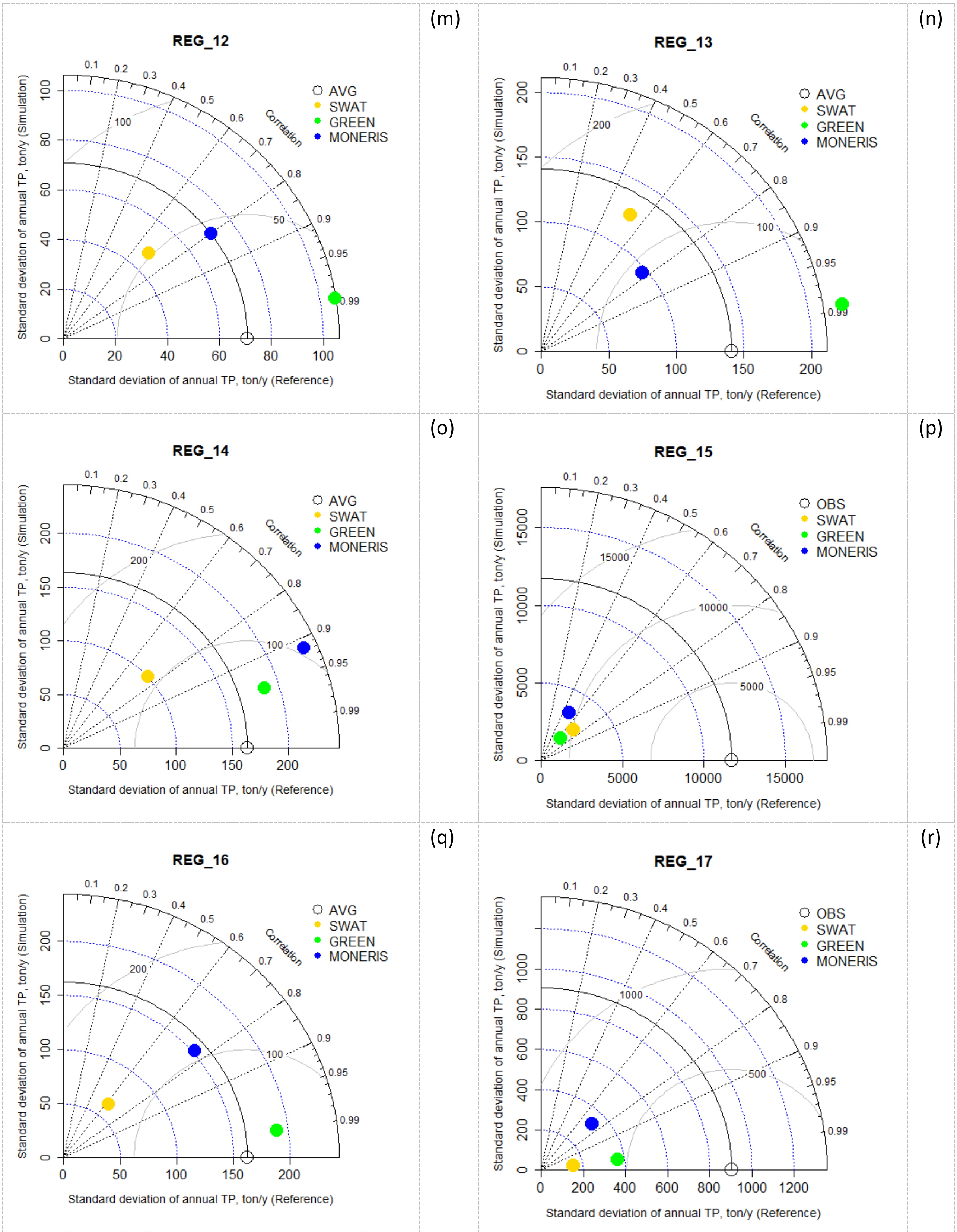

Fig.14. m-r. Taylor diagram of annual TP loads (period 2000-2009). AVG represents the standard deviation of average of the three model predictions as observations were not available.

