KIT

Karlsruher Institut für Technologie

iWG

Fachbereich Siedlungswasserwirtschaft und Wassergütewirtschaft

KIT - Die Forschungsuniversität in der Helmholtz-Gemeinschaft

Adrian Wagner

Event-Based Measurement and Mean Annual Flux Assessment of Suspended Sediment in Meso Scale Catchments

156

Event-Based Measurement and Mean Annual Flux Assessment
of Suspended Sediment in Meso Scale Catchments

Zur Erlangung des akademischen Grades eines

DOKTORS DER INGENIEURWISSENSCHAFTEN (Dr.-Ing.)

von der KIT-Fakultät für

Bauingenieur-, Geo- und Umweltwissenschaften

des Karlsruher Instituts für Technologie (KIT)

genehmigte

DISSERTATION

von

Adrian Wagner, M.Sc.

aus Freiburg im Breisgau

Tag der mündlichen Prüfung: 06.12.2019

Referent: PD Dr.-Ing. Stephan Fuchs

Korreferent: Prof. Dr. rer. nat. Andreas Lorke

Karlsruhe 2019

ABSTRACT

Humans substantially interfere with the sediment household of fluvial systems, which has negative environmental and economic consequences. Agriculture and deforestation more than doubled global natural erosion while large dams trap at least 25% of the sediment flux to the global ocean. Sediment is a vector for pollutants and nutrients, transporting large shares of their annual budget during short and often poorly monitored periods of extreme flow. Precise estimates of mean annual fluxes of suspended sediment and particle-bound substances are needed on the scale of meso catchments (50-250 km²) for supporting management decisions and area-wide emission model validation. For covering high-flow transport of sediment and phosphorous, large volume samplers (LVS, composite sample volume = 1 m³) were installed to in three meso catchments of different climatic, pedologic and land use characteristics to sample discharge-proportionally. A novel approach to construct event-based rating curves from LVS measurements was developed, based on event mean concentration and the highest discharge in the period. Subsequently calculated LVS fluxes were compared to grab sampling based assessments and modeled emissions. The LVS proved to be robust towards different flow conditions at two of the sites and failed due to technical problems at the third site, where the sediment stock of a downstream reservoir serves as an additional reference. Grab sampling assessments are biased low. The bias is driven by sediment yield, particle-association of the parameter and flow flashiness. The existing flow flashiness index (FFI) was modified into a load flashiness index (LFI) to generalize shortness and flux contribution of high-flow events. LVS sampling and LVS rating curves are recommended for load-flashy catchments and parameters; when cost of analysis must be minimized and when in-situ proxies are inadequate. By measuring and modeling fluxes, deficiencies in the emission model were revealed and largely attributed to the erosion pathway's static catchment connectivity assessment. To provide meaningful model validation, sampling strategies must be adapted to load flashiness, which is usually not considered. In addition, a new low-maintenance method for autonomous above-water reflectance measurements was established and compared to turbidimetry and rating curve assessments. In several reservoir campaigns, reflectance and partial least squares regression showed high predictive abilities for total phosphorous and suspended sediment surface concentration. River reflectance monitoring was limited by changing light conditions throughout the day and year and did not improve flux assessments in eleven LVS validation events. Turbidity and grab sampling rating curves were also unable to represent event fluxes, while LVS rating curves performed best. Reflectance monitoring is beneficial in terms of maintenance and multi-parameter retrieval. It may be employed with artificial lighting and at larger streams with lower flashiness and optically deep conditions. Optical methods can provide critical information, but require expertise and calibration samples across the desired range. Specifically sampling high flows provides this as well as better flux estimates and must therefore be integrated in sampling strategies. While the LFI guides site prioritization, LVS sampling and LVS fluxes provide a meaningful yet applicable observation scale required to represent particulate flux processes for meso catchments, and should therefore be considered for widespread implementation.

ZUSAMMENFASSUNG

Eingriffe durch den Menschen haben den Sedimenthaushalt fluvialer Systeme grundlegend verändert, mit negativen ökologischen und wirtschaftlichen Folgen. Landwirtschaft und Abholzung haben das Ausmaß natürlicher Erosion mehr als verdoppelt, während große Staudämme mindestens 25% des global transportierten Sediments auffangen. Sediment ist ein Transportvektor für Schad- und Nährstoffe. Große Anteile der mittleren Jahresfracht werden in kurzen und häufig unzureichend beprobten Zeiträumen mit hohen Abflüssen transportiert. Genaue Schätzungen der mittleren Jahresfracht von suspendiertem Sediment (abfiltrierbare Stoffe) und partikulär transportierten Stoffen werden für mesoskalige Einzugsgebieten (50–250 km²) zur Bewirtschaftungsplanung und zur Plausibilisierung von flächendeckenden Stoffeintragsmodellen benötigt. Um Hochwasserfrachten von abfiltrierbaren Stoffen und Phosphor zu erfassen, wurden großvolumige Sammler (LVS, Probevolumen = 1 m³) in drei mesoskaligen Einzugsgebieten unterschiedlicher klimatischer, pedologischer und landnutzungsbezogener Bedingungen zur abflussproportionalen Beprobung installiert. Aus dem höchsten gemessenen Abfluss im Beprobungszeitraum und der mittleren Ereigniskonzentration wurde ein neuartiger Ansatz zur Erstellung von Konzentrationsschlüsselkurven entwickelt. So konnten Frachten berechnet und mit Berechnungen auf Basis von Einzelbeprobungen verglichen werden. Der LVS erwies sich an zwei Standorten als robust gegenüber veränderlichen Abflussbedingungen, konnte am dritten Standort wegen technischer Mängel aber keine Hochwassermessungen durchführen. An diesem Standort wurde der Sedimentkörper des unterliegenden Stausees als Bezugswert verwendet. Die Berechnungen auf Basis von Einzelbeprobungen sind systematisch zu niedrig. Der Fehler wird durch die spezifische Sedimentfracht, den Grad der Assoziation zu Partikeln und die Volatilität der Abflüsse bestimmt. Ein existierender Abflussvolatilitätsindex (FFI) wurde zu einem Frachtvolatilitätsindex (LFI) angepasst um die Kürze von hohen Abflüssen und deren Bedeutung für die Gesamtfracht in einer Kennzahl zu vereinen. Die Beprobung mit dem LVS und die Anwendung der Schlüsselkurven wird für frachtvolatile Einzugsgebiete und Parameter, zur Begrenzung von Analysekosten sowie in Fällen empfohlen, wenn in-situ gemessene Näherungsparameter nicht angemessen sind. Durch Frachtmessungen und –modellierung wurden Defizite beim Stoffeintragsmodell deutlich, die größtenteils auf den Eintragspfad Erosion und dessen statische Konnektivitätsansätze des Einzugsgebiets zurückgehen. Entgegen der derzeitig üblichen Praxis müssen Beprobungsstrategien der Frachtvolatilität angepasst werden, um Eintragsmodelle aussagekräftig Plausibilisieren zu können. Zusätzlich wurde eine neue wartungsarme Methode zur autonomen, kontaktlosen Reflektanzmessung entwickelt und mit Trübungsmessungen und Schätzungen durch Schlüsselkurven verglichen. Während mehrerer Messkampagnen auf einem Stausee konnten durch Reflektanzmessungen in einer Regression der partiellen kleinsten Quadrate hohe Anpassungsgüten für Gesamtphosphor und abfiltrierbare Stoffe erreicht werden. Dauermessungen am Fluss wurden durch veränderliche Lichtbedingungen im Tages- und Jahresverlauf eingeschränkt und konnten die Frachtermittlung in elf Validierungsmessungen nicht verbessern. Trübemessungen und Schlüsselkurven auf Basis von Einzelbeprobungen konnten Ereignisfrachten ebenfalls nicht angemessen nachzeichnen, LVS-Schlüsselkurven erzielten die höchste Güte. Reflektanzmessungen haben Vorzüge hinsichtlich des niedrigen Wartungsbedarfs und der gleichzeitigen Bestimmung mehrerer Parameter.

Ihr Potential sollte mit künstlicher Beleuchtung sowie an größeren Gewässern mit niedrigerer Volatilität und optisch Tiefen Bedingungen weiter getestet werden. Optische Messmethoden können wichtige Informationen liefern, benötigen aber spezielle Fachkenntnis und Kalibrationsmessungen über die volle Breite des Messbereichs. Die explizite Beprobung von Hochwasserereignissen kann diese Daten liefern und zugleich Frachtabschätzungen verbessern, weshalb sie stets in Beprobungsprogramme integriert werden müssen. Während der LFI zur Priorisierung der Standorte genutzt werden kann, liefern LVS-Beprobungen und LVS-Frachten eine aussagekräftige und zugleich anwendbare Beobachtungsskala, die für die Wiedergabe der partikulären Frachtprozesse in mesoskaligen Einzugsgebieten benötigt wird. Die breite Implementierung dieser Beprobungstechnik sollte daher geprüft werden.

DANKSAGUNG

CONTENT

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS

adjCV	bias-corrected (adjusted) Cross-Validation
CV	Cross-Validation
FI	Flashiness Index
FFI	Flow Flashiness Index
LFI	Load Flashiness Index
LVS	Large Volume Sampler
METRIS	Modelling of Emissions and Transport in River Systems
MoRE	Modelling of Regionalized Emissions
MUSLE	Modified Universal Soil Loss Equation
NTU	Nephelometric Turbidity Units
OAC	Optically Active Constituents
o-PO$_4$-P	Orthophosphate-phosphorous
PLS	Partial Least Squares Regression
Q$_{10}$	The highest discharges contributing 10% of the cumulative annual discharge
R^2	Coefficient of determination. In plots: r2; r2adj (adjusted R^2)
RC	Rating Curve
RMSEP	Root Mean Square Error of Prediction
RUSLE	Revised Universal Soil Loss Equation
SS	Suspended Sediment
SDD	Secchi Disk Depth
SSC	Suspended Sediment Concentration
SWAT	Soil and Water Assessment Tool
TP	Total Phosphorous
TSS	Total Suspended Solids
WFD	Water Framework Directive

1 SUSPENDED SEDIMENT IN THE AQUATIC ENVIRONMENT

In the following, the role of suspended sediment in aquatic inland systems is given. The section begins with considerations regarding the historic global development and today's relevance of suspended sediment (1.1), followed by a review of processes contributing to the dynamics of suspended sediment across time and space (1.3). In sections 1.4 and 1.5, the composition of suspended sediment itself and its influence on water quality is outlined. Section 1.6 briefly introduces management options available for sediment-related issues in aquatic systems.

1.1 A KEY PARAMETER IN FLUVIAL SYSTEMS

Suspended sediment plays a key role for surface water quality and is related to various processes in river basins. The material transfer from the continents to the world's oceans by rivers is one of the most important dislocation processes on a global scale. Weathering and denudation constantly supply erosion material to be transported along the gravity gradient, in most cases towards the sea. Here, particles may be deposited, compacted to form bedrock and eventually raised by geological processes to be weathered down again. Deforestation, construction and mining activities and –above all– agriculture have greatly accelerated these processes throughout human history, with the sharpest increase following the industrial revolution (Owens *et al.* 2005: 694). Many places in the world have today reached a degree of erosion that endangers the functioning of terrestrial and aquatic ecosystems and the people dependent on it: About 80% of the world's agricultural land is affected by moderate or severe erosion, resulting in a global loss of at least 75 10^9 t a^{-1} of fertile soil from these systems (Pimentel 2006). Pimentel (2006) further states that eroded material typically contains several times more nutrients than the remaining soil, leading to a depletion of soil quality and productivity. Together with sediment from other sources, a relatively small share of this material reaches a stream network. Rivers are responsible for 95% of the sediment mass transport to the oceans which is estimated to amount to 15–20 10^9 t a^{-1} (Walling 2006). By weight and volume, suspended matter is therefore the most common pollutant in the earth's rivers– if it is to be regarded as a pollutant. Sediment is a natural and essential part for healthy fluvial systems (Brils 2008). For many rivers, high concentrations of suspended sediment alone do not indicate poor water quality, and low concentrations are not necessarily management goals for every system. Because of its close link to other water quality parameters and to hydromorphology, it is a key parameter needed to understand and manage fluvial systems. Like no other water quality parameter, the presence and absence of suspended sediment to a large part determines the erosive and depositive properties of rivers, and thereby forms the landscape and controls ground water levels. Suspended sediment is further associated with many hydrophobic pollutants such as polycyclic aromatic hydrocarbons, heavy metals and others, acting as a facilitator for pollutant transport. It is therefore an informative surrogate for these pollutants, as it can be monitored at relatively low costs. Many harmful particle-associated substances were introduced and increasingly emitted during the last 150 years.

This lead to the convergence of globally increasing suspended sediment availability and increasing pollutant attachment onto these solids. Special attention needs to be directed towards the macro nutrient phosphorous, which is also mostly transported in particulate form (Apitz and White 2003: 133). Its loss from agricultural land following erosion is undesirable for economic reasons. In the receiving water body, it often leads to ecological degradation. Being the limiting factor for primary production in most freshwater bodies, it can have adverse effects by enhancing the growth of algae and or cyanobacteria, some of which produce toxins. Enhanced growth eventually leads to oxygen depletion as well as a general poverty in species within the affected environments (Seitzinger *et al.* 2010). This may in turn affect the ecosystem services provided by these systems and can lead towards ecological tipping points, from which it is difficult to return to the original state (Dokulil and Teubner 2003). Apart from the direct ecological consequences in the rivers, reservoirs, lakes, deltas and coastal areas, the phosphorous flux to the oceans and its subsequent dilution represents a terminal loss of a non-renewable resource on which the global food production largely depends on (Liu *et al.* 2008).

Phosphorous and other particle-bound substances are permanently or temporarily stored in river-, channel- or reservoir sediments. They can be re-mobilized under certain flow or reductive conditions (Hupfer and Lewandowski 2008), sometimes affecting the water quality years or even decades after their emission. Connected to this is a severe disposal problem for the contaminated sediment, e.g. when dredging becomes necessary. The accumulation of sediment in reservoirs represents a disruption of the watershed's carbon cycle and may be accelerating climate change. The main reservoir mechanism affecting the global climate can be outlined as follows: Instead of being mineralized to CO_2 under aerobic conditions in flowing water bodies, organic carbon is more likely to be transformed to CH_4 (methane) in the anaerobic bottom sediment conditions when trapped in reservoirs. Mainly because methane has a 25 times greater global warming potential than CO_2, an annual emission of 0.8 10^9 tons of CO_2-equivalents can be attributed to man-made water surfaces globally. This represents 1.3% of the global anthropogenic CO_2-equivalent emissions (Deemer *et al.* 2016). In this estimate, only direct emissions are regarded. For net emissions, additional factors like the greenhouse gas balance of the pre-flooded area, the carbon burial rate in the reservoir and its long term fate have to be included. The few existing holistic studies point towards reservoirs being a net source of CO_2-equivalents (Teodoru *et al.* 2012; Faria *et al.* 2015). The overall effect of reservoir construction on the catchment's carbon cycle is complex and specific for climate and other factors. Lakes and reservoirs receive, transport and transform large amounts of carbon from their watersheds and have therefore a greater role in carbon cycling than indicated by their spatial extent. Annually, about 2.9 10^9 tons of carbon are imported to global inland waters, of which 2.0 10^9 tons are either stored in sediments or emitted to the atmosphere, while the rest is transported to the oceans (Tranvik *et al.* 2009). These amounts are just a small fraction of the 100–150 10^9 tons (Randerson *et al.* 2002) of carbon absorbed by photosynthesis annually, but they are in the same order of magnitude of e.g. the carbon uptake of the oceans (2.6 10^9 tons) (Tranvik *et al.* 2009). Considering the large number of dams completed in the last decades with many more under construction, sediment management is among the greatest challenges of present-day environmental engineering. Downstream of dams, the lack of suspended sediment can also become a problem (Owens *et al.* 2005: 696). Increased river bed erosion, sinking groundwater levels and the absence of substrate and nutrients for freshwater and coastal ecosystems are examples for this.

In recent decades, emission pathways and environmental effects of suspended sediment and pollutants have become better understood. Large-scale technical measures and management practices in the catchments to reduce, mitigate and adapt have been implemented in many countries in the world. However, environmental protection measures are not enforced everywhere and often fall short of the requirements of sustainability. Given the complex and at times counter-directional anthropogenic influences on the generation, composition and transport behavior of suspended sediment, watershed-specific evaluations on sediment budgets are needed. Modeling and monitoring techniques can complement each other in order to maximize spatial coverage, temporal resolution and to differentiate sources and emission pathways.

1.2 HUMAN INTERVENTIONS ON THE GLOBAL SEDIMENT HOUSEHOLD

"The ultimate destiny of all reservoirs is to be filled with sediments" (Jain and Singh 2003, p. 684)

All stages of fluvial transport dynamics have been heavily influenced by mankind on a global scale (Bravard and Petts 1996; Syvitski *et al.* 2005; Walling 2006; Grill *et al.* 2019). In terms of sediment loads, two opposed interferences stand out: An increase of sediment availability because of the disruption of the vegetation cover and an increase in sediment trapping behind flow-regulating structures (Kondolf *et al.* 2014). Agriculture and deforestation are the main causes for accelerated erosion, and about 60% of current soil loss is caused by human activity (Yang *et al.* 2003). This heightens the availability of sediment which may be transported towards the stream network. Yang *et al.* (2003) further found that cropland expansion in the last 100 years alone contributed to a 17% increase; that one third of the global arable land was lost to erosion in the last 40 years of the 20th century; and that erosion rates are expected to increase in the future. Secondly, the construction of dams and reservoirs at most large rivers in the world promotes a massive trapping of suspended sediment, leading to sediment starvation downstream of these constructions (Vörösmarty *et al.* 2003). Worldwide, about half of all large rivers (mean discharge > 1000 m³ s⁻¹) have a cumulative reservoir capacity to store at least 2% (some more than 300%) of the total annual flow and are therefore considered 'affected' by adverse effects of flow regulation (Lehner *et al.* 2011). About 40% of the total land surface excluding Antarctica and Greenland are flow regulated catchments. Reservoirs can be very effective in trapping sediment, especially when residence times are long. Vörösmarty *et al.* (2003) estimate a global trapping of 4-5 10^9 t a⁻¹, or 25–30% of the total sediment flux. This corresponds to an annual capacity loss rate of 0.5%–1% (Mahmood 1987; Schleiss *et al.* 2016) and resulted in 330 10^9 t of sediment being currently stored behind large dams globally. According to Wisser *et al.* (2013), more than half of this amount is trapped in about 600 large reservoirs (≥ 0.5 km³ max. storage capacity), whereas 44,000 smaller reservoirs are responsible for the other half. Like all global data bases on dams, an estimated 800,000 additional, small and unregistered reservoirs were not considered in the calculation, suggesting even higher trapping rates. Global reservoir storage was estimated to have peaked in 2006 (Wisser *et al.* 2013), which means that the provision of additional capacity in new reservoirs cannot keep up with storage loss due to siltation and dam removal (often also a management decision following severe siltation). In the study by Wisser *et al.* (2013), global population growth was

included as a proxy for water demand and it was found that global storage volume per capita has already been declining since 1980s.

Reservoirs are built for various purposes, with electricity generation and water storage for irrigation standing out as the main uses. Broadly speaking, in temperate regions of Europe and North America, dams are built for hydropower, drinking water generation and recreation with a tendency towards multi-purpose dams in the last decades. In regions with more irregular rainfall patterns like Southern Europe and the Mid-West of the USA, Central America, Africa, South Asia and Australia, irrigation dams are the main type. Irrigation dams are by far the most numerous, but they are usually small and have a large share of single-purpose structures. In contrast, flood control is more common in multi-purpose reservoirs. Due to the flow regulating and dampening effect of reservoirs, it can in many cases be best described as an auxiliary function of reservoirs. Flood control is an important function of dams, especially in catchments with high elevation headwaters e.g. around the Himalayans and Andes (Lempérière and Lafitte 2006; Biemans *et al.* 2011).

Next flow regulation, electricity and drinking water generation are the main uses of reservoirs in South America. Due to its topography and climate, the continent has the greatest freshwater resources in the world, of which Brazil has by far the largest share. In the north east of the country, the Amazon alone discharges 6,300 km³ a^{-1} or about 18% of the global total, more than the next seven largest rivers combined (Milliman and Farnsworth 2013: 14). With the water, it discharges only 1.2 10^9 t a^{-1} or 6% of the global sediment flux (Milliman and Farnsworth 2013: 28), partly due to its large floodplain areas, trapping much of the sediment before it reaches the oceans. Because of the low population density in the Amazon basin, Brazil has been building damns mainly on the population rich south east, the most notable exception being the Belo Monte Dam at the Rio Xingui, a main tributary to the Amazon. Compared to the global average, Brazil has a lower share of regulated large rivers (22.4% compared to 46.7% globally, Lehner *et al.* 2011). However, its reservoirs are large, providing 6% of the global reservoir capacity in just 179 structures. In comparison, China built almost 800 reservoirs to store the same amount of water (Lehner *et al.* 2011). Additionally, Brazil is largely dependent on electricity from hydropower (~80% of the total consumption) and is one of the main dam building countries in the world (Zarfl *et al.* 2015).

Fed from a catchment just 15% of the size of the Amazons', the Huanghe (Yellow River) in China throughout the 1950s to 1970s famously was responsible for almost the same sediment flux (1.1 10^9 t a^{-1}) while discharging less than a hundredth of the water (57 km³ a^{-1}) into the ocean. By the beginning of the 21st century, both water and sediment discharge decreased to just a fraction of these values. Annual sediment loads were as low as 0.1 10^9 t a^{-1} in the years 2000–2002, while annual water discharge due to a prolonged drought was just 5 km³ a^{-1} between 1997–2002 (Milliman and Farnsworth 2013). Given the former relevance of the Huanghe as a deliverer of sediment, this corresponded to a global sediment flux reduction of at least 4% (Walling 2006). It has been stated that by this reduction, the Huanghe catchment returned to the level of sediment export before the time of human intervention. The first of these interventions was massive deforestation of the loess catchment starting 1000 B.C. which led to a distinct increase of sediment transport over the centuries. Wang *et al.* (2007) studied the changes in the Huanghes sediment budget and found that its sediment is composed to 90% of eroded soil.

In the same study, it was further found that the recent decline in concentrations and loads could be attributed to reduced precipitation due to climate change (explains ca. 30% of the reduction);

increasing water abstraction for irrigation and soil conservation measures (combined 40%) and the construction of reservoirs (30%). There are over 3000 reservoirs in the Huanghe catchment, ranging from many small impoundments to 24 'large' dams, each storing more than 0.1 km³ of water. The largest four of them (3.1–24.7 km³ designed storage capacity; Chen *et al.* 2005) are accountable for two thirds of the Huanghe sediment trapping (Wang *et al.* 2007). The Sanmexia Dam (3.1 km³) already lost about 20% of its storage capacity during the 18-month filling phase after construction was completed in 1960, and was filled to over 60% just four years later (Wang *et al.* 2005). This implies a trapping about 0.55 10^9 t of sediment per year, or 93% of the inflowing particulate material and posed a serious threat to the structure and its purpose itself, to flood security and the proper functioning of transport mechanisms in the catchment. After a series of large-scale technical and management measures to flush out accumulated sediment and limit further siltation, the Sanmexia Dam today is operational with strong restrictions in regard of the original design (Wang *et al.* 2005). The situation in Sanmexia Dam is by no means unusual. There are numerous examples of reservoirs which have been filled with sediment within years of construction. Although the phenomena is present in other countries, many cases are described in China (cf. Jain and Singh 2003), presumably because of the pedologic-climatologic conditions, the large number of dams and the extensive scientific community in the country. In total, the loss rate of storage volume is several times higher in China (2.3%) than the global average (0.5–1% Wang and Hu 2009)

The Huanghe therefore is just one, well-documented example for complex interlinkages of catchment and water regulation on sediment loads, which have also been studied e.g. in the Mississippi in the USA (Mize *et al.* 2018), the Nethen in Central Belgium (Verstraeten *et al.* 2009), the Rhone in France (Loizeau and Dominik 2000) or the Paraná and Paraguay Rivers in Brazil and neighboring countries (Amsler and Drago 2009). These cases illustrate how human interventions can influence the transport system of a catchment in –at times– opposed directions. While for a heavily altered catchment like the Huanghes' the amount of sediment reaching the ocean today may be similar to pre-human levels, the condition of the catchment and the transport properties of the river are not. It must be acknowledged that the interventions are highly site-specific, and management tools differ from catchment to catchment.

1.3 SPATIO-TEMPORAL DYNAMICS OF SUSPENDED SEDIMENT TRANSPORT

Suspended sediment transport in a landscape follows complex and irregular patterns in time and space (Förstner and Owens 2007). For a given river, the suspended sediment concentration may vary over three orders of magnitude within one year (Chapman 1996). It is controlled by a number of factors, many of which link back to erosion processes in the catchment, usually the main source of suspended sediment in rivers (Vercruysse *et al.* 2017). Erosion is controlled by slope lengths and steepness, soil properties, land cover, land use and management practices (*catchment characteristics* in Figure 1) as well as meteorological parameters such as rainfall amounts and intensities. The amount sediment and of substances like phosphorous may vary due to flushing or dilution, the event magnitude, the type and number of preceding events, and the time of the year with respect to agricultural practices (Kulasova *et al.* 2012). Although there are other sources of suspended sediment like urban runoff or atmospheric deposition, the erosion parameters to a large part determine the general availability of material which may be entering the stream network. The eroded material may be transported to a stream or deposited temporarily or permanently

on the way, depending on the catchment connectivity (Fryirs 2013). Globally, about 90% of eroded material is not flushed out of the catchment, but stored within the basin (Wisser *et al.* 2013), a phenomenon termed 'sediment delivery problem' by Walling (1983). Connectivity is described by Vercruysse *et al.* (2017) as natural and man-made linkages (e.g. bank erosion) and blockages (e.g. road embankments) that control the movement of sediment within a catchment (*catchment connectivity* in Figure 1). This conceptual framework helps to explain why certain catchment areas prone to erosion are not contributing to suspended sediment loads, i.e. because they are not connected to a river. If particles do reach a water body, they may be carried away with the current or deposited at the bottom, mostly depending on the grain sizes of the material and varying, discharge-dependent flow velocities of the stream (*river transport capacity* in Figure 1). The subsequent variation of transport, sedimentation and resuspension in a fluvial system is termed sediment cascade to underline its non-linearity (Fryirs 2013).
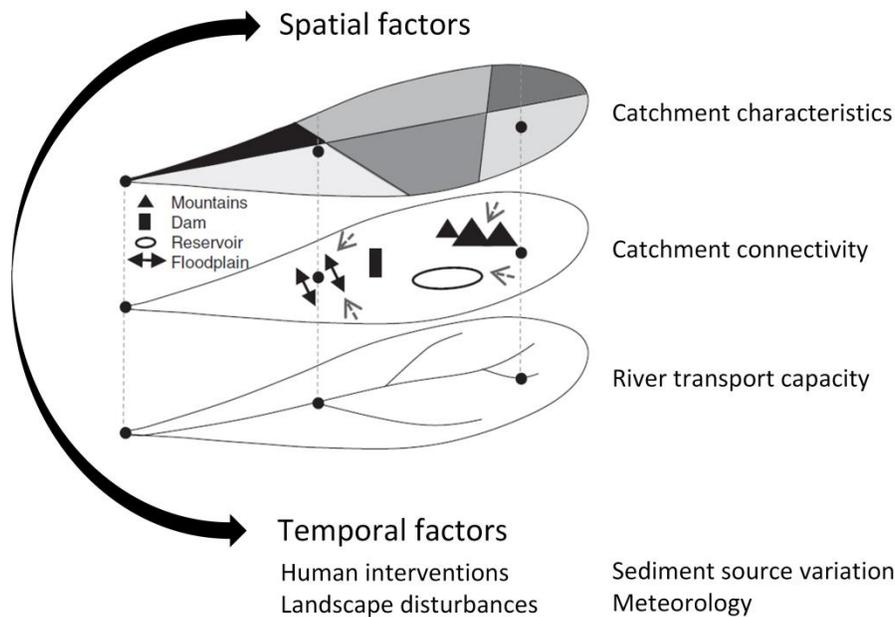
Figure 1: Conceptual framework to describe suspended sediment transport in a catchment. Superposed spatial and temporal factors induce a spatio-temporal complexity. Catchment characteristics: geology, land use, climate, topography. Catchment connectivity: Linkages and (artificial) blockages controlling sediment delivery to the stream network (modified after Vercruysse *et al.* 2017: 40).

The combination of factors controlling erosion outlined above is therefore unique to each catchment. Within a given area, it can change distinctly on a variety time scales from millennia (geological processes) to decades (e.g. anthropogenic land use change) to months (cropping) to minutes (rain events). To describe the transport behavior of suspended sediment, the spatial complexity therefore has to be extended by a temporal complexity (Vercruysse *et al.* 2017). Outside deserts, the amount and intensity of rainfall –and subsequently the height and shape of the hydrograph– usually control the suspended sediment concentration in rivers. Following the patterns of rainfall, in many systems SSC is in a relatively low range for extended periods of time, intermittent by comparatively short periods of high discharges and high concentrations (Horowitz 2008). This means that typically, a large share of the annual water and substance flux takes place during few 'hot moments'; isolated *events*, when discharges and substance

concentrations are both high (Kulasova *et al.* 2012). Horowitz (2003) stated that sediment rating curves in many cases underpredict high (and overpredict low) SSC and strongly argues for hydrology-based sampling schemes instead of the usual calendar-based sampling implemented by the authorities. For the Mississippi Basin at Thebes, above the confluence of the Ohio River, just eleven high-flow events within five years (1996-2000) contributed to 85% of the total sediment flux in that time (Horowitz 2013). Although the term 'event' is not closer defined here, it underlines the importance of elevated flows. Kentroti (2017) found for a selection of 44 gauging stations at a variety of streams and rivers in the state of Baden-Württemberg (Germany) that an average of 35% (min: 19%, max: 51%) of the annual water volume flux (without considering substance concentrations) was attributed to the highest 10% of discharges. This means that even if suspended sediment concentrations would be constant across discharge variations, high-flow events would be disproportionally responsible for the movement of material. In a similar study, Asselman (2000) included the positive correlation of SSC and discharge in the analysis. He created sediment rating curves for the Rhine River at several locations and for its main tributaries (total of 12 stations) and found that an average of 53% (min: 34%, max: 77%) of the annual sediment flux takes place during the highest 10% of the discharges. In both studies, the importance of high flow increases, as catchment sizes decrease. This is due to the fact that hydrographs in smaller catchments show a more 'flashy' behavior while rating curves become steeper, i.e. concentrations increase more for each increment of discharge (Asselman 2000). Overall, there is a clear tendency that transport processes for virtually all particle-associated substances are stormflow-driven, especially in smaller catchments (Figure 2).

On a narrower time scale, substance concentrations can follow patterns during the course of an event, depending on the stage of the hydrograph. Like in other disciplines, this dependence on pre-conditions is termed hysteresis. The concentration changes in a stream during the increase and later decrease of the discharge can be complex, because several at times antagonistic mechanisms are active (Gellis 2013). For instance, clear rain water can dilute substance-laden stream water. At the same time, erosion, urban wash-off and remobilized deposits from sewer systems can outmatch dilution even for non-particulate substances and can lead to peak concentrations during rising discharges. On a more general level, hysteresis can be classified along two gradients (Pohle *et al.* 2019). Firstly, there is the substance- and catchment-specific export behavior ranging from dilution to neutral and enrichment. Secondly, the hysteresis rotation pattern can be clockwise, counter-clockwise or without rotation. Therefore, nine classes of patterns concerning concentration changes during a discharge event can be identified (Figure 3).

In summary, suspended sediment transport related to many processes and highly dependent on catchment parameters and pre-conditions. Short and poorly monitored high-flow events have a strong tendency to be the main driver of suspended sediment transport. This is amplified by steep inclinations, high shares of arable land, small catchment sizes and high rainfall intensities. Monitoring programs are rarely adapted to these findings, which leads to underestimations of high-flow fluxes (Section 3.2).

Mean Stormflow/Baseflow Contributions to the AnnualFluxes of Suspended Sediments and Selected Constituents in City of Atlanta Streams
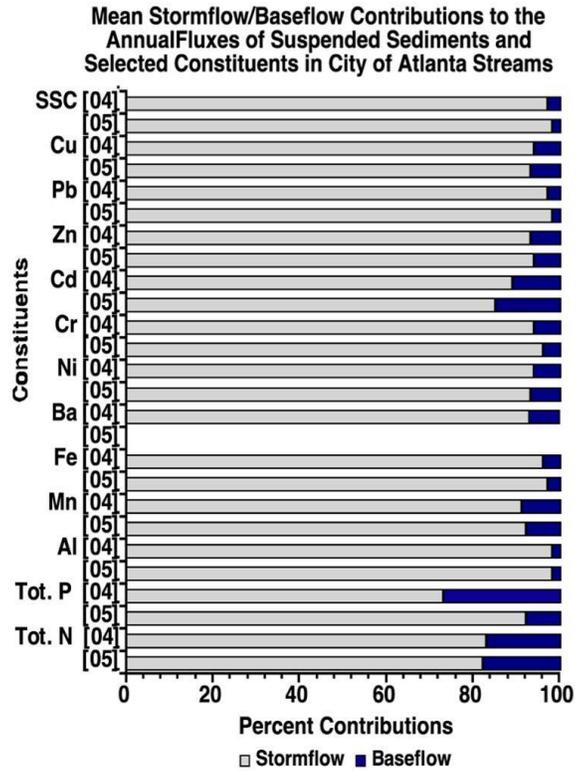
Figure 2: Contribution of base flow and storm flow to annual fluxes of SSC and particle-associated constituents in small urban streams in the City of Atlanta, GA, for the years 2004 [04] and 2005 [05] (Horowitz 2013: 2475).
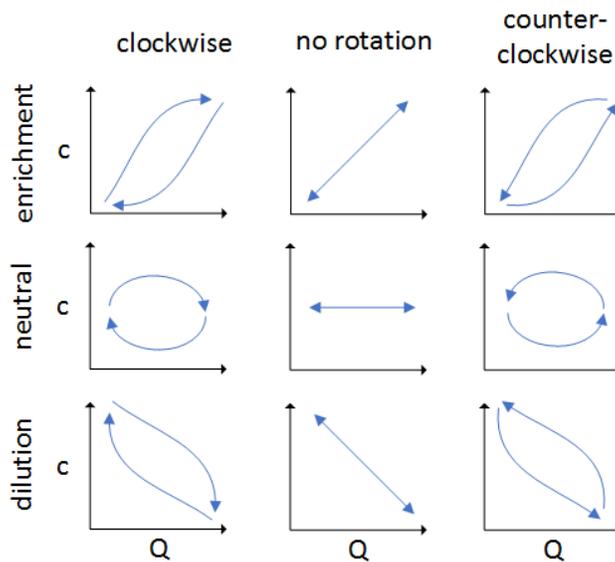
Figure 3: Idealized representation of nine concentration (c) – and discharge (Q) relationships during the rising and falling hydrograph of an event. The classes are grouped by export behavior (vertically) and hysteresis rotation (horizontally). After Pohle *et al.* (2019).

1.4 PROPERTIES AND SOURCES OF SUSPENDED SEDIMENT

Apart from being described by grain sizes and transport behavior, suspended particles can be described by their formation. Materials transported by rivers are to a large part attributed to external sources (Riggsbee *et al.* 2008). Examples for allochthonous material are calcite, clay or sand particles from eroded soil and bedrock; lignin, cellulose and other nonliving organic compounds (detritus) of terrestrial origin; or aerosols from combustion processes. The composition of suspended sediment in a water body therefore mirrors the composition of organic and inorganic solids present in the catchment (Evrard *et al.* 2011). More specifically, the composition of transported inorganic material is related the composition of the rock (and soil) types found in the catchment, whereby the least soluble elements like aluminum, iron, titanium and manganese are enriched in the particulate phase (Chapman 1996). Chapman (1996) further states that the silt fraction is often small in both the source material and the suspended sediment, leading to a clear separation between sand and clay particles. In loess catchments however, silt can be dominant in soils and suspended sediment alike. Autochthonous matter was formed in the water body itself; like phytoplankton and cyanobacteria (Gao *et al.* 2016). The relevance of autochthonous substances increases as flow velocities decrease, for instance in lowland rivers, natural lakes or reservoirs. Suspended particles can further be described by the degree of human influence ranging from naturally-formed matter like loess or desert dust, to combined, changed or polluted natural materials like tailings or cement, to entirely artificial materials like micro plastic. Lastly, suspended sediment can be characterized by its sources.

The topsoil of agricultural lands is in many systems a main contributor, from which ploughed land and overgrazed meadows are a more important source than permanent crop plantations. Under natural vegetation cover, average erosion rates are in the range of 0.001–2 t ha^{-1} a^{-1} for flat areas and between 1–5 t ha^{-1} a^{-1} in mountainous regions (Pimentel 2006). The same studies finds losses from pastures (in the US) to average at 6 t ha^{-1} a^{-1}, with the worldwide global average for cropland at about 30 t ha^{1} a^{-1}. Forestry and mining activity can also significantly increase suspended sediment yields of a catchment, and suspended sediment can also origin from the river bed and banks (Owens *et al.* 2005). The eroded bank and bed material can be subdivided into two types. The first type is genuine bank or bed erosion material from the underlying or adjacent (unconsolidated) rock and soil. The second type is material from outside the river channel itself that was temporarily deposited during low or medium flow conditions, and is resuspended during high flow conditions. When only measuring the river load, it is difficult to differentiate between the two types, which appear simultaneously to varying degrees depending on sediment input, catchment geology, river morphology, management history, flow conditions etc. (Dunne and Reid 2016). Steep river banks can indicate a history of net depth erosion. Methods like radionuclide analysis, determination of mineral magnetism or geochemical composition, as well as phosphorous and organic carbon content can help to determine the share of these and other suspended sediment sources; a technique called source fingerprinting (Walling *et al.* 1999; Carter *et al.* 2003; Förstner and Owens 2007).

In urban environments, construction sites, road wear, road dust and sewer systems are the main sources of suspended sediments (Owens *et al.* 2005). Well-managed waste water treatment plants are very effective in removing particles during treatment leading to effluent suspended solids concentrations of usually less than 10 mg L^{-1} (Mucha and Kułakowski 2016; Dehghani and Taleb Beydokhti 2018). Nevertheless, the effluent concentration can be higher than the concentration in the receiving water body.

In addition, in densely populated areas, treated waste water can make up a considerable share of the total river flow, especially during low-flow conditions (LUBW 2016). If untreated sewage is emitted into rivers via combined sewer overflows, substantial amounts of suspended sediment can be emitted with it in short periods of time (Riechel *et al.* 2016). More importantly, particles from sewage are more heavily laden with phosphorous (and other substances) than sediment from most other sources. For instance, Carter *et al.* (2003) found TP content in waste water treatment plant influent sediment of more than $20\,g\,kg^{-1}$. Compared to this, agricultural topsoil contents are low (around $1\,g\,kg^{-1}$ in Baden-Württemberg, LTZ 2015). Combined sewer overflows release a mix of domestic sewage and urban storm water, usually with a much greater share of rain water. Fuchs and Nickel (2018) determined a median overflow SSC of $52\,mg\,L^{-1}$ for about 100 events in Bavaria, indicating a potential stress for the receiving water body. Phosphorous contents in overflow sediment were about $10\,g\,kg^{-1}$. Storm water outlets from separate sewer systems should, by design, not emit domestic sewage, but contain polluted particles from the urban environment such as road wear, roofs and facades, leaves, litter, dust, etc. Storm water is not or not thoroughly treated everywhere, which is why it can be a substantial emission pathway for suspended sediment and pollutants (Fuchs *et al.* 2014). The contribution of all the sources discussed above to the total amount of sediment depends on the flow situation. While soil erosion occurs most prominently during high-intensity precipitation events, combined sewer overflows and storm water releases can be triggered from moderate and even low rainfall (Fuchs and Nickel 2018). This shifts the relative contributions of the different sediment during in the course of a rainfall event (Figure 4), depending on intensity, amount and pre-conditions.

Overall, suspended sediment can have a variety of sources. It is often dominated by eroded material from arable land or other areas where the natural vegetation cover is disturbed. In these cases, a large share of phosphorous is transported in particulate form and therefore covaries with suspended sediment concentrations. Phosphorous content in sediment from sewer systems, which locally represents another important source of suspended sediment, can be expected to be several times higher than in natural and fertilized soils. This can give indications about the origin of the sediment in a sample. The frequency and magnitude of overflow and release events are usually not monitored and poorly understood. Storm and combined sewer overflow structures are numerous and scattered along urban stream sections. Next to truly diffusive sources such as agricultural sites, they are therefore difficult to measure directly. For these reasons, suitable in-river monitoring systems and sampling schemes are needed to complement and validate robust modeling approaches.
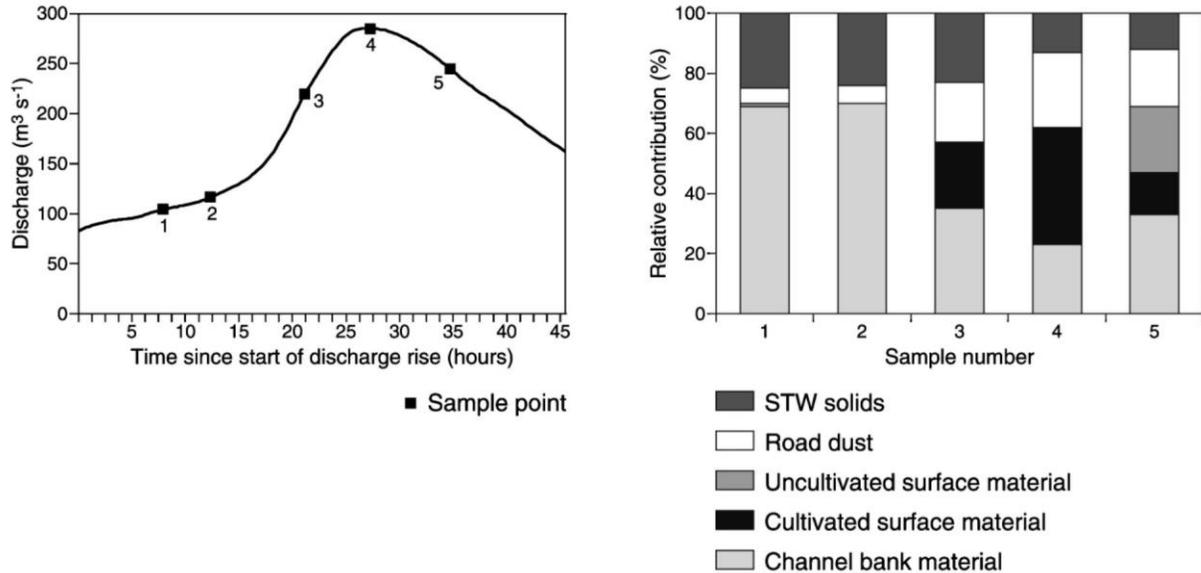
Figure 4: Relative contributions of different sources to suspended sediment sampled during a storm event in Beal, UK, estimated by source fingerprinting techniques. SWT: Sewage treatment works (Carter *et al.* 2003: 528).