4. Conclusion and recommendations

Models performance

The comparison of model outputs showed that the models SWAT, MONERIS and GREEN performed well in simulating water flow and nutrient loads at the outlets of the main Danube regions. The model results in general agreed well with each other and with the observed flow data. Similarly, the models showed good agreement in estimating total nitrogen loads. Differences among models estimates were more noticeable in Morava, Sava, Tisa, Jiu, Olt, Arges-Vedea, Buzau-Ialomita, Siret-Prut-Buzau (Regions *4, 8, 9, 12, 13, 14, 16, 17*, respectively, Fig. 2). Observations of total nitrogen concentration at the outlet of these regions would improve the understanding of the nitrogen fluxes in the Danube basin. Concerning phosphorus loads, SWAT and MONERIS resulted in similar long-term mean annual total phosphorus loads at the region outlets, but differed in the amplitude and variability of annual values. Estimations of GREEN were generally higher than of the other models, especially in the upstream part of the Danube basin.

The estimations of water and nutrient flows of SWAT and MONERIS are in good agreement with the available observations, which supports the use of these tools as Integrated Basin Models to describe the dominant processes affecting nutrient transfer from land to rivers and to the sea. Also the results of the model GREEN were in good agreement with the other models and the available observations, despite the model was calibrated at a coarser spatial scale (the whole Europe) and for a different temporal period (GREEN simulation covered only the period 2000-2005 that might not be representative of the period 2000-2009 considered in the comparison).

Implications for river basin management

The results of this comparative study show that, despite the differences in model approaches and input data, assessments from the three models are coherent; hence all three models may be confidently used as tools in river basin management. In particular, SWAT and MONERIS models might support the analysis of the impact of management measures in the Danube River Basin, as they were specifically set-up and calibrated for the region. Sharing the same baseline can help capturing the uncertainty when predicting the impact of measures in the river basins. Further analysis should focus on comparing sources and pathways of the nutrient fluxes, as this would improve support to the assessment of measures.

The inter-comparison allowed to identify areas where good agreement between models reinforced the independent assessments, for example in terms of water discharges across the basin, and areas where disagreements pointed to the need to collect more environmental data, such as nitrogen concentrations at some regions outlets and phosphorus concentrations in the Middle and Lower Danube regions.

Finally, the exercise provided the opportunity to promote dialogue and cooperation within the scientific community working in the region, enhancing the transparency of the modelling approaches and results, and improving the scientific support to the Danube river basin management.

5. References

Arnold, J. G., R. Srinivasan, R. S. Muttiah, and J. R. Williams (1998). Large area hydrologic modeling and assessment: Part I. Model development. J. American Water Resour. Assoc. 34(1), 73-89.

Behrendt, H. And D. Opitz, 2000: Retention of nutrients in river systems: Dependence on specific runoff and hydraulic load. – Hydrobiologia 410: 111–122.

Bouraoui, F., and Grizzetti, B., 2014. Modelling mitigation options to reduce diffuse nitrogen water pollution from agriculture Sci. Total Environ.468–469, 1267 –77.

European Commission, 2010. European Union Strategy for Danube Region. COM(2010) 715 final.

European Commission. 2000. Directive 2000/60/EC of the European Parliament and of the Council of 23 October 2000 establishing a framework for Community action in the field of water policy (The Water Framework Directove 2000/60/EC). Official Journal L327, 1-73.

Grizzetti, B., Bouraoui, F., and De Marsily, G., 2008. Assessing nitrogen pressures on European surface water Glob. Biogeochem. Cycles 22 GB4023.

Grizzetti, B., Bouraoui, F., Aloe A., 2012. Changes of nitrogen and phosphorus loads to European seas, Global Change Biology, 18, 769–782.

Grizzetti B., Passy P., Billen G., Bouraoui F., Garnier J., Lassaletta L. 2015. The role of water nitrogen retention in integrated nutrient management: assessment in a large basin by different modelling approaches. Environmental Research Letters, 10 (2015) 065008.

Habersack, H., Jäger, E., Hauer, C. 2013. The status of the Danube River sediment regime and morphology as a basis for future basin management. International Journal of River Basin Management 11, 153-166.

ICPDR, 2000. Water quality in the Danube river Basin, 2000 (TNMN-Yearbook), International Commission for the protection of the Danube River, Vienna, Austria, 2000.

ICPDR, 2002. Water quality in the Danube river Basin, 2002(TNMN-Yearbook), International Commission for the protection of the Danube River, Vienna, Austria, 2002.

ICPDR, 2003. Water quality in the Danube river Basin, 2003(TNMN-Yearbook), International Commission for the protection of the Danube River, Vienna, Austria, 2003.

ICPDR, 2004. Water quality in the Danube river Basin, 2004 (TNMN-Yearbook), International Commission for the protection of the Danube River, Vienna, Austria, 2004.

ICPDR, 2005. Water quality in the Danube river Basin, 2005 (TNMN-Yearbook), International Commission for the protection of the Danube River, Vienna, Austria, 2005.

ICPDR, 2006. Water quality in the Danube river Basin, 2006 (TNMN-Yearbook), International Commission for the protection of the Danube River, Vienna, Austria, 2006.

ICPDR, 2007. Water quality in the Danube river Basin, 2007 (TNMN-Yearbook), International Commission for the protection of the Danube River, Vienna, Austria, 2007.

ICPDR, 2008. Water quality in the Danube river Basin, 2008 (TNMN-Yearbook), International Commission for the protection of the Danube River, Vienna, Austria, 2008.

ICPDR, 2009. Water quality in the Danube river Basin, 2009 (TNMN-Yearbook), International Commission for the protection of the Danube River, Vienna, Austria, 2009.

Malagó, A., Pagliero, L., Bouraoui, F., Franchini, M. 2015. Comparing calibrated parameter sets for the Scandinavian Peninsula and for the Iberian Peninsula. Hydrological Sciences Journal 60(5), 949-967. doi:10.1080/02626667.2014.978332.

Moatar, F., Meybeck, M., 2005. Compared performances of different algorithms for estimating annual nutrients loads discharged by the eutrophic Loire. Hydrol. Process, 19, 429-444.

Pagliero, L., Bouraoui, F., Willems, P., Diels, J. 2014. Large-scale hydrological simulations using the Soil and Water Assessment Tool, Protocol development, and application in the Danube Basin. Journal of Environmental Quality, 145-154. doi:10.2134/jeq2011.0359

Silgram, M., Anthony,S.G., Collins, A.L., Strőmqvist , J., Bouraoui, F., Schoumans, O., Lo Porto, A., Groenendijk, P., Arheim, B., Mimikouf, M. and Johnssong, H., 2009. Evaluation of diffuse pollution model applications in EUROHARP catchments with limited data. J. Environ. Monit.,11, 554-571.

Taylor, K.E., 2001. Summarizing multiple aspects of model performance in a single diagram. Journal of Geophysical Research 106, 7183–7192.

Thunis, P., Georgieva, E., Galmarini, S., 2011. A procedure for air quality models benchmarking, Version 2. JRC Technical Report.

Van Gils, 2004: Revised Danube WQ Model: Analysis of available data (deliverable D5.9a). Report prepared in the framework of the daNUbs project (EU 5th Framework Programme), by Delft Hydraulics, Delft. The Netherlands, version 2, January 2004.

Venohr, M., Hirt, U., Hofmann, J., Opitz, D., Gericke, A., Wetzig, A., Natho, S., Neumann, F., Hürdler, J., Matranga, M., Mahnkopf, J., Gadegast, M., Behrendt, H., 2011. Modelling of Nutrient Emissions in River Systems - MONERIS - Methods and Background. International Review of Hydrobiology, 96(5) 435-483.