1.5 ENVIRONMENTAL IMPACT OF SUSPENDED SEDIMENT

The historic sequence of surface water quality deterioration by human interaction starts with fecal bacteria and viruses, organic pollution, salts, metals, nutrients and eutrophication, organic micropollutants, pharmaceuticals and hormones, ending with acidification, radioactive, thermal and light pollution (Meybeck and Helmer 1989). Suspended sediment is often missing from such lists of anthropogenically introduced pollutants, even though it is introduced to surface waters in great quantities and has a range of environmental impacts. These impacts may be attributed directly to the presence of particles in the water column; to the substances attached to the particles; or to interactions with the (temporarily) settled sediment on the bottom of a water body (Owens *et al.* 2005; Förstner and Owens 2007).

In Europe and other places in the world, distinct changes in the sediment regime and subsequently in river and floodplain morphology started 3000 years ago, when deforestation and agriculture started to become important landscape factors (Bravard and Petts 1996). Once rocky river beds and floodplain areas became increasingly covered with alluvial soils, changing the ecological boundary conditions for adapted plants and animals. Suspended sediment, even from uncontaminated sources, can clog fish gills and has adverse effects on the reproduction of bottom-dwelling organisms. The mortality of fish because of the presence of suspended sediment generally rises with concentration and exposure time and depends on the tolerance of the species towards suspended sediment, which can be low (Wilber and Clarke 2001; Kjelland *et al.* 2015). In addition, sediment is associated with organic pollutants and heavy metals and therefore a transport vector for these substances. These pollutants may origin from direct industrial discharge, treatment plants and waste dumps. In addition, pollutants from diffuse sources like atmospheric deposition, agricultural runoff, and urban environments may also attach themselves to suspended sediment in the water phase. Settled sediment can also limit their bioavailability by trapping

and burying contaminants for extended periods of time (Salomons and Förstner 1984; Owens *et al.* 2005). However, contaminated settled sediment can inhibit the growth of certain aquatic organisms, even when concentrations in the water column are low. In addition, flood events or dredging can lead to a release and downstream redistribution of buried substances in an unpredictable way. This situation is very common for European rivers, where water quality has improved significantly over the last decades, but large stocks of toxic substances from unregulated industrial production processes are still present in the settled sediment (Salomons and Förstner 2010). It is widely acknowledged that contaminated sediment has a strong influence on water quality, especially when it is resuspended. Pollutant and sediment loads often increase distinctly towards higher flow velocities and can stay at an elevated level after the flood subsides (Ciszewski and Grygar 2016). Multidisciplinary approaches are needed to gain a better understanding of the complex and often site-specific and timely interaction of cohesive sediment and flow conditions leading to sediment remobilization and redistribution, which ultimately controls the bioavailability of contaminants (Hollert *et al.* 2007).

Apart from ecological hazards emerging from the contaminant stock, there are other interferences of suspended sediment with water quality and surface water use. Drinking water generation from turbid water is more expensive, because suspended sediment interferes with treatment processes like coagulation, filtration and disinfection. Because of its appearance and interference with recreational activities (e.g. diving) and aesthetics, turbid water generally has a lower recreational value (Bilotta and Brazier 2008). In the water column itself, suspended sediment directly reduces light penetration, which reduces the growth rate and oxygen production of phototrophic organisms (Davies-Colley and Smith 2001). However, the emission of suspended sediment into a water body can also stimulate primary production because of its association with nutrients. Unlike coastal waters, which are limited by nitrogen (Paerl 2009), phosphorous is usually the nutrient limiting primary production in freshwater systems (Correll 1999). Eutrophication is generally more often caused by the inflow of untreated sewage which contains higher concentrations of plant available dissolved phosphorous species, but it can also be triggered by excessive suspended sediment input (Krasa *et al.* 2019). At low flow velocities encountered in reservoirs or the lower reaches of streams, the produced biomass sinks to the bottom, where organic matter is metabolized while consuming oxygen. Under the subsequent anaerobic conditions, iron-phosphorous compounds in the settled sediment are reduced, leading to an increase of dissolved inorganic phosphorous in the hypolimnion (Hupfer and Lewandowski 2008). This can in turn lead to a higher bioavailability of phosphorous, which has a positive feedback on primary production and is therefore self-enhancing. At the same time, turbid water absorbs more heat than clear water, which decreases the oxygen solubility with increasing temperatures. At low flow velocities and sufficient water depth (> 7 m) a warm surface layer can have a stabilizing effect on stratification, which limits mixing processes and may further decrease the exchange of oxygen and other substances with deeper layers. Subsequent eutrophication can therefore lead to hypoxia and poverty in biodiversity and fish production, as only few species are adapted to extreme conditions (Diaz 2001).

The sheer amount of sediment emitted into the stream networks around the globe deserves special attention. Next to the direct release of toxic substances as well as the destruction of habitats and ecological functions through river straightening, regulation and fortification; sedimentological degradation is one of the main reasons for the deterioration of inland water bodies (Dudgeon *et al.* 2006; Hauer *et al.*

2018). Sediment therefore has tremendous and long-lasting influence on the ecological, hydrological and aesthetic properties of flowing and standing waters. Another important aspect is the altered fluctuation patterns of suspended sediment concentrations along with discharge fluctuations. Fluvial ecosystems are well adapted to 'flow flashiness' (Baker *et al.* 2004), i.e. the (rapid) variation between low and high flow. In natural catchment conditions, only extreme discharges can increase sediment concentrations to levels where aquatic organisms are forced to apply specific survival strategies. In catchments altered by humans, increased sediment concentrations can occur more frequently and therefore alter the composition of benthic communities (Hauer *et al.* 2018: 159). Short-term suspended sediment pulses trigger alarm reactions in certain fish species and thereby impedes the feeding and reproductive behavior (Wilber and Clarke 2001). For very high SSCs, very short exposure times can have severe effects on aquatic organisms (Bilotta and Brazier 2008), and can lead to a decrease in species diversity. If 'load flashiness' (Moatar *et al.* 2019) is increased by human interventions, this may therefore endanger associated ecosystem services. Structural interferences in the natural river course intensify these problems (Figure 5, a and b). For these reasons, quantity and quality problems cannot be regarded separately for suspended sediment. While pollutant-related problems can be catchment-specific and are more often related to point sources which are comparatively easy to measure, model and manage, the massive diffusive sediment and phosphorous input from agricultural land is an almost globally universal problem. The unsustainable cultivation of soils is occurring in virtually all agricultural areas, leading to a global soil loss rate 10–40 times faster than the rate of soil renewal (Pimentel 2006). The globally universal character also applies to phosphorous, which is applied in great quantities for the global food production and is transported along with the eroded soil into the water bodies.

In summary, four main adverse effects are connected to suspended sediment: The loss of arable land due to erosion (Figure 5, c); the loss of ecological functions following degradation of the receiving inland water bodies; the loss of hydraulic functions of rivers and water-regulating structures due to siltation (Figure 5, d); and the loss of non-renewable phosphorous to the global ocean. It is especially the diffuse and inconstant character of these inputs, which make in-river high-frequency sampling such an important aspect. With proper monitoring and plausible modelling based on these measurements, the most critical areas to implement emission-reducing measures can be identified.

Figure 5: Examples illustrating the suspended sediment environmental impact in the Upper Rhine Valley. a) Clear low-flow conditions at the Kriegbach/Kraichbach spillway channel in November 2018. b) The same channel during a two-year flood in May 2019 with a SSC of ~460 mg L^{-1}. c) Accumulation of sediment at the foot of a farmed loess slope (potatoes) near Menzingen, Kraichbach catchment after heavy rainfall in December 2018. d) Advanced siltation of the Wagbach near Bruchsal, Saalbach catchment in March 2017. Pictures a), c) and d) taken by Stephan Hilgert. Picture b) taken by the author.

1.6   MANAGING QUALITY AND QUANTITY OF SEDIMENTS

Sediment quantity issues have been in existence for millennia, while the magnitude of today's quality issues is a comparatively recent phenomenon. In order to provide or maintain the right amount of sediment at a quality which does not interfere with ecosystem services, many site-specific parameters have to be considered. Integrated studies on hydrodynamic, biological and geochemical processes are required. This makes the assessment of sediment quality usually more complex than the assessment of water quality. For these reasons, stating universal standards is difficult (Brils 2008). This is indicated for example by the fact that the Water Framework Directive 2000/60/EG (WFD, European Parliament and Council 2000) does not specifically deal with the quantity and quality of sediments. In recent years, sustainable sediment management is increasingly included into transboundary WFD management plans, for instance for the Rhine and Danube. Also, if sediment quality is too low to reach 'good ecological status', the EU state is required to find solutions (Brils 2008). In the light of a poor monitoring basis and ambiguities regarding the toxicity of polluted sediments, the role of sediments as a long-term secondary source of pollutants has been neglected by the WFD (Förstner 2015). Apitz and White (2003) suggest a holistic, catchment-based approach to for site prioritization and risk assessment concerning sediments. This

includes bedded sediments as well as material that may be transported in the future, e.g. soils. It is based on the energetic position and the quality of the material, indicated using key parameters. In short, risk is highest for sites with high probability of mobilization and poor quality, and low for uncontaminated solids in low-lying areas. Based on such an assessment, limited resources to control sediment can be used more effectively.

Sediment can be flushed or dredged and disposed or treated, which in many places is necessary to maintain waterways, to mitigate reservoir siltation, or to reduce the risk emerging from highly contaminated sites (Owens 2005). Another example is the artificial addition of sediment into (straightened) rivers to prevent bed erosion, which for instance is done in the sediment-starved Rhine River downstream of the last water barrage. Considering the enormous quantities of sediment in the earth's river systems, these expensive and disruptive measures can only be local and/or temporal solutions. The main part of sediment management must therefore take place in the catchment . There are measures concerning the design of landscape features which can help to control sediment quality and quantity (Hauer *et al.* 2018: 162). Examples for these are the construction or support of buffering features; such as margins of natural vegetation along rivers; wetlands; or designated flooding areas in order to re-establish river-floodplain connectivity (Ward and Stanford 1995). Another powerful intervention is the design of the water course itself, because it has direct impact on flow velocities and transport behavior. In the past, the consequences for sediment transport of river straightening and dam construction were often not deemed important. Measures to return to a more natural flow of rivers can therefore play a crucial role in the management concept of a basin (Batalla 2003).

Regarding particle-associated pollutants, the importance of substituting problematic substances, the source control and the consideration of the entire pollutant transport pathway are generally emphasized (Apitz and White 2003). At rivers with a heritage of unregulated industrial pollutant emissions, quality management is often limited to monitoring the stability of the layer of more recent and less contaminated sediment. Following great floods, the remobilization of these substance is then monitored (FGG Elbe 2014). Measures aiming at the control of diffuse sources are mainly covered by soil-conserving agricultural practices, forest management and storm water treatment. Point sources of sediment and sediment-bound pollutants such as wastewater treatment plants, industrial discharges and mining sites can be equipped with various physical, chemical and biological treatment steps to limit risks emerging from their effluents.

In more general terms, controlling the sources and processes in the catchment provides far greater leverage over the quantity and quality of the sediment than fighting the symptoms in the receiving water bodies. For catchment management aiming to reduce sediment or pollutant emissions, it is therefore crucial to know the amount and locations of these substances entering the stream network. While emission modelling can provide valuable pathway-specific data, suitable monitoring techniques are needed to assess the plausibility of these models and to identify elements in need of revision.

2 FRAMEWORK OF THIS STUDY

This Chapter is directed at describing the scopes, aims and working hypotheses of this study. For this purpose, the selected temporal and spatial resolutions are established at first (2.1). With respect to these scales, a brief summary of the relevance and research gaps from Chapter 1 is given in Section 2.2. From this, research goals and hypotheses are defined in Section 2.3. Section 2.4 summarizes the technical and scientific findings of this study. Thereafter, by giving an overview of the terminology surrounding suspended sediment and justifying the terms used here, the research area is further specified (2.5). Lastly, the structure of the thesis is presented (2.6).

2.1 SYSTEM CONCEPT AND SCALES

Fluvial systems are defined here as rivers and their catchments, explicitly including dammed rivers and the water bodies created by the impoundments. Dammed rivers share characteristics with natural lakes, e.g. low flow velocities. There are however distinct differences such as an active flow management, alternating water levels, complex morphometry as well as usually larger catchments and smaller capacity to inflow ratios. Here, reservoirs are therefore regarded as parts of rivers (i.e. 'fluvial') whose free-flowing nature and transport continuum have been interrupted – a situation which is becoming the norm in the Anthropocene (Lehner *et al.* 2011). It must be acknowledged that interferences (e.g. dam construction, sediment removal…) have an impact on quantity and quality (e.g. exposure of contaminated sediment) especially in downstream parts of the system (Apitz and White 2003). However, changes in the continuum have their major impact downstream. Rivers are therefore treated here as well-mixed, unidirectional, longitudinal transfer systems. Vertical, lateral and upstream interactions recognized in ecology-focused concepts (Petts and Amoros 1996) and in-river transformation processes are excluded from this work.

Monitoring and modeling are primarily technical problems that can be narrowed down when sufficient resources are provided. There are at the same time many possible scales to approach substance fluxes (Blöschl and Sivapalan 1995) in fluvial systems ranging from global annual budgets to e.g. the study of the rapid change of instantaneous loads from a combined sewer overflow during an event. While the global relevance is undoubtedly given and small-scale intensive monitoring can give valuable insight about emission processes, a spatial *meso* scale approach is pursued here. This is loosely defined here as the smallest spatial unit which is typically represented by emission models and monitoring programs nationwide, e.g. by environmental agencies. This way, tools are assessed here which already are either state-of-the-art or could be used in the future without having to spatially increase the monitoring network. For Germany and other densely populated countries, a typical *meso* catchment for which e.g. emission model results and discharge gauges are available are in the broad range of 10–1000 km² (Blöschl and Sivapalan 1995; Uhlenbrook *et al.* 2004; Berlekamp *et al.* 2007; Fuchs *et al.* 2017). For this thesis, the spatial scale for a meso catchment is narrowed to 50–250 km². Although some models support monthly resolution, substance fluxes from such meso scale catchments are often modeled in yearly time steps to support management decisions (Borah and Bera 2004; Bouraoui and Grizzetti 2014; Fuchs *et al.* 2017). Because fluxes, but also emissions are closely connected to the variable total annual discharge in a meso

catchment, mean annual fluxes and emissions for several years are pursued e.g. in WFD reports to the European Union in order to dampen the hydrologic characteristics of individual years. This scale is also addressed here. To establish these values, the scale of events must be considered for monitoring (cf. Section 2.5).

2.2 SUMMARY OF THE RATIONALE

Suspended sediment transport follows irregular patterns in time and space, whereas short periods of high flow are disproportionally relevant. A variety of promoting and inhibitory catchment factors control the transport. These factors are usually heavily altered by human interventions, which typically leads to an increase of sediment availability combined with flow regulation separating sedimentation and erosion in space. Suspended and bedded sediment interferes with eco- hydrological functions and therefore with usage requirements and ecosystem services. Particle-associated phosphorous plays a key role in these interferences and is globally relevant as a pollutant and as a non-renewable resource. For these reasons, source assessment via emission modeling is needed to advise emission-reducing measures. Representative in-river concentration measurements are needed to evaluate the quality of the model across catchments of different properties. Official calendar-based monitoring does not comply with the flashy transport behavior of particle-associated substances, which is the reason representative, adaptive, affordable and low-maintenance monitoring equipment is needed. High-frequency and hydrology-based flux monitoring at the event scale must be considered as a pragmatic solution for adequately covering transport processes and to obtain precise mean annual flux estimates. Mean annual substance fluxes from meso scale catchments (50–250 km²) allow for full-scale applicability in populated areas.

2.3 GOALS AND HYPOTHESES

In this thesis, different substance flux estimation methods for rivers are evaluated. Traditional and novel approaches are explored; and three main goals are pursued.

Firstly, the ability of traditional grab sampling rating curves to approximate mean annual fluxes are compared to a newly developed assessment based on rating curves from discharge-proportional large volume sampling. In order to explore the certainty of these assessments for different boundary conditions, three catchments with differing land cover, climatic conditions and data quality are compared. To create annual budgets, data with higher temporal resolutions ranging from hourly discharge data to multi-day mean substance concentrations are processed. Attention is placed on the relationship between event flow magnitude and event substance flux. For the goal of assessing the validity of mean annual flux estimation methods, it is expected that traditional grab sampling rating curves underestimate results provided by novel rating curves based on mean event concentrations sampled discharge-proportionally with a Large Volume Sampler (LVS, Fuchs *et al.* 2014). The relevance of measuring discharge-proportionally to provide adequate mean annual fluxes increases, as the yields from diffusive sources and the importance of particulate substances increase. This means that efficient sampling schemes for a river depend on the substance of interest, catchment size and rainfall patterns, but also on land use and hydro-pedological parameters. Because calendar grab sampling based assessments ignore such relationships, they are not

suited to support management decisions.

Secondly, for the two catchments in Germany where pathway-specific emission model results are available, the estimated fluxes are further investigated in terms of the different origins of the substances. With reference to the mean annual fluxes established by rating curve approaches, the plausibility of the total emissions is assessed and limitations in meso scale emission modeling are explored. The study catchments have a similar size which makes them roughly comparable in terms of mean flow and hydrologic response time. They are placed in the same climatic region, but differ in terms of land use, pedology and terrain. This way, it can be explored how well sediment and phosphorous fluxes can be represented by emission modeling with attention placed on diffusive sources. It is expected that discharge-proportional large volume sampling can reveal emission patterns difficult to assess by grab sampling. It is further explored under which boundary conditions calendar sampling schemes can favor false low emission model validation. It is expected that the greater the load flashiness of a parameter (Section 5.3), the greater the risk of flux underestimation by emission modeling.

Lastly, the abilities of above-water reflectance measurements, turbidimetry and grab sample based rating curves to represent discharge-proportionally measured event fluxes of suspended sediment and total phosphorous are assessed. For this purpose, a novel approach for autonomous above-water measurements is implemented and tested at the Kraichbach River. It is hypothesized that during an discharge event, above-water reflectance measurements can provide sufficiently accurate concentration measurements in order to estimate event fluxes at a higher precision than pre-defined rating curves or submerged turbidimeters.

2.4 CONTRIBUTIONS OF THE THESIS

In compliance with the relevant scales needed for area-wide implementation, this thesis proposes and evaluates novel approaches to objectively assess the need for hydrology-based sampling; to adequately measure event fluxes of suspended sediment and associated substances; and to predict long-term fluxes from these measurements. It is further exemplified how conclusions about the importance of different emission pathways and the validity and scope of emission modeling in general can be drawn from these measurements.

The use of discharge-proportional large volume sampling was extensively tested in three meso catchments to capture event mean suspended sediment and phosphorous concentrations. In a novel approach, the concept of rating curves is adapted to the characteristics of LVS sampling based on flow parameters from the sampling period. Discharge-proportional sampling, although not a widely used method, has some relevance in the monitoring of fluvial systems (King 2003; Ackerman *et al.* 2011). The approach developed here fuses the representativeness of the sampling procedure for the precise assessment of individual events with the modeling properties of conventional rating curves i.e. for the establishment of mean long-term fluxes. The setup for river monitoring has been presented orally at the European Geoscience Union General Assembly 2019 (Wagner *et al.* 2019).

Further, the potential and limitations of a novel setup for close-range above-water reflectance measurements to measure suspended sediment and total phosphorous concentrations are explored. The subsequently developed boat-based application was described in a peer-reviewed journal article, which

has been integrated in this thesis (Wagner *et al.* 2018). The prototype spectrometer used was specifically developed for this purpose in exchange with the author. In comply with the requirements of applicability and robustness, the use of partial least squares regression as a rapidly implemented empirical procedure was tested for parameter retrieval. This represents one of the first applications of partial least squares on close-range hyperspectral water reflectance measurements.

In order to evaluate the need for hydrology-based sampling, the concept a flow flashiness index (also called the 'Richards-Baker Flashiness Index') introduced by Baker *et al.* (2004) was modified to create a novel (normalized) load flashiness index. It combines the two essential parameter to determine the need of hydrology-based sampling –the importance fluxes of high flows and their duration– and makes them therefore comparable across catchments and substances in order to guide the selection of appropriate monitoring programs. Relevant factors influencing load flashiness are identified. For the examined reservoir it is demonstrated how the uncertainties associated with different monitoring techniques can dramatically impact the lifetime assessment of the structure.

In addition, the author was substantially involved in the description and publication of the MoRE model framework (Fuchs *et al.* 2017) as well as in the development and implementation of the MoRE version METRIS applied for Baden-Württemberg, the results of which are presented here for the Alb and Kraichbach catchments (Kiemle *et al.* 2019). At the example of these two meso scale catchments it is further shown how dissolved and particulate phosphorous emissions and fluxes from point and diffusive sources are related to catchment characteristics; how composite sampling can have advantages to identify emission patterns; and how different sampling strategies could lead to different model evaluations.

2.5 TERMINOLOGY

Particles in natural waters can be defined by physical and (bio-)chemical properties. Further, particles are defined according to their state of transport, a property which is linked to the physico-chemical properties of the particles, but also to flow conditions. A general description is given with the terms, 'suspended solids', 'seston' and 'suspended particulate matter', which can be defined as fine insoluble mineral or organic particles in the water phase. When the term 'suspended sediment' (SS) is used instead, a reference to the particles ability to be transported is made. Sediment, in its broadest sense, is a conglomerate of organic and inorganic materials that can be carried away by water, wind or ice. In flowing water bodies, the mass sum of transported sediments in a given time (i.e. the sediment discharge) is called 'sediment load' or 'total load'. The term 'sediment yield' is defined as the load divided by the drainage basin area; a normalization which allows it to compare loads of rivers of different sizes (Milliman and Farnsworth 2013). 'Flux' is a sometimes more ambiguous term used to describe a mass moving through a cross section per unit of time. In the context of river transport, flux is usually used to describe mass flows over longer periods, such as the yearly total mass of sediment delivered to the ocean by a given river. The total load can be further distinguished by its settling behavior. Sediment which has periodic contact to the bottom of a water body during transport is classified as 'bed load', while particles floating in the water column with a transport speed equal to the water flow velocity are called 'suspended load'. Suspended sediment is usually sand-sized (< 2 mm) or smaller, with the largest share in the silt and clay fraction (< 63 µm) (Owens *et al.* 2005: 693). In the context of water quality, available data on transported particles

usually refers to suspended load. Measuring bed load is difficult, but nonetheless important to determine hydrodynamic parameters, which control hydromorphology and therefore link back to water quality. Globally, bed load is roughly estimated to make up about 10% of the total load, with great differences between mountainous and slow-flowing rivers (Milliman and Farnsworth 2013: 24). In this study, bed load is not considered. Particles from the suspended load small enough to stay afloat in non-moving water bodies are called 'wash load' and therefore form a subset of the suspended load. Wash load cannot be distinguished from suspended load in a moving water body. Contrary to this, 'settleable solids' are moving towards the bottom of a water body to form 'bedded sediments'. When using these overlapping definitions, it is important to be aware that particle transport is highly dependent on flow conditions which can vary distinctly; in longitudinal, lateral and transversal directions of a water body; or in small time steps for a given location. The term 'load' is used in this work in the context of instantaneous loads, often expressed in grams per second. 'Flux' is preferred when referring to longer periods time, which may be one year, days or events of different length.

In contrast to the conceptual definitions above, particles can also be defined by the methods of quantification. To measure the mass of suspended sediment in a water sample, dissolved and particulate water constituents have to be separated. This is usually done by filtering the sample through weighted glass fiber filters. By drying and re-weighting the filters, the parameters 'total suspended solids' (TSS) or 'suspended sediment concentration' (SSC) are measured. While TSS is measured in a sub-sample from a larger sample, SSC is determined on an entire sample container. Several sources of error arise from these practices. Firstly, physically filtered water still contains colloids, which do not comply with the chemical definition of a dissolved substance, or solute. Solutes are charged ions of molecular size, gases or organic compounds; while colloids are sub-micrometer solids. In further analysis of the filtrate, colloidally associated components are therefore quantified together with truly dissolved constituents. Secondly, different filter pore sizes (e.g. 0.45 µm, 0.6 µm or 1.5 µm) are used to determine the same parameter. Thirdly, representative subsampling of a larger sample container is virtually impossible and leads to lower values of TSS relative to SSC; especially for samples with a higher share of larger particles (Horowitz 2013). While 'suspended sediment', 'suspended solids' and 'suspended particulate matter' have largely overlapping definitions, the term *suspended sediment* (SS) is preferred in this dissertation. Reasons for this are the widespread use of the term and its clear reference to aquatic systems. In addition, most of the presented laboratory particle concentration measurements were made without subsampling. The main water quality parameter evaluated here is therefore SSC. With respect to the concept of suspended load and the filtration methods applied, this work focuses on fine-grained particles likely to be transported in suspension in a river channel (> 0.45 µm and < 2 mm). Due to the important role in biochemical cycling processes, emphasis is put on the silt and clay fraction (> 0.45 µm and < 63 µm).

The fine particles which are in the focus here are regarded as a proxy for sediment-related water quality and management problems. They are further considered evenly distributed by turbulence within the sampled cross-sections, much like a truly dissolved substance. Next to suspended sediment, total phosphorous (TP, particulate and dissolved) and its main dissolved form orthophosphate-phosphorous (o-PO$_4$-P, in the following also simply 'orthophosphate') are investigated. As outlined above, phosphorous has a global relevance for agriculture and an important role for aquatic ecosystems. Its transport is closely

linked to particle transport, does not have an atmospheric sink and its sources and emission pathways are modeled for many watersheds. Reliable, measured substance budgets in terms of annual river fluxes for these catchments are therefore required.

The term 'flashiness' is used here to describe the rate of change of a parameter within a river. The concept of flow flashiness (Baker *et al.* 2004) was adapted for substance loads; the mathematical definition is given in Section 5.2. An event is usually defined as a section of the hydrograph which covers its complete increase-decrease reaction to isolated, uninterrupted periods of precipitation, without considering changes in base flow. For meso catchments, the length of an so-defined event can range typically from hours to days. In this thesis, the term event has a strong reference to the volume-constant and time-variable discharge interval sampled by the large volume sampler (Section 5.1). The interval was selected in each catchment order to be cost-efficient while maintaining the required temporal resolution for meaningful post-processing. This can include high-flow periods of few days, which largely fit the general definition above, but also precipitation-free periods of up to several weeks.

2.6 STRUCTURE OF THE THESIS

In Chapter 1, an introduction to the topic is given, the principles and interdependencies between and the factors influencing suspended sediment transport are outlined and research gaps are pointed out. From there, corresponding goals and hypotheses are derived in Chapter 2 while defining the scope and research area of the study. Chapter 3 outlines the state of the art methods used for assessing particle fluxes in rivers and provides the necessary technical background. Chapter 4 briefly introduces the three study sites in terms of land use and other boundary conditions for water quality and hydrology.

In Chapter 5, all methods and materials used in this study are introduced. This includes sampling tools like the LVS, turbidimeters and the above-water spectrometer as well as emission modeling and rating curves. The development of the above-water spectrometry included mesocosm experiments, boat-based campaigns at Passauna Reservoir, and finally a permanent monitoring station at the Kraichbach. All steps are described in Section 5.5.

In Chapter 6, all results regarding the assessment of mean annual fluxes and emissions from the three catchments are presented. The chapter ends with a site-specific discussion about the plausibility of the different approaches and subsequently identifies the best available flux estimates to be used as a reference. On this basis, the sites are compared in terms of their mean annual substance export behavior. Chapter 7 presents all results concerning the development of above-water reflectance measurements and compares different methods for event flux assessments at the Kraichbach.

Chapter 8 concludes on the hypotheses using all available findings, summarizes the potential and limitations of different methods, gives recommendations for their use and explores implications for water resource management. Chapter 9 draws final conclusions and gives a brief outlook.

3 QUANTIFYING PARTICULATE RIVER FLUXES

This chapter provides the required background for the methodological aspects of this work. First, sediment rating curves (3.1) and manual grab sampling (3.2) as the standard method to assess fluxes are introduced. Because of their potential to better represent substance dynamics, automated samplers and passive samplers are briefly outlined (3.3). Section 3.4 deals with the measurement principle, applications and limitations of optical systems places in or outside the water. Lastly, background information on emission modeling is given (3.5)

3.1 SEDIMENT RATING CURVES

Annual flux is defined as the mass of substances transported through the cross-section of a river within one year. The substance can be suspended sediment, but the definition and approaches can be transferred to any other substance of interest. By coupling concentration measurements with discharge, fluxes can be estimated. For this purpose, flux can be defined as the integral of the product of discharge and concentration (Hilden 2003).

$$F = 10^{-6} * \int_{year} Q(t) * c(t) dt \tag{1}$$

where

F	is the annual flux in t a^{-1}
Q(t)	is the discharge in $m^3\ s^{-1}$
c(t)	is the substance concentration in mg L^{-1}

Estimates of (e.g. annual) substance fluxes are often made using sediment rating curves. To build rating curves, continuous stream flow data is usually coupled with (e.g. monthly) concentration measurements from grab samples (Walling 1977; Harrington and Harrington 2013). It is assumed that a measured concentration is representative for the entire cross-section and that measured concentrations and corresponding daily mean discharges are representative samples of the load. The resulting regression is then used to estimate suspended sediment concentrations from discharge, which is a parameter available for many streams that can be measured reliably with relatively little maintenance. Many rating curves show a positive linear, log-linear or logistic relationship between discharge and SSC, suggesting that higher river flows are also increasingly laden with particles. Others are better represented by second- or third order polynomial regression, some of which first increase and then decrease again when moving towards extreme discharges (Horowitz 2003). Irrelevant of the type of regression, rating curves often suffer from poor predictive ability due to high scattering. The scattering is a result of the transport complexity (Walling 1977) outlined in Section 1.3. By grouping observations of concentration and discharge according to similar boundary conditions which demonstrably influence the concentration-discharge relationship (Figure 6), scattering can be reduced. Most commonly, separate rating curves are created for summer and winter; for individual years in longer time series; or for measurements collected during raising and falling stage of the hydrograph, respectively (Walling 1977: 532).
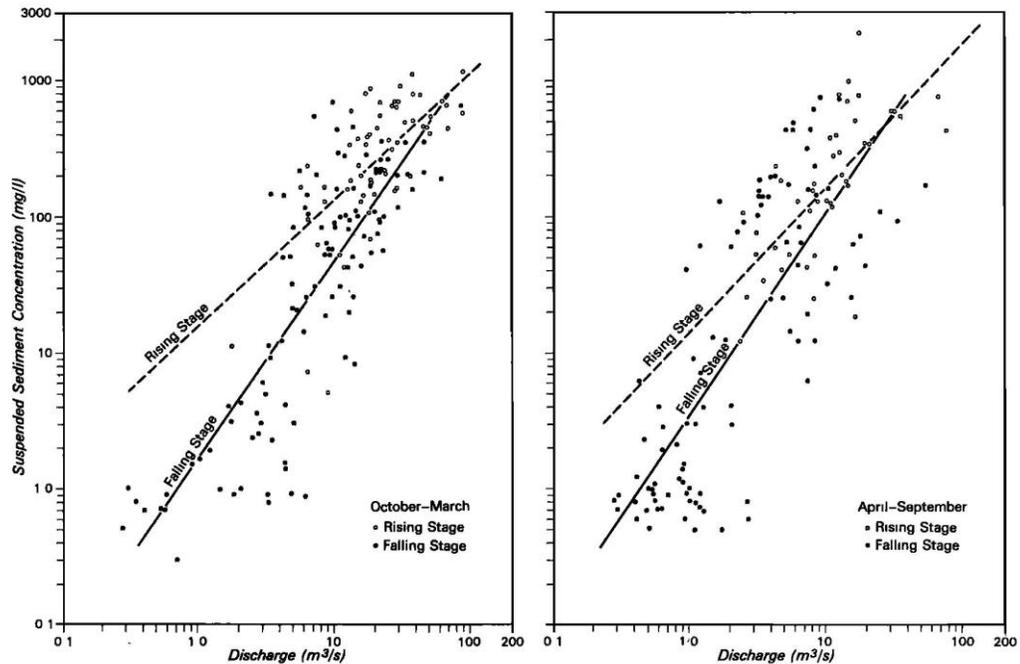
Figure 6: Log-linear sediment rating curves for the River Creedy, England (Walling 1977: 535). Grouping rating curves by stage and season is still common today, as are high scattering and subsequently poor predictive ability of the regressions.

3.2 MANUAL GRAB SAMPLES AND SAMPLING SCHEMES

River quality monitoring is still widely dependent on manual grab sampling. This is usually done weekly, biweekly, monthly or quarterly at specific sampling sites. Specifically sampling high-flow events is not foreseen in most schemes, even though budgets can be improved when hydrology-based schemes are implemented (Horowitz 2013). Kentroti (2017) found that in 14 years of monitoring 44 rivers in Baden-Württemberg, only ten out of ~6700 water samples (0.15%) were sampled during conditions exceeding mean annual high flow. Although the exceedance of mean annual high flow is, by definition, a rare event, it illustrates that there is no incentive in specifically monitoring these periods of high transport. The implementation of equal-interval sampling schemes leads to the highest flow decile being sampled as often as any other decile, although it is representing for a much greater water volume and usually shows greater scattering. The tendency of rating curves is to overestimate dissolved substances and low flow concentrations and to underestimate particulate substances and high flow concentrations (Stevens and Smith 1978; Horowitz 2013). This results in at times grossly underestimated annual loads of sediment and associated substances (Kulasova *et al.* 2012). By comparing to a reference data sets composed of daily samples in several large (> 900 km²) catchments, it was found that with the same amount of samples (i.e. one sample per month versus one sample per flow decile, plus two samples at high-flow conditions), hydrology-based sampling schemes can distinctly improve sediment flux estimates (Horowitz *et al.* 2015). One example is given in Figure 7, where the lack of samples in the highest flow deciles on average leads to substantial underestimations of annual fluxes. One such approach is 'stratified sampling' where the highest 10% of discharges are sampled at a much higher rate as the remaining discharges (Kulasova *et al.* 2012). These schemes are also more robust against basin scale, although it is stated that the smaller the

24

catchment, the more frequent the sampling has to occur. Streams in small catchments react quicker to precipitation and are more driven by local, short-termed rain events than large rivers. Hydrology-based sampling may not cover all spatio-temporal variations of sediment transport, especially in smaller streams. Here, more intensive monitoring techniques are needed.
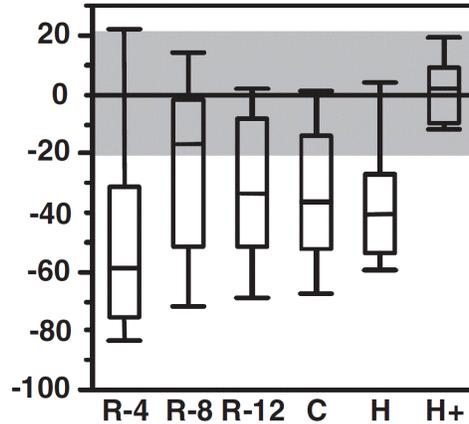
Figure 7: The range of percent error between annual SS flux estimated of different sampling schemes relative to the value calculated from daily measurements for Schuylkill River, USA (catchment 919 km²). R-4, R-8 and R-12 plots refer to 4, 8 and 12 random samples; C refers to 12 calendar-based samples; H refers to 10 samples, one randomly selected from each annual decile discharge range; and H+ refers to 12 samples, one randomly selected from each annual decile discharge range plus two additional randomly selected high-flow events (Horowitz *et al.* 2015).

3.3 AUTOMATED AND PASSIVE SAMPLERS

Automated water samplers are considered an adequate technique to comply with concentration variability, and are applied e.g. in storm sewers (Lee *et al.* 2007) or rivers (Facchi *et al.* 2007). They can take time paced or flow paced individual or composite samples (King 2003; Ackerman *et al.* 2011). One major limitation of many available systems is that individual samples are too small to yield sufficient particulate material for the analysis of particle-bound parameters with a low limit of quantification, or many parameters at once. Also, the costs of analysis can become inhibitive if a high temporal resolution is pursued. Time-based composite samples can create larger samples, but are only representative for the sampled medium when the flow is constant in the sampling period, which is rarely the case for rivers. Discharge-proportionate sampling creates samples with a mean concentration equal to the mean in-river concentration during the composition of the sample.

Passive samplers are unpowered devices placed in a water body to collect time-integrated samples of sediment and other substances by natural processes such as sedimentation and adsorption. The integration time is usually days or weeks. Sediment traps are upward-facing funnels or cylinders to capture sinking particles in order to quantify the rates of sedimentation in a water body. Flow-proportional adsorbent-based passive samplers work without electricity and create flow-integrated samples, but have been criticized for their high costs, unreliable results and low comparability to other methods (Audet *et al.* 2014). An alternative approach to passive sampling is the use of biofilm. Biofilm is the living layer composed of fungi, algae, protozoa, and other microorganisms abundant on most surfaces in natural (e.g. on bed sediment and rocks) and artificial (e.g. in pipes, canals etc.) aquatic environments. A biofilm

sampler is a transparent or opaque plastic container opened on two sides to allow water to flow through (Fuchs *et al.* 1996). It contains removable glass plates positioned laterally to the flow direction as a standardized growth surface for biofilm. The biofilm can be scraped off biweekly or monthly and analyzed for parameters of interest. Biofilm represents a low end of the aquatic food chain. Pollutants measured within it are therefore not only event-integrated with respect to the specific ability of organisms to adsorb them or seclude themselves from them, but can indicate the starting point for further accumulation along the food chain. When the relation between in-river substance concentration dynamics and biofilm concentrations is established more clearly, it could therefore become a standardized tool to assess the biological relevance of pollutant fluxes

3.4 OPTICAL SENSORS

This section gives the necessary background water-light interactions from a measurement perspective (3.4.1). The measurement principles and applications of submerged sensors (3.4.2) and above-water sensors (3.4.3) are introduced. Finally, possibilities for parameter retrieval from remote sensing are outlined (3.4.4).

3.4.1 OPTICAL CONSTITUENTS AND PROPERTIES

Optical water quality measurement devices make use of the fact that constituents present in the water column can alter the way light interacts with the water. A water constituent which changes the optical properties of the water is called an optically active constituent (OAC). Suspended sediment scatters and absorbs light; and represents one major group of OACs. The other major groups are dissolved substances which have specific absorption properties like colored dissolved organic matter; and algae, which absorb blue and red light while reflecting the green (Figure 8).
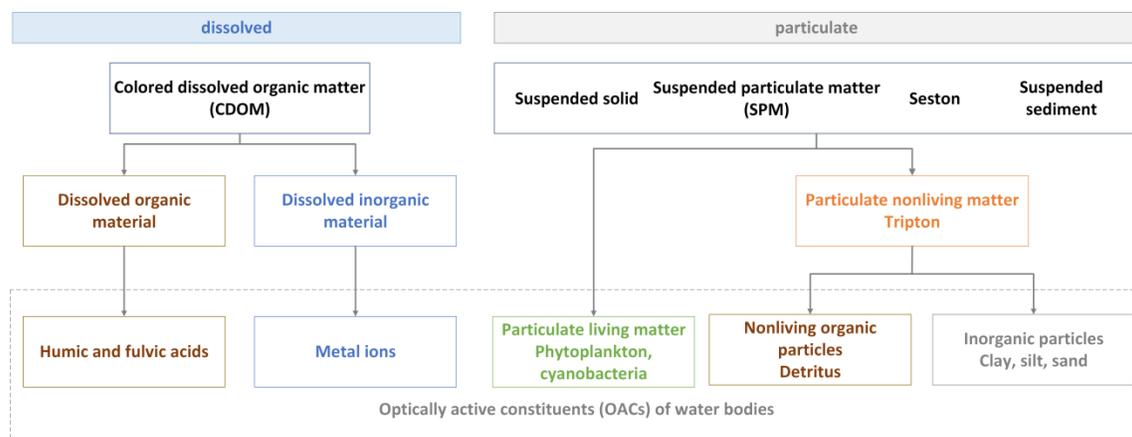
Figure 8: Classification of optically active constituents of water bodies. After Frauendorf (2002: 34).

Absorption and scattering are independent from intensity and direction of the light and can therefore be seen as material properties, or inherent optical properties. Absorption coefficient and volume scattering coefficient can be summed up to an extinction coefficient describing the loss of light per meter:

$$c = a + b \qquad (5)$$

where

c	is extinction coefficient in m^{-1}
a	is absorption coefficient in m^{-1}
b	is volume scattering coefficient in m^{-1}

The scattering coefficient can be obtained by integrating the volume scattering function over all directions. The volume scattering function describes all types of scattering like Raleigh scattering, Mie scattering and refraction/reflection on larger particles quantitatively (Frauendorf 2002: 37).

$$\beta(\theta) = \frac{dI(\theta)}{E_d\,dV} \tag{6}$$

where

$\beta(\Theta)$	is the volume scattering function in $m^{-1}\,sr^{-1}$
$dI_2(\Theta)$	is the radiant intensity in $W\,sr^{-1}$
E_d	is the incident light scattered on the volume dV in $W\,m^{-2}$
dV	is the volume on which the light is scattered in m^3

Pure water shows very little absorption for short wavelengths and very strong absorption for longer wavelengths: Above 850 nm (near infrared), pure water absorbs almost all electromagnetic radiation. Scattering on the other hand is high for short wavelengths and low for long wavelengths, which is the reason clear water appears blue (Frauendorf 2002: 39). The optical absorption and scattering coefficients are proportional to the concentration of a substance. Coefficients of several substances can be added to form a total optical coefficient (k), which largely determines the 'color' of a water body. Sensors placed above the water surface can record the overall signal emerging from water surfaces, for instance as remote sensing reflectance. Remote sensing reflectance can be defined as the proportion of the sunlight reflected into the upper hemisphere by the water body across the electromagnetic spectrum (Formula 7, Dierssen and Randolph (2012)).

$$R_{rs} = \frac{L_w(I)}{E_d(I)} \tag{7}$$

where

R_{rs}	is the remote sensing reflectance in sr^{-1}
$L_W(I)$	is the spectral water-leaving radiance in $W\,m^{-2}\,nm^{-1}\,sr^{-1}$
$E_D(I)$	is the downward planar irradiance in $W\,m^{-2}\,nm^{-1}$

Because it is –next to its dependency on properties of the water itself and of the constituents and their quantity in the water– to a small degree dependent of the ambient light field, remote sensing reflectance is a so-called apparent optical property. For above-water measurements, light scattered into the upper hemisphere is more important than light scattered downwards. The scattering coefficient from Formula (5) can therefore be divided into forward scattering and a backward scattering by separately integrating $\beta(\Theta)$ for $0 \leq \Theta < \pi/2$ and for $\pi/2 \leq \Theta < \pi$, respectively. The backscatter fraction B_b is then defined

as the quotient of backscattered light b_b and forward scattered light b. For natural waters, the backscatter fraction is typically below 5%, meaning that the majority of the sunlight is scattered downwards into the water column (Dierssen and Randolph 2012).

3.4.2 SUBMERGED SENSORS

The most common type of measurement device for the quantification of suspended sediment are turbidimeters, which make use of the reflective properties of particles in the water column. A red or near infrared light beam is emitted by the device into the water column. A detector positioned at a 90° angle towards the light source measures the fraction of the emitted light scattered towards it. 90°-scattering of a red or near infrared light beam is reasonably linearly correlated to SSC (Figure 9). Local calibration of turbidimeters is nevertheless recommended (Rügner *et al.* 2014), as the scattering properties vary by particle type and size, which varies from catchment to catchment.