Venohr, M., Mahnkopf, J., Bremerich, V., Birk, S., Duell, H. and Tockner, K., 2015. Spatial distribution and source apportionment of nutrient emission in Europe, Journal of Environmental Management.

Vigiak O., Malagó A., Bouraoui F., Vanmaercke M., Poesen J., 2015a. Adapting SWAT hillslope erosion model to predict sediment concentrations and yields in large Basins. Science of the Total Environment. 538, 855–875.

Vigiak O., Malagó A., Bouraoui F., Pastori, M., Borrelli, P., Pistocchi, A. 2015b. Impact of current conservation practices on sediment load reduction in the Danube River Basin. In Cau et al. (eds) SWAT 2015 Conference, 24th-26th June 2015, Pula (Italy), Book of Abstracts, pp 66. http://swat.tamu.edu/media/114758/d3-2-vigiak.pdf.

APPENDIX A

Table A 1. Mean annual water discharge (m3/s) in different periods in each region (OBS: observations with at least 5 years of data)

REGIONS	Periods	OBS	SWAT	GREEN	MONERIS
0	2000-2009	6,681	6,477	6,371	6,814
0	1996-2005	6,901	6,685	6,635	6,827
0	1995-2009	6,805	6,612	6,706	6,814
1	2000-2009	685	578	506	652
1	1996-2005	694	631	490	671
1	1995-2009	688	600	504	652
2	2000-2009	751	612	734	606
2	1996-2005	769	657	744	610
2	1995-2009	759	618	762	606
3	2000-2009	1,959	1,628	1,570	1,806
3	1996-2005	1,994	1,706	1,587	1,814
3	1995-2009	1,969	1,647	1,628	1,806
4	2000-2009	102	136	140	112
4	1996-2005	110	140	150	109
4	1995-2009	109	143	151	112
5	2000-2009		232	268	277
5	1996-2005		262	276	279
5	1995-2009		257	276	277
6	2000-2009	2,322	2,196	2,138	2,458
6	1996-2005	2,367	2,330	2,208	2,455
6	1995-2009	2,350	2,271	2,261	2,458
7	2000-2009	490	507	592	496
7	1996-2005	509	530	643	483
7	1995-2009	504	504	656	496
8	2000-2009	1,403	1,337	1,091	1,292
8	1996-2005		1,350	1,204	1,324
8	1995-2009	1,403	1,355	1,225	1,292
9	2000-2009		1,000	890	1,038
9	1996-2005		990	923	1,023
9	1995-2009		1,012	908	1,038
10	2000-2009		270	231	214
10	1996-2005		247	246	210
10	1995-2009		257	248	214
11	2000-2009	5,301	5,522	5,192	5,667
11	1996-2005	5,526	5,657	5,482	5,666
11	1995-2009	5,457	5,612	5,558	5,667
12	2000-2009	97	79	81	100
12	1996-2005	92	76	80	97

12	1995-2009	97	78	79	100
13	2000-2009	139	156	179	148
13	1996-2005	153	164	179	155
13	1995-2009	139	164	178	148
14	2000-2009		110	113	104
14	1996-2005		123	108	109
14	1995-2009		116	107	104
15	2000-2009	6,059	6,013	5,807	6,338
15	1996-2005	6,346	6,185	6,067	6,366
15	1995-2009	6,221	6,118	6,139	6,338
16	2000-2009		52	112	62
16	1996-2005		67	112	60
16	1995-2009		62	112	62
17	2000-2009	299	331	416	343
17	1996-2005	333	340	426	329
17	1995-2009	333	346	426	343

Table A 2. Mean annual total nitrogen loads (TN ton/y) in different periods in each region (OBS: observations with at least 5 years of data)

REGIONS	Periods	OBS	SWAT	GREEN	MONERIS
0	2000-2009	484,290	453,210	442,614	527,157
0	1996-2005		461,007	460,827	545,607
0	1995-2009	484,290	459,895	465,256	527,157
1	2000-2009		75,129	80,173	86,129
1	1996-2005		77,742	79,922	91,209
1	1995-2009		76,472	81,781	86,129
2	2000-2009		38,643	40,144	40,265
2	1996-2005		41,200	41,618	42,532
2	1995-2009		39,962	42,487	40,265
3	2000-2009		146,140	140,898	157,650
3	1996-2005		149,218	143,566	164,848
3	1995-2009		148,087	146,602	157,650
4	2000-2009		14,988	26,325	16,714
4	1996-2005		14,890	27,818	16,702
4	1995-2009		15,380	28,108	16,714
5	2000-2009		23,908	24,966	28,340
5	1996-2005		24,693	26,116	28,564
5	1995-2009		25,362	26,344	28,340
6	2000-2009	195,714	209,282	201,953	227,330
6	1996-2005	191,090	213,370	210,362	234,022
6	1995-2009	194,455	212,788	214,554	227,330
7	2000-2009	25,271	26,986	34,521	28,875
7	1996-2005	28,564	26,784	37,386	28,939
7	1995-2009	27,690	26,697	37,896	28,875
8	2000-2009		69,486	74,514	77,031
8	1996-2005		69,398	82,885	79,856
8	1995-2009		69,558	84,168	77,031
9	2000-2009		59,184	85,802	68,497
9	1996-2005		60,076	88,543	69,798
9	1995-2009		60,305	88,235	68,497
10	2000-2009		23,040	25,125	23,373
10	1996-2005		22,382	26,702	23,588
10	1995-2009		22,822	26,980	23,373
11	2000-2009		391,863	389,326	436,705
11	1996-2005		396,968	411,513	448,452
11	1995-2009		396,797	416,836	436,705
12	2000-2009		4,317	8,081	6,254
12	1996-2005		4,267	8,269	6,405
12	1995-2009		4,316	8,264	6,254
13	2000-2009		10,157	16,532	10,291
13	1996-2005		10,073	16,806	10,874

13	1995-2009		10,357	16,819	10,291
14	2000-2009		12,448	19,000	16,092
14	1996-2005		12,952	19,420	16,745
14	1995-2009		12,705	19,440	16,092
15	2000-2009	424,168	428,121	406,922	502,216
15	1996-2005		434,443	426,297	519,294
15	1995-2009	424,168	433,467	430,923	502,216
16	2000-2009		3,065	12,862	4,825
16	1996-2005		3,596	13,048	4,857
16	1995-2009		3,454	13,079	4,825
17	2000-2009		21,021	44,745	20,054
17	1996-2005		21,312	45,046	20,428
17	1995-2009		21,666	45,212	20,054

Table A 3. Mean annual total phosphorous loads (TP ton/y) in different periods in each region (OBS: observations with at least 5 years of data)

REGIONS	Periods	OBS	SWAT	GREEN	MONERIS
0	2000-2009	23,930	24,697	23,030	23,063
0	1996-2005	24,240	25,088	23,988	23,235
0	1995-2009	23,280	25,231	24,195	23,063
1	2000-2009		1,580	3,373	1,629
1	1996-2005		1,619	3,481	1,602
1	1995-2009		1,596	3,617	1,629
2	2000-2009	655	738	2,079	1,590
2	1996-2005	998	772	2,281	1,615
2	1995-2009	857	749	2,370	1,590
3	2000-2009	4,046	3,020	6,296	3,638
3	1996-2005	6,228	3,052	6,713	3,640
3	1995-2009	5,009	3,035	6,957	3,638
4	2000-2009	402	509	1,662	412
4	1996-2005	547	488	1,759	411
4	1995-2009	505	507	1,776	412
5	2000-2009		1,268	1,497	632
5	1996-2005		1,282	1,527	639
5	1995-2009		1,327	1,533	632
6	2000-2009	9,539	8,388	9,899	6,680
6	1996-2005	10,818	8,597	10,497	6,573
6	1995-2009	10,190	8,712	10,748	6,680
7	2000-2009	1,474	1,289	2,134	1,106
7	1996-2005	1,812	1,275	2,383	1,090
7	1995-2009	1,689	1,298	2,426	1,106
8	2000-2009		3,732	6,052	5,051
8	1996-2005		3,729	6,565	5,107
8	1995-2009		3,752	6,640	5,051
9	2000-2009		4,121	5,161	2,630
9	1996-2005	7,600	4,147	5,309	2,541
9	1995-2009	7,600	4,177	5,299	2,630
10	2000-2009		955	1,797	1,739
10	1996-2005		926	1,889	1,749
10	1995-2009		940	1,903	1,739
11	2000-2009	23,637	19,065	21,131	19,391
11	1996-2005	21,769	19,271	22,438	19,341
11	1995-2009	23,169	19,466	22,727	19,391
12	2000-2009		424	614	355
12	1996-2005		425	628	355
12	1995-2009		426	628	355
13	2000-2009		862	1,143	356
13	1996-2005		864	1,161	366