Next to scattering-based devices, there are submerged absorption-based spectrometers which transmit a light beam through a short distance directly in the water and measure the amount of light per wavelength band reaching the detector placed opposite of the transmitter. They have the advantage that they can simultaneously detect substances without reflective properties, but substance-specific absorption properties like nitrate and chemical oxygen demand. Similar to turbidimeters, real-time measurements without sampling and chemicals can be performed at high temporal resolutions, for which local calibrations are recommended (Gruber *et al.* 2006; Rieger *et al.* 2008; Torres and Bertrand-Krajewski 2008).

One disadvantage of submerged sensors is their difficult installation. The sensors have to withstand a variety of flow velocities and possibly the impact from objects floating in the river, sewer or canal. When installed at a boat during a measurement campaign, cruising speeds are typically slow to avoid the formation of air bubbles in front of the measuring window. Furthermore, objects placed underwater for extended periods of time are almost always subject to biofilm growth and for the precipitation of minerals, which leads to sensor drift. Figure 10 exemplarily shows the sensor response of two turbidimeters before and after manually cleaning the measurement window. Automated wipers or pressured air systems can extend the measurement periods for optical submerged sensors, but still need regular maintenance and/or connection to the electrical power grid (Gruber *et al.* 2006; Martínez-Carreras *et al.* 2016).
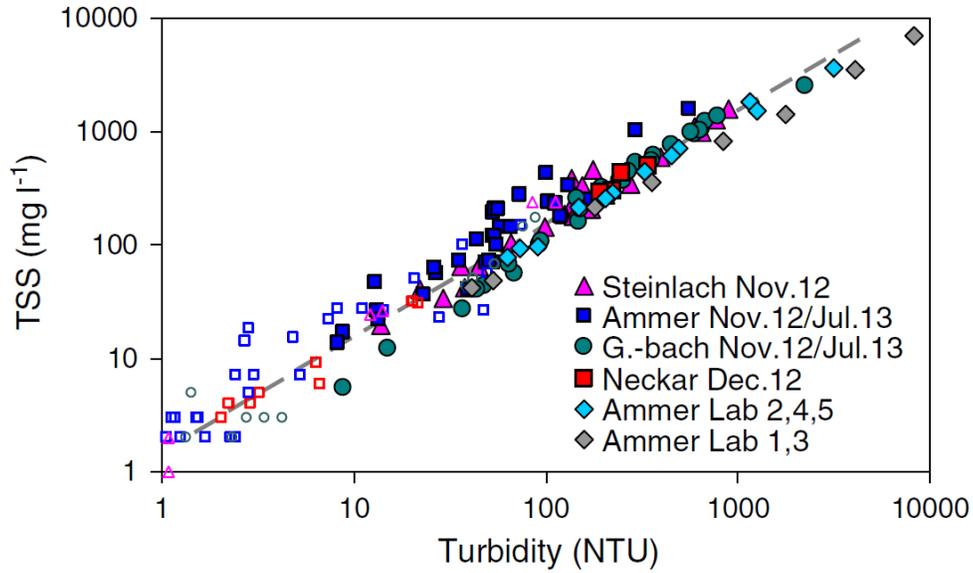
Figure 9: Turbidity measured in Nephelometric Turbidity Units (NTU) plotted against Total Suspended Solids (TSS) for various catchments in south-west Germany. The dashed grey line indicates is given by TSS [mg l^{-1}] = 1.5 × turbidity [NTU]. Note the log-log scale and the uncertainties of the measurement especially at the low end of the concentration range (Rügner *et al.* 2014).
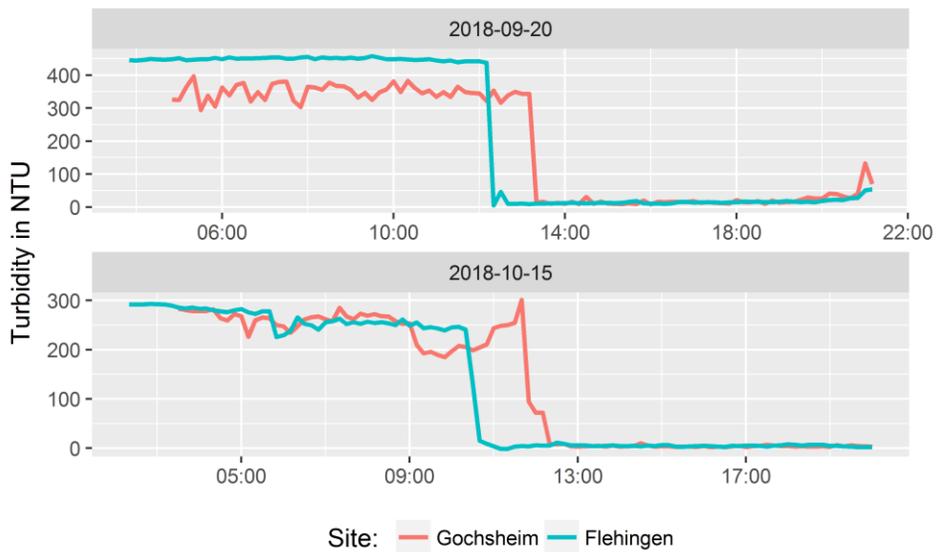
Figure 10: Turbidity measurements at two different sites at the Kraichbach River. The drop in measurement values is caused by manually cleaning the measurement window and underlines the problems of optical devices submerged in natural waters. NTU: Nephelometric Turbidity Units.

3.4.3 ABOVE-WATER SENSORS

Next to submerged sensors equipped with an active light source, there are satellite-based, airplane based and ground-based sensors which are designed measure remote sensing reflectance. Water remote sensing already supplements traditional water quality monitoring for the study of many water bodies across the globe. Water color has been used since the 1970's to measure and map suspended sediment and other OACs in surface waters (Jerlov 1976; Ritchie *et al.* 1976; Gordon and Morel 1983). In the beginning, ocean waters were in the focus of the research. As spectral and spatial resolution have increased during the last decades, the analysis of smaller (inland) water bodies with higher precision has become possible (Dekker and Donze 1994; Karaska *et al.* 2004; Chipman *et al.* 2009; Olmanson *et al.* 2013). Lately, studies using unmanned aerial vehicles (UAVs) and hyperspectral sensors can be increasingly found (Zang *et al.* 2012; Honkavaara *et al.* 2013; Su and Chou 2015). Hyperspectral data, i.e. measured spectra consisting of many narrow spectral bands, can be used to derive more accurate results especially in optically complex waters containing several different OACs which can vary independently from one another (Govender *et al.* 2007; Matthews 2011).

Suspended sediment strongly increases reflection across a broad range of wavelengths in the visible and near infrared (Figure 11). Because clear water appears dark for longer wavelengths, the reflectance increase in the near infrared is particularly helpful for retrieving SSC (Schalles *et al.* 2001). This is the reason SSC is a parameter which is relatively easy to assess by remote sensing techniques. Total phosphorous itself is optically inactive, but often highly covariant with suspended sediment and/or chlorophyll-a. It can therefore be detected through its association to optically active parameters (Gholizadeh *et al.* 2016).
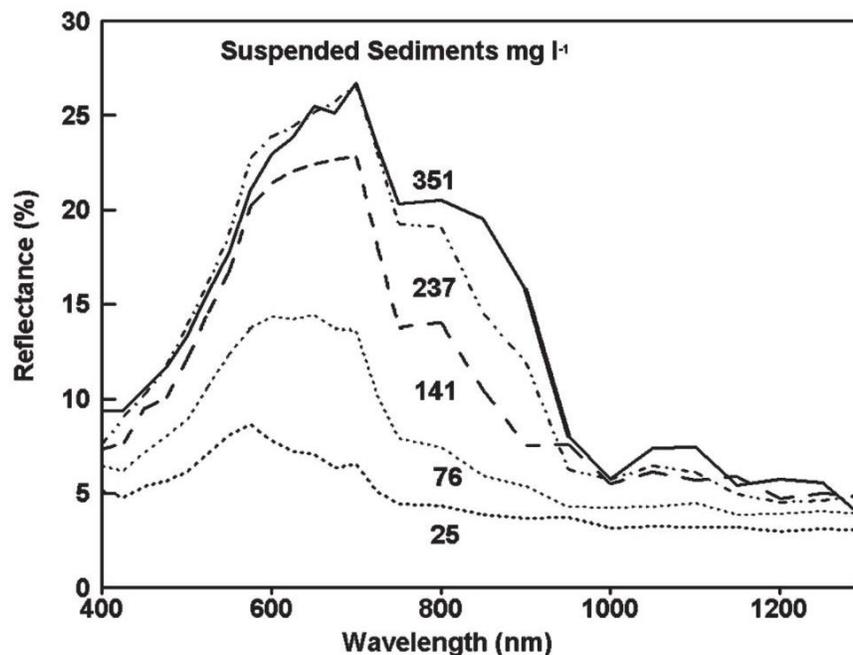
Figure 11: The relationship between remote sensing reflectance and wavelength for different concentrations of suspended sediment (Ritchie *et al.* 2003: 696).

Because the above-water detected signal emerges from different depths of the water column, it is a specific feature of water remote sensing that an integrated volume is measured. Increasing concentrations of suspended matter or other OACs decrease transparency and therefore the attenuation depth from which 90% of the water-leaving radiance emerges (Gordon and McCluney 1975), which can be approximated by attenuation measurements or measurements of transparency (Padial and Thomaz 2008). The measured volume is therefore inconstant and depends not only on the current concentration of OACs, but also on wavelength. Blue light in clear ocean water can have attenuation depth of more than 50 m, while red light in coastal or inland water bodies is attenuated within few meters (Gordon and McCluney 1975). Next to this, the signal from the bed of the water body can play a role in optically shallow water bodies, where the attenuation depth is greater than the water depth (Dierssen and Randolph 2012). Bottom reflectance is dependent on the water level, growth state of vascular plants and substance concentrations (Olmanson *et al.* 2013). Up to SSC of ~100mg L^{-1}, bottom color was shown to have a strong influence on the measured reflectance spectra in the visible domain, while the near infrared is less affected (Tolk *et al.* 2000). Specular reflection at the water surface is often specifically excluded in the definition of remote sensing reflectance. In practice, light reaching a sensor may in part be sunlight from mirror-like reflection at waves (glint) (Dierssen and Randolph 2012). In order to measure water-leaving radiance, e.g. for calibration purposes, the measurement device is usually placed just underneath the water surface to avoid the influence of glint (Chipman *et al.* 2009: 14).

Another challenge for above-water measurements is that the rate of change of the target parameters can be high, for instance, when a river becomes heavily laden with sediment after a rain event. Ground truthing these peak concentrations has therefore to be performed in a close time span, which can be difficult when relying on satellite overpasses. In addition, cloud cover can inhibit satellite-based assessments (Tyler *et al.* 2006). Using portable (imaging) spectrometers close to the target do not have the same spatial coverage, but are more flexible from cloud cover and overpass times and can reveal spatial pattern in greater detail (Dierssen and Randolph 2012; Zang *et al.* 2012).

3.4.4 Linking Reflectance and Concentration

Water remote sensing requires ground truthing, usually in the form of individual water quality grab samples. The number of reference samples needed for meaningful parameter maps varies according to algorithm and number and complexity of water constituents. Such an algorithm can be defined as "a calculation method to derive concentrations of (…) water constituents from spectral observations" (Laanen 2007: 47). Algorithm types range from analytical to empirical, with hybrid types called semi-analytical or semi-empirical (Knaeps *et al.* 2010: 2041). Analytical algorithms inverse a bio-optical model to obtain a given (reflectance) signal by altering the constituent concentrations until a sufficient match between observed and modeled spectra is made. All groups of water constituents (colored dissolved organic matter, suspended sediment, algae) can be retrieved simultaneously. The main advantage of established analytical procedures is the high potential for transferability between sensors, targets and seasons with relatively small numbers of additional reference measurements. The main disadvantages are that the signal must be obtained with a calibrated instrument, corrected for atmospheric and other factors to obtain true remote sensing reflectance, that all optical properties have to be known and entered and the high computational effort needed to model the propagation of light through water (Keller 2001: 56).

Semi-analytical algorithms simplify the underlying bio-optical model to its most relevant parts and make use statistical relationships, which is a widely used method due to its applicability and transferability (Malthus *et al.* 2012). For semi-empirical methods, parts of the spectrum of highest efficiency are identified before applying statistical relationships, such as simple band ratio algorithms.

Purely empirical methods do not include any knowledge about the underlying physics. The measured signal is simply correlated against the target parameters by statistical means. They can therefore be relatively quickly implemented using radiometrically uncalibrated sensors, or even on raw data (Laanen 2007: 48). They also require training samples, preferably evenly distributed across the entire concentration range in which the measurement results are to be expected. Unlike analytical procedures is the established relationship entirely dependent on the input data provided. Such empirical relationships are usually site- and sensor-specific, which limits their transferability especially for complex Case 2 waters (Dekker *et al.* 2001). If such a transfer is not in the focus of the research, for instance because only a small number of water bodies is studied with the same instrument, empirical methods are nevertheless very popular because of their robust and reliable results and their relative simplicity and transparency (Matthews 2011). Furthermore, to model optically inactive parameters like total phosphorous, empirical algorithms are considered the most applicable approach (Gholizadeh *et al.* 2016).

Once calibrated, optical methods have the advantage to quickly provide measurement values in high spatial and/or temporal resolutions. For monitoring dynamic or extensive surface waters, these properties are of particular importance. A current, extensive review of the use and limitations of remote sensing techniques for the determination of numerous parameters and available sensors is given by Gholizadeh *et al.* (2016). It is emphasized there that hyperspectral sensors are offering an increasingly higher precision for the determination of water quality parameters. Suspended sediment and total phosphorous are named among the most commonly measured quality parameters. One major limitation of these techniques is that the data processing is usually complex and currently requires too much training to be easily implemented in monitoring programs. Simple and robust solutions are therefore needed. Devices placed closer to the target can provide frequent measurements fast and at low cost. At the same time, these techniques to at least some degree dependent on traditional sampling and will never fully replace them, as linking measured spectra to concentration values requires training data.

3.5 EMISSION MODELLING

The application of a sediment rating curve as described in Section 3.1 can be regarded as simple empirical model to predict suspended sediment and other substances from discharge and subsequently calculate substance fluxes for the watershed of the gauging station. To assess substance inputs into water bodies on a catchment level, conceptual and semi-empirical models are useful (Borah and Bera 2004). Such models are used to assess the substance-specific relevance of different sources and pathways as well as the total inputs. The empirical and conceptual environment Modelling of Regionalized Emissions (MoRE, Fuchs *et al.* 2017) can calculate emissions via different pathways on the level of catchments in yearly time steps. Because catchment-based models like MoRE are designed to cover great areas, they must work on relatively few input data which are available for the entire model region. Examples for input data are land use/land cover classifications from satellite imagery, digital elevation models, monitoring data from waste water treatment plants and water volumes per flow compartment of water household models. Their scope typically ends at the water body itself, meaning that in-stream retention and degradation is not considered. In some studies, emission models have been coupled with water quality models to further assess the impact on in-stream conditions (Krysanova *et al.* 1998; Berlekamp *et al.* 2007; Gebel *et al.* 2018). In other cases, retention and degradation were implemented by means of an empirical comparison between emissions and monitored loads; or the assumption, that all material which enters a river is entirely transported further downstream (Kiemle *et al.* 2019). This assumption becomes less valid when looking at shorter time intervals and for non-dissolved substances. Emissions into water bodies were deemed a suitable interface and legislative reporting levels, for instance for the Water Framework Directive. Members of the EU are obliged to identify high-impact areas and pathways and reduce emissions by means of measures in order to achieve good ecological status (European Parliament and Council 2000).

Water quality monitoring stations with nearby discharge monitoring stations can serve as balance points to assess the validity of the modeled upstream emissions. Because of the lack of high-frequency monitoring data, the conceptual nature of the models, the manifold input data demand and the large areas they cover, a straightforward calibration and validation of mesoscale models is challenging (Abbaspour *et al.* 2015). Even if an offset between flux measurement and emission model (or coupled emission/water quality model) is detected, there is uncertainty about which element of the model is responsible for the over- or underestimation, and 'expert judgement' is required (Arnold *et al.* 2012). To validate a single pathway is not feasible (Tetzlaff and Wendland 2007). Furthermore, because of the transport complexity outlined in Section 1.3 and the monitoring bias described in Section 3.2, the offset may also be attributed to over- or underestimations from the measurement. Lastly, substance emissions into water bodies can be higher than in-river fluxes because auf retention and degradation processes within the stream. For these reasons, some emission models not calibrated at all, but simply compared to flux estimates from monitoring (Kiemle *et al.* 2019). To improve pathway-oriented emission models, comparisons with other models, hints from monitoring data and experience of the modeler have to be combined to identify the pathway most likely to cause an error. The more accurate and spatially distributed the annual measured fluxes are, the more suitable the basis for such an assessment will be.

4 STUDY AREAS

Substance fluxes were studied in three catchments. The German catchments of Kraichbach (4.1) and Alb (4.2) are in close proximity to one another, but have differing land use and terrain properties. Also, modeled TP and orthophosphate emissions are available here. The gauged part of the Passauna Reservoir catchment (4.3) in South Brazil is offers the possibility to transfer the applied methods into a different hydro-climatic region and to validate the results using the sediment stock in the reservoir.

4.1 KRAICHBACH CATCHMENT

The Kraichbach is a small, north-west flowing tributary to the Rhine River with a length of about 55 km. It originates at altitudes of just around 300 meters above sea level in the relatively densely populated and intensively farmed Kraichgau area in the north-west of Baden-Württemberg. After flowing about half of its total length through the loess-covered gentle slopes of the Kraichgau and losing an altitude of about 170 m, it enters the gravel and sand dominated Upper Rhine Valley. For the remaining 30 km of flow length before it enters the Rhine, only 40 m difference in altitude are left, i.e. bed slope and flow velocities are comparatively small. This transition zone is where the gauging station Ubstadt is located and where a mean flow of 1.1 m³ is measured. The catchment size at this point is about 160 km², about half of the catchment of the junction with the Rhine. Half of the Kraichbach catchment at Ubstadt is used as arable land with the main crop types corn, grains and potatoes (Figure 12). This reflects the fertile, erodible soils (luvisols, colluvic soils) found in this area (RP Freiburg, LGRB 2015b), which have formed from silt-rich calcareous aeolian deposits (loess) during the last ice age. The loess cover is up to 15 m thick and therefore have a large water retention capacity, resulting in low rainfall-runoff activities (Zehe *et al.* 2001). Pastures and meadows play comparatively a small role (6% of the total area). The steepest slopes in the catchment are used as vineyards (4%). Deciduous forests and areas with sparse trees make up another third of the catchment. The area covered by settlements and roads makes up 8% (LGL 2013). There are two waste water treatment plants serving combined 55,000 people (DWA 2009-2014). Their effluent is 2.5 and 3.3 mio m³ per year respectively, which means that on average around 15% of the flow at Ubstadt is treated waste water. The annual precipitation of 700 mm a^{-1} is relatively evenly distributed around the year. The Kraichbach is a highly regulated river, especially after entering the Upper Rhine Valley. Here, it becomes part of a complex canal network characterized by slow flow velocities, tall artificial river banks and regulating structures such as weirs. Compared to this, the river in the upper (gauged) catchment has a rolling, more natural appearance. However, the river has been in part straightened and constrained to its current course, which presumably lead to bed erosion the formation of the tall river banks visible in many parts today.

4.2   ALB CATCHMENT

About 40 km south, the Alb River is comparable to the Kraichbach in length (total of 52 km) and catchment size (448 km²). Like the Kraichbach, it is a tributary to the Rhine flowing from south-east to north-west, and its gauging station in Ettlingen is also located after about half of its flow length in the transition zone between the Upper Rhine Valley and the surrounding higher terrain. In the case of the Alb, the surrounding mountains belong to the sand stone segment of the northern Black Forest. The 150 km² gauged headwater catchment however is distinctly different to the Upper Kraichbach catchment. The terrain is more mountainous; the sources of the Alb are to be found at elevations around 750 meters above sea level. After a flow length of 23 km at the gauging station, the Alb has a mean flow of 2.39 m³ s^{-1} and has lost an altitude more than 600 meters. Annual precipitation is at around 833 mm a^{-1} with a relatively even distribution around the year. Due to the higher elevation, snow and snowmelt (March-May) have a larger impact on the hydrology compared to the Kraichbach catchment. The soils (cambisols, podzols, regosols, pelosols) in the Alb catchment (RP Freiburg, LGRB 2015b) have formed primarily from sand stone and do not allow for intensive agriculture. Almost three quarters of the catchment is forested with deciduous and coniferous trees (Figure 12). Pastures make up 12%, orchards 4% and arable land 2% of the catchment. Settlements and roads are at around 9% (LGL 2013). There is one waste water treatment plant in the catchment serving 40,000 inhabitants and emitting close to 7 mio m³ a^{-1} of treated wastewater annually into the Alb (DWA 2009-2014). At the Ettlingen gauge, the mean flow share of treated wastewater is therefore at around 10%. The Alb has been straightened and its beds and banks have been fortified, with an increasing degree of interventions within settlements and towards the junction with the Rhine.
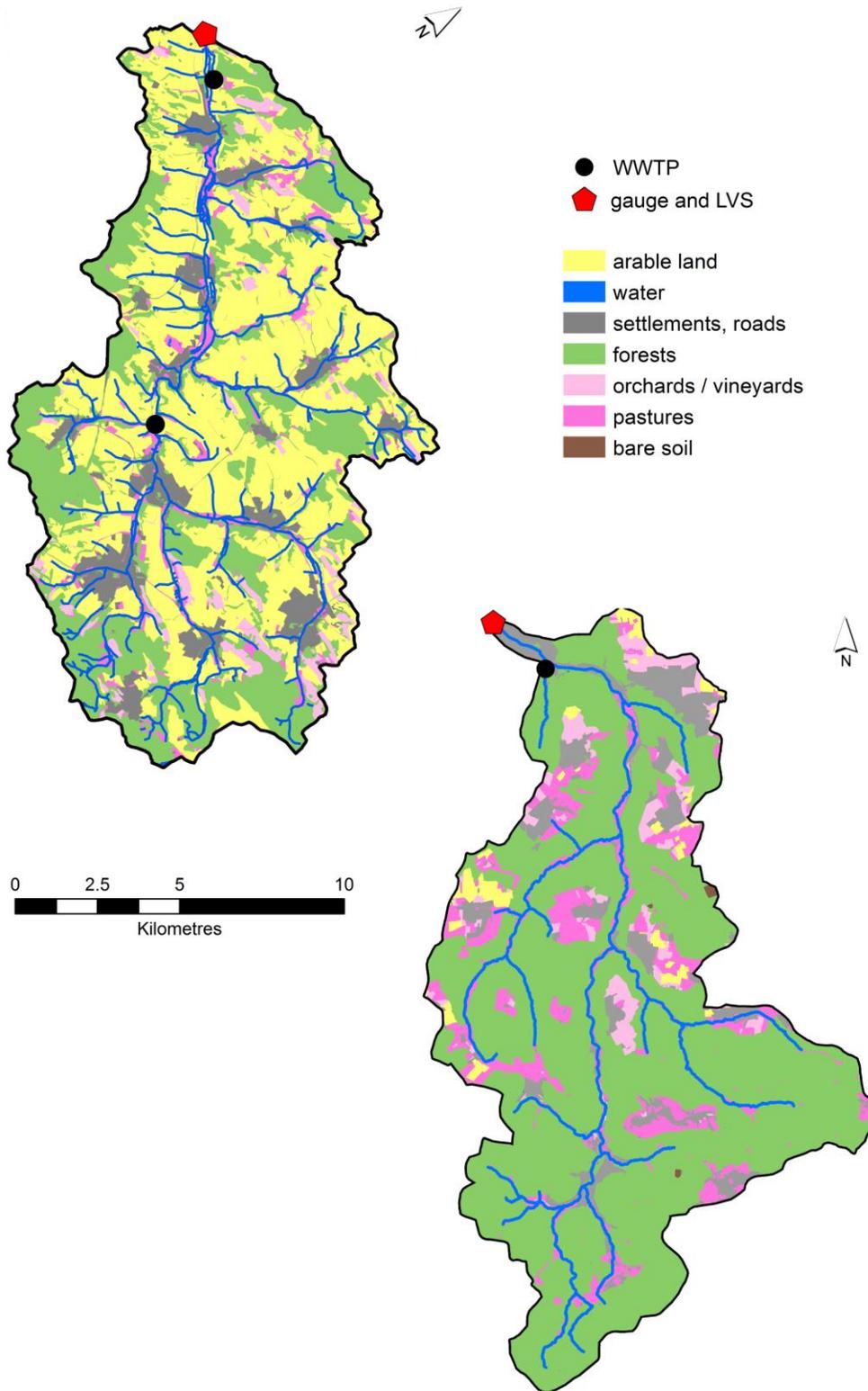
Figure 12: Land use/land cover of Kraichbach and Alb catchments. WWTP: Waste water treatment plant, LVS: Large volume sampler.

4.3 PASSAUNA CATCHMENT AND RESERVOIR

The Passauna catchment is located in the south of Brazil in the state of Paraná. It covers parts of the state capital Curitiba with about 3 mio inhabitants. The catchment is located in a sub-tropical climate at a mean elevation of 925 m and receives annual rainfall of about 1600 mm a^{-1}. The summer months between September and March tend to be wetter than the cooler winter months. Temperatures can drop close to the freezing point, but snow almost never falls. The Passauna River is impounded near Curitiba since 1989 and forms the Passauna reservoir, which is used to provide drinking water for approximately 1 mio inhabitants. The total catchment of the reservoir is about 152 km², whereas the gauged catchment is 81 km² (Figure 13). The main form of land cover of the gauged catchment is forests with 43% of the total area. Extensively used meadows and pastures and fallow land cover 20% of the catchment, whereas non-permanent crops make up 22%. Urban areas cover 8% of the catchment. The remaining areas are covered by shrub land (5%) and other land uses.

The reservoir is about 10 km long between inflow and dam and only 1 km wide. Its mean depth is 9.5 meters; the maximum depth is at 17 meters. The storage volume is estimated 69 hm³. Before entering the main body of the reservoir, Passauna River flows into a small (1000 m x 350 m) and shallow (~0.5 m) sediment basin. This so-called 'buffer' is connected to the main reservoir at the site opposite to the inflow by a constriction of about 20 m, followed by an S-shaped bend. From here on, the reservoir becomes gradually deeper before reaching its greatest depth right before the earthen dam in the south. The drinking water extraction point is located on the east side of the reservoir, about 3 km from the dam and 7 km to the inflow, respectively. A gauging station for Passauna River is located 2.7 km upstream of the inflow, where a mean flow of 2 m³ s^{-1} is recorded. There are several other, much smaller inflows to the Passauna Reservoir, the largest one being Rio Ferraria in the North West with about one fifth of the Rio Passauna's discharge. Rio Ferraria flows through a partly formal, partly informal settlement. Road runoff is directed into urban streams without further treatment and usually shows elevated concentrations of nutrients and heavy metals. It is acknowledged that only a part of the domestic sewage from the catchment is collected and transferred to a wastewater treatment plant.

Using hydro acoustic methods, manual sediment sampling and penetration experiments, the sediment stock in the Passauna Reservoir is estimated 4.14 10^6 tons (unpublished, MuDak-WRM 2019). In 30 years between construction and measurements, this indicates an annual increase of 138,000 t a^{-1}. Based on differences in carbon contents between catchment soils and reservoir sediments, about 15% of the material is considered to be from autochthonous production. When further correcting for catchment sizes, an estimated ~63,000 t a^{-1} of sediment have passed the monitoring station between 1989 and 2018.
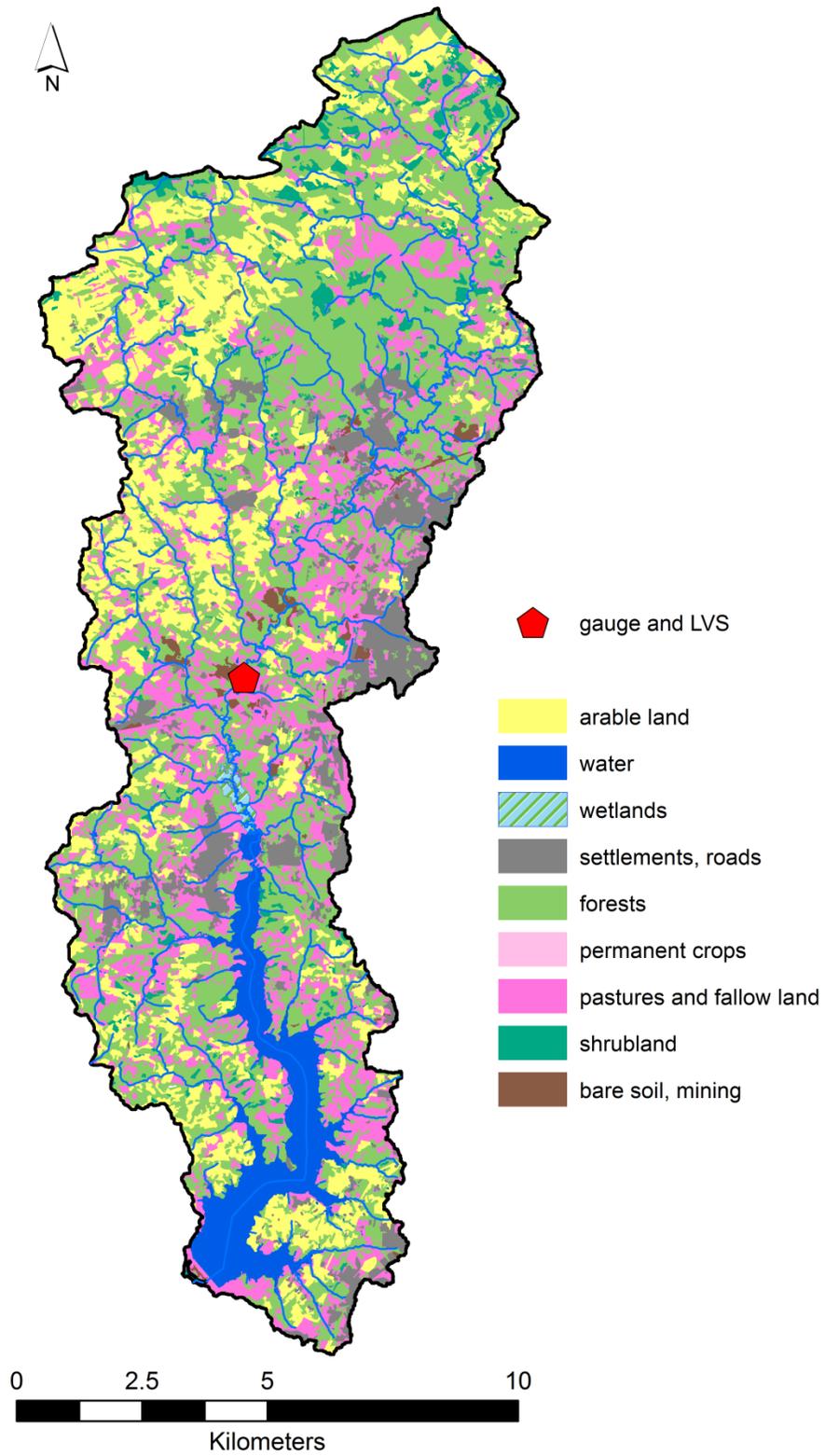
Figure 13: Land use/land cover of Passauna catchment. LVS: Large volume sampler.

5 METHODS

In this chapter, the methods and tools are described. The Large Volume Sampler (LVS) was installed in all three study areas for automated discharge-proportional sampling (Section 5.1) and different rating curves were established (5.2). The novel load flashiness index is introduced (5.3). Emission model results were produced for Alb and Kraichbach catchment for phosphorous and orthophosphate (Section 5.4.). Proximal water spectroscopy (5.5) was developed in mesocosm experiments and on Passauna Reservoir and set up for continuous river monitoring at Kraichbach. Turbidimeters (5.6) were also exclusively installed at the Kraichbach. The different goodness-of-fit-measures used are recapitulated in Section (5.7).

5.1 LARGE VOLUME SAMPLING

In this section, the Large Volume Sampler (LVS) is briefly described in terms of its components, setup, sampling procedure and subsequent processing of flow and concentration data. A similar setup has been briefly described in the context of storm water treatment in Fuchs *et al.* (2014). Further, there is a publication submitted to peer-review by Nickel and Fuchs (2019) about the use of the LVS for the monitoring of combined sewer overflows. The setup for river monitoring as described here has been presented by the author at the EGU General Assembly 2019 (Wagner *et al.* 2019).

5.1.1 DESIGN AND OPERATION

Central piece of the LVS is a stainless steel wine tank with a volume of about 1 m³. Ad- or desorption effects are minimized by the material. The tank is produced with a tapered bottom with bottom outlet and two faucets in different heights to facilitate the collection of the sediment and cleaning. The tank and a pump control unit were placed inside a small wooden shed to protect it from the weather and theft.

In order to create flow-proportional large volume composite samples, the LVS needs live discharge measurement input data. Here, a pump control unit HP Technik µCon AQUA processed analog signals scaled to 4–20 mA from the discharge measurements (Alb, Kraichbach) or the water level sensor (Passauna), respectively. The control unit calculated discharge based on polynomic functions. Based on cumulated volume, a 220 V pump was activated in an adjustable constant interval. The desired aliquot volume was defined by setting a site-specific pumping time depending on pumping capacity and height. Data was saved every minute for subsequent evaluation of hydraulics and the sampling process. Furthermore, the control unit was able to trigger a 220 V 3-way valve (e.g. Wiltec DQ320) to flush the sampling hose before each aliquot was taken. Steel spiral enforced 3/4–1" PVC suction hoses were used to connect the valve to the pump and to the 1 m³ container. A float switch prevented overflowing of the tank (cf. Nickel and Fuchs 2019).

5.1.2 SETUP FOR RIVER MONITORING

In the presented case at the rivers Kraichbach and Alb, discharges based on minute-by-minute water level measurements from an ultra sound based, official monitoring stations were used. The goal was to create a flow-proportional composite sample every 10–20 days to calculate suspended sediment and total phosphorous fluxes. This interval is short enough interval to include only a small number of (or no) rainfall events in the given climatic region for catchments of the given sizes, and long enough to sustain a manageable maintenance interval and moderate costs of analysis. In Brazil at Passauna River, similar considerations were made. Here, the official gauge was several kilometers downstream of the LVS location, so a pressure sensor measuring water height was installed to control the pump.

5.1.3 SAMPLING PROCEDURE

Once the container was full, the sample was withdrawn from the LVS. Separating particulate material and water phase by sedimentation was desired to limit the sample volume to be processed. Suspended particles were allowed to settle for at least several hours. A sample of supernatant water was then taken a) to determine the efficiency of sedimentation by quantification of remaining suspended solids, and b) for laboratory analysis of dissolved substances. Another ~4 L of water were saved for cleaning the tank. The first flush of sediment was collected into a 15 L sampling flask. Then, the remaining supernatant water was drained into the river using the upper and bottom faucets, taking care that no sediment particles were flushed out by adjusting drain speed respectively. After most of the supernatant water was drained, the saved supernatant water was used to mobilize all particles attached to the surface of the tank and to flush the entire sediment into the sampling flask.
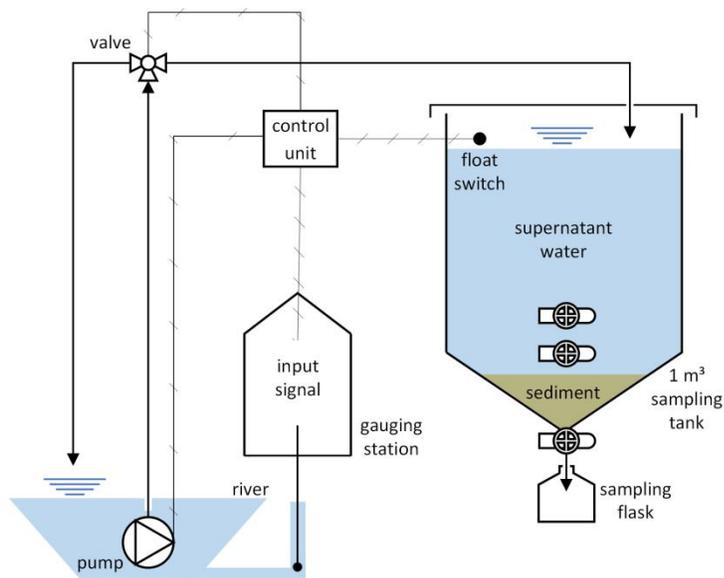
Figure 14: Left: Schematic drawing showing the basic functionality and the construction of the large volume sampler (LVS). Right: Photo of the LVS tank at the River Kraichbach.

Table 1: Overview of analyses performed with large volume samples.

	Kraichbach	Alb	Passauna
Sediment			
Total dry mass	x	x	x
Loss-on-ignition at 550°C of whole sample	x	x	
Silt and clay content (< 63 µm)	x	x	
Loss-on-ignition at 550°C of < 63 µm	x	x	
Total phosphorous content	x	x	
Supernatant water			
Suspended sediment concentration	x	x	x
Total phosphorous concentration	x	x	
Orthophosphate-phosphorous concentration	x	x	

The sediment and supernatant water subsample were analyzed in a laboratory. At Passauna River, analyses were limited to total suspended sediment (Table 1). The settled sediment was homogenized, sub-sampled and wet sieved using a standardized 63 µm sieve. Both fractions were then dried at 105°C for several hours and weighted to determine dry mass. Both fractions were sub-sampled then heated to 550°C for 120 minutes and weighted again to determine loss-on-ignition. Total phosphorous content was determined in a retain sample by disintegration in perchloric acid and subsequent colorimetric measurement. The supernatant water was analyzed for SSC by filtering through pre-weighted 0.6 µm glass-fiber filters by means of a vacuum pump, drying them for several hours at 105°C and weighing them again. Total phosphorous and orthophosphate-phosphorous in liquid samples were analyzed colorimetrically using the LCK 349 test in a Hach-Lange DR 2800 cuvette photometer.

5.1.4   LVS Data Processing

Because of the discharge-proportional sampling setup, the concentration in the LVS is, by design, a close representation of the mean in-river concentration for the sampling period. In order to determine SSC of the entire LVS, the total mass of the sediment, whether settled or still suspended in the supernatant water, had to be divided by the total water volume of the sample. To determine the total sediment mass of the sample, the total dry mass of the settled sediment was added to the total mass of the suspended sediment and then divided by the sample volume to get mean SSC. Similarly, total phosphorous concentrations were calculated from the phosphorous mass present in the settled sediment added to the phosphorous suspended or dissolved in the supernatant water. The general formula is as follows:

$$C_{LVS} = \frac{(C_{SW} * V_{LVS}) + M_S}{V_{LVS}} \tag{8}$$

where

C_{LVS} is the mean concentration of a substance in the LVS in mg L^{-1}

C_{SW} is the substance concentration in the supernatant water in mg L^{-1}

V_{LVS} is the LVS volume in L

M_S is the mass of the substance in the sediment in mg

The mean substance concentration C_{LVS} was then used as the target variable to build rating curves.

Here, the concentration measurements represent a mean concentration for an extended monitoring period with varying flow, referred to as event mean concentration. The concentrations were therefore correlated against discharge parameters for this period, i.e. the mean hourly discharge within the respective period, the maximum and minimum discharge, the median and the 5th and 95th percentile. For the Passauna River, discharge measurements were aggregated to daily mean values, which is the temporal resolution of available longer time series of discharge measurements on which the resulting rating curves were applied. For Alb and Kraichbach, hourly values were used. Scatter plots and goodness-of-fit-measures of linear and polynomial least squares regressions were compared and the flow parameter with the highest predictive power was selected. The same was performed with mean loads rather than mean concentrations, which were derived by multiplying mean flow and mean concentrations. The resulting linear or polynomial rating curves were only used, when a sufficiently strong correlation was found for a given parameter ($R^2 > 0.5$; $p < 0.05$, see Section 5.7). If these conditions were met in concentration-discharge and load-discharge regressions for the same data set, the concentration-discharge regression was used. Because discharge occurs in both terms in load-discharge regressions, there is an implicit relationship present. As suspended sediment concentration is usually positively correlated to discharge, scattering in sediment load-discharge plots is in most cases lower than the corresponding concentration-discharge plots. This does not mean however that the predictive ability of load-discharge regressions are higher (Milliman and Farnsworth 2013: 25).

When these criteria were met, the rating curves were applied to longer available time series of discharge measurements. For this purpose, the respective discharge curve was divided into continuous segments of constant cumulative flow volume equal to the flow volume represented by the sample (Figure 15). Then, the flow parameter from the rating curve in each segment was extracted and used as the explanatory variable in the rating curves to gain mean substance concentrations for each segment. By multiplying with mean substance concentrations, total event loads were determined, which could then be summed up to annual loads. This was done to gain long-term mean annual fluxes across dry and wet years. In addition, the dependency of fluxes on event magnitude was studied by ordering events by exceedance probability and calculating their cumulative long-term percentage on substance fluxes. When the criteria were not met, loads were calculated by multiplying annual discharge with mean measured concentration to calculate mean annual fluxes.
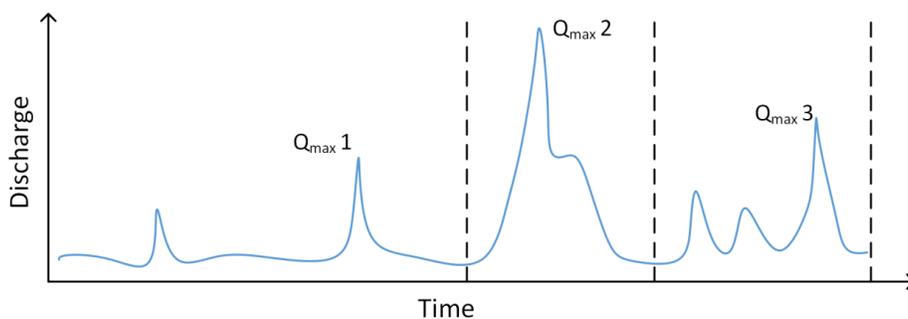
Figure 15: Three flow segments of different duration, but equal cumulative volume (area under the curve). The flow parameter used as explanatory variable (Q_{max}) is indicated for each segment. After Zheng (2018).