13	1995-2009		889	1,161	356
14	2000-2009		1,477	1,679	1,362
14	1996-2005		1,505	1,715	1,418
14	1995-2009		1,495	1,718	1,362
15	2000-2009	21,862	22,956	21,345	22,194
15	1996-2005	22,978	23,248	22,412	22,300
15	1995-2009	22,215	23,414	22,638	22,194
16	2000-2009		290	908	254
16	1996-2005		326	922	271
16	1995-2009		317	924	254
17	2000-2009	1,351	1,129	2,989	1,027
17	1996-2005	1,275	1,163	3,008	1,020
17	1995-2009	1,275	1,167	3,017	1,027

APPENDIX B

Table B 1. Mean annual water discharge (m3/s) in the period (2000-2009) ("average" column) and the statistics derived from Taylor diagram. The table includes where available the observations (and correlated statistics) in black, otherwise the average of the three model simulations (and correlated statistics) in blue.

	Water Discharge average (m3/s)				standard deviation (m3/s)				Coeff. Of Correlation			RMS error (m3/s) [2]		
Region	Reference	SWAT	GREEN	MONERIS	Reference	SWAT	GREEN	MONERIS	SWAT	GREEN	MONERIS	SWAT	GREEN	MONERIS
REG_0	6,681	6,477	6,371	6,814	1,211	1,061	1,824	1,103	0.84	0.81	1.00	656	1,102	144
REG_1	685	578	506	652	140	180	151	136	0.95	0.78	0.92	64	97	56
REG_2	751	612	734	606	93	122	111	86	0.89	0.68	0.85	58	84	50
REG_3	1,959	1,628	1,570	1,806	288	337	336	271	0.90	0.87	0.95	147	166	90
REG_4	102	136	140	112	21	37	34	23	0.88	0.56	0.98	21	28	4
REG_5	257	232	268	277	46	50	68	41	0.93	0.97	0.89	19	26	21
REG_6	2,322	2,196	2,138	2,458	382	373	452	316	0.94	0.81	0.94	130	263	137
REG_7	490	507	592	496	85	80	89	88	0.89	0.71	0.92	39	67	34
REG_8	1,403	1,337	1,091	1,292	204	256	338	229	0.69	1.00	0.78	186	133	144
REG_9	1,004	1,000	890	1,038	273	291	393	280	0.90	0.96	0.81	126	148	171
REG_10	242	270	231	214	64	72	102	54	0.95	0.89	0.84	23	54	35
REG_11	5,301	5,522	5,192	5,667	877	790	1,293	807	0.83	0.74	0.97	494	878	208
REG_12	97	79	81	100	32	28	49	30	0.83	0.91	0.95	18	23	10
REG_13	139	156	179	148	59	52	106	61	0.72	0.79	0.65	42	69	50
REG_14	108	110	113	104	55	57	82	52	0.97	0.98	0.93	13	30	20
REG_15	6,059	6,013	5,807	6,338	985	934	1,569	970	0.84	0.79	0.99	538	1,001	150
REG_16	68	52	112	62	39	28	75	31	0.96	0.98	0.87	14	39	19
REG_17	299	331	416	343	115	119	184	114	0.85	0.63	0.74	65	142	83

[2] The RMS error was derived from the geometric relationship between standard deviation of reference and simulation (σ_r and σ_s respectively) and the coefficient of correlation (R): $RMS^2 = \sigma_s^2 + \sigma_r^2 - 2\,\sigma_s \cdot \sigma_r \cdot R$

Table B 2. Mean annual total nitrogen (TN, ton/y) in the period (2000-2009) ("average" column) and the statistics derived from Taylor diagram. The table includes where available the observations (and correlated statistics) in black, otherwise the average of the three model simulations (and correlated statistics) in blue.

	TN average (ton/y)				standard deviation (ton/y)				Coeff. Of Correlation			RMS error (ton/y)[3]		
Region	Reference	SWAT	GREEN	MONERIS	Reference	SWAT	GREEN	MONERIS	SWAT	GREEN	MONERIS	SWAT	GREEN	MONERIS
REG_0	484,290	453,210	442,614	527,157	112,737	60,677	79,472	87,068	0.85	0.98	0.89	68,609	37,138	53,681
REG_1	79,565	75,129	80,173	86,129	15,196	15,803	17,908	17,411	0.96	0.98	0.94	4,311	4,484	5,992
REG_2	39,125	38,643	40,144	40,265	5,275	7,180	5,057	5,306	0.94	0.88	0.93	2,893	2,499	1,972
REG_3	148,433	146,140	140,898	157,650	22,867	26,428	26,709	25,943	0.95	0.98	0.92	8,894	6,619	10,055
REG_4	17,965	14,988	26,325	16,714	3,794	2,791	4,042	4,016	0.78	0.92	0.83	2,362	1,582	2,282
REG_5	25,889	23,908	24,966	28,340	4,071	5,221	3,141	4,302	0.88	0.97	0.86	2,558	1,285	2,255
REG_6	195,714	209,282	201,953	227,330	43,078	30,845	30,989	33,507	0.63	0.81	0.27	33,750	25,797	46,911
REG_7	25,271	26,986	34,521	28,875	4,638	3,613	3,582	4,843	0.83	0.72	0.76	2,578	3,252	3,286
REG_8	73,195	69,486	74,514	77,031	9,812	9,157	12,445	11,450	0.94	0.97	0.94	3,388	3,774	3,969
REG_9	68,241	59,184	85,802	68,497	14,666	13,310	20,620	18,378	0.71	0.99	0.83	10,734	6,212	10,395
REG_10	23,627	23,040	25,125	23,373	3,899	4,053	6,417	3,576	0.93	0.95	0.81	1,446	2,954	2,309
REG_11	407,723	391,863	389,326	436,705	51,481	51,412	61,039	65,457	0.85	0.97	0.89	28,370	15,909	31,106
REG_12	5,849	4,317	8,081	6,254	1,206	893	2,336	1,120	0.69	0.97	0.79	869	1,194	767
REG_13	11,475	10,157	16,532	10,291	3,041	2,137	5,417	2,901	0.71	0.99	0.91	2,154	2,458	1,239
REG_14	15,140	12,448	19,000	16,092	2,999	1,918	4,952	3,226	0.83	0.98	0.90	1,768	2,091	1,400
REG_15	424,168	428,121	406,922	502,216	138,315	56,210	68,301	79,105	0.71	1.00	0.87	105,755	70,018	79,884
REG_16	5,725	3,065	12,862	4,825	2,352	1,105	4,132	1,340	0.71	0.99	0.69	1,746	1,831	1,735
REG_17	25,482	21,021	44,745	20,054	7,147	5,975	9,887	5,992	0.63	0.96	0.80	5,760	3,586	4,260

[3] The RMS error was derived from the geometric relationship between standard deviation of reference and simulation (σ_r and σ_s respectively) and the coefficient of correlation (R): $RMS^2 = \sigma_s^2 + \sigma_r^2 - 2\,\sigma_s \cdot \sigma_r \cdot R$

Table B 3. Mean annual total phosphorous (TP, ton/y) in the period (2000-2009) ("average" column) and the statistics derived from Taylor diagram. The table includes where available the observations (and correlated statistics) in black, otherwise the average of the three model simulations (and correlated statistics) in blue.