5.2 Rating Curves and Flux Scaling from Grab Samples

For the Kraichbach and Alb rivers, conventional flow rating curves based on grab samples were created. The basis at Kraichbach were 69 suspended sediment, 58 total phosphorous and 56 orthophosphate measurements performed between January 2017 and May 2019 (Table 6, Appendix). For the Alb River, data from the routine water quality monitoring could be used, because discharge gauge and quality monitoring station are in close proximity. Data was downloaded at the website http://jdkfg.lubw.baden-wuerttemberg.de (accessed 2019-09-01) for the years 2006–2017. While total phosphorous is not routinely measured at the given station, 65 suspended sediment and 148 orthophosphate measurements were coupled with discharge data (Table 7, Appendix). Linear or polynomial least squares regression was performed for all parameters. Similar to the establishment of rating curves for the LVS data, the resulting linear or polynomial rating curves were only used, when a sufficiently strong correlation was found for a given parameter ($R^2 > 0.5$; $p < 0.05$, see Section 5.7). While sediment rating curves can be applied to discharge curves from other time periods, flux scaling after Hilden (2003) requires monitoring data from the year of interest. Consequently, flux scaling could only be realized for orthophosphate at the Alb River for the emission model years 2009–2014. After suitable regressions between concentration and discharge were created, fluxes were estimated using the formula below:

$$F_{QC} = \frac{60*60\,s}{10^6} * \sum_{j=1}^{365*24} c_Q(Q_j) * Q_j \tag{2}$$

where

F_{QC}	is the estimated annual flux in t a^{-1}
$C_Q(Q_j)$	is the substance concentration in mg L^{-1} expr. as a pre-defined function of Q_j
Q_j	is the mean hourly discharge in m³ s^{-1}

For Passauna River, where from 2010–2015 only daily discharges were available, Formula (2) was slightly adapted in so that daily mean loads were calculated and then summed up to annual fluxes. In Germany, scaling techniques are often used instead of sediment rating curves, especially for (dissolved) parameters with relatively low dynamics. To estimate annual loads from actual measurements and applying the principle of Formula (1), the formula below is used. It is extended by a correction factor which is built from the mean discharge condition during sampling and the mean discharge in the year of interest (Hilden 2003).

$$F_Q = \frac{MQ_{year}}{Q_M} \frac{365 * 86400\,s}{10^6} * \frac{1}{N}\sum_{i=1}^{N} Q(t_i) * c(t_i) \tag{3}$$

where

F_Q	is the estimated annual load in t a^{-1}
MQ_{year}	is the mean discharge of the year of interest in m³ s^{-1}
Q_M	is the mean sampled discharge in m³ s^{-1}
$Q(t_i)$	is the mean daily discharge at time t_i in m³ s^{-1}
$c(t_i)$	is the substance concentration at time t_i in mg L^{-1}
N	is the number of water quality samples

At Passauna River, no regular monitoring data is available. Instead there is one intensively

measured high-flow event from October 2018 (MuDak-WRM 2019). Three stage-dependent linear rating curves were built following the clockwise dilution hysteresis. Hydrographs were then classified according to flow stage before applying the rating curves for all segments. This was performed on hourly discharge measured 2016–2018 as well as for daily discharges measured 2010–2015. In addition, an existing rating curve based on daily mean discharges to predict daily mean loads was used (Rauen *et al.* 2017, Figure 16) by applying Formula (4) to daily discharges for the years 2010–2015.

$$F_{QF} = \frac{86400\,s}{10^6} * \sum_{j=1}^{365} F(Q_j) \tag{4}$$

where

F_{QF}	is the estimated annual flux in t a^{-1}
$F(Q_j)$	is the daily load expressed as a pre-defined function of Q in m³ s^{-1}
Q_j	is the mean daily discharge in m³ s^{-1}
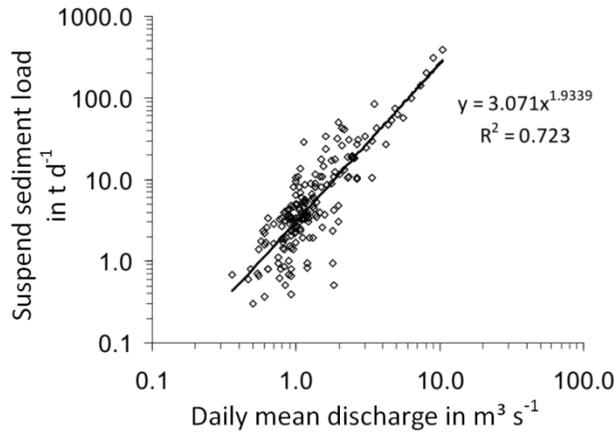
Figure 16: Sediment rating curve based on 115 SSC measurements from 1985-2015 for the Passauna catchment. Note the log-log scale (Rauen *et al.* 2017, translated).

5.3   FLASHINESS INDICES

Flow flashiness indices (FFI) were computed from the resulting grab sampling rating curves. A flashiness index is a mean representation of how quickly flow changes between two measurements. Small, mountainous catchments with frequent thunderstorms have high FFIs while large watersheds are usually well-buffered and therefore have low FFIs. The formula by Baker *et al.* (2004) for hourly measurements was used:

$$FFI = \frac{\sum_{i=1}^{n} |Q_i - Q_{i-1}|}{\sum_{j=1}^{m} Q_j} \tag{5}$$

where

FFI	is the flow flashiness index for, dimensionless
Q_i	is the mean hourly discharge for n hourly measurements in m³ s^{-1}
Q_j	is the mean daily discharge for m (= n/24) daily measurements in m³ s^{-1}

The formula was slightly adapted used to calculate the load flashiness index (LFI) based on loads

from grab sampling rating curves. Load flashiness represents the rapidity of short-term alterations of instantaneous loads.

$$LFI = \frac{\sum_{i=n}^{n}|L_i - L_{i-1}|}{\sum_{j=1}^{m} L_j} \tag{6}$$

where

LFI   is the load flashiness index, dimensionless

L_i is the mean hourly substance load for n hourly measurements in g s^{-1}

L_j is the mean daily substance load for m (= n/24) daily measurements in g s^{-1}

For the transport of suspended sediment and associated substances in most catchments, it can be assumed that concentrations are positively proportional to flow, or constant for any flow. Under this premise, the LFI can be regarded as a measure of probability, that calendar sampling is coinciding with periods of high transport: If it is low, changes in loads are slow and can therefore have a high chance to be covered by low-frequency sampling. If it is high, the rate of change of loads is high, which means short-termed pulses of substance transport control the annual flux. The index fuses the relevance of flux contributions from high flows –which can be large independently of catchment size and flow flashiness (Horowitz *et al.* 2015)– with the shortness of their duration. The LFI shares some characteristics of an approach proposed by Stelzer and Likens (2006), which also attempts to couple discharge-dependent concentration variations with flow flashiness, in their case for dissolved silicate which is usually negatively proportional to discharge. The simple approach presented here however is designed for particle-associated enrichment transport; it is purely driven from the measurement data and does not require prior knowledge for model parametrization.

From the indices from Formula (5) and (6), the normalized load flashiness index (nLFI) can be computed:

$$nLFI = \frac{LFI}{FFI} \tag{7}$$

Under the above assumptions, the nLFI indicates the degree of detachment of load flashiness from flow flashiness. It reflects on the inclination of the underlying rating curve with respect to the base load. For nLFI = 1, concentrations are constant across all flow stages. Load flashiness may still be high because of high flow flashiness, but specifically sampling high flows would not be required. Both indices must be inspected to conclude on the sampling strategy.

5.4   EMISSION MODELING

Modelling of Regionalized Emissions (MoRE) has been implemented in several regions in and outside Germany. It is a conceptual mesoscale model designed to assess pathway-specific substance emissions in yearly time steps (Fuchs *et al.* 2017). In the state of Baden-Württemberg, phosphorous and orthophosphate emissions are modeled for six diffuse pathways and four point source types (Kiemle *et al.* 2019). In this model version "METRIS-BW Version August 2019", the analytical units for which model results are computed are only ~4 km² large. These small sub-catchments were combined to spatially match the Kraichbach catchment at Ubstadt and the Alb catchment at Ettlingen. The mean annual emissions for the years 2009–2014 for orthophosphate and total phosphorous were computed according to the input data and calculation approaches described in Kiemle *et al.* (2019). One characteristic of METRIS is that it calculates diffuse phosphorous emissions from agricultural areas in two separate pathways.

Dissolved orthophosphate emissions from agricultural areas are assessed in the surface runoff pathway. It is calculated by multiplying the surface runoff volume from the water household model with an orthophosphate concentration, which is derived by applying an empirical regression to phosphorous topsoil concentrations. For total phosphorous in particulate form, the pathway erosion is used, which also uses area-wide measured mean P_2O_5 contents of agricultural soils (LTZ 2015) as an input date. This data set is updated every five years using recent soil measurements on the spatial level of districts (5-10 km²) for arable land, pastures and orchards. For the Kraichbach catchment a mean TP-content of agricultural soils of 1016 mg kg^{-1} was calculated which could be validated by five grab samples of eroded material collected during a field visit (min: 855 mg kg^{-1}, max 1209 mg kg^{-1}, mean 956 mg kg^{-1}).

Erosion is modeled in a three steps. First, soil loss from agricultural areas is assessed based on the Revised Universal Soil Loss Equation (RUSLE) (RP Freiburg, LGRB 2015a). Second, using land cover data and a digital elevation model, the degree of connectivity of agricultural areas to the stream network is determined and expressed as a constant called sediment delivery ratio. Third, sediment input is calculated by multiplying soil loss with sediment delivery ratio (GALF 2018). The phosphorous content of the soil data set described above is then used to compute phosphorous emissions via erosion. An enrichment ratio based on soil properties (typically between 1 and 6) is used to account for the fact that fine particles are disproportionally responsible for phosphorous transport, and that these fine particles have a higher probability of reaching a water body (Kiemle *et al.* 2019).

A spatially distributed sediment emission is only available on the basis of sheet erosion from arable land. In order to assess sediment emissions from other sources, constant emission values for forests (5.95 kg ha^{-1} a^{-1}), pastures (3.26 kg ha^{-1} a^{-1}) and orchards and vineyards (281 kg ha^{-1} a^{-1}) based on recommendations by GALF (2018). The much higher value for vineyards compared to the other two land uses can be attributed to the steepness of the terrain which is usually used for this purpose in Saxony, where these values have been established. This is also true for the Kraichbach catchment and soil properties are also comparable to Saxony. For these reasons and because of the small overall area in the study catchments, these values were adopted. Gully- and rill erosion as well as in-channel bank and bed erosion are however not considered. Sediment from urban sources was considered by multiplying a mean SSC of 52 mg L^{-1} (Fuchs and Nickel 2018) by modeled annual combined sewer overflow discharges (Kiemle *et al.* 2019). Waste water treatment plant emissions were included by multiplying an effluent SSC of

10 mg L^{-1} and the total annual effluent volume of the plant (DWA 2009-2014).

5.5 PROXIMAL WATER SPECTROSCOPY

The use of close-range, above-surface reflectance measurements in the visual and near infrared (VIS-NIR) domain to measure water quality parameters in surface water bodies was investigated. Three setups were tested: Mesocosm experiments (5.5.1), a boat-based method on the Passauna reservoir (5.5.2) and a permanent river monitoring station at the Kraichbach (5.5.3).

5.5.1 MESOCOSM EXPERIMENTS[1]

In order to define the potential range of SSC detection from above-water measurements, remote sensing reflectance was measured in a water tank (LxWxH 148x105x80 cm), painted flat black. Silt and clay sediment was collected from a wet detention basin of the Weingarter Bach the upper Rhine valley shortly after a series of high flow events. The material was then suspended in municipal tap water using an immersion pump to create gradually increasing concentrations of SSC between 0 and 2200 mg L^{-1}.

Hyperspectral data were acquired with AISA+ Eagle imaging spectrometer; a push broom scanner with an instantaneous field of view (IFOV) of 0.648 mRad and a field of view (FOV) of 36.04°. The AISA+ Eagle was equipped with a mirror scanner (Specim Limited, Oulu, Finland) to enable the recording of hyperspectral frames. The scanner was mounted on a scaffold at a height of 1.3 m above the water tank. Scans were recorded in four times spectral binning mode to achieve an appropriate trade-off between signal-to-noise ratio and spectral resolution (Silván-Cárdenas and Wang 2010). The corresponding data features 61 spectral bands between 398 and 957 nm with a full width at half maximum (FWHM) of 8.5–9.5 nm. The final dataset was restricted to 55 spectral bands, as bands above 900 nm were removed due to very low reflectance.

Two separate data sets were acquired, i.e. one during cloud free conditions and the other during homogenous cloud cover. Spectral acquisitions involved internal dark current measurements before each scan to account for systematic scan artefacts. During the scans we recorded the incoming radiance data, which enabled an absolute radiometric calibration to reflectance. The calibration was performed in the ENVI plugin CaliGeoPro (Specim Limited, Oulu, Finland). The signal from a 50x50 cm section of the water surface area free of sun glint or shading was averaged. From each curve, a baseline was subtracted measured in clear water in the shaded inside of the tank.

5.5.2 BOAT-BASED RESERVOIR SPECTROMETRY[1]

The approach was transferred and expanded to estimate concentrations during five campaigns in the Passauna Reservoir. A 'RoX' instrument (JB Hyperspectral Devices, Düsseldorf, DE) was used to measure reflectance of the reservoir. The device is based on a FLAME-S-VIS spectrometer (Ocean Optics, Dunedin FL, USA) and is an advanced development of a spectroscopy setup described in Burkart *et al.* (2014) to perform unsupervised reflectance measurements. The two ends of a bifurcated glass fiber leaving the spectrometer point at nadir and zenith, respectively. The downward-looking fiber is measuring upwelling radiance within an opening angle of about 20°. Shutters alternately close one and both fibers to

---

[1] This section partly consists of modified excerpts of a peer-reviewed journal article by the author (Wagner *et al.* 2018).

measure up- and downwelling radiance as well as dark current in each measuring cycle. After subtracting the dark current, reflectance is calculated as the quotient between downwelling and upwelling radiance. Compared to the AISA+ Eagle imaging spectrometer which can be e.g. mounted in aircrafts, the RoX is designed for a mobile and/or autonomous use . It is much smaller, lighter (~3 kg), and requires less than 5 W from various 12 V power sources. In the case of the reservoir campaign, a rechargeable 2600 mAh lithium polymer battery pack was used. The RoX is non-imaging and measures mean reflectance within the footprint, which is a function of the 20° glass fiber opening angle and the height above the water surface (Figure 17). The prototype used here is also not radiometrically calibrated, although calibration is available for recent versions. For empirical approaches as pursued here, radiometrically correct spectra are not needed. The built-in spectrometer features about ~1500 spectral bands between 400 and 900 nm with a full width at half maximum (FWHM) of ~0.3 nm. However, the JB Hyperspectral shell already averages the response of two neighboring bands, so that the accessible raw data consists of ~750 bands. For the reservoir campaigns, the spectrometer was mounted to a boom stand on a small motor boat. From about 1.2 m above the water surface, the device pointed at the water surface outside the boat shadow and measured reflectance curves in short time intervals (about 10 seconds), while driving transects across the impoundment. This way, water surface reflectance curves with corresponding locations and time information were recorded. In the presented case, water surface is defined as the part of the water column from which sunlight is reflected back into the upper hemisphere; or the water column above the attenuation depth. Only optically deep water was measured.
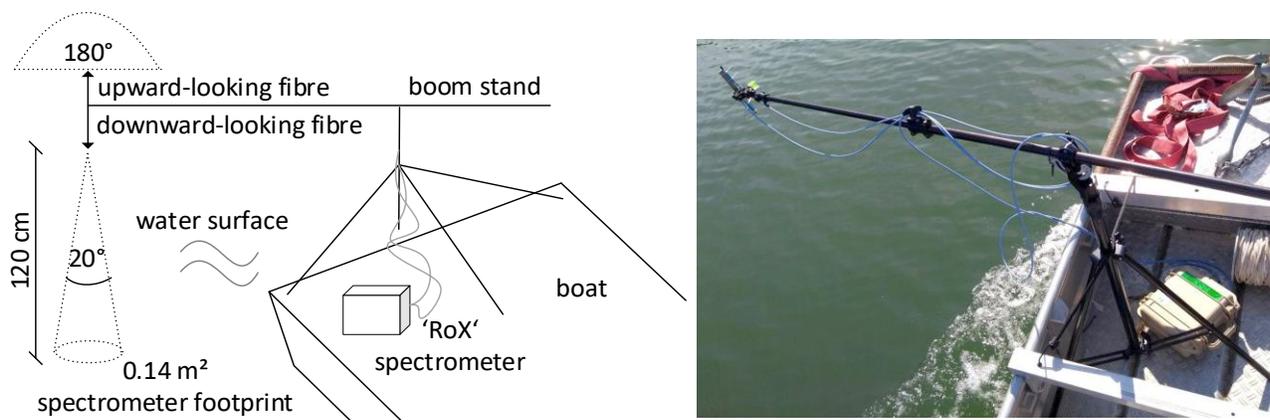
Figure 17: Left: Diagram of the boat setup used to acquire reflectance measurements at Passauna drinking water reservoir. Right: Photo of the system during measurement.

5.5.3 Permanent River Water Quality Monitoring

The RoX spectrometer was installed near the Kraichbach gauging station to regularly (every 10 minutes) monitor water reflectance between 10:00 and 15:00. For this purpose, the spectrometer was placed inside a metal switch box directly at a weir, where the Kraichbach is comprised by vertical walls. From the inside, the optical cables were passed through a 1.4 m aluminum pipe, so that the end of the pipe was directly above a section of the river where even during low flow conditions, the water level was about 30 cm. A vertically placed threaded bar was used to decrease the distance between the optical cables and the water surface to about 1.5 m, depending on flow height. This corresponds to a measured footprint of 0.17 m². The amplitude of the measured foot print due to water level fluctuations is estimated 0.06 m² to 0.2 m². The system was powered by a 100 W solar panel which charged a 12 V led car battery connected to the spectrometer. The measurement station was visited regularly to retrieve the data and to take reference samples. Maintenance was limited to occasionally removing spider webs around the sensor.

Figure 18: Kraichbach water reflectance monitoring station with a solar panel and a metal switch box containing a car battery and the 'RoX' spectrometer. One downward- and one upward-looking optical cable are installed ~1.5 m above the water surface by means of an aluminum boom.

5.5.4   Spectral Analysis[2]

All spectral data was processed in R (R Core Team 2017). Raw spectra were smoothed using a Savitzky-Golay filter (Savitzky and Golay 1964). Filter order was 1, filter width for the AISA mesocosm data was 3; for the RoX data filter width was 51. Only for the reservoir data, boundary reflectance values for each band were defined based on visual inspections to eliminate a small share of strongly distorted spectra. From each reservoir reflectance value, the mean value of the corresponding curve was subtracted to minimize observed shifts of the entire reflectance curve across wavelengths. This was presumably caused by rapidly changing light conditions during the measurement or varying degrees of sun glint within the footprint. In the case of the river monitoring station, subtracting the mean did not improve the later parameter estimations. Instead, the daily mean of the six spectra recorded between 11:00 and 12:00 (CET, ignoring daylight saving time in summer) were used. Given the south-east facing wall of the weir where the instrument was installed, the measured water surface was not shaded throughout seasons during this time of the day. For the reservoir, the first derivative of each reflectance curve was determined. For the case of the and mesocosm and Kraichbach monitoring station, the inclusion of the first derivative did not improve later parameter estimations.

The processed reflectance curves (and their first derivatives) were used as predictors for substance concentrations in partial least squares (PLS) regression models using the orthogonal scores algorithm. Partial least squares regression (PLS) is an empirical method especially well-suited for data with many spectral bands as ('explanatory') predictor variables and few concentration measurements as ('dependent') response variables (Wold *et al.* 2001). 'Small N large P problems' cannot be solved using multiple linear regression methods because of the risk of collinearities and singularity. PLS is a generalization of the better-known principal component analysis. Unlike a principal component analysis which attempts to explain the variance of the response variable, PLS selects the latent variables in order to explain as well as possible the covariance between predictors and response variables. It circumvents the problems emerging from predictor covariance by decomposing the entire feature space into orthogonal scores and loadings. The regression itself is then performed only on a small number of latent variables contained in the scores and loadings matrices (Torres and Bertrand-Krajewski 2008). In many disciplines, modern measurement like chromatographers and spectrometers devices produce many values per observation. PLS is thus well suited for the high dimensionality and redundancy in hyperspectral data. PLS and related methods which are able to make use of the entire data set have become standard statistical tools and are therefore well established and described (Mevik and Wehrens 2007). Because of these features, PLS is a suitable method for relating hyperspectral data with water quality measurements. The R package 'pls' (Mevik and Wehrens 2007) was used. Pairs of concentration values and corresponding (mean) spectra were randomly split in a training and test groups of five as recommended by Li *et al.* (2002) to calibrate and validate the empirical models using cross-validation. In this process, the model is calibrated with one group left out. Subsequently, prediction accuracy was tested on the group not used for calibration (Reyna *et al.* 2017). To gain more stable assessments of the overall cross-validation, this process was iterated 100 times to extract mean goodness-of-fit measures. The models with one to five

---

[2] This section partly consists of modified excerpts of a peer-reviewed journal article by the author (Wagner *et al.* 2018)

(mesocosm: seven) latent vectors were then evaluated by their root mean square error of prediction (RMSEP) and their R^2 for the test and training sets. The setup with the smallest number of latent vectors and the lowest RMSEP was deemed optimal. For the reservoir and Kraichbach monitoring data set, one of the PLS models for each parameter was used to estimate concentrations from the remaining spectra which were not used to train or test the model.

5.6 TURBIDITY MEASUREMENTS

At the Ubstadt gauging station of the Kraichbach, an automated Turner Designs Cyclops-7 submersible logger was installed to measure turbidity every 10 minutes. The measurements are based on the voltage response of the detection from a 90° scattered light beam, which is calibrated to a turbidity standard (formazine). A linear regression between the measured nephelometric turbidity units (NTU) and available simultaneous SSC measurements was made to transform the turbidity in concentrations. Based on typical NTU-SSC relationships reported by other studies, unreasonably high NTU measurements were excluded from the regression (Rügner *et al.* 2014; Martínez-Carreras *et al.* 2016) Cleaning and data recovery was performed about every 2–3 weeks. An automated wiper was installed to regularly clean the measuring window, because previous attempts of using turbidity data failed due to biofouling and sensor drift. The time frame where quasi-continuous turbidity data is available is November 2018 to September 2019.

5.7 GOODNESS-OF-FIT MEASURES

To evaluate whether a rating curve is suitable for predicting concentrations (or loads) from discharge, a threshold coefficient of determination $R^2 \geq 0.5$ at a significance level of $p \leq 0.05$ was selected. The degree of polynomial approaches was determined by increasing the degree until the improvement of the model became insignificant ($p \geq 0.05$). For polynomial regressions, the adjusted R^2 was used, to adapt to the number of explanatory variables. R^2 was also used to evaluate PLS models along with the root mean square error of prediction (RMSEP) to select the number of latent vectors used (Mevik and Cederkvist 2004). While the R^2 indicates how well the model explains the variance of the response, the RMSEP indicated the mean deviation of model from the observation with respect to the original value range. The local maximum/minimum for R^2/RMSEP was used to select the optimal number of latent vectors. To evaluate event fluxes at the Kraichbach for different modes of estimation, the commonly applied Nash-Sutcliffe-Efficiency (NSE) was used (Nash and Sutcliffe 1970). It ranges from $-\infty$ to 1, whereas values > 0.5 are deemed acceptable for applying the model (Moriasi *et al.* 2007).

6 MEAN ANNUAL FLUXES

First, the LVS results of all three sites are presented in terms of flow conditions, event mean concentrations and their relationships (Section 6.1). Section 6.2 presents the conventional rating curves and flux scaling results established by coupling discharge and substance concentration measurements at Alb and Kraichbach. Pathway-specific emission model results for Alb and Kraichbach are shown in Section 6.3. All techniques estimating annual fluxes are then compared in Section 6.4.

6.1 LARGE VOLUME SAMPLING RESULTS

In this section, the results of the two LVS monitoring campaigns at Kraichbach and Alb River as well as Passauna River are presented. Each equal-volume sampling interval is regarded in terms of flow parameters and mean LVS (i.e. mean in-river) substance concentrations. An overview about the two stations and the sampling is given in Table 2.

Table 2: Overview of characteristics and general set-up of the large volume sampling at Kraichbach and Alb. SSC: Suspended sediment concentration; TP: Total phosphorous; o-PO$_4$-P: Orthophosphate-phosphorous; LOI: loss-on-ignition at 550°C

	Unit	Kraichbach	Alb	Passauna
Station name	-	Ubstadt (76182)	Ettlingen (03301)	BR277 – Campo Largo
Catchment area	km²	161	149	81
Easting	m	3473357	3456016	662423
Northing	m	5446669	5423428	7187133
Meters above sea level	m	113.71	122.81	892
Monitoring period	-	03/2017–12/2018	02/2017–04/2019	02/2018-05/2019
Number of (complete) samples	-	32 (26)	27 (18)	33 (27)
Mean duration of comp. sampling	days	10.3	17.9	9.1
Mean flow (MQ)	m³ s⁻¹	1.10	2.39	~2.0
Mean low flow (MNQ)	m³ s⁻¹	0.62	0.8	~0.67
Two-year flood (HQ 2)	m³ s⁻¹	7.9	26.4	-
10-year flood (HQ 10)	m³ s⁻¹	14.2	54.6	-
Mean sampled flow	m³ s⁻¹	1.29	2.57	1.7
Highest sampled flow	m³ s⁻¹	7.78	12.5	12.4
Flow interval aliquot	m³	13,500	21,700	15,498
Volume aliquot	L	12.05	8.5	15.5
Flow per composite sample	mio m³	1.11	2.39	0.94
Mean SSC	mg L⁻¹	104	13.1	48.4
Mean TP	mg L⁻¹	0.32	0.13	-
o-PO$_4$-P share of TP	%	45	75	-
Mean share silt and clay (<63µm)	%	94	81	-
Mean LOI of silt and clay	%	13	31	-

6.1.1 KRAICHBACH

At Kraichbach River, 32 large volume samples were collected between February 2017 and December 2018 (Table 2). For 26 of them, all three parameter concentration measurements are available.

Discharge conditions during the 6–24 day sampling periods varied widely (Figure 63, Appendix). One sampling interval in April 2017 was characterized by constant flow conditions around the long-term mean flow with almost no variation. Three sampling periods in June/July and September represented almost rainfall-free intervals of a very dry summer with discharges around the mean low discharge (Figure 69, Appendix). The pumping intervals varied between 27 minutes for high discharge conditions (7.8 m³ s^{-1}) and about 6 hours for low flow conditions (0.6 m³ s^{-1}). The conditions with very low base flow could be observed even in autumn (Figure 66, middle, Appendix) and is contrary to the wet conditions of the preceding spring. Between autumn 2017 and spring 2018, soil water stores were filled which is apparent from elevated discharge conditions above mean annual flow even during sampling periods that were largely unaffected by rain. Unlike the year before, autumn and winter 2018/2019 stayed too dry to recuperate the exhausted water storages in the catchment. The large May 2018 event was a two year flood while the December 2018 event (Figure 66, right, Appendix) stayed just below this threshold.

Mean in-river substance concentrations also varied distinctly. The lowest SSC recorded was 11.6 mg L^{-1} in a 19 day low flow period in November 2018 and reached 827 mg L^{-1} as an eight day average during the two-year flood in May 2018 (Figure 69, Appendix). Throughout the monitoring period, phosphorous concentrations at Kraichbach were above their respective regional management goals of 0.1 mg L^{-1} for TP and 0.07 mg L^{-1} for orthophosphate (BMJV 2016). Figure 19 shows the rating curves produced from coupling the highest discharge in the monitoring period with substance concentrations and loads. Like SSC, mean TP concentrations covary with event magnitude. Two samples from July 2017 and June 2018 do not fully fit in this pattern and are showing elevated TP concentrations around 0.5 mg L^{-1} for moderate peak discharges around 3 m s^{-1} and moderate mean event SSC compared to other samples with similar TP levels. These events also show the highest orthophosphate concentrations in the observation periods (0.24 and 0.21 mg L^{-1}), making up around 50% of TP. The results suggest a steep increase of both SSC and TP concentrations and loads for events with discharges higher than 4 m³ s^{-1}. Higher discharges result in higher SSC, with a lower share of phosphorous close to the mean level found in agricultural soils in the catchment. For orthophosphate, no relationship between the highest discharge in each sampling interval and concentrations or loads was found which complies with the pre-defined criteria for rating curves (R² > 0.5; p < 0.05, see Section 5.7).
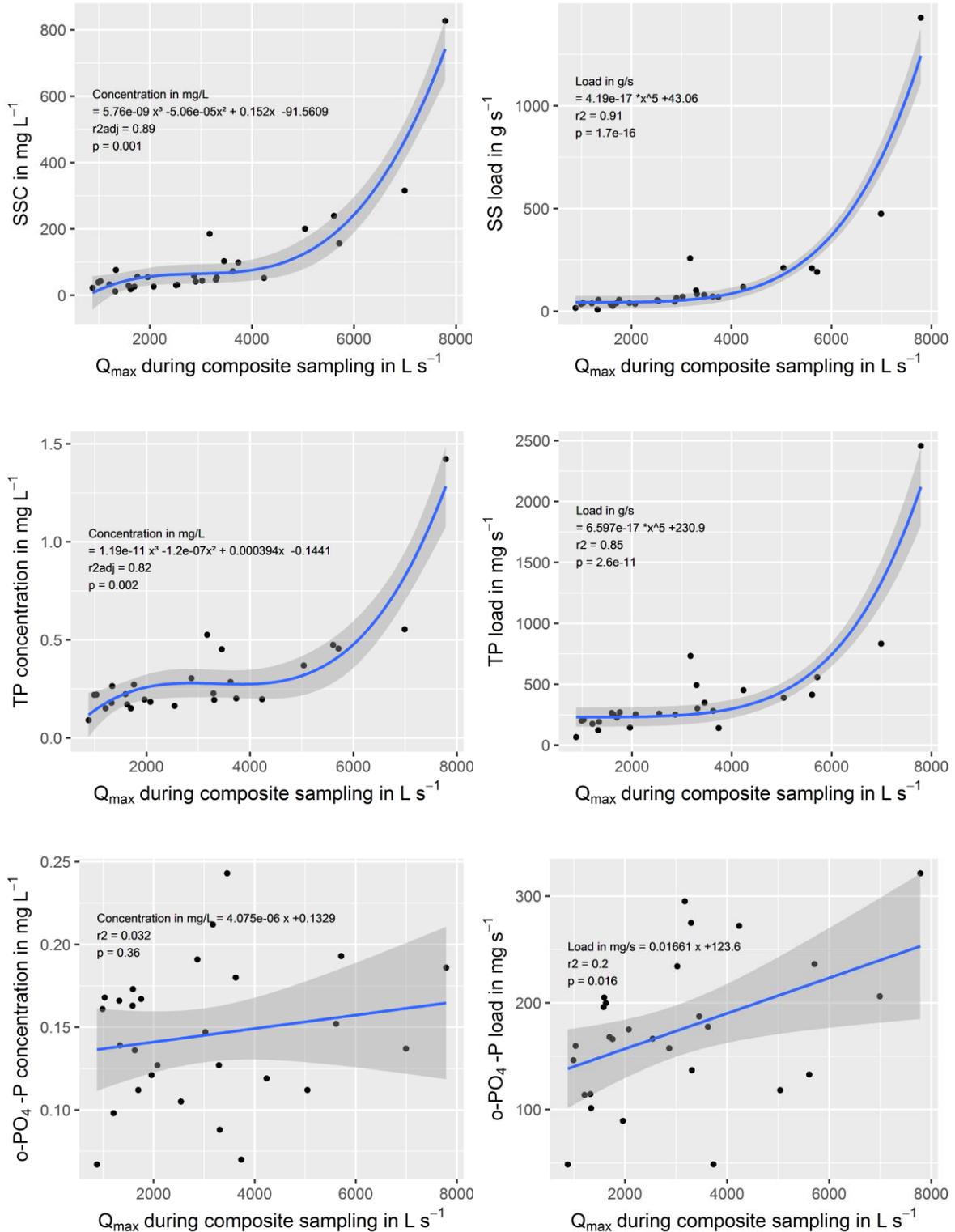
Figure 19: Adapted rating curves (blue) for for suspended sediment (SS), total phosphorous (TP) and orthophosphate (o-PO_4-P) at the Kraichbach River, based on the highest discharge recorded in the sampling period and the corresponding event mean concentration (left colum) or event mean load (right column). Curve equations, (adjusted) R^2 and p-values of regressions indicated. Grey area: 95% confidence level interval for predictions

It was further examined how LVS substance concentrations relate among one another. Figure 20 shows a scatterplot of TP and orthophosphate, Figure 21 of SSC and TP. While there is a significant linear relationship between SSC and TP, TP is only weakly correlated to orthophosphate in the Kraichbach samples. Phosphorous transport in the Kraichbach is therefore controlled by suspended sediment transport, which is also indicated by the low contribution of orthophosphate to the actual (measured) TP mass fluxes (Table 2). The relationship between TP content of the LVS sediment and the loss-on-ignition is not significant (Figure 73), but it can be observed that sediment from certain low-SSC samples tend to coincide with elevated TP and loss-on-ignition contents. Taking into consideration Figure 22 and Figure 23, it is suggested that events with a maximum flow with a maximum discharge around 2.5 m³ s^{-1} show elevated phosphorous contents in the LVS sediment of up to 2,500 mg kg^{-1} and higher shares of organic matter, indicated by elevated loss-on-ignition values. Combined sewer overflows and storm-water overflows, which discharge into the river under these non-erosive flow conditions, can be identified as the likely sources for material richer in phosphorous and organic matter. The higher availability of autochthonous and allochthonous organic matter may be responsible for the higher TP and loss-on-ignition values at low flows during the growing period (Figure 23). For periods without flows exceeding the overflow threshold, the suspended matter base load of about 30 mg L^{-1} (Figure 19), is likely to be dominated by relocated river deposits, a process that can be substantially facilitated by bioturbation (Bosworth and Thibodeaux 1990). Suspended sediment mass and properties for flows higher than 4 m³ s^{-1} are increasingly dominated by erosion processes in the catchment. This is indicated by the convergence of phosphorous and organic matter contents towards the mean agricultural soil content for erosive events (Figure 22). It is noteworthy that the mean TP content of 1016 mg kg^{-1} for arable lands in the Kraichbach catchment itself is rather high, representing an 'optimal' phosphorous supply (class C, range A–E) for crops according to the responsible association in Germany (VDLUFA 2018). Such events can be identified by a high mean suspended sediment concentration and high maximum flows. In other words, the high phosphorous concentrations associated with urban sources are diluted by eroded soil with relatively low TP and organic matter contents. These high mean event SSC measured with the LVS for high-flow events are likely to be controlled by extreme concentrations during a short time span at peak flow, preceded and followed by a dilution with clearer water. This can be concluded from smaller events with lower suspended sediment concentrations (Figure 69). However, hysteresis cannot be fully assessed with the LVS method due to the integrating mechanism.
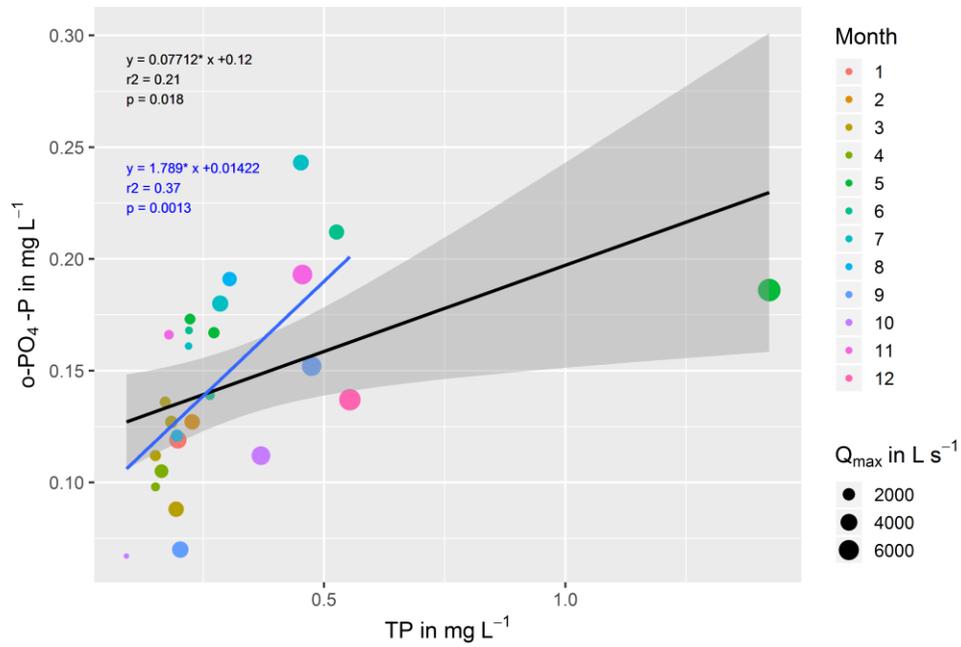
Figure 20: Mean composite sample total phosphorous (TP) against mean orthophosphate concentration (o-PO$_4$-P) at the Kraichbach River. Linear regression line (black) indicated with equation, R² and p-value. Blue line: linear regression without the highest TP value. Grey area: 95% confidence level interval for predictions. Colors indicate the end date of sample collection from January (1) to December (12). The highest measured discharge in the respective period is represented as point size.
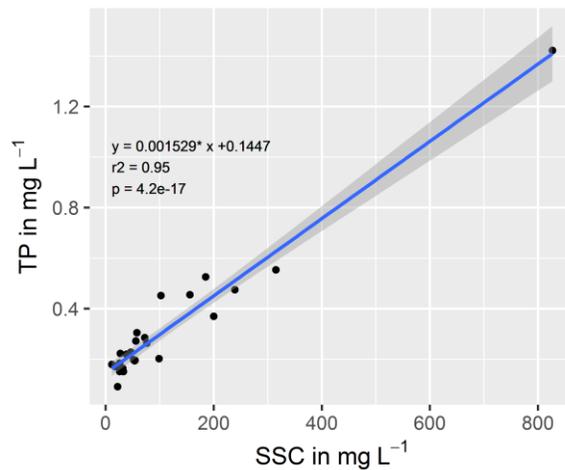
Figure 21: Left: Mean composite sample suspended sediment concentration (SSC) against total phosphorous concentrations at the Kraichbach River. Linear regression line (blue), equation with R² and p-value indicated. Grey area: 95% confidence level interval for predictions.
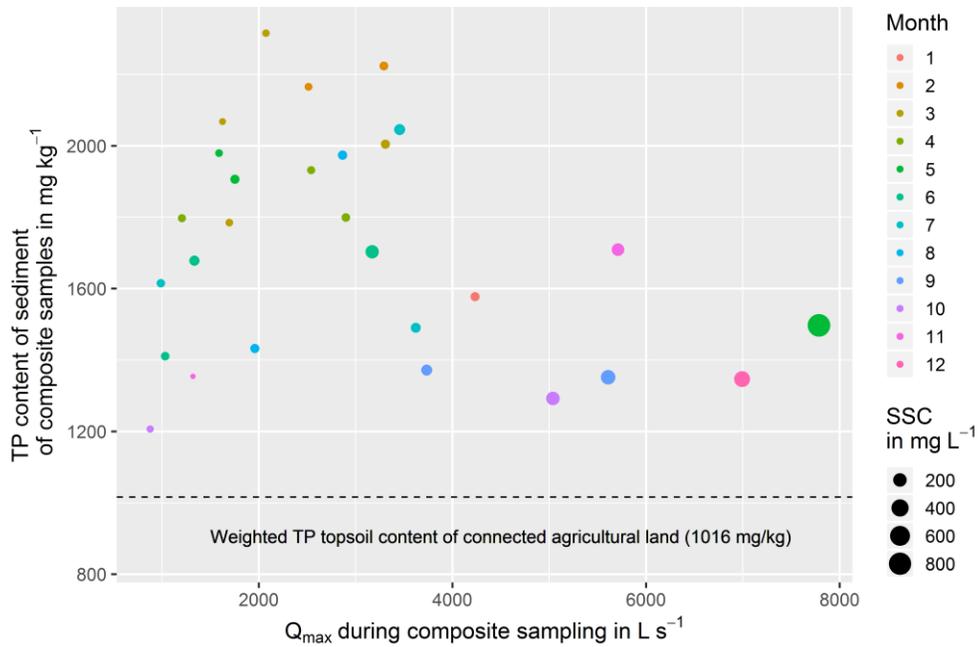
Figure 22: Total phosphorous (TP) concentrations of large volume sampler sediment against the highest discharges (Q_{max}) recorded in each sampling period for the Kraichbach River. Colors indicate the end date of sample collection from January (1) to December (12). The mean suspended sediment concentration (SSC) is represented as point size. The topsoil total phosphorous content of arable land in the catchment is indicated (dashed line).
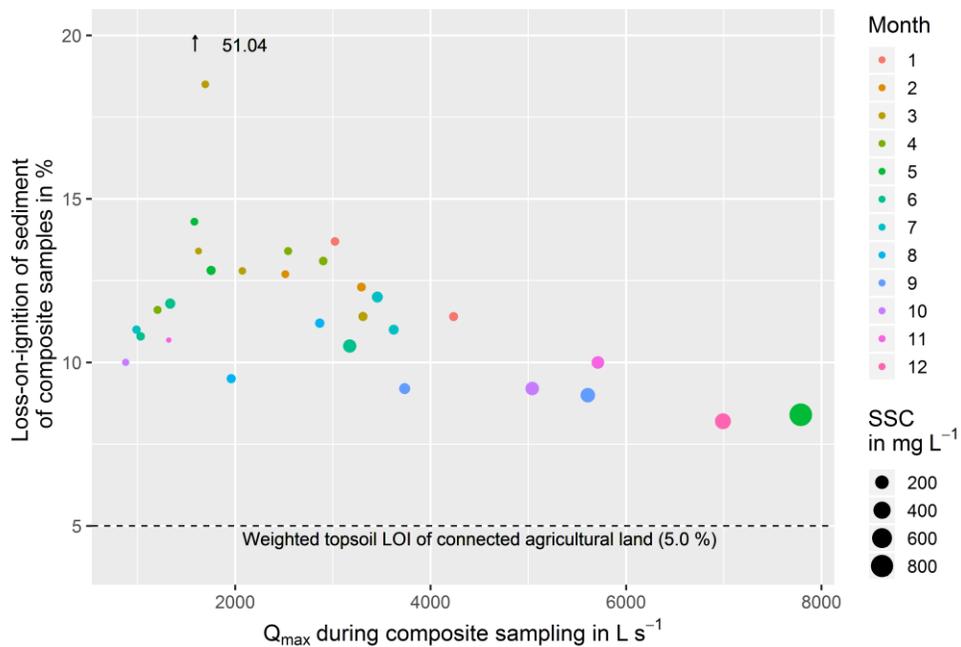
Figure 23: Loss-on-ignition of large volume sample sediment against the highest discharges (Q_{max}) recorded in each sampling period for the Kraichbach River. Colors indicate the end date of sample collection from January (1) to December (12). The mean suspended sediment concentration (SSC) is represented as point size. The topsoil loss-on-ignition (LOI) of soil from arable land in the catchment is indicated (dashed line).