	TP average (ton/y)				standard deviation (ton/y)				Coeff. Of Correlation			RMS error (ton/y)[4]		
Region	Reference	SWAT	GREEN	MONERIS	Reference	SWAT	GREEN	MONERIS	SWAT	GREEN	MONERIS	SWAT	GREEN	MONERIS
REG_0	23,930	24,697	23,030	23,063	20,489	3,165	2,476	4,250	0.70	0.82	0.75	18,396	18,523	17,532
REG_1	1,954	1,580	3,373	1,629	508	226	990	255	0.91	0.99	0.63	316	495	399
REG_2	655	738	2,079	1,590	307	106	391	132	0.87	0.30	0.81	221	419	214
REG_3	4,046	3,020	6,296	3,638	2,770	400	1,620	446	0.55	0.13	0.41	2,573	3,020	2,618
REG_4	402	509	1,662	412	108	154	162	92	0.23	0.60	0.72	167	130	76
REG_5	1,060	1,268	1,497	632	171	225	115	119	0.73	0.94	0.77	153	74	110
REG_6	9,539	8,388	9,899	6,680	2,515	1,311	1,428	864	0.01	0.74	0.68	2,820	1,752	2,033
REG_7	1,474	1,289	2,134	1,106	406	161	219	160	0.26	0.33	0.38	396	393	375
REG_8	4,721	3,732	6,052	5,051	578	334	489	766	0.64	0.99	0.87	444	118	384
REG_9	3,754	4,121	5,161	2,630	738	826	708	849	0.79	0.98	0.82	515	151	486
REG_10	1,438	955	1,797	1,739	164	131	195	210	0.71	0.82	0.80	116	112	127
REG_11	23,637	19,065	21,131	19,391	12,377	2,597	2,043	3,314	0.53	0.51	0.17	11,209	11,461	12,267
REG_12	435	424	614	355	75	50	116	75	0.69	0.99	0.80	54	44	47
REG_13	718	862	1,143	356	149	131	247	101	0.53	0.99	0.77	137	103	95
REG_14	1,465	1,477	1,679	1,362	172	105	205	246	0.75	0.95	0.92	117	66	112
REG_15	21,862	22,956	21,345	22,194	12,352	2,959	2,079	3,748	0.71	0.64	0.49	10,454	11,145	11,024
REG_16	397	290	908	254	171	67	209	160	0.62	0.99	0.76	140	45	115
REG_17	1,351	1,129	2,989	1,027	1,045	163	402	352	0.99	0.99	0.72	885	649	827

[4] The RMS error was derived from the geometric relationship between standard deviation of reference and simulation (σ_r and σ_s respectively) and the coefficient of correlation (R): $RMS^2 = \sigma_s^2 + \sigma_r^2 - 2\,\sigma_s \cdot \sigma_r \cdot R$

List of figures

List of tables

Appendix A

Appendix B

LB-NA-27676-EN-N

JRC Mission

As the Commission's in-house science service, the Joint Research Centre's mission is to provide EU policies with independent, evidence-based scientific and technical support throughout the whole policy cycle.

Working in close cooperation with policy Directorates-General, the JRC addresses key societal challenges while stimulating innovation through developing new methods, tools and standards, and sharing its know-how with the Member States, the scientific community and international partners.

Serving society
Stimulating innovation
Supporting legislation

Publications Office

doi:10.2788/156278
ISBN 978-92-79-54239-8