57

The rating curves for SS and TP were applied to discharge data from the Kraichbach gauge from 1976-2019. For this purpose, the discharge curve was divided into continuous segments of a constant cumulative flow volume (1.11 mio m³), equal to the cumulative volume represented by each composite sample. Then, the maximum discharge in each segment was extracted and used as the explanatory variable in the rating curves to gain mean substance concentrations for each segment. For orthophosphate, no meaningful relationship as defined in Section 5.1.4 was developed between concentrations and flow parameters (Figure 19). Load estimates were therefore made by multiplying mean concentrations with the flow volume. LVS load calculations were made for 2009–2014, where modeled emission data is available (see Section 6.3), and for 1976–2019. The latter analysis was made to assess the effect of event magnitude and its relationship to the total loads transported out of the catchment. Figure 24 shows capped and uncapped cumulative curves for the long-term modeled substance fluxes of the separated segments for SS and TP. In the capped version, concentrations for discharges higher than the highest one observed during the LVS monitoring stay constant and therefore represent a conservative estimate. In the uncapped scenario, the (exponential) rating curve is maintained for all discharges (maximum: 26 m³ s^{-1}), leading to extreme substance concentrations (max SSC: 70,000 mg L^{-1}) and subsequently to high loads attributed to individual events. The discharge or water volume flux is equal to the 1:1 line, because each segment delivers by definition the same quantity of cumulative discharge. Rating curves with positive inclinations like the ones applied here result in cumulative flux curves below the water flux curve. The closer a cumulative substance flux curve is to the water flux curve, the smaller the dependency of that substance to discharge.
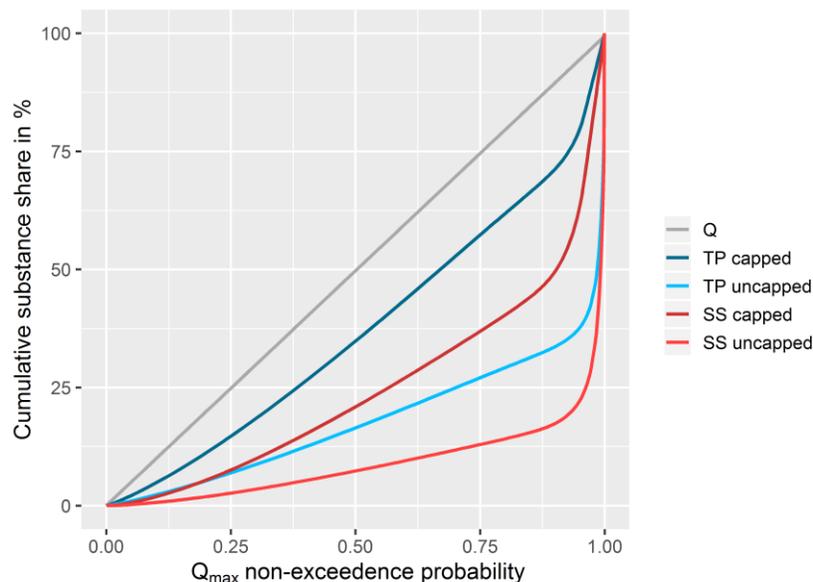
Figure 24: Cumulative curves (colored lines) of modeled long-term flux contributions of equal-volume flow intervals at the Kraichbach River, sorted by magnitude of the highest discharge (Q_{max}) per event. Q: Water flux. TP capped: Application of Q_{max}-Total phosphorous (TP) rating curve with TP_{max} = 1.42 mg L^{-1}. TP uncapped: Same, without upper concentration boundary. SS capped: Application of Q_{max}-suspended sediment concentration (SSC) rating curve with SSC_{max} = 742 mg L^{-1}. SS uncapped: Same, without upper concentration boundary.

6.1.2   ALB

At the Alb River, 27 composite samples were taken between February 2017 and April 2019. For 18

of them, all three substance concentration measurements are available. The aliquot volume was at times slightly lower than expected, leading to longer average sampling durations of 18 days (Table 2). One sample in June and July 2017 required over a month to complete. In order to maintain comparability of the individual samples, reconfigurations were kept to a minimum, but cleaning and maintenance of the system lead to small alterations in the pump performance. Compared to the Kraichbach, the Alb River has higher discharges as well as greater flashiness due to the more mountainous catchment (Figure 64). Even during the very dry summer of 2018, local thunderstorms led to sharp increases and decreases of the hydrograph (Figure 67, middle). The highest discharge recorded was 12.5 m³ s⁻¹, which is more than five times higher than mean flow (2.39 m³ s⁻¹) but well below a two-year flood (26.4 m³ s⁻¹). Several events between 15 and 20 m³ s⁻¹ did occur in the monitoring period, but were not sampled because the LVS was inactive. Similar to the Kraichbach, a dry summer in 2018 was followed by a dry autumn and winter which lead to a distinct decrease of mean flows.

Substance concentrations were comparatively constant across discharge conditions. Several hydrological parameters were tested as explanatory variables for SSC, TP and orthophosphate. For the Alb River, only one correlation in comply with the criteria ($R^2 > 0.5$; $p < 0.05$, see Section 5.7) was found. The maximum discharge in each sampling period was linearly correlated to SSC ($R^2 = 0.52$, $p < 0.001$, Figure 25). If loads rather than concentrations are the target variables, the relationship is better approximated by an exponential curve. Neither TP and orthophosphate show significant relationships to flow parameters, concentrations are rather constant across different event magnitudes. Due to the geology of the black forest with phosphorous-rich minerals, the mean groundwater concentration is at elevated 0.11 mg L⁻¹ (KUP 2018), 36% higher than the management goal for surface water bodies (BMJV 2016). This concentration is reflected during low flow in the river with a tendency of lower concentrations towards higher flows – a dilution with rain water.

Unlike the Kraichbach catchment, SSC and TP are unrelated in the Alb River (Figure 28, left), while TP is highly covariant with orthophosphate (Figure 28, right). Phosphorous loads are therefore controlled by the dissolved fraction, which makes up 75% of the total monitored load at the Alb River (47% for the Kraichbach, see Table 2). The loss-on-ignition of LVS sediment is weakly positively correlated to the sediment TP content (Figure 72). The relationship towards higher loss-on-ignition remains unclear, because for the two samples with loss-on-ignition values around 70% from Figure 27, there are no TP measurements available. For both, TP and loss-on-ignition, the pattern of highest contents for medium-sized events observed at Kraichbach is, if at all, much less apparent for the Alb River (Figure 26 and Figure 27). It must be noticed that sediment in the Alb has about twice the loss-on-ignition and TP content compared to the Kraichbach, indicating a higher share of organic components in the sediment. For both parameters, the mean arable land topsoil contents within the catchment are indicated (LGRB 2015; LTZ 2015), which are well below the contents found in the sampled sediment.
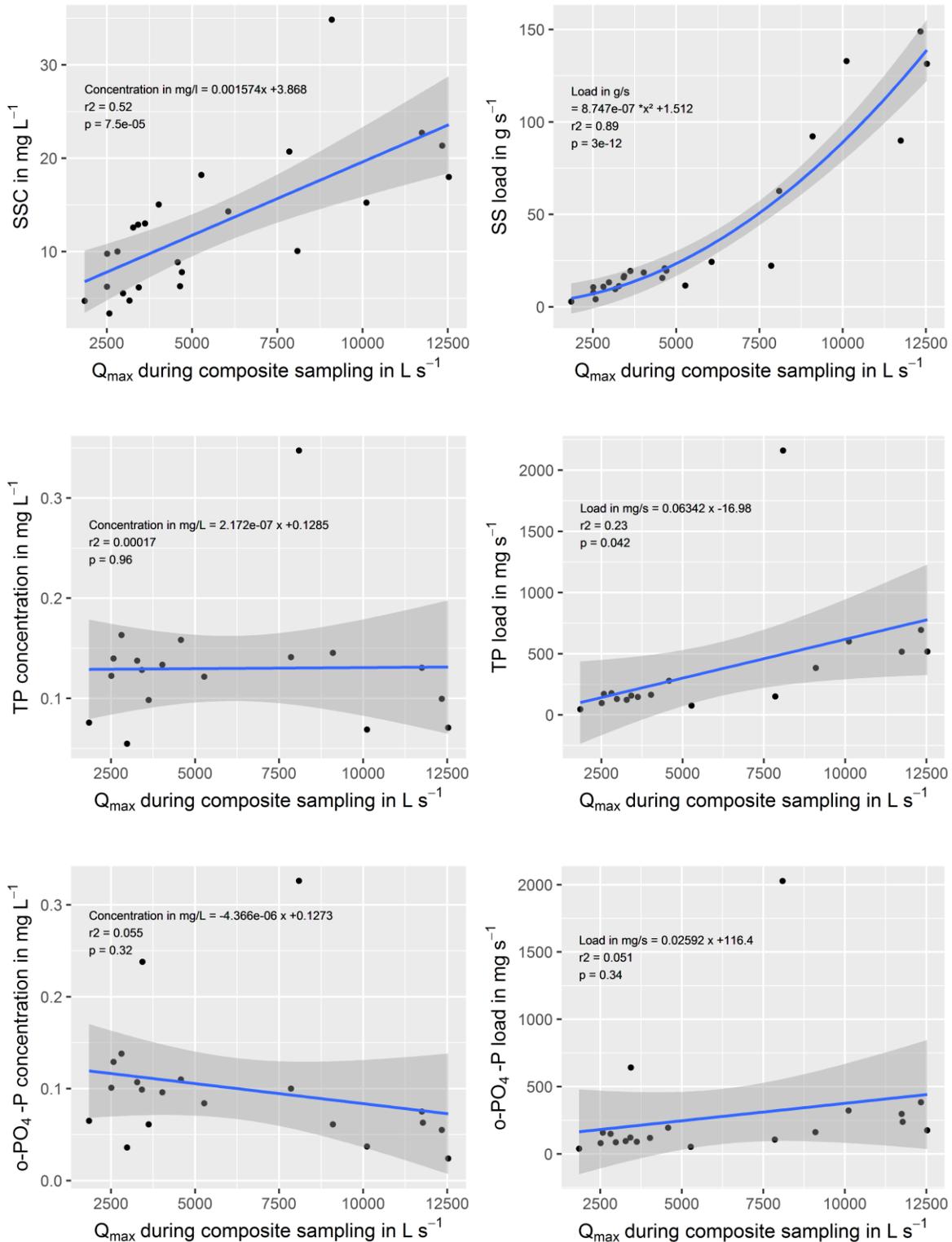
Figure 25: Adapted rating curves (blue) for suspended sediment (SS), total phosphorous (TP) and orthophosphate (o-PO$_4$-P) at the Alb River, based on the highest discharge recorded in the sampling period (Q$_{max}$) and the corresponding event mean concentration (left column) or event mean load (right column). Curve equations, R² and p-values of regressions indicated. Grey area: 95% confidence level interval for predictions.
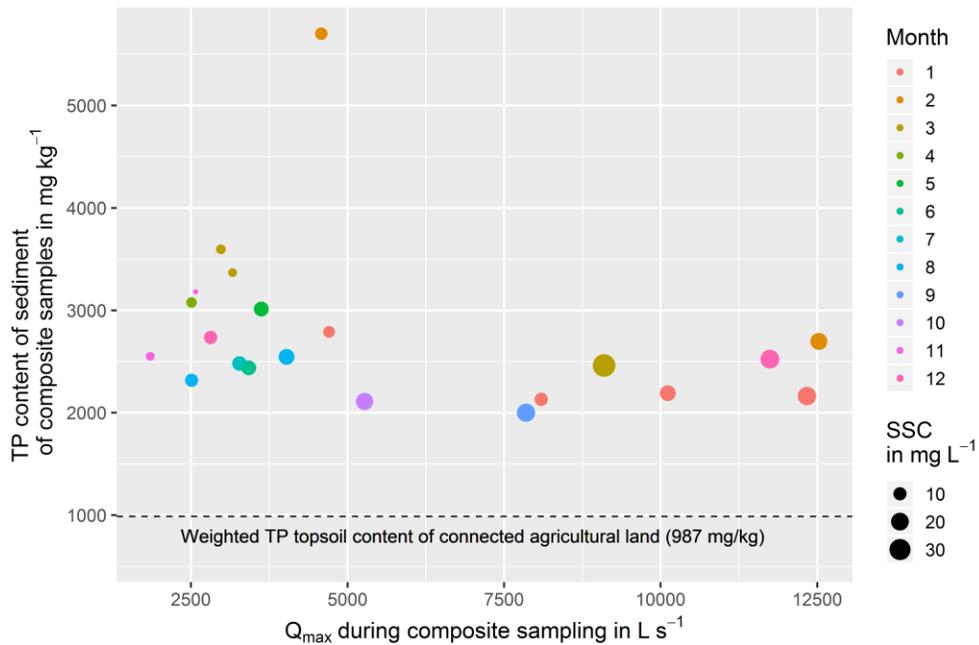
Figure 26: Total phosphorous (TP) concentrations of sediment from large volume composite samples versus the highest discharges recorded (Q_{max}) in each sampling period for the Alb River. Colors indicate the end date of sample collection from January (1) to December (12). The mean suspended sediment concentration (SSC) is represented as point size. The topsoil total phosphorous content of arable land in the catchment is indicated (dashed line).
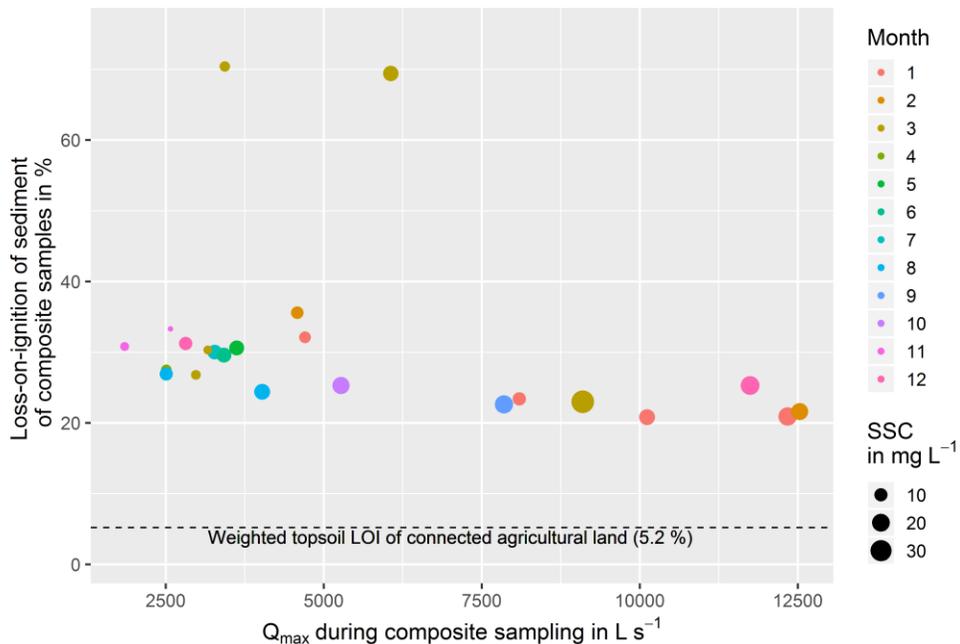
Figure 27: Loss-on-ignition (LOI) of sediment from large volume composite samples versus the highest discharges recorded hourly in each sampling period for the Alb River. Colors indicate the end date of sample collection from January (1) to December (12). The mean suspended sediment concentration (SSC) is represented as point size. The topsoil loss-on-ignition (LOI) of soil from arable land in the catchment is indicated (dashed line).

For TP and o-PO$_4$-P parameters, the criteria to flow parameter correlations ($R^2 > 0.5$; $p < 0.05$, see Section 5.7) were not met to build rating curves. Loads were calculated by multiplying mean concentrations with the flow volume. For the Alb River, only the SS rating curve was applied to discharge data with the same procedure as for the Kraichbach River (Sections 5.1.4 and 6.1.1). The cumulative flow volume of each LVS period for the Alb is 2.39 mio m³. LVS load calculations were made for the emission model period 2009-2014 and for 1976–2015 to examine long-term relevance of event magnitude to SS fluxes. For the Alb River, the difference between the capped (Figure 29) and uncapped rating curve are relatively small because of the non-polynomial and non-exponential rating curve used. Cumulative curves for the Alb have a lower variance in inclination, suggesting a less flashy transport behavior than the Kraichbach.
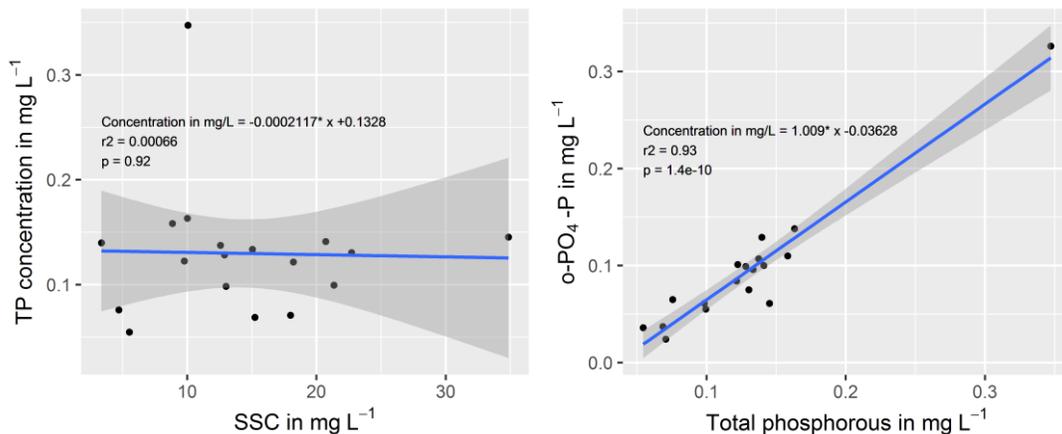
Figure 28: Mean composite sample suspended sediment concentration (SSC) and total phosphorous (TP) at the Alb River. Right: Mean composite sample TP and orthophosphate (o-PO$_4$-P) at Alb. Linear regression lines (blue), curve equations, (adjusted) R² and p-values of regressions indicated. Grey area: 95% confidence level interval for predictions.
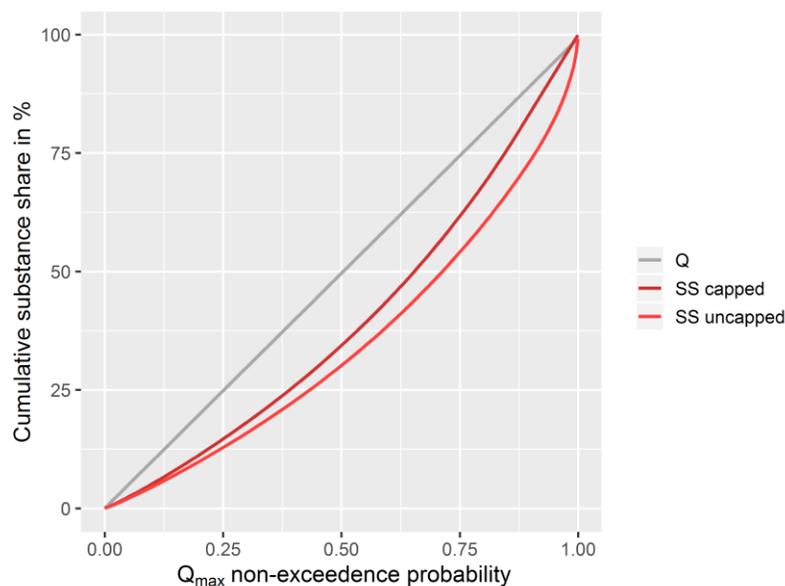
Figure 29: Cumulative curves (colored lines) of modeled long-term flux contributions of equal-volume flow intervals at the Alb River, sorted by magnitude of the highest discharge (Q$_{max}$) per event. Q: Water flux. SS capped: Application of Q$_{max}$-SSC rating curve with SSC$_{max}$ = 23.6 mg L^{-1}. SS uncapped: Same, without upper concentration boundary.

6.1.3 Passauna

At Passauna River, 33 large volume samples were collected between February 2018 and July 2019 (Table 2). At this site, several problems occurred that did not appear at the other two sites. For 27 of the 33 events, complete SS data is available. Measurements of TP and orthophosphate were not available at all until the time of writing. Furthermore, the gauge operated by the responsible agency in Paraná suffered from technical errors and did not produce a discharge signal for most of the monitoring period. The problem was circumvented by relating the water level information from the pressure sensor from a period of overlap of several months. Two regression curves were fitted to have a close approximation of flow from the pressure sensor data (Figure 30). Flows exceeding 12.5 m³ s⁻¹ could not be assessed from the pressure sensor at the LVS. Sampling periods differed between 2 and 23 days, depending on the flow conditions. Hydrological parameters were not provided by the gauge operator. However, a mean high flow of 16.4 m³ s⁻¹ was determined for the years 2010–2015, where daily discharge data is available and does not show the considerable and sudden shifts in discharge apparent in other periods. The highest flow event recorded where SS data is also available is 12.4 m³ s⁻¹ and therefore lower than the mean high flow. The lowest and highest mean SSC of a sampling period were 9.8 and 165 mg L⁻¹, respectively (Figure 71, Appendix). Lastly, a declining aliquot size for increasing flows was observed, apparent from the need of more pumping intervals per total sample. The three panels in Figure 68 in the Appendix show three sampling periods of different flow conditions. The size of the sample in all cases was 940 L, but the number of pumping intervals (orange lines) differed between 41 and 204. In Figure 30, the relationship between pumping intervals and maximum flow is shown for all 33 events, with unpredictable behavior for maximum flows exceeding 5 m³ s⁻¹. This means, that the aliquot volume is not constant, which is a requirement for discharge-proportional sampling.

Despite the uncertainties in relation with the uneven aliquot sample, adapted sediment rating curves were built. A relationship between maximum flow and suspended sediment transport that met the pre-defined criteria ($R^2 > 0.5$; $p < 0.05$, see Section 5.7) could be established only for SS loads, not concentrations. Even though the flow was recorded every minute by the pressure sensor, rating curves based on daily discharge measurements were created (Figure 31). This is due to the fact that longer time series of flow measurements are only available as mean daily discharges. In order to assess substance fluxes for longer time periods, the regression model must therefore be constructed using data with the same temporal resolution as the input data used later. The resulting cumulative flux curve for an assessment based on the Q_{max}-load rating curve and 2010–2015 daily mean discharges can be inspected in Figure 32. Unlike the cumulative curves presented for the Kraichbach, this curve is flattening towards the highest discharges. This can be observed because multi-day discharge events in the high range have a greater cumulative volume than the $0.94\ 10^6$ m³ cumulative volume the LVS was programmed to sample. During the flux modeling, these events are therefore split into several events of short duration, counterintuitively resulting in small mass gains per high flow event.
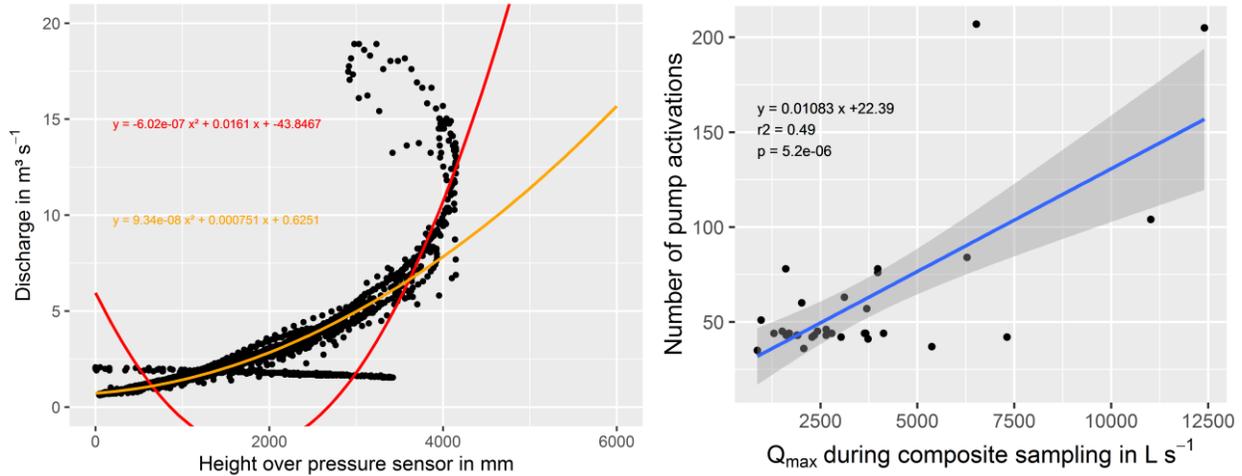
Figure 30: Left: Relationship between the large volume sampler pressure sensor and discharge data recorded at the official gauging station BR277 of Passauna River. For measured heights between 0–3600 mm, the orange curve is applied. For heights > 3600 mm, the red curve is applied. Curve equations indicated. Right: Number of pump activations per sampling event versus the highest hourly discharge (Q_{max}) in the sampling interval. Linear regression line (blue), equation with R^2 and p-value indicated. Grey area: 95% confidence level interval for predictions.
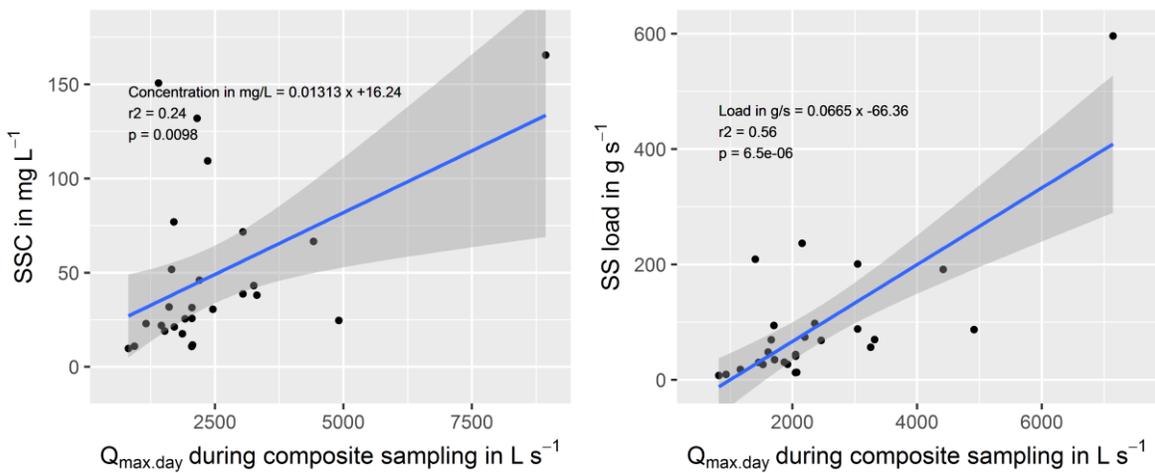
Figure 31: Adapted rating curves (blue) for suspended sediment (SS) at Passauna River, based on the highest discharge recorded in the sampling period ($Q_{max.day}$) and the corresponding SSC. Curve equations, (adjusted) R^2 and p-values of regressions indicated. Left: Concentration as target variable. Right: Load (concentration x discharge) as target variable. Grey area: 95% confidence level interval for predictions.
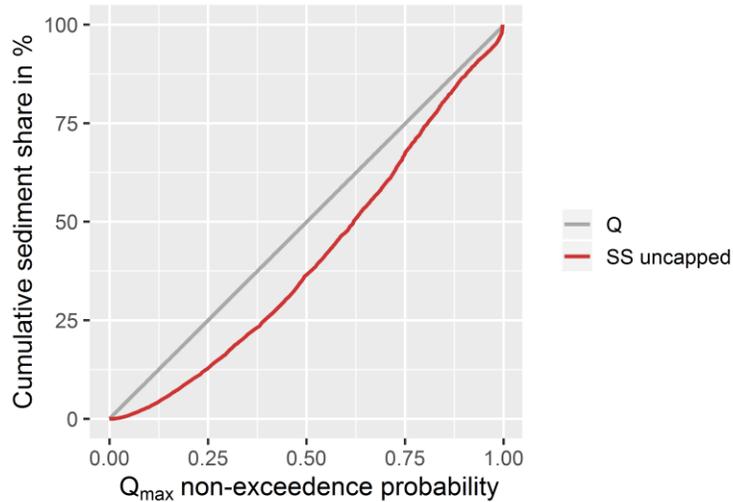
Figure 32: Cumulative curves of modeled long-term flux contributions of equal-volume flow intervals at Passauna River, sorted by magnitude of the highest discharge (Q_{max}) per event (colored lines). Q: Water flux; SS: Suspended sediment flux. Application of Q_{max}-SSC rating curve without upper concentration boundary.

6.1.4 SYNOPSIS

Event mean concentrations and loads as established by large volume sampling were related to the highest flow in the respective monitoring period to construct adapted rating curves. Rating curves with a coefficient of determination of $R^2 > 0.5$ at a significance level of $p > 0.05$ were deemed acceptable for subsequent predictions (Section 5.7). Concentration-based assessments were preferred over load-based assessments. These criteria are met at the Kraichbach for SS and TP, at the Alb for SS and at Passauna for SS (load-based and daily discharges). The remaining parameters were deemed unrelated to flow. At Passauna, technical problems were noted. TP is covariant to SS at the Kraichbach and to orthophosphate at the Alb. SSC ranges at Kraichbach exceeds the range at the Alb more than twenty-fold. Sediment contents of TP and LOI at Kraichbach are roughly half of the Alb sediment and show a maximum for moderate events. Long-term flux contribution estimates of the highest events differ with respect to the inclinations of the underlying rating curves in their capped and uncapped versions.

6.2 GRAB SAMPLING RESULTS

Rating curves were created from manual grab sampling (Kraichbach; SSC, TP and o-PO$_4$-P) and water quality data from the official monitoring program (Alb; SSC and orthophosphate). Water quality data was coupled with the closest measured hourly mean discharge.

6.2.1 KRAICHBACH

Rating curves for the Kraichbach are shown in Figure 33. For SSC, a second-order polynomial showed the best fit with an adjusted coefficient of determination of $R^2 = 0.87$ at a p-value close to zero. For TP, a linear relationship was established. Both rating curves rely on less than six measurements for discharges greater than 1.5 m³ s^{-1}, with a long-term mean flow of 1.1 m³ s^{-1}. This can be attributed to the very dry summer of 2018, when most of the sampling took place. For orthophosphate, no relationship between flow and concentration was found which complies with the criteria ($R^2 > 0.5$; $p < 0.05$, see Section 5.7). Figure 34 shows cumulative flux curves for discharge, suspended sediment and phosphorous. Note that unlike the corresponding plots for LVS data, the water volume flux is deviating from the 1:1 line. This can be attributed to the phenomenon explained in Section 1.3; that higher flows contribute disproportionally to water fluxes. In the case of the Kraichbach for instance, 25% of the mean annual cumulative discharge can be attributed to the highest 12% of discharges. For TP, capped and uncapped curves are almost identical due to the linear nature of the rating curve as opposed to the polynomial curve used for the LVS data. Contrarily, capped and uncapped suspended sediment curves are similar to the corresponding curves in Figure 24.
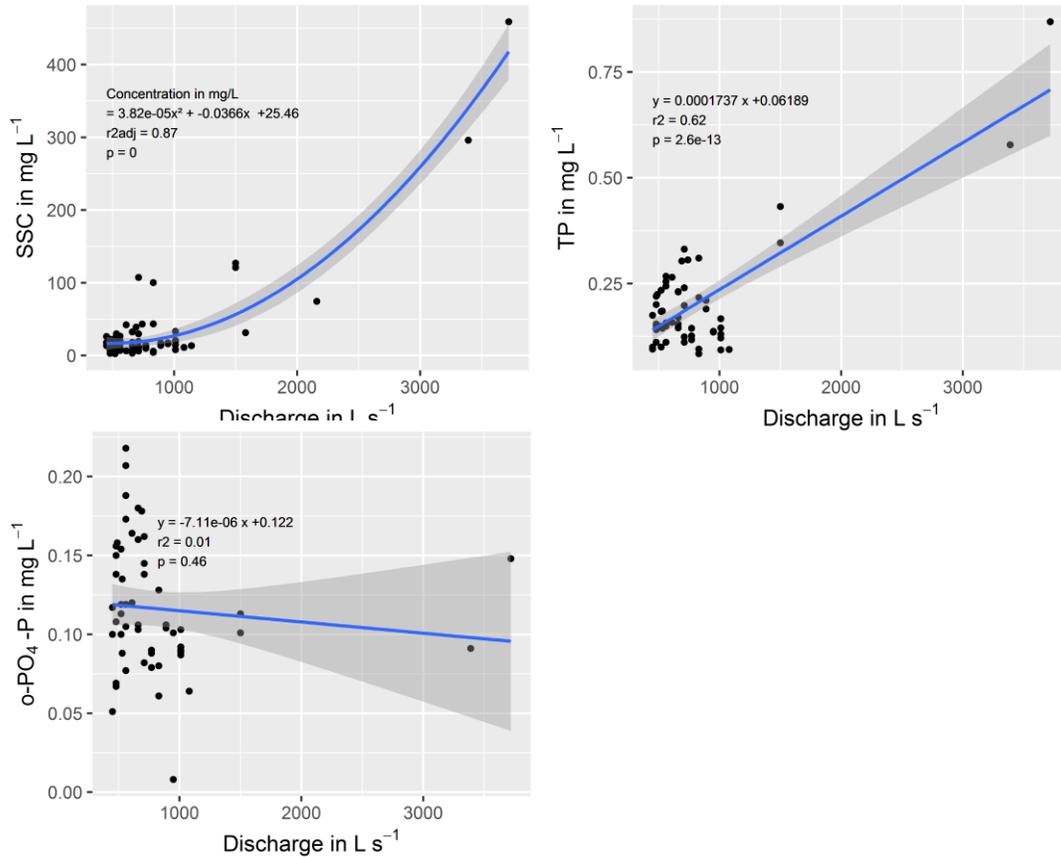
Figure 33: Grab sampling based rating curves (blue) for suspended sediment concentration (SSC), total phosphorous (TP) and orthophosphate (o-PO$_4$-P) for the Kraichbach River. Regression equations, (adjusted) R^2 and p-values indicated. Grey area: 95% confidence level interval for predictions.
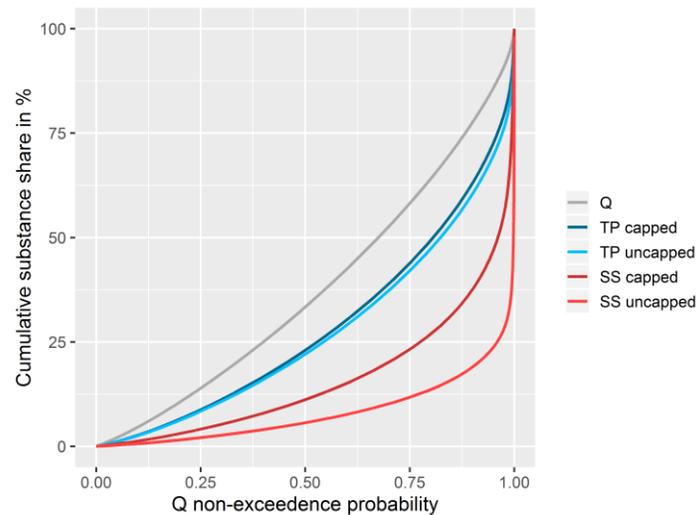

Figure 34: Cumulative curves of modeled long-term (1976–2018) fluxes of water (Q), suspended sediment (SS) and total phosphorous (TP) for the Kraichbach River, sorted by discharge magnitude (colored lines). Rating curves with ('capped') and without ('uncapped') the highest measured value as the upper concentration boundary were applied.

6.2.2 ALB

Rating curves based on official monitoring data for SSC and o-PO$_4$-P for the Alb River are shown in Figure 35. No meaningful relationships between flow and concentration were found that comply with the defined criteria ($R^2 > 0.5$; $p < 0.05$, see Section 5.7). In order to calculate mean annual fluxes, mean concentrations were simply multiplied with annual flows (Section 6.4.3). Flux scaling after Hilden (2003) was performed for orthophosphate for the years 2009–2014 to compare to emission model results from the same period. SSC measurements were only sporadically available in this period, and no TP measurements were performed.
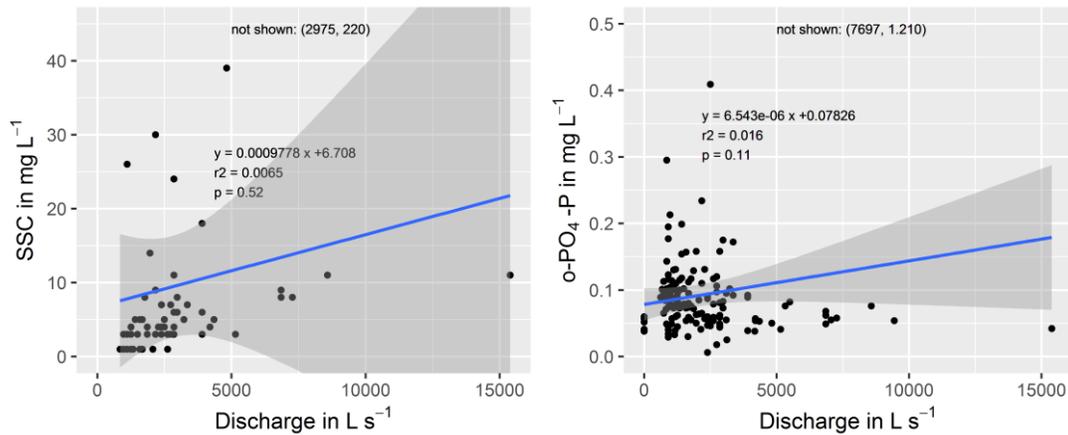
Figure 35: Left: Linear suspended sediment (SSC) rating curve (blue) for the Alb River based on grab samples and discharge. Right: Same, for orthophosphate-phosphorous (o-PO$_4$-P). Regression equations, R^2 and p-values indicated. Grey area: 95% confidence level interval for predictions.

6.2.3 PASSAUNA

At Passauna River, there are SSC measurements from one intensively measured high-flow event from October 2018 available. The highest concentration of ~750 mg L^{-1} was measured before the peak discharge was reached (Figure 36). Three simple stage-dependent linear rating curves were built following the clockwise dilution hysteresis observed during the event. The stage classification of the hydrographs for quarter hourly measurements can be seen in Figure 37 (data 2016–2018), for daily time steps in Figure 38 (data 2010–2015). For flows exceeding the highest event flow (~15 m³ s^{-1}), a constant concentration of 385 mg L^{-1} is presumed. Although there is no overlapping period available allowing direct comparison, daily mean discharge measurements appear to be able to represent some of the dynamics recorded with higher-frequency data. It must be noted however that the highest event discharge observed as daily mean values must be lower than the same event value observed with higher-frequency data. The validity of the application of a high-frequency rating curve on low-frequency data remains doubtful. Nevertheless, fluxes were computed and the two resulting cumulative curves for long-term suspended sediment fluxes are shown in Figure 39. They are similar and suggest that 75% of suspended sediment flux can be attributed to the highest 12% and 17% of discharges, respectively. This supports the notion that the general pattern of high-frequency discharge measurements is comparable to daily mean measurements.
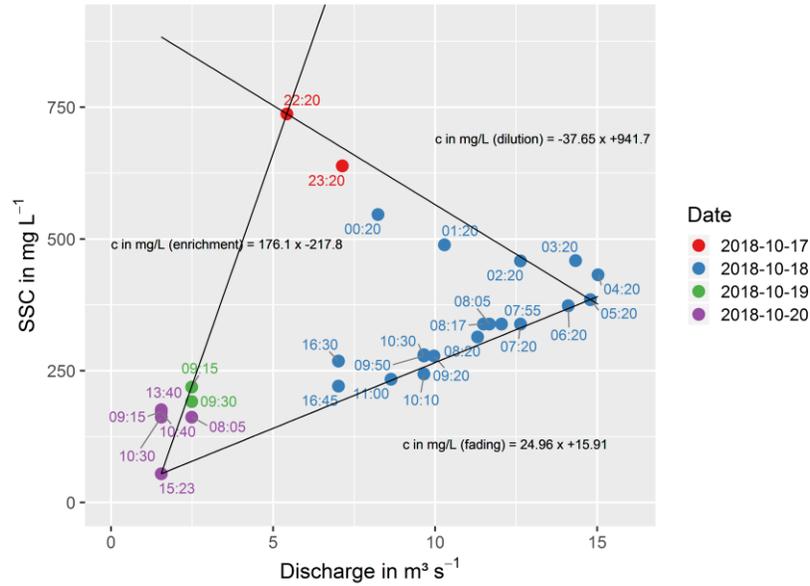
Figure 36: Relationship between suspended sediment concentration (SSC) and discharge measured frequently during a high-flow event in October 2018 at Passauna River. Colors indicate date, numbers the time of sampling. Three simple linear rating curves depending on flow stage are indicated (black lines). Curve equations are indicated. Data courtesy of the MuDaK-WRM project team (MuDak-WRM 2019).
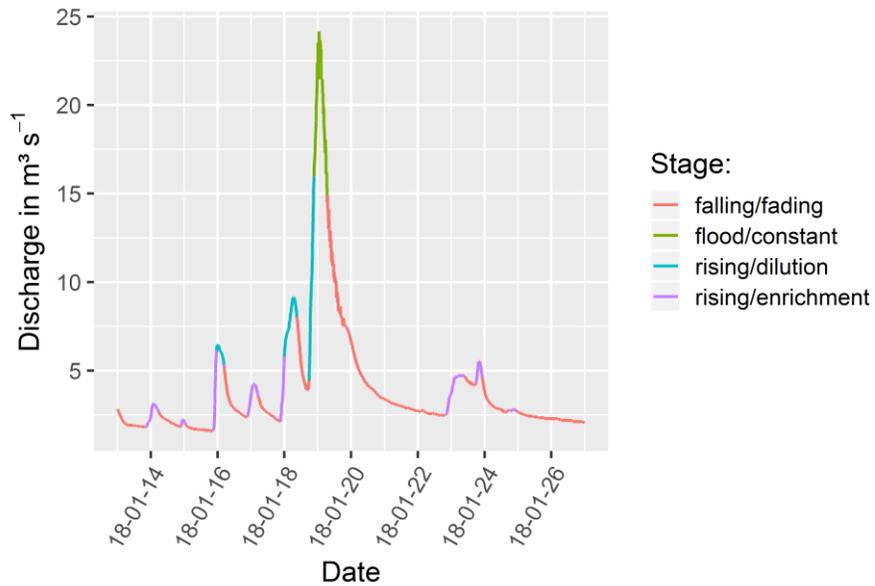
Figure 37: Example for the classification of flow stage for the application of stage-dependent rating curves of the Passauna River hydrograph (colored line) based on quarter-hourly discharge data available from 2016-2018. Colors indicate the stage.
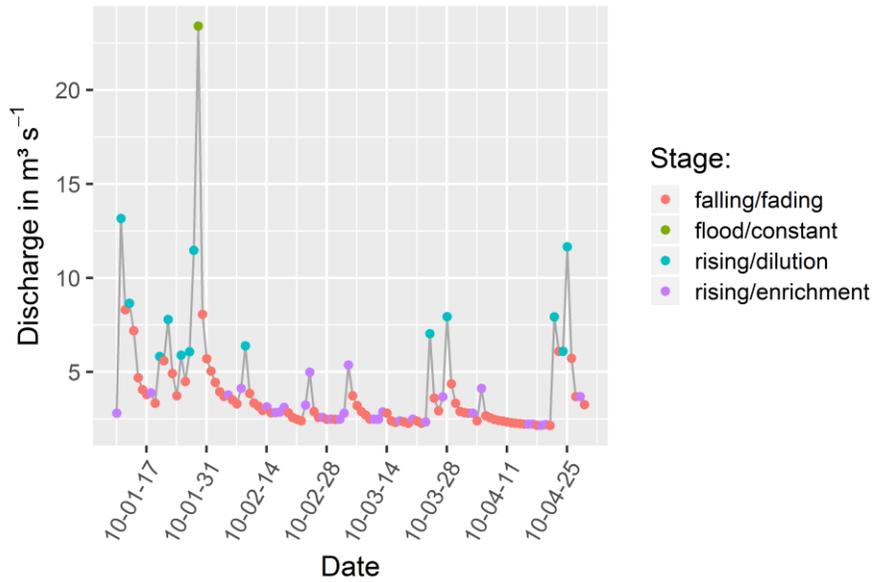
Figure 38: Example for the classification of flow stage for the application of stage-dependent rating curves of the Passauna River hydrograph (grey line) based on daily discharge data available from 2010-2015. Colored circles indicate the stage.
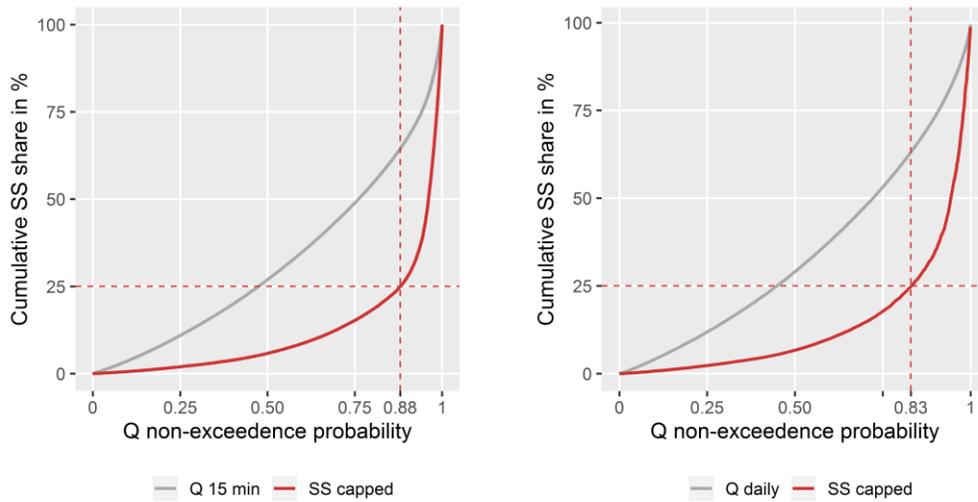
Figure 39: Cumulative curves (colored lines) of modeled long-term suspended sediment (SS capped) and water (Q) fluxes of Passauna River, sorted by discharge magnitude. Application of Q-SSC rating curve to quarter-horly discharge data from 2016–2018 (left) and daily discharge data from 2010–2015 (right).

6.2.4 SYNOPSIS

Conventional rating curves were established for the three catchments based on different data sets. At Kraichbach, significant relationships were found for SS (polynomial) and TP (linear). At the Alb, no relationship to flow was established, including SS. The Passauna data was restricted to the high-flow monitoring of one event, which shows a distinct hysteresis pattern and required a stage-dependent application to predict concentrations. The grab sampling rating curves for TP and SS at the Kraichbach are later used to compute load flashiness (Section 6.4.4).

6.3 EMISSION MODELLING RESULTS

Mean annual TP and orthophosphate emissions for the years 2009–2014 were modeled for the Alb and Kraichbach catchment using the pathway-oriented emission model METRIS-BW. The contributions of each emission pathway are shown in Figure 40. During the six years, modeled annual TP (orthophosphate) emissions at the Kraichbach vary between 4.3 t a^{-1} (2.6 t a^{-1}) and 5.8 t a^{-1} (3.5 t a^{-1}); at the Alb between 6.9 t a^{-1} (4.9 t a^{-1}) and 10.3 t a^{-1} (8.6 t a^{-1}) according to the wetness of the years.

For suspended sediment, the sediment input from the underlying RUSLE-based erosion model for connected arable land was extracted from METRIS-BW. Erosion from other land uses was estimated using constant soil loss factors per area and land use type. Sediment from waste water treatment plants were included by assuming an effluent SSC of 10 mg L^{-1} and the total annual effluent volume. Combined sewer overflow sediment was estimated by multiplying a mean SSC of 52 mg L^{-1} (Fuchs and Nickel 2018) by the modeled annual overflow discharge (Figure 41). According to METRIS, waste water treatment plants dominate phosphorous emissions in both catchments with sewer system as another important urban contributor. Erosion is only significant in the Kraichbach catchment. Basis for the phosphorous erosion pathway is the sediment input, which is dominated by sediment from arable land in the Kraichbach catchment according to the underlying sediment input model.
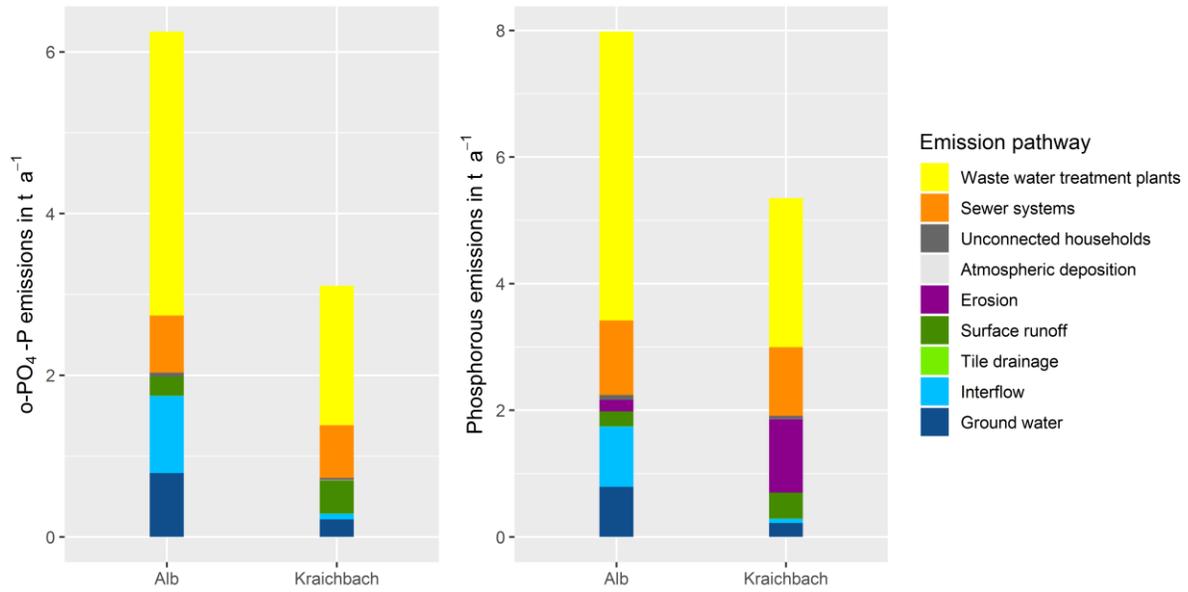
Figure 40: Modeled mean annual pathway-specific substance emissions for Alb and Kraichbach catchments from the METRIS emission model for the years 2009–2014. o-PO$_4$-P: orthophosphate-phosphorous.
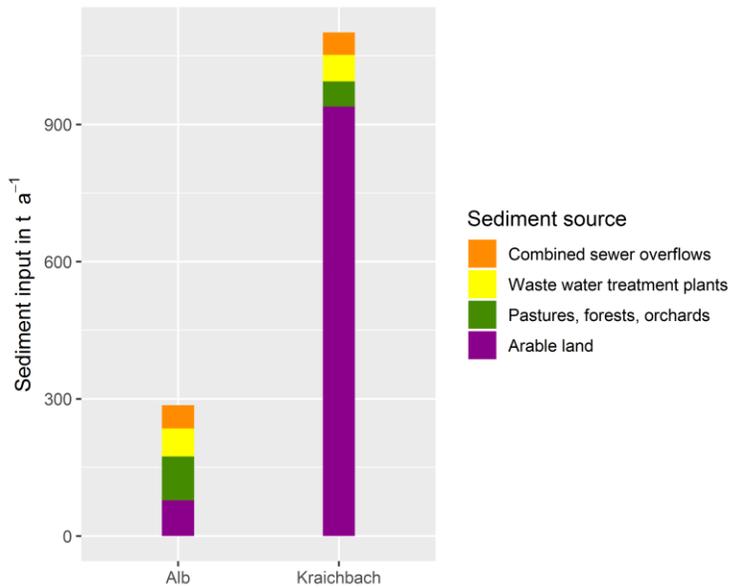
Figure 41: Modeled mean annual source-specific sediment input from the RUSLE-based sediment input modeling embedded in the METRIS emission model for Alb and Kraichbach catchments for the years 2009–2014.

6.4 SITE-SPECIFIC DISCUSSION OF MEAN ANNUAL FLUXES

In this section, mean annual flows for different calculation approaches and for all three study catchments are presented and compared to METRIS emissions (Alb and Kraichbach) and reservoir sediment stock (Passauna). Not all methods and parameters are available at all sites.

6.4.1 KRAICHBACH

Of the three catchments, the Kraichbach catchment has the most erodible soils and a the highest share of agriculture, but only about half of the annual discharge compared to Passauna and Alb. Suspended sediment and associated total phosphorous fluxes are driven by high flows. Estimates vary widely by a factor of up to four, depending on the method and period of interest.

Mean Flux Comparison

For the Kraichbach River, there are considerable differences between the flux estimates for different calculation methods. Three patterns can be identified. Firstly, flux assessments based on discharge-proportional sampling are generally higher than for assessments of the same parameter and period based on grab sampling. The only exception are TP fluxes for the year 2009–2014, where the capped grab sampling rating curve provides slightly higher results than the capped LVS rating curve. This can be linked back to the capping of the curves, which was independently done for both rating curves and depends on the highest available measurement in each data set.

Secondly, capped rating curves provide –by design– lower flux estimates than the corresponding uncapped curves. Differences between capped and uncapped versions are largest for suspended sediment (Figure 42). Uncapped rating curves generate flux estimates about twice as high as the capped versions. This can be attributed to the polynomial nature of the rating curves. For TP, the differences are less pronounced (Figure 43), with the uncapped long-term LVS calculation standing out as a very high estimate. Again, these flux results reflect the inclination of the underlying rating curves in combination with the amplitude of the discharges they were applied to. Orthophosphate fluxes (Figure 44) are unaffected by this, because they were not assessed by means of rating curves. Their differences are small and purely attributed to the changes in mean annual discharge.

Lastly, METRIS emissions are underestimating the modeled fluxes in all measurement approaches. Suspended sediment shows the greatest offset, followed by TP and orthophosphate. This means that pathways associated with sediment transport must be responsible to a higher degree for the offset of TP than other pathways. The erosion pathway in METRIS is only responsible for the emission of TP, not for the dissolved fraction of phosphorous (Section 5.4). Deficiencies in the erosion pathway are therefore a plausible cause for a large share of the offset.
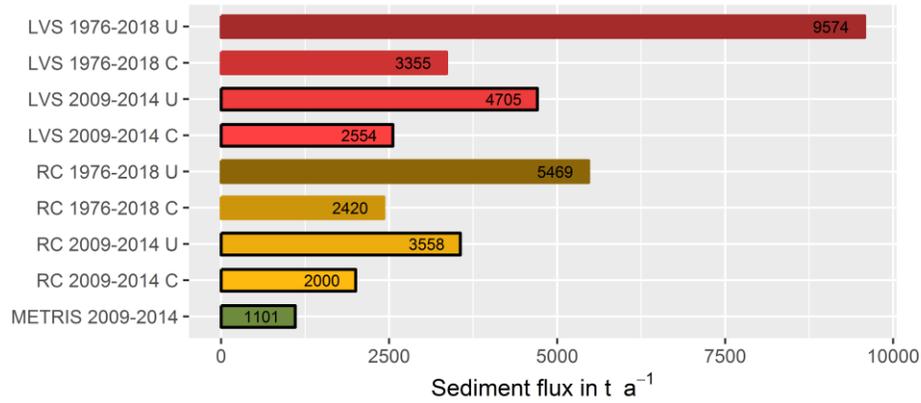
Figure 42: Mean annual suspended sediment fluxes/emissions for different periods and calculation approaches for the Kraichbach River. LVS: Large Volume Sampler approach, RC: rating curve from grab samples, METRIS: modelled emission, U: uncapped, C: capped.
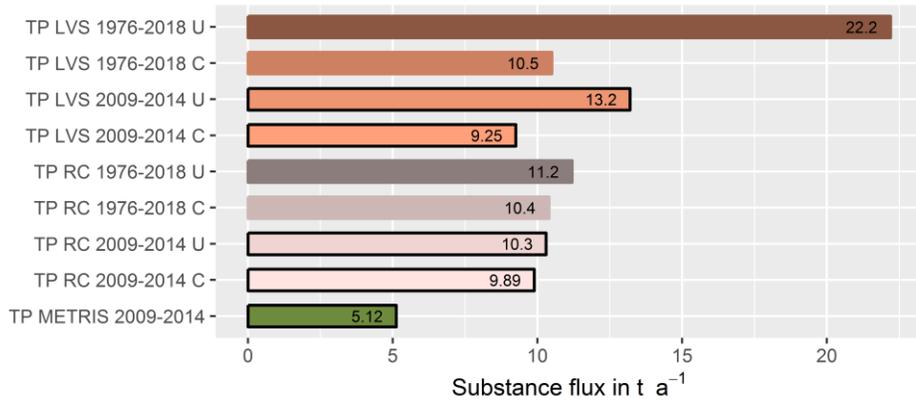
Figure 43: Mean annual total phosphorous (TP) fluxes/emissions for different periods and calculation approaches for the Kraichbach River. LVS: Large Volume Sampler approach, RC: rating curve from grab samples, METRIS: modelled emission, U: uncapped, C: capped.
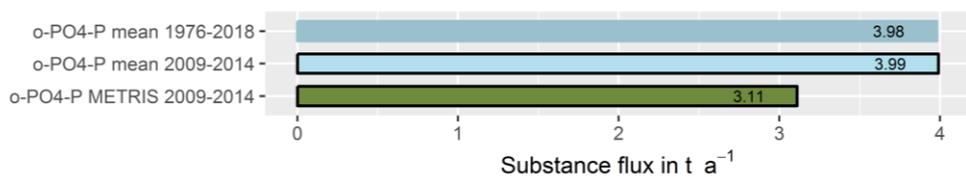
Figure 44: Mean annual orthophosphate-phosphorous (o-PO$_4$-P) fluxes/emissions for different periods and calculation approaches for the Kraichbach River. Mean: mean concentration x cumulative annual discharge, METRIS: modelled emission.

Capped and Uncapped Rating Curves

At the Kraichbach, differences between capped and uncapped rating curves are particularly large and their applicability needs to be discussed. For discharges greater than the highest one with corresponding SSC measurements, a further increase of concentrations is assumed in the uncapped rating curves. If these curves are applied to extreme discharges, very high concentrations are modeled, especially when the curves are non-linear. The year 2002 brought severe floods to Central Europe, where the Elbe at Magdeburg exported more than a third of its regular annual sediment flux within 5 days (FGG Elbe 2014). The Kraichbach was also affected, although not during the great Elbe and Danube floods in August, but in March and May. When applying the uncapped LVS rating curve to the long-term hydrograph, two consecutive events of a total duration of 36 hours in March 2002 with a 26 m^3 s^{-1} peak deliver 95,000 tons of sediment (and ~190 tons of phosphorous) or 21% of the total of 42 years of monitoring (TP: 20%). This event therefore explains a large part of the 'hockey stick' increase in the cumulative curve seen in Figure 24. Peak concentrations would be >70,000 mg L^1 (TP: 139 mg L^{-1}) for the first 13 hours of the event. Regarding SSC, these peak concentration are values that can be reached in rivers; for instance the Huanghe, which is draining an extensive loess area showed an annual mean SSC of 25,000 mg L^{-1} in the years 1950–2010 (Milliman and Farnsworth 2013: 157). During extreme conditions like the ones for the Kraichbach at the end of March 2002 (~100 year flood), these values are therefore not unrealistic per se. For comparison, Fuchs *et al.* (2004: 183) found concentrations of around 40,000 mg L^{-1} (TP: 35 mg L^{-1}) at a discharge of just 5 m^3 s^{-1}. However, the modeled result is still far out of the calibration range (three times the highest observed discharge, 80 times the highest observed concentration) and therefore carries uncertainties.

Uncapped rating curves are conflicting with the hypothetical model assumption, that from a certain water level, concentrations would not further increase due to limits in sediment availability. There are however indications that in the Kraichbach, sediment availability is not a limiting factor. First of all, the loess layer is very thick, reaching 15 m in certain areas (Zehe *et al.* 2001). Secondly, catchment connectivity increases, as rainfall and discharges increase, breaching blockades and creating transport pathways (Fryirs 2013). Further, in a study of the 6.3 km^2 Weiherbach sub-catchment of the gauged Kraichbach catchment (4% of the total area), high variabilities of SS and TP fluxes for individual years have been reported by Beudert (1997). For the years 1991–1994, SS (TP) fluxes from high-flow events differed between 4 t a^{-1} (0.005 t a^{-1}) and 1883 t a^{-1} (1.2 t a^{-1}). The maximum SSC was observed in the Weiherbach was ~16,000 mg L^{-1}. 1994 was a wet year with the largest events around 13.2 m^3 s^{-1}, which is close the highest value recorded in the 2009–2014 METRIS period (13.9 m^3 s^1). The Weiherbach catchment is ~90% arable land, as opposed to ~50% of the entire Kraichbach catchment. Scaling the arable land fluxes to the entire catchment by first approximation, the 1994 SS (TP) fluxes of the gauged Kraichbach catchment would amount to ~26,500 t a^{-1} (16.5 t a^{-1}). The uncapped LVS rating curve for the same year suggest ~16,800 t a^{-1} (34.7 t a^{-1}), while the capped versions amount to ~4,500 t a^{-1} (12.06 t a^{-1}). While the simple upscaling is assuming representativeness of the Weiherbach catchment in all related hydro-pedological parameters, this indicates that the uncapped LVS rating curve does indeed provide SS flux estimates in a plausible or even conservative range for discharges far out of the calibration range.

For TP, the comparison is pointing towards an over-estimation of fluxes by means of the uncapped LVS rating curve. This notion is further supported by the following considerations. For phosphorous, it

must be considered that during high-flow events, TP concentrations cannot rise much higher than determined by the TP content of the agricultural soils, because the suspended sediment in the river is dominated by eroded soils during floods. Taking the extreme example of the 2002 flood with peak discharges at 26 m³ s^{-1} with the modeled 70,000 mg L^{-1} peak SSC and a topsoil concentration of ~1000 mg kg^{-1}, the peak TP concentration would be expected to have a theoretical upper limit of 70 mg L^{-1} (Fuchs *et al.* (2004: 181) measured 35 mg L^{-1}) as the importance other sources and relative phosphorous enrichment would decrease. Again, these conditions are somewhat speculative and out of the observed value range. But the calculation shows that for TP, the uncapped LVS rating curve cannot be accurate for extreme discharges, where peak concentrations of 140 mg L^{-1} are calculated.

Overall, only an adequate number of measurements during these events can reveal the true magnitude of sediment flux. As long as such measurements are not or only sporadically available, capped and uncapped rating curves may be regarded as a corridor of fluxes, unless there is a clear indication that one or the other is biased. This corridor for the Kraichbach and suspended sediment is not very narrow, even when only taking into account the years 2009–2014. From the clues outlined above and the results presented in the event flux assessment Section 7.3 (Table 5), the uncapped LVS rating curves is considered the best available estimate for mean annual SS flux 2009–2014 (4,705 t a^{-1}). The highest modeled concentration during the March 2013 flood of 13.9 m³ s^{-1} is 7,850 mg L^{-1} and therefore plausible. For TP the mean of the uncapped grab sampling rating curve and the capped LVS rating curve is considered (9.57 t a^{-1}), because the extrapolation of the uncapped LVS rating curve leads to irrational values for the highest flows during the 2009–2014 period. For the primarily dissolved form of phosphorous, orthophosphate, there was neither a relationship to flow parameters found for the LVS data (Figure 21) nor for the grab sample data (Figure 33). The best available estimate for mean annual fluxes is the multiplication of the mean measured value with the annual flow, i.e. 3.99 t a^{-1}.

In comparison with measurements conducted for the Kraichbach gauge around the year 2000 by Fuchs *et al.* (2004), the relevance of high flow events increases for the transport of TP. In the named study, discharges < 2.5 m³ s^{-1} were responsible for 85% of TP transport. High-flow discharges > 2.5 m³ s^{-1} which occurred on 7 days per year were attributed to the remaining 15%. From the long-term data used in this thesis, discharges > 2.5 m³ s^{-1} occur on about 10 days per year and are responsible for 10% of the mean cumulative annual water flux. The rating curves selected here would suggest that those flows contribute 25% (grab sampling rating curve) and 27% (LVS rating curve) to the annual TP flux (Figure 24 and Figure 34). This can be attributed to the reduction of TP emission from the constantly emitting waste water treatment plants, which are today about 50% lower than in the year 1997 (Fuchs *et al.* 2004: 122).

Evaluation of Emission Model Results

Considering the best available estimates from above, modeled annual suspended sediment emissions are underestimated by ~77%, TP emissions by half, and orthophosphate by one fourth in the Kraichbach catchment. Modeled METRIS TP emissions are at just 31% of the values established in a similar emission model evaluation for the years 2001–2003 by Fuchs *et al.* (2004). While some of this offset can be explained by actual reduction of emissions, most prominently by the two waste water treatment plants, there are indications that the current modeled emissions are biased low. As argued above, this is an indication that METRIS underestimates an emission pathway controlled by sediment transport in the Kraichbach catchment. Three of the METRIS emission pathways consider particulate emissions: Waste water treatment plants, sewer systems and erosion (Kiemle *et al.* 2019). Treatment plant emissions are computed by multiplying mean effluent concentration with cumulative annual discharge, which are well-monitored parameters. Emissions from combined sewer overflow and storm overflows are modeled by establishing a plant-specific urban water balance using a variety of hydrological and substance-related parameters. Although validated in terms of modeled overflow concentrations and volume, this pathway can indeed carry uncertainties for individual catchments. For the Kraichbach (and Alb) catchment, the urban water balance is in a plausible range (unpublished data).

The erosion pathway however may explain a large share of the offset. The RUSLE-based sediment input model considers only sheet erosion, not rill or gully erosion and is therefore inherently biased low, which is acknowledged by the modelers (GALF 2018). Linear erosion has been observed in the Kraichbach catchment (Fuchs *et al.* 2004). Furthermore, as shown in Section 6.1.1, high-flow events are associated with high concentration of suspended sediment, which additionally has properties found in the agricultural soils of the catchment. The notion that the erosion pathway is underestimating sediment emissions was further confirmed during a field visit after the second-largest event (< two-year-flood) in the Kraichbach catchment in December 2018. Great quantities of erosion material from farmed slopes could be found in drainage channels, on roads and settled on grass strips between arable land and streams. This was also observed for similar events by Fuchs *et al.* (2004). Several locations which were deemed 'unconnected' by the sediment input model showed sediment plumes directed towards the nearest streams (Figure 77, Appendix). Small drainage channels which are passing under roads cannot be considered as a connectivity feature in the sediment input model, because they do not appear on the digital elevation model. To the contrary, roads and grass strips are generally regarded as barriers. In the case of high-intensity events at least, the barrier function of these features was obviously no longer applicable. Fryirs (2013) argues, that barriers and buffers can contribute only to disconnectivity, as long as their breaching capacity is maintained. Events of sufficient stream power can therefore induce sudden sediment transport. Combining these observations, it can be concluded that the share of connected agricultural areas is assessed too low for high event magnitudes (Figure 76, Appendix). This demonstrates, that connectivity is better described dynamically magnitude-dependent, e.g. at the scale of events.

Although there are uncertainties with respect to the contribution of different land uses and pathways, arable land is identified as the dominant and underestimated contributor. It can be helpful to use simplifications in order to carve out benefits and limitations of measurement and model approaches: If we assume that a low biased agricultural sediment input model is fully responsible for the 'missing' ~3,600 t a^{-1} of suspended sediment (measured LVS flux − modeled emission, Figure 42) and apply the

~1,000 mg kg^{-1} top soil phosphorous content, this could explain (without enrichment) about ~3.6 t a^{-1} of non-orthophosphate phosphorous missing currently from the emission model and narrow the TP measurement-model gap to ~1 t a^{-1}, which is also very close to the offset between modeled and measured orthophosphate and could therefore be explained by other pathways. The annual TP emissions via erosion would then amount to ~4.8 t a^{-1}, which is also much closer to the ~5.9 t a^{-1} assessed by Fuchs *et al.* (2004) in a previous study based on the USLE and different input data. This way, the sediment underestimation and the underestimation for particulate phosphorous could largely solved simultaneously. This would mean that the current sediment input from arable land is less than one fourth of the true value, and that not only the share of connected arable land of ~50% is too low, but also the specific sediment immission from these areas.

Further, these considerations are made under the assumption that there is no net bank or bed erosion in the Kraichbach, which could also explain some of the difference between suspended sediment measurement and model results. Fuchs *et al.* (2004) conclude however that there is neither bed erosion nor retention in the gauged Kraichbach catchment. Sediment fingerprinting approaches to distinguish eroded soil from eroded river bed material are not available for the Kraichbach yet. While there are sections in the Kraichbach showing steep banks as an indication for erosion, thick sediment layers as indicators for depositive conditions can also be found. Further, the high-flow samples of SSC show TP contents close to those found in fertilized arable soils. Origin bank and bed material is expected to have much lower contents. Overall, net bank and bed erosion are unlikely to explain the missing ~3,600 t a^{-1}.

It has further been argued by the sediment input modelers that due to the implementation of soil conservation measures in the Kraichbach during the last years and an subsequent reductions of sediment input the Kraichbach currently is in a state of hydraulic imbalance and therefore releasing and redistributing bedded sediment within the channel, which is measured in the LVS and explains the large gap between sediment input model and sediment flux estimates (GALF 2018). This is however contrary to the assessment by Fuchs *et al.* (2004) that because of the hydromorphological characteristics there is a low potential for in-channel sediment retention in the Kraichbach. In addition, it has been at least qualitatively shown by the field visits that great quantities of fresh erosion material is reaching the Kraichbach during extreme events. Again, it is plausible that connectivity and sediment yield are underestimated for high-flow events by the input model.

Independently of this, orthophosphate emissions are also underestimated by ~1 t a^{-1} compared to the measurements. Because of the lack of concentration correlation to flow parameters for LVS and grab samples as well as many potential causes, the responsible pathway is more difficult to assess. The contribution of orthophosphate to the total phosphorous concentration varies between 13 and 92% for low flows. This means, that there is an inconstant source of orthophosphate for low flows. This noteworthy, because low flows should be mostly controlled by rather constant sources like groundwater (METRIS model concentration: 0.039 mg L^{-1}, KUP 2018), interflow 1.5 x C_{GW} = 0.059 mg L^{-1}), and waste water treatment plants (0.39 mg L^{-1}). However, there is a wide variety of low flow orthophosphate concentrations (Figure 19, Figure 33), most of them above the management goal of 0.07 mg L^{-1} (BMJV 2016), with a mean around 0.10 mg L^{-1}. This mean value is in good agreement with a assumed low-flow mixing of 50% ground water, 36% interflow and 14% treatment plant effluent, which can be regarded as a fair estimate of the runoff composition. The treatment plant effluent share was taken from a regionalized

discharge data set (LUBW 2016). Due to their relatively high effluent concentration, treatment plants determine the orthophosphate base load in the Kraichbach. They are however well monitored and carry little uncertainties with respect to annual emissions, which make them an unlikely error source (Fuchs *et al.* 2004). Elevated runoff concentrations from recently fertilized agricultural fields are a possible explanation for the fluctuating orthophosphate levels during moderate flows, which would point towards an underestimation of emissions via surface runoff. When establishing the surface runoff approach in METRIS, fertilization was identified as a key driver for surface runoff concentrations (Kiemle *et al.* 2019). Data about timing and amounts of fertilizer application however was is not available area-wide, which is the reason it could not be used for regionalization. Another source may be phosphorous release from bedded sediment. It was shown that during hot summer months and dry winters, oxygen levels in the free water of the Kraichbach can fall near or below the 2 mg L^{-1} threshold needed to sustain the reduction of iron in the water column and subsequently facilitate the release of phosphates, which was validated in incubation experiments (unpublished data). There is a cluster of low-flow events with high orthophosphate concentrations which could be affected by this process (Figure 20). In other words, phosphorous which is emitted in particulate form may deposit in the channel and then measured (years) later in dissolved form. If this effect is indeed relevant in the Kraichbach, model validation would be complicated because of the time lag between emission and measurement. Measurements of the phosphorous release are needed for clarification.

Summary

In summary, there is ample evidence that the large offset between measurement and emission model for both, suspended sediment and particulate phosphorous, can be traced back to the underlying sediment input model. This poses questions about the scale issues in hydrology, connected to the claim of the input model to represent mean conditions, which are discussed in Section 8.4. The remaining measurement-model gap can partly be attributed to underestimated emissions of orthophosphate during low flow conditions, with ambiguous sources. Time-lagged phosphorous emission from the sediment stock may complicate model validation. Sediment fingerprinting and phosphorous concentration measurements along the sediment-water interface are needed.

6.4.2 ALB

Although it has the steepest terrain and the greatest discharge, the Alb catchment has the lowest suspended sediment flux. Unlike the other two catchments, high flows are not extremely overrepresented in their relevance of sediment transport. The main reason for this is the lack of open soil in the catchment.

Mean Flux Comparison

At the Alb River, there are fewer calculation methods available than for the Kraichbach, because of the lack of monitoring data or meaningful concentration-discharge relationships. Variability for suspended sediment flux calculations is lower (Figure 45). This is due to the non-exponential relationship and the generally low SSC in the Alb. Unlike the grab sampling rating curves, the LVS rating curve revealed that there is a positive correlation of SSC with event magnitude. This raises the question whether the regression can be extrapolated to discharges outside the calibration range. In the period 2009–2014, for which a comparative value is needed, the maximum discharge at the Alb was at ~62 m³ s⁻¹, which is a 20 year flood (at the same day as the nearby Kraichbach). For these extreme conditions, the uncapped LVS rating curve suggests a mean event concentration of ~100 mg L⁻¹. Even though there are no concentration measurements available for such high discharges, this moderate value is not a clear indication that the uncapped LVS rating curve should be rejected. Therefore, the uncapped LVS rating curve value of ~1,132 t a⁻¹ will be used as a reference. The flux scaling approach after Hilden (2003) was also used for orthophosphate and is overestimating the actual flux. The reason is one individual orthophosphate measurement out of the 13 values available in the year 2011 showing 1.2 mg L⁻¹ measured during a flow of 7.7 m³ s⁻¹. Following the flux scaling method, this value, which is three times higher than the second-highest in 155 measurements, leads to a 2011 flux of 28 t a⁻¹. This underlines the sensitivity of flux scaling towards outliers and makes the scaled flux questionable.
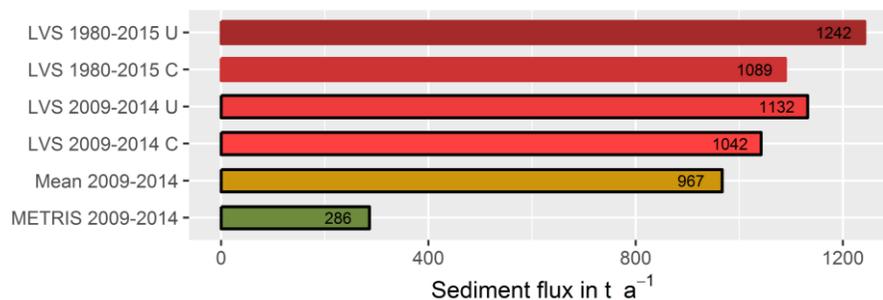
Figure 45: Mean annual suspended sediment fluxes/emissions for different periods and calculation approaches for the Alb River. LVS: Large Volume Sampler approach, mean: mean concentration x cumulative annual discharge, METRIS: modelled emission, U: uncapped, C: capped.
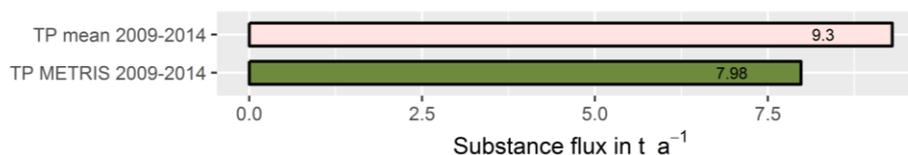
Figure 46: Mean annual total phosphorous (TP) flux/emission for the Alb River. Mean: mean concentration x cumulative annual discharge, METRIS: modelled emission.
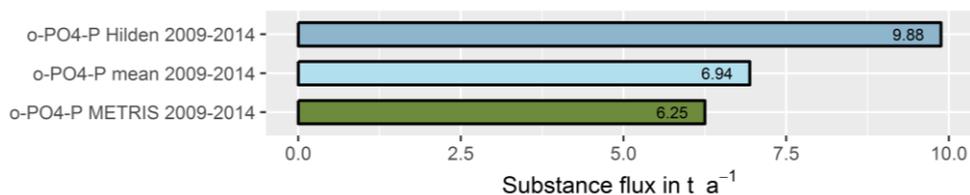
Figure 47: Mean annual total phosphorous (TP) flux/emission for the Alb River. Hilden: flux scaling method, Mean: mean concentration x cumulative annual discharge, METRIS: modelled emissions.

Evaluation of Emission Model Results

Modeled TP and orthophosphate emissions are in good agreement with the measurements (Figure 46 and Figure 47), both slightly underestimating mean annual fluxes. Sediment fluxes are however underestimated by ~850 t a^{-1} or 75%. Similar to the Kraichbach, this is suggesting deficiencies in the sediment input model. In the Alb, this has less dramatic consequences for the modeled TP emissions, because TP transport in the Alb is controlled by the dissolved fraction.

In the Alb catchment with just ~300 ha of cropland and currently just ~78 t a^{-1} of modeled sediment input from these areas, it is unlikely that the missing ~850 t a^{-1} are solely from underestimations in the sediment input model for arable land. If that was the case and we assume, like for the Kraichbach, that this material has a phosphorous content equal to the mean topsoil concentration of agricultural lands in the catchment of 987 mg kg (LTZ 2015), this could explain about 0.84 t a^{-1} of total phosphorous, or more half of the difference between measurement and model (Figure 46). The remaining difference (0.5 t a^{-1}) would be close to the measurement-model gap of orthophosphate (0.7 t a^{-1}), which may be explained by underestimations in other emission pathways like surface runoff. However, the Alb catchment is dominated by deciduous forests, where the mean phosphorous topsoil content is just 410 mg kg. Topsoil is defined here as the A horizon without leaf litter. For forests, no area-specific sediment input model is available. Instead a constant of 5.95 kg ha^{-1} a^{-1} was applied, which is a mean value for Saxony and is likely to be an underestimation because of the higher inclinations in the Alb catchment. If we consider forests to be the dominant factor in the underestimation and attribute the 850 t a^{-1} fully to forests soils with a phosphorous content of 410 mg kg^{-1}, the sediment input from forests could explain an additional 0.35 t a^{-1} total phosphorous. The comparatively small amounts of sediment in LVS samples are however richer in phosphorous and organic matter than forest or agricultural soils in the catchment (Figure 26 and Figure 27). A reason for this could be relative enrichment of the material during the transport into the stream network. Another explanation is that the sediment exported from forests, especially in areas with steep inclinations, contains higher shares of leaf litter and other organic material than is represented by the (mineral) topsoil concentrations used above. Overall, the underestimation of suspended sediment emissions can explain the total phosphorous underestimation, although there is uncertainty about the origin and the magnitude. The general agreement between emission model and measurement for the Alb is fair and a deviation of 10–15% may have numerous reasons. It must be considered that METRIS was not calibrated to these specific catchments and conditions, but has to provide mean long-term estimates for a model region of more than 35,000 km². In addition, fluxes are compared to emissions with no accounting for in-stream processes.

6.4.3   PASSAUNA

The Passauna River in Brazil generally shows the greatest sediment flux, followed by the Kraichbach while the Alb exports relatively small amounts of sediment. This result was expected, even though mean flow and inclination are smaller than the Alb River and the share of arable land is smaller than in the Kraichbach catchment. Due to the sub-tropical climate, the Passauna catchment receives frequent heavy thunderstorms with high rainfall intensities and high erosive power, after which the rivers in the region are colored brightly brown from high sediment loads. The Passauna sediment regime is dominated by high-flow events, which is evident from the cumulative curves from grab samples (Figure 39).

For the Passauna catchment, only suspended sediment fluxes are available (Figure 48). There are large differences between the calculation approaches. Especially the annual input calculated from the sediment stock stands out and is more than six times higher than the second-highest estimate. Mean annual fluxes from the stage-dependent rating curve derived from the single high-flow event are similar for quarter hourly discharge measurements (2016–2018) and daily data (2010–2015).
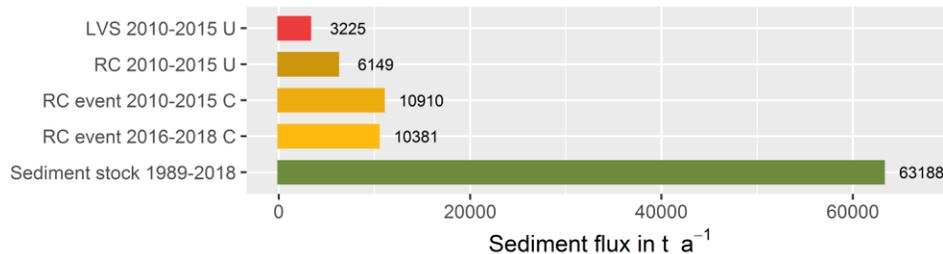
Figure 48: Mean annual suspended sediment fluxes/emissions for different periods and calculation approaches for the Passauna River. LVS: Large Volume Sampler rating curve, RC: rating curve from grab samples, RC event: stage-dependent rating curve from single high-flow event, Sediment stock: mean annual input calculated from reservoir sediment stock, corrected for catchment size and autochthonous matter, U: uncapped, C: capped.
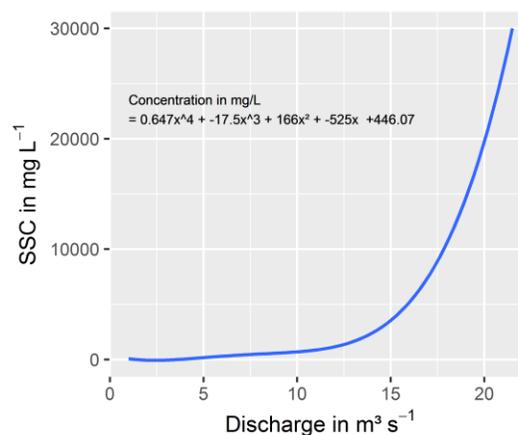
Figure 49: Hypothetical example of a sediment rating curve (blue curve) for the Passauna River which could explain the current sediment stock. SSC: Suspended sediment concentration. Curve equation indicated.

Rating Curve and Technical Issues in Passauna

The rating curve derived from the intensely monitored single high-flow event in October 2018 was applied to daily and quarter hourly discharges of different time periods, which produce very similar results in terms of mean annual fluxes. Under the premise that the hysteresis of the single event is representative, the application of the stage-dependent rating curves on data with a similar temporal resolution is legitimate. The similarity of the results can be regarded as an indication that with a daily resolution, the discharge dynamics of Passauna can still be sufficiently recorded, although the validity of this approach remains questionable. The highest SSC recorded during the event in October 2018 was 750 mg L^{-1}, the first measurement on the rising limb of the hydrograph. It remains unclear if in the hours before the SSC was even higher, and if the event itself can be regarded as a blueprint for all high-flow events in the catchment. Because there are no further measurements available, a constant SSC of 385 mg L^{-1} had to be assumed for flows exceeding the highest observed. Compared to the >800 mg L^{-1} reached at the Kraichbach as event mean concentration for several days, this is a moderate value. This could be attributed to the differences in pedology, with very silty and easily erodible loess soils in the Kraichbach catchment and clay soils in the Passauna catchment (MuDak-WRM 2019). In any case, the event rating curve method is capped, i.e. it has an upper boundary, which makes the assessment a conservative one. In this light it is surprising that the uncapped long-term rating curve taken from Rauen *et al.* (2017) yields much lower mean annual fluxes when applied to the same discharge data. This long-term curve was established based on data from an extended time period and can be regarded as more robust against special characteristics of individual events. However, daily mean discharges greater than 8 m^3 s^{-1} were rarely observed and included in the rating curve. The highest SSC included to create the curve was just 294 mg L^{-1} and therefore much lower than the SSC maximum of the October 2018 event. Even though it is the highest SSC measurement available, it is unlikely that the October 2018 concentration peak around 750 mg L^{-1} actually was the highest SSC within the last 30 years. Much more plausible is the notion that even though the data base is large, the long-term rating curve is biased low for high-flow events, which are rarely sampled.

The LVS is designed to capture the 'hot moments' of sediment transport. There is ample evidence that the LVS was not operating in the correct manner at Passauna. High-flow events generally required many more pumping intervals than low-flows (Figure 68). A peristaltic pump was used instead of the submerged pumps used in Germany. It is now assumed that rising flow velocities induce an underpressure at the sampling pipe in the water, outmatching the suction capacity of the pump. This means, that the aliquot volume is not constant, which is a requirement for discharge-proportional sampling. Further, it means that high-flows are underrepresented in the sampling scheme, which is contrary to the process understanding and sampler design. After the LVS is filled eventually, the concentration in the tank does not represent event mean concentrations, but is to a larger (and unknown) share determined by low-flow conditions. This is reflected in the low SSC measured in the LVS even for high-flow events (maximum of just 165 mg L^{-1}), which does not comply with SSCs measured from grab samples. In addition, the LVS pressure sensor used to determine discharges could only be used for flows below 13 m^3 s^{-1}. Overall, the LVS approach for Passauna is biased low to an unknown degree, which renders the data useless, at least for the study of high-flow events for which the device was designed.

Sediment Stock as Long-Term Validation

From the sediment stock in Passauna Reservoir indicating an annual flux of ~63,000 tons a^{-1} at the gauging station, all other approaches appear to create far too low estimates. Several explanations for this large offset can be considered. A large part of the sediment stock of the reservoir may have formed in a relatively short periods of extremely high sediment availability in the catchment, periods in which even low or moderate flows were heavily laden with sediment. These time spans must have been short in order to not be represented by the long-term rating curve. For example, one of these periods could have taken place shortly after reservoir construction when the soil was disturbed from tree removal in the inundation area. This would mean that much of the reservoir's today sediment is a colluvium of the direct proximity or the bed slopes themselves. Another example of short-term high sediment availability could be attributed to activities at the upstream mine. While sediment availability is certainly inconstant, such hypothetical events are unlikely to explain most of the 3.1×10^6 tons of bedded sediment in the reservoir. Furthermore, the sediment stock assessment could be a gross overestimation. This can be considered very unlikely, because the assessment is based on several multi-method campaigns combining echo sounding, traditional sampling and penetrometry at high spatial resolution. In addition, a RUSLE-based erosion model seems to confirm sediment stock data (unpublished), also indicating a constant or episodic sediment immission. Based on the sediment stock data, the annual storage volume loss of Passauna Reservoir is estimated 7% since construction or 0.23% a^{-1} (unpublished). This value is neither uncommon for reservoirs nor particularly high; it about half of the mean siltation rate of Central European reservoirs (0.51% a^{-1}). It must be emphasized that the 'regular range' stretches across at least two orders of magnitude from 0.02% a^{-1} to > 2% a^{-1} (Rahmani *et al.* 2018). Nevertheless, we can assume that the measured sediment stock can be used to derive a realistic sediment flux which is least influenced by the hydrology of individual years.

The most likely explanation for the tremendous differences of grab sample rating curve and reservoir stock data is that a large part of the sediment stock has formed during frequent, but short and poorly monitored high-flow events. While low and moderate flows are well covered by the sampling strategies and rating curves, there is still a lack of data for the highest of discharges. In other words, the proper rating curves would be steeper towards extreme discharges, and the contribution of extreme events to annual fluxes would be even higher than shown in the existing cumulative curves, where the highest 12% of discharges transport ~75% of the annual flux, the highest 1% of discharges ~15% of the annual sediment flux (Figure 39). These numbers may already appear to be reflect a flashy transport behavior, but roughly the same ratios has been determined for much larger catchments with much lower rainfall intensities such as the Lahn, the Main, Mosel and the Neckar (Asselman 2000). In very broad study including long-term data from 1,314 catchments of different sizes in the entire USA, the highest 1% of discharges was attributed to a mean of 30% of the total sediment flux (Gonzalez-Hidalgo *et al.* 2010). Horowitz (2013) found similarly asymmetrical relationships even for the much larger Mississippi at Thebes, Illinois, and states that there suspended sediment transport generally is dominated by high-flow events, regardless of river basin area. Catchment size however determines concentration times and durations of high-flow events, which are shorter and therefore more difficult to monitor in smaller catchments, especially when applying calendar sampling.

The Passauna catchment is relatively small, has sufficient relief to allow for considerable movement of sediment, receives high-intensity rainfall and has a relatively large share of agriculture, all of which are conditions favoring high load flashiness. In summary, there are various indications that long-term rating curve is biased low for high-flow events, and that the rating curve based on just one event cannot be extrapolated to higher flows in a representative manner. If we conclude that the ~63,000 t a^{-1} derived from the sediment stock are a valid assessment, this would correspond to a mean SSC of about 1000 mg L^{-1}. This twice the average found in 760 rivers around the world, which showed an amplitude of 1–45,000 mg L^{-1} (Milliman and Farnsworth 2013: 27). A mean value however does not give insight in the amplitude of the actual in-river concentration, which is of particular interest here. Under the assumption that the grab sampling rating curve flux gives reasonable results for the lowest 99% of daily mean discharges (i.e. < 8 m³s^{-1}), the highest 1% of discharges would be responsible for ~54,000 t a^{-1} or 85% of the total mean annual sediment flux (mean of USA: 30%; Gonzalez-Hidalgo *et al.* 2010), transported with just 5.2% of the mean annual cumulative discharge. The mean SSC during these rare events (i.e. < 4 days per year) would then be ~16,700 mg L^{-1}. Considering that many rivers in the world have higher *annual* SSC averages (Milliman and Farnsworth 2013: 27), this is a plausible flood concentration for a small sub-tropical, high-precipitation and agricultural catchment. For comparison, Minella *et al.* (2009) determined a peak concentration of 11,420 mg L^{-1} during a high-flow event in a small 1.19 km² catchment in Rio Grande do Sul, i.e. in a neighboring region with similar properties. This supports the view that the sediment stock data is adequate and that other attempts to compute fluxes failed, including the discharge proportional sampling of the LVS. Figure 49 shows how a sediment rating curve able to explain the current sediment stock could potentially look like.

6.4.4 Synopsis and Site Comparison

From the available flux estimates for each catchment, specific yields, flashiness indices and (if applicable) offsets towards modeled emissions were computed (Table 3). Even though Passauna has the smallest catchment, it also has the highest annual cumulative discharge. It is also most stormflow-driven, indicated by the highest FFI, followed by the mountainous and shallow-soiled Alb and the strongly buffering Kraichbach, where thick loess layers dampen the response of the hydrograph to precipitation (Zehe *et al.* 2001). Few studies published FFIs based on hourly measurements, so the FFI based on daily measurements (Baker *et al.* 2004) was also computed for better comparability. With respect to this and to their catchment size class, the Kraichbach and Alb would be considered flow-stable, and also Passauna would be classified as rather stable. Their FFI increases however strongly when hourly values are applied, indicating that the daily resolution is underestimating flow oscillations and intra-day variability (Baker *et al.* 2004). In comparison to the limited values available for FFI based on hourly data and under consideration of their size, Passauna can be regarded as flashy, the Alb moderately flashy and the Kraichbach as stable (Deelstra and Iital 2008). The sediment yield in the Passauna catchment based on the best available flux estimate is two orders of magnitude higher than in the Alb catchment. Sediment yields are indeed highly variable around the world. All established values are plausible, and the Passauna sediment yield is not particularly extreme neither on a global perspective (Walling 1996: 15; Milliman and Farnsworth 2013: 31) nor in comparable geographical regions (Minella *et al.* 2009).

Next to FFI, the LFI and the ratio between the two indices, the nLFI were established based on the

instantaneous loads calculated from grab sampling rating curves. LVS rating curves cannot be used to compute instantaneous loads at the same temporal resolution of the discharge measurement, therefore they cannot be used to compute load flashiness. The grab sampling rating curves are biased low, therefore are the LFI and the nLFI conservative estimates. If no relationship between flow and substance concentration was established from grab sampling rating curves, the nLFI is considered to be one.

Fluxes, yields, indices and model offsets were are all derived from essentially the same data. Nonetheless, some helpful comparisons can be made. The substance flux contribution of the highest discharges contributing 10% of the mean annual total flow (Q_{10}) varies widely between 10% (i.e. constant concentration for all discharges) and 95% (extreme SSC for extreme flows at Passauna). The mean number of days per year during which this flux takes place varies between ~5 at Passauna and ~10 at the Kraichbach. While the first parameter illustrates the importance of sampling these events, the second points up the importance of sampling timing. Both aspects are fused into the LFI and nLFI, which express the relevance of high flow sampling with respect to the catchment-specific fluxes in the order SS Passauna > SS Kraichbach > TP Kraichbach, whereas the remaining parameters showed no relationship to discharge in grab sampling rating curves. If only flow flashiness had been used for this assessment, high-flow monitoring would be falsely regarded as more relevant to correctly assess substance fluxes of the Alb than at the Kraichbach. For SS, load flashiness is also covariant with sediment yields across the three catchments, indicating that for the presented case, general erosivity is related to the shortness and magnitude of the events and subsequently to the relevance of monitoring high flows. This must not necessarily be the case in other catchments. Especially large loess catchments can be expected to have high yields at a low load flashiness. Regarding TP, Alb and Kraichbach have comparable yields. The relevance of the flow-independent dissolved form of TP, orthophosphate, is however far greater in the Alb catchment. Therefore is the relevance of high-flow monitoring to assess phosphorous is greater at the Kraichbach, but there is no relationship to absolute yields. The flux offset to the emission model is inversely related to the LFI and nLFI at the Kraichbach, while there is no relationship at the Alb.

In summary, the magnitude of particle-associated substance yields from diffusive sources control the relevance of high-flow monitoring but must be regarded together with flow flashiness. The proposed LFI and nLFI can be used to generalize the relationship between shortness and importance of high flows and can therefore be used to prioritize the implementation of high-flow sampling schemes for each catchment and substance. In other words, they expresses the probability that flux assessments based on calendar sampling are accurate: For nLFI > 1, a high LFI signals that calendar sampling is going to be biased, with a clear tendency towards a low bias for particle-associated substances (Coynel *et al.* 2004). Each catchment and substance is evaluated on its own. If the prioritization is to be made with respect to the impact of the catchment within a larger area, total fluxes have to be also considered. It is plausible that among similar-sized catchments, the sediment load flashiness is indeed systematically related to sediment yield. Broader studies comparing a variety of catchments and boundary conditions are needed.

Table 3: Summary of variables comparing substance fluxes and transport behavior in the three study catchments. Q: Discharge; SS: Suspended Sediment; TP: Total Phosphorous; o-PO_4-P: Orthophosphate-phosphorous; FFI: Flow Flashiness Index; LFI: Load Flashiness Index; nLFI: normalized Load Flashiness Index; Q_{10}: The highest flows contributing 10% of the annual total discharge; C: concentration; offset: difference between measurement and emission model. Fluxes, yields and offset are computed from best available estimate, while Flashiness Indices were computed from (low-biased) grab sampling rating curves only. *= Estimates based on hypothetical rating curve.

Parameter	Variable	Kraichbach	Alb	Passauna
	Catchment size in km²	161	149	81
Q	Mean total annual runoff in mio m³ a⁻¹	34.3	73. 8	62.1
	Specific runoff in mm a⁻¹	213	495	767
	Duration of Q_{10} in days per year	10.2	6.6	5.4
	FFI (hourly), dimensionless	0.49	0.76	1.62
	FFI (daily), dimensionless	0.14	0.16	0.24
SS	flux in t a⁻¹	4,705	1,132	63,188
	yield in t km⁻² a⁻¹	29.2	7.60	780
	flux contribution of Q_{10}	**66%**	**10%**	**95%***
	Flow-weighted C in mg L⁻¹	137	15.3	1017
	Model offset	-77%	-75%	-
	LFI, dimensionless	2.66	0.76	11.0*
	nLFI, dimensionless	5.43	1	6.79*
TP	flux in t a⁻¹	9.57	9.3	-
	yield in kg km⁻² a⁻¹	59	62	-
	flux contribution of Q_{10}	**26%**	**10%**	-
	Flow-weighted C in mg L⁻¹	0.28	0.13	-
	Model offset	-44%	-14%	-
	LFI, dimensionless	1.08	0.76	-
	nLFI, dimensionless	2.20	1	-
o-PO_4-P	flux in t a⁻¹	3.99	6.94	-
	yield in kg km⁻² a⁻¹	25	47	-
	flux contribution of Q_{10}	**10%**	**10%**	-
	Flow-weighted C in mg L⁻¹	0.12	0.09	-
	Model offset	-22%	-10%	-
	LFI, dimensionless	0.49	0.76	-
	nLFI, dimensionless	1	1	-

7 EVENT FLUXES

This chapter presents the development of the proximal sensing setup (Section 7.1) and shows how turbidity measurements were related to SSC measurements (Section 7.2). Subsequently, event fluxes are calculated from a variety of methods and compared to the measured fluxes (Section 7.3).

7.1 PROXIMAL SENSING

PLS models were built from reflectance spectra and corresponding reference data for mesocosm experiments, campaign-based measurements at Passauna Reservoir and from the mean of six automated reflectance measurements performed between 11:00 and 12:00 CET at Kraichbach River.

7.1.1 MESOCOSM RESULTS[3]

Mesocosm spectra were measured in two sets, during sunny and overcast conditions, respectively. Like also observed by Lodhi *et al.* (1998), an increase of SSC leads to an increase of reflectance most pronounced in the wavelength regions between 450–550 nm and especially 700-900 nm, the latter with a local maximum at around 815 nm (Figure 50), the increase in reflectance through the increase in SSC affects the entire measured wavelength range, whereas the near infrared shows the most uniform increase. The illumination conditions were found to have a strong impact on the shape of the reflectance curves. Reflectance from diffusive sunlight shows an about three times steeper increase between 400 and 500 nm. For direct sunlight, the spectra appear flatter, even though downwelling irradiance was measured several times during the acquisitions. This negatively impacted the quality of the prediction. However, a regression model built from a combined sunny and overcast data set (Table 4), is still well within the performance range of linear or second-order regressions used in similar setups (Lodhi *et al.* 1998). From the experiences made, frequent measurements of the downwelling irradiance are recommended, ideally one for each measurement of the upwelling radiance. By subsetting the data set, PLS calibrations using smaller SSC ranges (0–200 mg L^{-1} and 0–600 mg L^{-1}) were made (Table 4) in order to test if high goodness-of-fit-measures could be attributed to the broad value range. The high predictive ability was maintained for all value ranges. Figure 51 visualizes the performance of a mesocosm PLS model. Due to the controlled environment, values are distributed without major gaps across the entire value range. Five latent vectors (components) were deemed optimal for this, as the root mean square error for predictions remains constant when adding more components. Adding components should then be avoided, as this would induce overfitting of the model (Wold *et al.* 2001). Overall, above-water reflectance measurements in the mesocosm experiments suggests that a broad range of SSC can be measured at high precision, even when measuring at a relatively low spectral resolution.

---

[3] This section partly consists of modified excerpts of a peer-reviewed journal article by the author (Wagner *et al.* 2018)
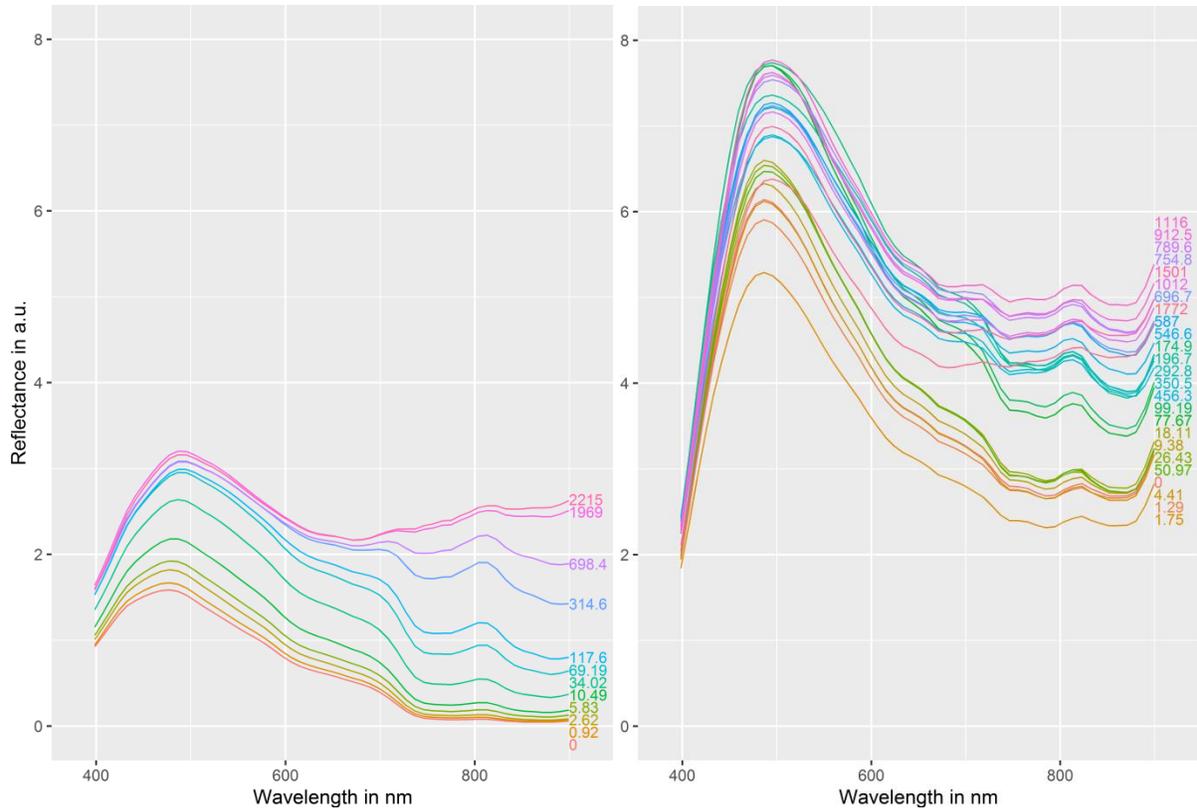
Figure 50: Reflectance spectra measured in the mesocosm during conditions with direct sunlight (left) and during overcast conditions (right). Colored numbers indicate suspended sediment concentrations in mg L^{-1} measured for the individual spectra.
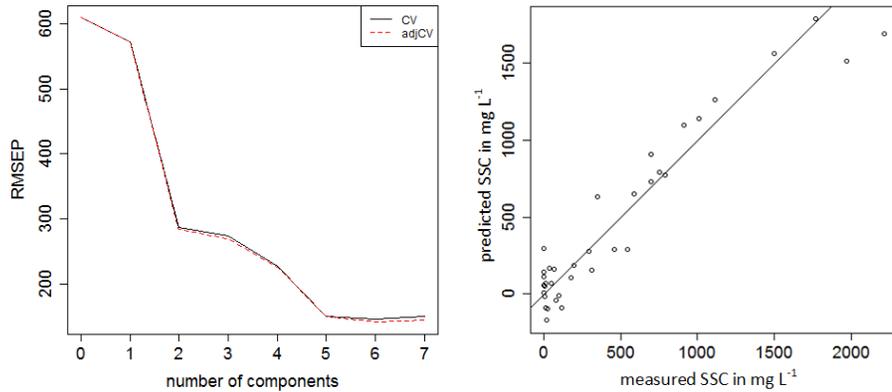
Figure 51: Left: Root mean square error of prediction (RMSEP, referring to suspended sediment concentration (SSC) in mg L^{-1}) against the number of latent vectors (components) used in cross-validated (CV) partial least squares regression SSC prediction from two sets of reflectance measurements in mesocosm experiments. Right: Measured against predicted SSC for a model with 5 latent vectors. adjCV (red line): Bias-corrected CV estimate.

7.1.2 PASSAUNA RESERVOIR RESULTS[4]

At Passauna Reservoir, the field applicability of above-water reflectance measurements was tested. Further, temporal transferability, sometimes considered a weak point of empirical regression techniques (Keller 2001: 56) was tested in five campaigns during different seasons. In the examined case, the transferability is given. The predictive abilities of the PLS regression are high for both target parameters (Table 4) using a reasonable number of ground truthing samples. The range of the values is much smaller than in the mesocosm experiments, with a highest measured SSC of 52 mg L^{-1} (TP: 0.2mg L^{-1}). Also, the values are not as evenly distributed across this range, with many samples in a low range and few samples above 20 mg L^{-1} (SSC) and 0.05 mg L^{-1} (TP), respectively (Figure 52 and Figure 53). From these figures it is also apparent that two latent vectors (components) are sufficient to obtain a reasonably low root mean square error of prediction.

Unlike rivers, the reservoir cannot be considered fully mixed. Therefore, there need to be considerations for which part of the water column the modelled concentrations are valid. One way to approach this issue is to use the attenuation depth as the lower boundary of the water column layer which contributes to the spectral measurement. It is defined by the depth from which 90% of the water-leaving radiance emerges and can be calculated from attenuation measurements. If such measurements are not available, Secchi Disk Depth (SDD) can be used as a proxy to calculate the attenuation coefficient and subsequently attenuation depth. Using the nonlinear equation for neotropical systems introduced by Padial and Thomaz (2008), and several SDD measurements, the attenuation depth in Passauna reservoir ranges from 0.12 m (SDD: 0.15 m) in the highly turbid buffer to 1.37 m (SDD: 3.7 m) near the dam. By this definition of surface, Figure 54 and Figure 55 show the resulting concentrations estimates of SS and TP based on hundreds of spectral point measurements from a moving boat. Inverse distance weighted interpolation (Shepard 1968, weighting power = 25) was performed between the points to represent the entire water surface. The underlying estimates for each point for these and other measurement days can be examined in Figure 84 and Figure 85 in the Appendix. While there are some artefacts apparent as horizontal lines, the setup provides a reasonable representation of the concentration gradient gradients from the reservoir inflow area towards deeper parts of the impoundment. At the narrow point between the shallow inflow area (mean depth of about 1.5 m) to the north and the deeper main reservoir to the south, the concentration changes rapidly. Between the 11th and the 14th of February 2017, a series of heavy rainfalls occurred in the catchment, triggering the SSC increase in the inflow area of the reservoir (Figure 83). Contrarily, the three-day-series in February 2019 (Figure 54 and Figure 55) show the subsiding influence of heavy rainfalls on the 3rd of February. From these figures, it is further apparent that TP and SSC are highly covariant. This supports the notion that the optically inactive TP is detected optically by its proxy SSC (Gholizadeh *et al.* 2016). This further indicates that TP transport in the Passauna catchment is controlled by sediment transport.

The results overall demonstrate the field applicability of the system and the transfer to total phosphorous as a target parameter. At boat speeds of up to 20 km h^{-1} with the influence of waves, sun glint and changing illumination conditions during five campaigns in two years during different seasons, SSC

---

[4] This section partly consists of modified excerpts of a peer-reviewed journal article by the author (Wagner *et al.* 2018)

results were obtained well in the accuracy range of designated turbidimeters (Rügner *et al.* 2014) and setups using empirical approaches in airborne hyperspectral measurements (Fraser 1998; Knaeps *et al.* 2010). Ship-based approaches are rarely reported. Khalid A. and Ortiz (2016) achieved similar precision for SSC detection in Lake Erie. Keller *et al.* (2018) determined turbidity from ship-based hyperspectral above-water reflectance measurements using PLS at a comparable precision as established at Passauna, using hundreds of in-situ calibration samples. Their value range was however extremely small, ranging from 0–4 NTU. They also found that for their data support vector machines are a slightly more accurate machine learning tool than PLS, indicating that testing other empirical retrieval methods could improve the estimate. Because of the very high covariance of SSC and TP in the Passauna Reservoir, there is little difference between the established PLS models. The setup for boat-based campaigns can therefore be recommended as an effective and non-obstructive method to accurately trace surface concentration gradients of SSC and TP.
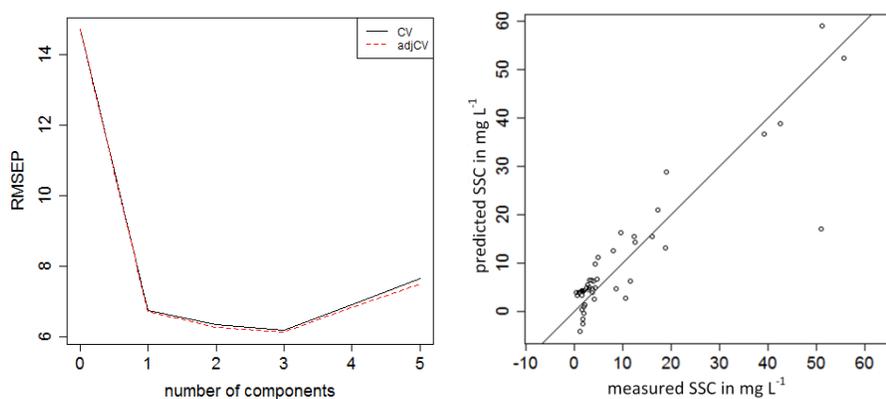
Figure 52: Left: Root mean square error of prediction (RMSEP, referring to suspended sediment concentration (SSC) in mg L^{-1}) against the number of latent vectors (components) used in cross-validated (CV, black line) partial least squares regression for SSC prediction from reflectance in reservoir measurements. adjCV (red line): Bias-corrected CV estimate. Right: Measured against predicted SSC for a model with 2 latent vectors.
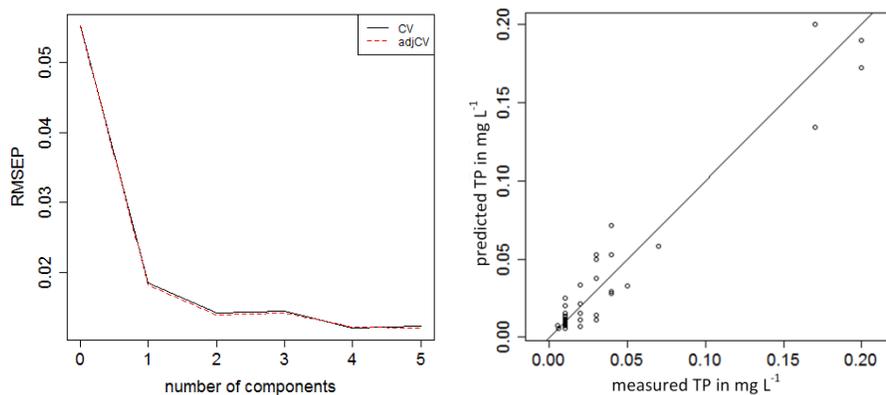
Figure 53: Left: Root mean square error of prediction (RMSEP, referring to total phosphorous (TP) in mg L^{-1}) against the number of latent vectors (components) used in cross-validated (CV, black lien) partial least squares regression for TP prediction from reflectance in in reservoir measurements. adjCV (red line): Bias-corrected CV estimate. Right: Measured against predicted TP for a model with 2 latent vectors.
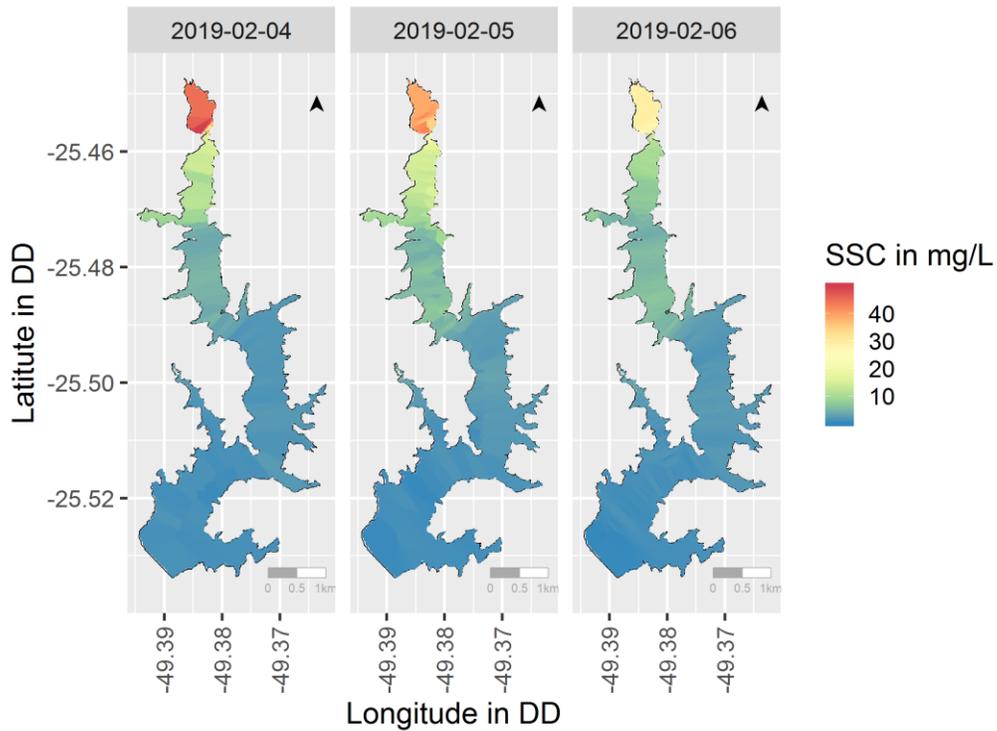
Figure 54: Inverse distance weighted interpolation of suspended sediment concentration (SSC) estimates from reflectance measrements at Passauna Reservoir on three consecutive days after heavy rainfall in February 2019. DD: Decimal degrees.
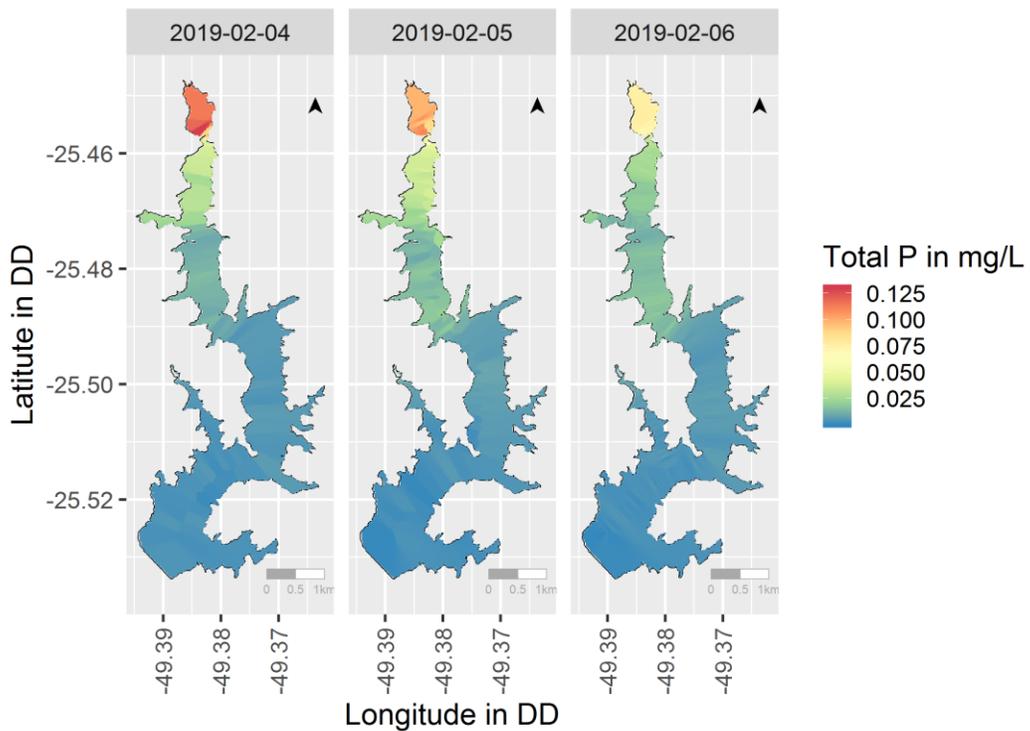
Figure 55: Inverse distance weighted interpolation of surface total phosphorous (P) estimates from reflectance measrements at Passauna Reservoir on three consecutive days after heavy rainfall in February 2019. DD: Decimal degrees.

7.1.3 KRAICHBACH MONITORING STATION RESULTS

Like for the other two spectral data sets, the performances of the selected PLS models can be examined in Figure 56 (SSC) and Figure 57 (TP). For the Kraichbach setup, the predictive ability of the PLS model in training and test for SSC is very good and moderate for TP. Both parameters reach a local minimum in terms of root mean square error of prediction for five and three latent vectors, respectively, before increasing again for more components. Reyna *et al.* (2017) report a similar behavior for near-infrared spectroscopy data and state that this could indicate overfitting and inadequate data sets. The data collective used for calibrating the model is indeed not optimal. Most values are in a low range with only few values for high concentrations. This stretches the validation plot and makes the comparison of the goodness-of-fit measures to the two other setups difficult. For TP, the validation set achieves much better goodness-of-fit measures than the calibration data set. This is a results of the uneven division of the training and test data and implies that there are insufficient data points spread across the desired measurement range. The 'RoX' spectrometer is automatically correcting for the illumination conditions for each measurement, which facilitated processing. However, there still are considerable differences in the shapes of the measured reflectance curves for similar concentrations, but different illumination conditions. Mean reflectance of low-flow measurements (discharges below 0.8 m³ s⁻¹) show a drift, which influences the SSC predictions (Figure 82, Appendix). Further, due to the very dry summer in 2018, it was not guaranteed that the river bed was not visible. This means that for model calibration, a mixture of optically deep and optically shallow water was used. Considering these difficult boundary conditions, the predictions from unsupervised reflectance measurements (Section 7.3) are credible inside the calibration range.
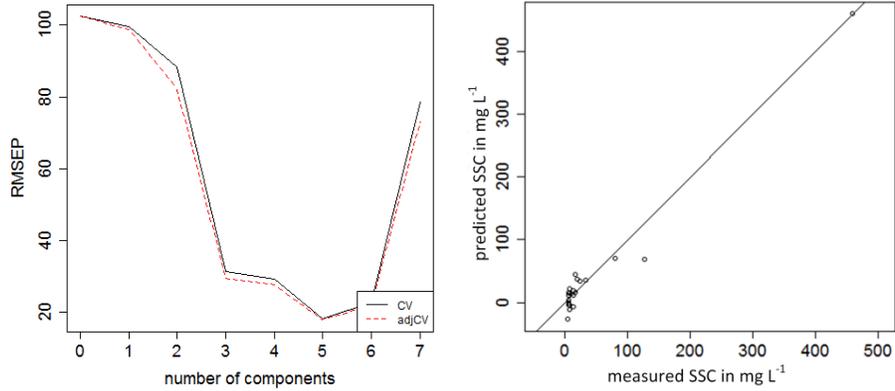
Figure 56: Root mean square error of prediction (RMSEP, referring to suspended sediment concentration (SSC) in mg L^{-1}) against the number of latent vectors (components) used in in cross-validated (CV, black line) partial least squares regression for SSC prediction from reflectance from the Kraichbach monitoring station. adjCV (red line): Bias-corrected CV estimate. Right: Measured against predicted SSC for a model with 5 latent vectors.
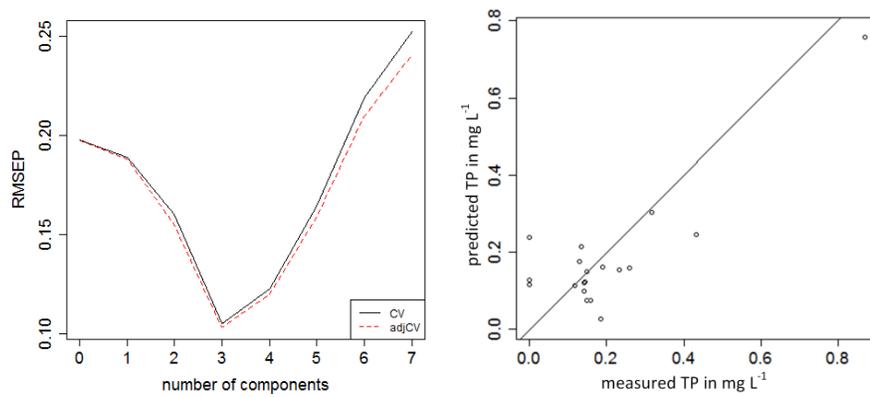
Figure 57: Root mean square error of prediction (RMSEP, referring to total phosphorous (TP) in mg L^{-1}) against the number of latent vectors (components) used in cross-validated (CV, black line) partial least squares regression for TP prediction from reflectance from the Kraichbach monitoring station. adjCV (red line): Bias-corrected CV estimate. Right: Measured against predicted TP for a model with 3 latent vectors.

7.1.4   SUMMARY OF PROXIMAL SENSING

An overview of their performance with their R² and root mean square error of prediction in given in Table 4. As expected, above-water reflectance measurements are well suited to predict surface water concentrations of suspended sediment in a broad range. This is in accordance to previous findings which highlighted the importance of the near infrared to detect suspended matter concentrations (Lodhi *et al.* 1998; Knaeps *et al.* 2010; Wu *et al.* 2014). The better the measurement setup is controlled, the better the predictions are, which can be derived from the decreasing predictive abilities for mesocosm, reservoir and river setup. Post-processing techniques like dividing by the mean, the elimination of implausibly high measurements, smoothing and the inclusion of the first derivative as a predictor variable were able to mitigate these problems.

Table 4: Mean goodness-of-fit-measures of 100-fold iterated PLS regression setups for mesocosm, Passauna Reservoir and the continuous monitoring station at Kraichbach. RP% Root mean square error of prediction, given as percentage of the maximum subtracted by the minimum measured value; LV: number of latent vectors; SD: standard deviation; M: mesocosm with subscripts S (sunny), Cl (cloudy) and Co (combined) and numbers indicating SSC range; cal: calibration data set; val: validation data set. Updated and modified version published in Wagner *et al.* (2018).

Data set	Param.	Unit	n	Mean	Min	Max	SD	LV	R^2_{cal}	RP%$_{cal}$	R^2_{val}	RP%$_{val}$
$M_{S,2300}$	SSC	mg L^{-1}	12	453	0	1,772	793	5	0.99	4.5	0.99	1.7
$M_{Cl,2300}$	SSC	mg L^{-1}	25	458	0	2,215	503	5	0.97	4.0	0.98	3.0
$M_{Co,2300}$	SSC	mg L^{-1}	37	456	0	2,215	601	5	0.96	5.5	0.97	4.4
$M_{Co,600}$	SSC	mg L^{-1}	25	114	0	547	158	4	0.96	5.3	0.99	3.3
$M_{Co,200}$	SSC	mg L^{-1}	20	45.1	0	197	59.9	4	0.98	3.7	0.99	2.1
	SSC	mg L^{-1}	45	9.86	0.37	55.7	14.6	2	0.81	11	0.87	9.5
Reservoir	Turb.	NTU	47	21.0	1.10	131	32.4	2	0.89	8.1	0.93	6.6
	TP	mg L^{-1}	36	0.06	0.005	0.20	0.054	2	0.93	7.5	0.95	5.9
	SSC	mg L^{-1}	21	42.3	4.20	459	99.9	5	0.96	3.9	0.98	2.0
Kraichbach	TP	mg L^{-1}	21	0.18	0	0.87	0.19	3	0.55	20	0.84	8.0

7.2 TURBIDITY RESULTS

Automated turbidity measurements were correlated against 29 SSC samples taken at flows between 0.5 and 3.7 m³ s⁻¹. A linear relationship was found between turbidity and suspended sediment concentration (Figure 58). However, turbidity measurements fluctuated and showed unrealistically high values for extended periods of time. Two NTU-SSC pairs were affected by this and lay far outside the range of the majority of the points and typically reported regressions (Rügner *et al.* 2014) and were excluded from the regression (2413 NTU for 8.20 mg L⁻¹ and 1965 NTU for 7.20 mg L⁻¹, not shown in the figure).
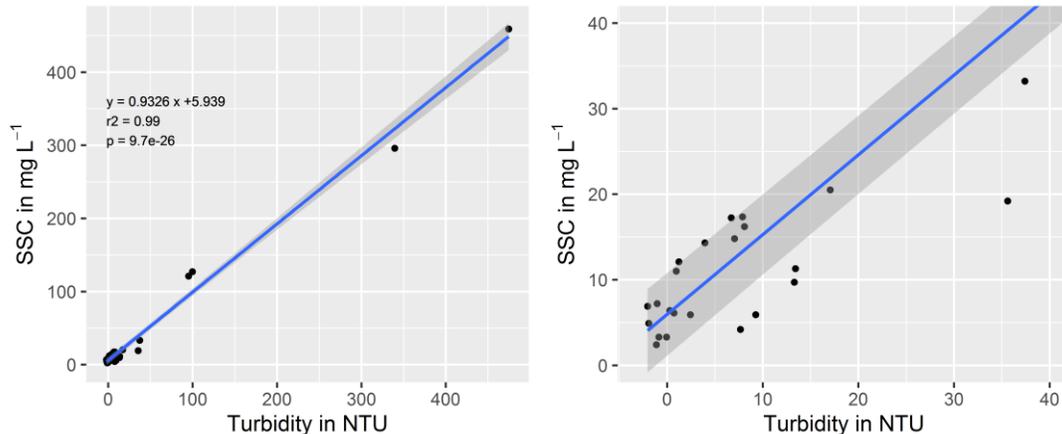
Figure 58: Left: Turbidity measured by an automated logger against SSC from individual grab samples. Regression equations, R² and p-values indicated. Grey area: 95% confidence level interval for predictions. Right: Enlarged view of the lower value range. Grey area: 95% confidence level interval for predictions.

7.3 COMPARISON OF KRAICHBACH EVENT FLUXES

In order to examine the performance of grab-sample based rating curves, reflectance measurements and turbidity measurements to represent event mean concentrations and fluxes, the different approaches are compared to the event fluxes measured discharge-proportionally by the LVS. At the Kraichbach (and Alb), there is a high degree of certainty towards the proper functioning of the setup. The pump, hose and valves were routinely checked and cleaned before and after each measurement cycle and the performance of the pump was tested. The large volume of the sample limits the random error of individual sampling. Although not systematically tested, the representativeness of the sampling for the entire cross-section can be regarded as high. At the measurement location, the Kraichbach is shallow, channelized and slightly constricted with flow velocities of 70 cm s⁻¹ at mean flow (LUBW 2016), ensuring a complete mixing of the water and small particles. This is supported by the fact that the sampled sediment consists almost exclusively of particle sizes smaller than 63 µm in diameter (Table 2). Based on the Hjulström diagram and the underlying formulas (Hjulström 1935; Miedema 2010), spherical quartz particles of this size would require flow velocities of less than 0.5 cm s⁻¹ to begin depositing.

Figure 59 shows the available time period where simultaneous data is available. Eleven LVS sampling periods with varying flow conditions can be used as validation. The three periods taking place in 2018 were also used to construct the adapted LVS rating curves (Section 6.1.1), the remaining eight periods

are additional measurements that were not presented in the preceding chapters. The rating curve based concentration estimates shown are by definition directly linked to discharge (cf. Section 6.2.1) and therefore parallel the peaks of the hydrograph. Contrarily, turbidimeter and reflectance-based concentrations are derived from measurements. The turbidimeter shows the highest SSC peaks, sometimes coinciding with peak measurements of other approaches, sometimes not. The spectrometer concentration estimates show peaks during discharge peaks, although discharge peaks of similar magnitude coincide with different concentration estimates. Figure 60 and Figure 61 show detailed views of two selected high-flow events and the applied concentration estimates from the different methods. In these figures, the lower temporal resolution of the spectrometer becomes apparent, providing only one value per day and representing conditions at noon. Further, the deviating measurements of the turbidimeter can be inspected. While in Figure 61, the turbidimeter records the peaks in agreement with the other measurements, there are two additional and distinctly higher concentration peaks apparent in Figure 60 before and after the actual discharge peak.

Event fluxes were subsequently computed by multiplying measured or modeled concentration by event discharges. In addition, the rating curves based on large volume sampling were also applied to model the event fluxes, i.e. providing a direct model validation. Because the three events in 2018 were also used for constructing the rating curve, there is an inherent relationship present. To maintain comparability, all model performance calculations were additionally made for the non-calibration events in 2019. The results of the computed fluxes by all methods are compared to measured LVS fluxes in Figure 62 as well as in Table 5 and in Figure 74 and Figure 75 in the Appendix. While low flow conditions are well represented in most assessments, events with higher flows are underestimated in almost all cases. LVS-based rating curves achieve the highest NSEs and low deviations from the cumulative masses of several events. The turbidimeter showed very high values for several extended periods which are neither represented in the reflectance data nor in the hydrograph. For three low-flow events, this lead to distinct overestimations of fluxes, and generally poor performance and low NSEs. Reflectance-based measurements highly underestimated event fluxes as well, and also reached low NSEs. Even though a 450 mg L^{-1} calibration sample was available for the SSC grab sampling rating curve and an uncapped polynomial increase of concentrations towards higher discharges were assumed (Figure 33), high flow events could not be adequately covered by rating curves based on grab sampling. This suggest that the increase of rating curve concentrations towards higher discharges is not steep enough in the existing rating curve, or that changing catchment conditions led to a decreasing validity of the established relationship.

The event flux comparison demonstrates that pre-established relationships between discharge and flux must not necessarily be valid when applied to other periods. Even when fluxes on longer time scales can be closely approximated by applying pre-established rating curves, individual events can deviate substantially from the behavior predicted from discharge. The optical methods applied were unable to improve event flux estimation. Turbidimetry for certain periods can be used to qualitatively evaluate the changes in concentration at a high frequency.

Table 5: Measured Large Volume Sampler (LVS) event fluxes, sorted by the highest observed discharge in the sampling period (Q_{max}) and their percentage deviations by different assessment methods. RoX: Reflectance-based flux, RC: Uncapped rating curve from grab sampling; SS: Suspended sediment; TP: Total phosphorous. LVS TP/SS mod: Based on rating curve from LVS sampling (i.e. method validation). The mass sum of measured substance fluxes of all 11 events (Total) and of the 8 events not used for LVS rating curve calibration (Total 2019) and the total deviation of the methods are stated, along with the respective Nash-Sutcliffe-Efficiencies (NSE; NSE 2019).

Event	LVS SS meas. in t	LVS TP meas. in kg	Q_{max} in m³ s⁻¹	LVS SS mod.	RC SS	RoX SS	Turbidity	LVS TP mod.	RC TP	RoX TP
Dec 18	409	292	7	+26%	-45%	-39%	-39%	+27%	-47%	-90%
Aug 19	358	119	5.2	-40%	-55%	-61%	-69%	-38%	-46%	-80%
Sep 18	142	204	3.7	-35%	-78%	-77%	-	+51%	-28%	-40%
Sep 19	20.2	-	2.9	+237%	+171%	+212%	+26%	-	-	-
Jul 19	46.8	719	2.4	+37%	-50%	-34%	+792%	-24%	-50%	-58%
Jan 19	37	231	2.1	+49%	-23%	+9%	-60%	+15%	-3%	-54%
May 19	30.9	159	1.7	+66%	-29%	-70%	+1405%	+12%	-21%	-52%
Jun 19	31.2	226	1.6	+64%	-3%	-42%	-29%	-1%	-12%	-47%
Feb 19	23.1	254	1.4	+85%	+11%	-43%	+204%	+44%	+36%	-18%
Nov 18	13.2	382	1.3	+190%	+32%	+67%	+1758%	+6%	-27%	-31%
Oct 18	29.1	840	0.9	-73%	-34%	-38%	-	+9%	-	+29%
Total	1141	3426		+6%	-48%	-58%	+60%	+3%	-34%	-55%
NSE				0.81	0.52	0.35	-1.5	0.66	0.26	-0.85
Total 2019	547.2	1708		±0%	-37%	-42%	+106%	+6%	-3%	-41%
NSE 2019				0.71	0.28	-0.17	-3.8	0.52	0.046	-0.99
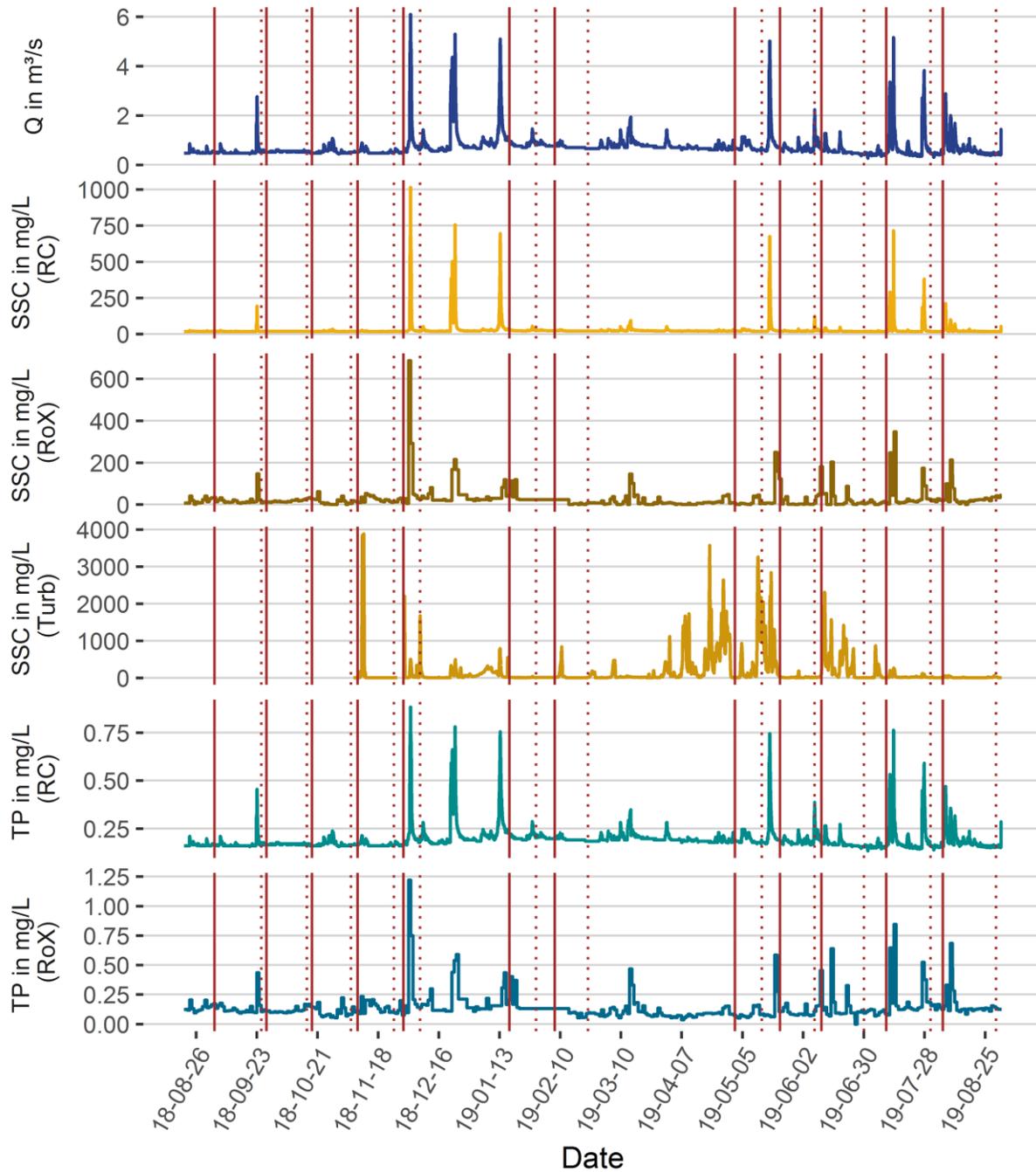
Figure 59: Hydrograph at the Kraichbach with according substance concentrations according to different methods. RC: Rating curve from grab samples; RoX: Reflectance-based concentrations; Turb: Turbidity-based concentration. Vertical red solid and dashed lines indicate beginning and end of an LVS sampling period, respectively.
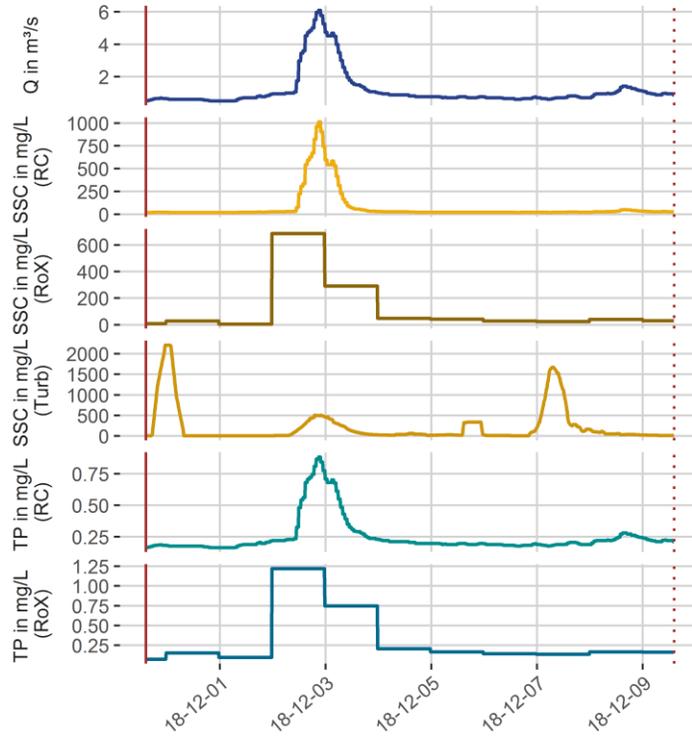
Figure 60: Detailed view of a high flow event in December 2018 with different substance concentration assessments.
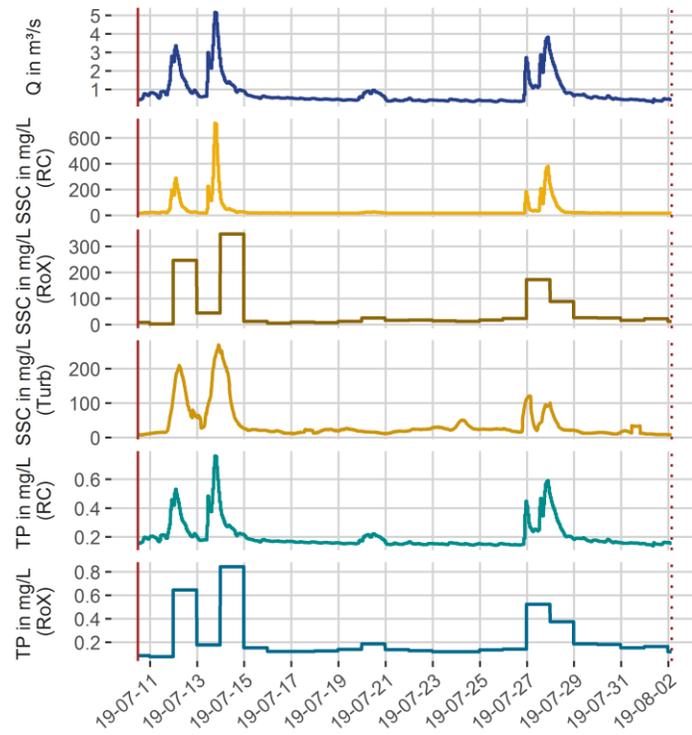
Figure 61: Detailed view of a double peak high flow event in July/August 2019 ('Aug19') with different substance concentration assessments.
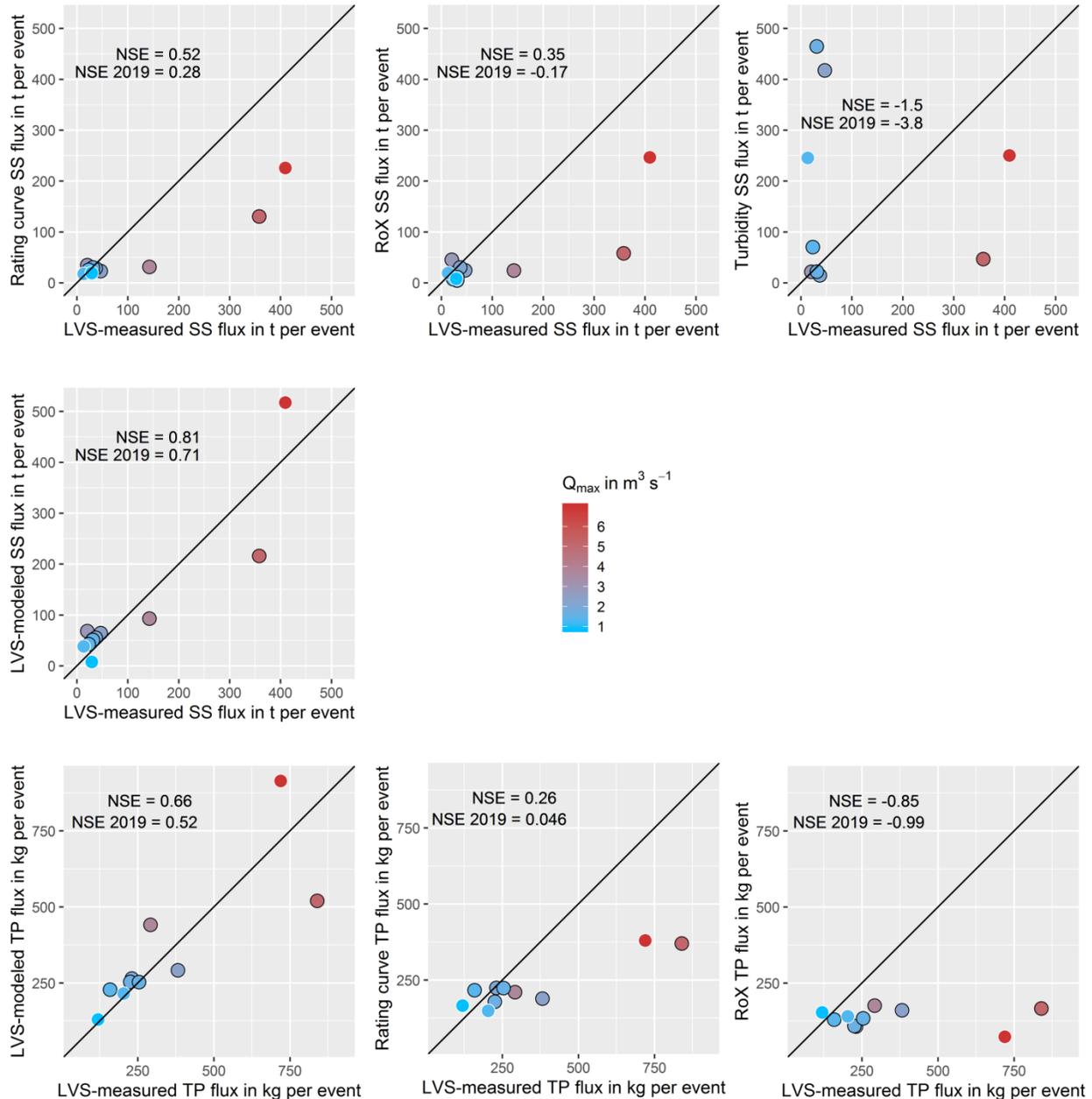
Figure 62: Measured Large Volume Sampler (LVS) suspended sediment (SS) and total phosphorous (TP) event fluxes against estimates based on other approaches. Rating curve SS/TP fluxes: Estimate based on rating curves from grab sampling. RoX SS/TP fluxes: Based on above-water reflectance measurements. Turbidity SS flux: Based on high frequency in-situ turbidimetry. LVS-modeled SS/TP fluxes: Based on rating curve from LVS sampling (i.e. method validation). NSE: Nash-Sutcliffe-Efficiency. NSE 2019: Nash-Sutcliffe-Efficiency for events of the year 2019 (black circle outline), which were not used to calibrate the LVS rating curve. 1:1 line (black) and highest discharge in the respective monitoring period (Q_{max}; fill color of circles) are indicated.

8 SYNTHESIS

The main goal of this work was to explore the potential and the limitations of established and novel methods to provide reliable substance monitoring and, subsequently, reliable mean annual flux estimates for meso scale catchments. These temporal and spatial scales were identified as relevant for area-wide water resource management. For this purposes and to explore the relationship of flux estimations to modeled emissions, discharge-proportional large volume sampling was compared to other means of quantification in three meso catchments of different climatic, pedologic and land use characteristics. By moving from the evaluation of the LVS on the event scale (8.1) and the optical methods (8.2) as well as the sensitive parameters for sampling scheme selection (8.3) to the implication for area-wide emission modeling (8.4), this chapter combines all findings in order to conclude on the hypotheses.

8.1 THE USE OF LARGE VOLUME DISCHARGE-PROPORTIONAL SAMPLING

The measured LVS fluxes serve as a reference for the comparison to other methods, because the discharge-proportional nature of the sampling is offering an cost-effective and representative method for capturing event mean concentrations, as was confirmed by other studies (King 2003; Ackerman *et al.* 2011). Because of its high sampling frequencies, it is an especially helpful tool when the rate of change in loads exceeds the rate of change in discharges, which is usually the case for SS. The large volume of the sample limits the random error of individual sampling and provides sufficient material for a broad chemical analysis of the sediment, even in low-yield catchments. The reported problem that a discharge-proportional sampler unnecessarily samples low flows and subsequently has not enough volume left for capturing a storm event (Lopez *et al.* 2000) is also mitigated by the large volume, leading to active times of several weeks during low flow. The presented system is therefore relatively robust towards different flow conditions throughout the year, and unlike smaller systems with a smaller volume did not require regular reconfigurations, e.g. regarding the expected magnitude of a future event (King 2003; Ackerman *et al.* 2011). To further minimize downtime, a second tank can be installed, thereby ensuring complete coverage. Alternatively, and less costly also in terms of laboratory analyses to be performed, threshold values can be set to start sampling only at a certain discharge or flow change rate (Lopez *et al.* 2000). The LVS is by design not suited for the investigation of hysteresis patterns. Further, the system requires usual maintenance and inspections as well as discharge measurement, protection from external influences and a connection to the electrical grid. This limits the possible sites and increases operation costs. The relevance of regular maintenance and the difficult installment have been demonstrated at the Passauna River, where the approach failed entirely due to an inadequate pump.

It was assumed that discharge-proportional composite sampling could reveal emission patterns difficult to assess by grab sampling. In the Kraichbach catchment, events dominated by emissions from sewer systems could be separated from low-flow events, where overflows are inactive, and high-flow events, where overflows do take place, but become non-relevant compared to the dominant sediment emissions via erosion from arable land. This was indicated by higher sediment contents of organic material and phosphorous (Brombach *et al.* 2005; Soonthornnonda and Christensen 2008) for medium-sized

events. Because of the short duration of these events, only automatized sampling is an efficient method of capturing sufficiently many events of different magnitudes in order to reveal such a pattern. This offers the opportunity to evaluate the impact of this pathway without specifically sampling the often numerous individual sewer overflow structures of a catchment. At the Alb, the LVS revealed that there is indeed a positive linear correlation of sediment concentration with event magnitude, while grab sampling indicated no relationship. Only the frequent, automated sampling of the highest flows allowed to draw this conclusion.

Further, by adequately measuring event fluxes, it was exemplified how the characteristics of a single events can be controlled by the preceding sequence of events. The Kraichbach high-flow event ending in August 2019 ('Aug 19') delivered twice the sediment estimated by the uncapped grab sampling rating curve, and was also significantly underestimated by the LVS adapted rating curve. The following monitoring period in turn showed a much smaller sediment flux than expected from the rating curves. While factors like rainfall intensity, soil moisture, harvesting activities and excavation works are important, this may also be seen a consequence of the episodically sediment transport pattern described by Ferguson (1982) as a 'jerky conveyor belt'. This episodically transport is not only a function of flow magnitude; but also the state of the temporary sediment storage elements in the fluvial system; namely bars, islands, benches, slopes, blankets and the river bed; as well as their degree of connectivity to the stream network (Fryirs and Brierley 2001; Fryirs 2013). Both, connectivity and storage, are dynamically altered by each flow event:

"The amount of sediment moved by a single flow is linked to the magnitude and duration of that event, whereas the amount of sediment that is moved over a period of time is linked to the frequency of flows of different sizes." (Fryirs 2013: 36)

Applied to the example of the Kraichbach, it can be argued that through the configuration of flow and catchment pre-conditions, the series of high flow events observed in July to August 2019 ('Aug 19') coincided with catchment conditions of high sediment availability; or high synchronicity within the sediment-controlling state of landscape features. The July/August high flow series exceeded the breaching capacity of the blockages and reworked barrier and blankets which led to an over-proportional release and flushing of sediment. In a flood relief channel of the Kraichbach offering depositive conditions during and after the events, substantial amounts of sediment had to be removed afterwards (Figure 78, Appendix). After the event, the state of the sediment storage elements in the catchment were altered. This lead to a decrease of readily available sediment in the catchment and river channel, which in turn resulted in an under-proportionally low flux during the August/September event, with one of the lowest mean event concentration of the entire data set, despite the elevated flow conditions. To study phenomena like these on the event scale, discharge- proportional large volume sampling can be considered an adequate method regardless of additional benefits such as higher predictive abilities of LVS rating curves compared to grab sampling rating curves (Section 8.3).

8.2 OPTICAL METHODS TO MEASURE SUBSTANCE CONCENTRATIONS

Above-water reflectance measurement were explored as a low-maintenance method to retrieve daily concentration data. As described in Section 7.1, the accuracy of above-water reflectance and PLS to

determine SSC and associated TP can be very high, when illumination and other boundary conditions are stable. The established boat-based setup is suitable to determine surface concentrations during measurement campaigns. Despite the automatically correcting measurement principle of the spectrometer employed as a permanent monitoring station at Kraichbach, the setup led to substantially low-biased results. Overall, the hypothesis that unsupervised above-water reflectance measurements can improve event flux budgets must be rejected for the procedure and catchment tested here.

One reason for this lies in the calibration, as was outlined in Section 7.1.3. PLS models suffered from an uneven distribution of training samples and presumably from changing illumination conditions throughout the year and varying degrees of bottom reflectance, leading to overfitting. Dust accumulation on the irradiance sensor may be responsible an observed drift and suggests the installation of an air-based cleaning system. Even though the highest available training sample was at around 450 mg L^{-1}, it was shown that the calibration range does not encompass the true concentration amplitude for high-flow events at Kraichbach. While the predicted changes in concentrations largely followed the, the high concentrations that must have been present to explain the measured reference fluxes were not covered. Because of their sensitivity to the training data, the lack of training samples at the extreme end is an unfavorable situation for empirical procedures like PLS. Therefore more samples in the high range are required (Torres and Bertrand-Krajewski 2008), and analytical procedures must be considered.

Another major limitation of the setup is that only one concentration estimate per day is available based on reflectance recorded between 11:00 and 12:00 am. In the current setup using natural sunlight, the extraction of more than one value per day in a seasonally robust way is not feasible, even though the instrument is designed to correct for downwelling irradiance. In order to be representative in subsequent processing, the concentration at this time would need to be equal to the flow-weighted mean concentration of that day. From the periods with presumably unbiased turbidimetry results, it is clear that concentrations change on the scale of minutes to hours at the Kraichbach. Under these premises, the probability of a daily snap shot to be representative for the entire day is therefore too low for a river of these hydrological properties; as well as for the study of individual events. Given the short durations of peak concentrations, these effects may only in part be compensated by occasionally covering peak concentrations during the hour of measurement when monitoring mass fluxes on longer time scales. In short, the load flashiness of the Kraichbach is too high for daily noon measurements to obtain a representative concentration estimate. An hourly resolution must be pursued. As long as only one measurement per day is available by this method, a monitoring device relying on natural sunlight should therefore be applied for larger streams with response times in the range of several hours to days and subsequently lower load flashiness. Alternatively, the current setup could be used to provide concentration-discharge measurements in order to gain more data for rating curves, rather than the approach presented here, where the daily measured concentration is multiplied by daily flow. As the application to long-term monitoring is new, comparison to other studies is difficult. In an terrestrial vegetation application with a comparable setup, Cogliati *et al.* (2015) reported sensor stability for several weeks. Torres and Bertrand-Krajewski (2008) report high predictive abilities for submerged spectrometers and PLS approaches, which must be considered as an option independent from ambient light.

The dependence to calibration data is also present for turbidimetry, although the measurement principle is well-established. The NTU-SSC relationship found here is near 1:1. In laboratory and field

analysis in the same geographical region than the Kraichbach, Rügner *et al.* (2014) found ratios of 0.93:1 to 2.35:1 with a mean of 1.5:1. Therefore, the established relationship used here can be considered in a normal range. Turbidimeters show a very linear behavior, which is the reason the confidence for SSC estimations from NTU measurements outside the calibration range is high. Nevertheless, the turbidimeter was the least reliable method to asses concentrations. Already during calibration, a number of automatically recorded NTU values were far off the usually reported lines and were removed for the establishment of a the NTU-SSC regression. Other researchers reportedly did the same (Martínez-Carreras *et al.* 2016). From the extended periods with very high and erratic measurements it must be concluded that the device was not working properly, presumably because a dirty or otherwise blocked measurement window. As the device was equipped with a wiper and there is no relationship to the maintenance intervals, the true nature of the disturbance remains unclear. It underlines however the difficulties of unsupervised long-term measurements in natural rivers. Martínez-Carreras *et al.* (2016), even though they used a pressured air cleaning system, report similar problems encountered with an submerged spectrometer and trace it back to a dirty measuring window. While there are studies demonstrating the benefits of high-frequency turbidimetry (Torres and Bertrand-Krajewski 2008; Rügner *et al.* 2014), high maintenance efforts and restricted spatial transferability and appropriateness for stating environmental standards are identified as limitations (Davies-Colley and Smith 2001). In the presented case, turbidimetry to monitor SSC proved to be highly unreliable.

The use of automated high-frequency measurements is appealing, because it offers a temporal observation scale in agreement with the scale of dynamic processes. In order to measure the conditions in natural or man-made environments, the devices must also be exposed to them. This demonstrably has unwanted effects on the long-term measurement stability. Because it is much easier to protect a device from atmospheric weathering compared to the more demanding conditions in a river, contact-free monitoring can provide a low-maintenance alternative. In the examined case, the dependency on natural sunlight clearly set the boundaries with respect to temporal resolution and accuracy of such a system. If these technical problems are overcome, the proposed monitoring station can be beneficial especially at remote sites, where regular maintenance is difficult to accomplish.

8.3 FLASHINESS AND SAMPLING SCHEMES

It was expected that explicitly sampling high flows would lead to higher flux estimates compared to methods based on sporadic grab sampling. This can be confirmed for the Kraichbach and Alb catchments, were LVS estimates are in almost all cases higher than the corresponding estimate from grab sampling. Integrated fluxes of multi-day, volume-constant segments of the hydrograph show stronger correlation to the highest discharge of the respective period than to any other flow parameter. This indicates that concentrations are driven by event magnitude. In the Passauna catchment, there are distinct indications that the LVS was not sampling discharge-proportionally. By means of including the sediment stock of the downstream reservoir, which is giving an integrated long-term sample over all flow conditions, it is implied however that especially in this catchment, actual fluxes are tremendously underestimated by grab sampling.

Parameters to impact the relevance of measuring concentrations during high flows were expected

to be the yield from diffusive sources and the degree of association to particulate transport of the substance. Flux underestimation, when defined as the difference between grab sampling rating curves and the best available estimate, is greatest in the high-yield, particle transport dominated Passauna catchment. The Kraichbach has much lower sediment yields and a much smaller offset between grab sampling and LVS assessments. The low-yield Alb catchment shows the smallest deviations between different modes of computation. The relationship to the degree of association to particles was also confirmed: For the Kraichbach, SS underestimation by grab sampling is greater than for TP, which is greater than for orthophosphate. At the Alb, TP transport is driven by orthophosphate, which is largely independent from flow conditions. Discharge-proportional sampling for this dissolved parameter becomes therefore obsolete, which supports the hypothesis. Of these two catchments, the relevance of automated and discharge-proportional sampling to represent fluxes is therefore highest for suspended sediment from the agricultural Kraichbach catchment and lowest for orthophosphate from the forested Alb catchment. In conclusion, the hypothesis can largely be confirmed with the present example, but only when joining he two statements: The relevance is controlled by the particulate yield from diffusive sources. At the Alb, absolute TP yields are similar to Kraichbach, but only because of the high base load with non-particulate and flow-independent orthophosphate. Regarding only SS and particulate phosphorous, the relevance-dependency is given and correctly indicated by the newly proposed flashiness indices, as described in Section 6.4.4.

The comparison of the different methods applied at the Kraichbach demonstrated how the spatio-temporal complexity of sediment transport can lead to substantial deviations of rating curve assessments from actual concentrations and fluxes of individual events. Rating curve scattering and the dependency on calibration data has been studied for decades (Walling 1977), with solutions aiming at hydrology-based sampling schemes, i.e. frequent high-flow sampling (Kulasova *et al.* 2012; Horowitz *et al.* 2015). If grab sampling is pursued, for instance due to budget restrictions, this must be the minimum requirement for systems with a high sediment availability such as the Kraichbach and Passauna catchments. Due to the short duration of these events, the timing of such sampling is a challenge and can make grab sampling an inefficient method for covering mean concentrations, especially for catchments with variable loads (Facchi *et al.* 2007). When only few measurements are available in the high range, the addition or removal of one measurement can lead to distinct changes in the inclinations of the rating curve. The alternative perspective offered here is the fusion of the representativeness and robustness of discharge-proportional large volume sampling with predictive properties of rating curves. It was demonstrated that for the understanding of individual event characteristics actual monitoring is needed. For longer time periods, there are indications that the predictive abilities for the LVS rating curve outmatch grab sampling rating curves, apparent by the respective Nash-Sutcliffe-Efficiencies, the close approximation regarding the flux from several combined events and generally a lower scattering of LVS rating curves as opposed to grab sampling rating curves.

Closely connected to the conclusions above is the hypothesis that calendar based grab sampling is not suited to support management decisions for high-yield, particle-dominated catchments. For the Kraichbach, grab sampling may provide slightly low-biased results. Due to the large quantity of samples (exceeding the usually pursued monthly monitoring by far) used to construct the rating curves in this study, the relationship between high flows and high fluxes would be sufficiently clear to identify arable land as

the main contributor of the sediment and to support the implementation of measures to reduce soil loss from these areas. In fact, according measures have been implemented in the catchment before discharge-proportional sampling was implemented, presumably because the amounts of displaced sediment after each storm event were obvious. Nevertheless are precise estimates needed, e.g. to evaluate the effectivity of the implemented measures, to assess annual soil loss and to calculate the amount of sediment that will deposit in the downstream canal network, where lower flow velocities predominate. This is critical information for the financial planning concerning e.g. the maintenance of relief channels and therefore the protection from floods. For Passauna, this becomes even more evident. Based on the least conservative estimate from grab sampling, the sedimentation rate in the reservoir would be estimated to just 0.04% a^{-1}, indicating a theoretical lifetime (initial capacity reduced to half, Sultana 2017) of 1,250 years. The actual rate is 0.23% a^{-1} indicating a lifetime of 217 years, of which 30 have already passed. Clearly, current grab sampling schemes are not suited for supporting management decisions in the Passauna catchment. A robust high-flow sampling scheme and a reinstated discharge measurement station are urgently needed. These measures can be complemented by regular acoustic measurements of the sediment volume in the reservoir.

It was demonstrated here that flow flashiness alone can be misleading when selecting an adequate monitoring scheme: The Kraichbach has more stable flows than the Alb, yet is the importance of high-flow monitoring undoubtedly higher. This is reflected by the load flashiness indices proposed in this thesis. They therefore have a higher informative value than the flow flashiness index alone for the planning of sampling schemes and can be used as a comparative value for the transport behavior for rivers around the world. The limitation of the LFI and nLFI is, that while flow flashiness can be derived directly from the hydrograph, high-flow monitoring is needed to establish these indices in the first place. A possible solution is a short-termed monitoring period, e.g. by few peak flow measurements, to roughly determine load flashiness and to select further long-term monitoring schemes on this basis. Deviations of the nLFI from 1 indicate a detachment of changes in loads from changes in discharges. If that is the case and the LFI is above a yet to be established threshold (e.g. > 0.5), hydrology-based sampling is increasingly required. For a complete picture, all three indices must be regarded, together with the overall relevance of the catchment emissions within a larger area.

For the Kraichbach and similarly erodible catchments, it is recommended to maintain the discharge-proportional large volume sampling to study long-term changes of sediment and phosphorous fluxes. This way, it can be clarified if ongoing soil-conserving measures in the catchment are effective. In addition, short-termed emissions from sewer systems can be covered. Given the short duration of these events, monthly sampling as employed several kilometers downstream is unlikely to sufficiently account for these events. Despite being more flow-flashy, monthly sampling with two additional high flow samples per year as suggested by Horowitz *et al.* (2015) could provide cost-effective flux estimates for the Alb. Given the short duration of the high flows, a certain degree of flexibility is required. At Passauna, the relevance of sampling discharge-proportionally is highest. This is indicated by the fact that the downstream reservoir provides drinking water for a large city and siltation must be monitored and reduced. Further, the flash flood periods of high transport are very short and accompanied by rainfalls and discharges so extreme, that manual sampling can become unsafe for the person carrying out the sampling, which may be the reason why no such events have been sampled within decades of official monitoring. It is doubtful

that submerged optical measurement equipment can persist in these conditions, or be able to cover the expected range of SSC. Satellite imagery to determine SSC are likely to be obstructed by clouds during the high-transport events. Robust, automated sampling solutions are therefore needed.

8.4   THE RELATIONSHIP TO EMISSION MODELING

Modeled sediment and phosphorous inputs via erosion at the Kraichbach are substantially lower than measured and extrapolated from a small number of high-flow events. Sediment input is a highly erratic process. How can we be sure that these few events should be used to describe the target of 'mean' conditions? This can be approached by the scale concept of Blöschl and Sivapalan (1995): The Kraichbach example shows that large volume sampling of events represents a convergence of the process scale and the observation scale. In other words, LVS monitoring revealed significant magnitude-input relationships, because its resolution is small enough to describe the process of erosion from regularly recurring events. The remaining noise in the data could be limited by further decreasing resolution, i.e. measuring or modeling in smaller spatial and temporal units. The sediment input model however has a much broader observation scale. It is also designed to describe a broader process, i.e. mean long-term conditions. Because it is applied on a broader scale (e.g. different landscape regions), a relatively high level of noise is observed (and often accepted) when zooming into specific areas. The study of two catchments at the extreme ends of the most sensitive process parameter, the share of arable land, indicates however that this noise may not be scattered around the mean, but far below it. In the presented cases, the process scale revealed that intra-annual transport has not sufficiently been integrated into the working (modeling) scale.

In this light, it was to be explored under which circumstances calendar sampling schemes can favor false positive low-biased model validation. The offset between emission modeling and actual fluxes was expected to be related to the load flashiness of a parameter in a catchment. From the limited data provided by this study, no final conclusions can be drawn. For both cases, grab sampling was frequent enough to reveal a low-biased sediment input model. One could argue that if the highest two SSC measurements had not been made at the Kraichbach due to less frequent calendar sampling, then a low biased validation could have happened. Flashiness indices were computed based on grab sampling rating curves. For the Kraichbach, it can be confirmed that the load flashiness for the three substances is covariant with the degree of underestimations of the model (Table 3). For the Alb catchment, the load flashiness indices are constant and have no explanatory power towards the differing model offsets.

It has been widely demonstrated that calendar sampling leads to inaccurate sediment flux estimates. Although there are examples where sparse sampling can also lead to the overestimation of fluxes (Johnes 2007; Horowitz *et al.* 2015), there is a general tendency towards flux underestimations, especially in smaller (i.e. flashier) catchments (Horowitz 2003; Coynel *et al.* 2004; Kulasova *et al.* 2012; Horowitz *et al.* 2015). This could be confirmed by this study, where almost all grab sampling approaches for SSC and TP were lower than fluxes assessed by discharge-proportional sampling (Sections 6.4 and 7.3). Calendar sampling is also the usual method to validate or even calibrate emission models (Borah and Bera 2004; Bouraoui and Grizzetti 2014; Kiemle *et al.* 2019). Therefore, there is indeed a substantial risk that emission models are eventually validated with or even calibrated to low-biased flux estimates, especially

in small catchments and for particle-associated parameters. In other words, when low-biased calendar-based fluxes are the only available measure to compare emission model results to, a proportional low-bias in the emission model would not be noticed.

Usually, deviations between emissions and fluxes are tolerated in a broad range with references to the uncertainties attributed to the large model areas, lack of input data or necessary simplifications; or to retention and degradation processes. Larger offsets are observed in smaller catchments, and the possibility of a measurement bias is acknowledged (Berlekamp *et al.* 2007; Fuchs *et al.* 2010: 113). Borah and Bera (2004) investigated the potentials and limitations of a variety of emission models and setups. They found that most models are calibrated and validated by monthly concentration samples. Peak flows, individual events and short time intervals were generally associated with poor model performance, which supports the hypothesis that emission models are usually inadequate to represent episodes of high fluxes. Even though there are strong indications that the requirements for misleading model validation are met for flashy catchments, there is no study demonstrating whether low-biased measurements can systematically feedback on low-biased emission modeling. Facchi *et al.* (2007) demonstrate at the example of diurnal emission peaks from a waste water treatment plant, how infrequent sampling can eventually loop back on a water quality model and obscure the actual environmental impact of the emission.

From another point of view, the large deficiencies of the sediment input model identified in this study may be very specific for the selected investigation areas. Zehe *et al.* (2001) showed that a physically based model can accurately predict SSC concentration changes during an event in the Weiherbach. Due to their input data demands, such models are however not suited to be applied area-wide. The empirical and much less data-demanding Universal Soil Loss Equation (USLE) and its modifications RUSLE and Modified Universal Soil Loss Equation (MUSLE) have been extensively validated by field plot measurements (Hudson 1993). It is emphasized that the USLE provides long-term averages of soil loss from arable land and may not be applied to individual storm events (which can be modeled with MUSLE, Djoukbala *et al.* 2019). Based on USLE soil losses, sediment loads have been also successfully validated by frequent concentration measurement, e.g. using the Soil and Water Assessment Tool (SWAT) in very large catchments in Tanzania (Ndomba *et al.* 2008). Qiu *et al.* (2012) applied MUSLE and SWAT to a small 8.3 km² hilly loess catchment in the Shaanxi Province, China, to compare modeled and measured (manual high-flow grab sampling) sediment fluxes during 18 stormflow events, with event fluxes of up several thousand tons. Even though the model also slightly underestimated actual fluxes, this exemplifies that empirical sediment input models can adequately represent rather extreme erosive conditions. In summary, in by using other approaches, low-biased flux estimates could also be detected by well-parameterized sediment input modeling and the subsequent offset in the flux comparison. Discharge-proportional sampling; in all cases with high variability of concentrations, whether attributed to erosion or oscillating point sources; increases the chances drastically that these effects are covered.

Despite the fact that other sediment input models may be more adequate, the considerations from above raise the question where the scope of emission modeling ends, and what role flashiness plays in this regard. The scope of flux estimation by coupling discharge and concentration is easier to determine; it ends when either or both cannot be determined within a the requirements of minimum certainty. As an extreme example, for catastrophic landslides or mudflows, other methods of assessing the material movement have to be applied. Emission models usually do not deal with these conditions but rather aim

to represent mean long-term conditions, thereby fulfilling WFD requirements (Fuchs *et al.* 2010). The WFD puts a strong emphasis on the long-term ecological state of the water bodies (Voulvoulis *et al.* 2017), for which very rare, short events of high transport are not necessarily crucial. For the emission and transport of sediment and phosphorous, these events are over proportionally important and can therefore control the 'mean' flux. In what way is the mean an adequate measure to determine fluxes, especially if long periods of low transport are intermittent by very short periods of extreme transport? A pragmatic view on this issue is to exclude extraordinary events from such assessments because the intention of emission modeling and the WFD is to represent an adequate picture of the predominant environmental state (European Parliament and Council 2000; Voulvoulis *et al.* 2017). Under these conditions, special commissions and reports take over to determine fluxes and environmental impact, as is exemplified by the Elbe floods of 2002 and 2013 (FGG Elbe 2014).

By moving from the macro towards the meso scale, the boundaries between regular condition (to be represented by emission models) and extraordinary events can however become more difficult to define. One could argue, the annuality of an event can be used to determine this boundary independently of catchment size: If the annuality is greater than the period, for which the general boundary parameters (e.g. land use, topsoil concentrations etc.) of the emission model were deemed applicable, the event is not to be included in regular emission modeling but must be covered by other means. The problem of this perception is that for flashy catchments, the duration of high-flux events decreases, thereby lowering the probability of adequately covering fluxes in calendar-based monitoring programs. This would mean that for an event of a certain annuality, the probability of adequate monitoring is greater for larger streams and stable catchments, while smaller rivers and flashy catchments are increasingly low-biased. For flashy catchments, this leads either to a false low validation of the emission model; or validation fails in case the emission model does in fact represent the short high flux periods correctly. Both is undesirable because area-wide emission modeling and its validation are required to be robust towards load flashiness within the pre-defined boundary conditions. Consequently, to maintain the annuality threshold across catchments of different flashiness, sampling strategies have to be adapted to provide a common basis of comparison. In other words, the scale of observation has to reflect on the process scale. With the concept of load flashiness and the LVS system, the tools to achieve this are available.

9 CONCLUSIONS AND OUTLOOK

The spatio-temporal complexity of suspended sediment transport and associated substances impedes its integration into the generalized concepts of mean long-term system descriptions demanded by watershed managers and policy makers on a spatial meso scale. Two main drivers controlling this complexity can be identified: The flow flashiness, which reflects on the height, frequency and –most importantly– shortness of discharge events with respect to the base flow; and soil-related catchment management parameters. Together, they control the load flashiness of particulate matter and particulate phosphorous, which is related to the probability that common calendar-based sampling schemes provide low-biased flux assessments. Discharge-proportional large volume sampling solves this by representatively capturing multi-day sections of the hydrograph, and, disregarding the technical failure at one site, proved to be the most reliable method of doing so. It was also demonstrated that sampling discharge-proportionally requires higher expenses, maintenance and expertise compared to grab sampling. For site prioritization, it is therefore suggested to manually sample a number of high flow events to create a preliminary substance rating curve. Based on this, the proposed LFI can be computed and compared scale-independently to identify the catchments and parameters with the highest requirements of high-frequency sampling. Naturally, the large volume sampler method has additional benefits such as the provision of sufficient sediment in low-yield catchments, which can be important for the study of particle-bound pollutants. Further, this study showed it is able to capture erosion-independent point source and sewer system emissions. The system can be equipped with a second tank, ensuring complete coverage, or discharge thresholds to only include actual high-flow events to limit costs. A combination of these approaches is currently implemented in a follow-up project at the Kraichbach. Based on the findings of this study, LVS sampling can be recommended wherever temporal complexity of the target parameter is given, where optical and other in-situ proxies do not suffice, where cost of analysis must be minimized, and where expertise and attendance is provided.

The newly proposed, adapted rating curves based on discharge-proportional sampling revealed flow dependencies of substance concentrations and made them more explicit than grab sampling rating curves, apparent by the higher coefficients of determination and the lower scattering. This is connected to the integrating and somewhat buffering mechanism of composite sampling which limits the random error of individual sampling, but also suggests that event-based monitoring provides an adequate scale for the associated erosion processes. High predictive abilities of the highest measured discharge in the sampling period towards mean SS and TP concentrations further support that notion. The adapted rating curves performed better to predict actual fluxes compared to conventional rating curves and optical methods in a small set of validation measurements. However, substantial measurement-model deviations for individual events underline the fact that every period is different and also depends on the sequence and magnitude of preceding discharges. As for all types rating curves, the extrapolation for extreme events outside the calibration range must be carefully deliberated. Automated sampling however increases the chances drastically that within relatively short observation times of 1-2 years, sufficient high-flow samples are available for establishing rating curves that are valid for a broad range of discharges.

The requirements for false positive low-biased emission model validation from calendar-based

grab sampling are generally met, and the phenomenon is poorly studied especially with respect to a systematic feedback on the model. The potential risk of it is highest for small and mountainous catchments, high shares of agriculture and particle-associated substances, which can again be expressed as the LFI. As every emission model setup is unique, the LFI and nLFI should however only be used to evaluate the probability that calendar sampling is providing an accurate reference. Broader studies are needed to assess if the LFI can also help to interpret model validations, e.g. with respect to constant and inconstant emission pathways.

Contrasting the current surge of attention and new developments regarding optical high-frequency equipment for monitoring, this study primarily revealed limitations of these systems. Turbidimetry to monitor SSC showed to be highly unreliable with alternating periods of erratic and plausible measurements, yet a substantial low bias even for events where concentration estimates, as would be expected, paralleled the hydrograph. This may be a site- or device-specific issue but underlines the difficulties of underwater optical measurements in natural environments, even at short maintenance intervals. To circumvent these issues, a novel, low-maintenance and autonomous setup for above-water reflectance measurements was developed. During short-termed reservoir measurement campaigns, high predictive abilities for SSC and covariant TP were achieved while showing advantages with respect to the retrieval of several parameters at once, simultaneous capturing of reference spectra and non-obstructive handling. PLS regression was confirmed as a suitable empirical method for parameter retrieval. Between the reservoir campaigns, temporal transferability was demonstrated, sometimes considered a weak point of empirical approaches. This was however not the case for permanent autonomous monitoring of in-river SS and TP concentrations, which lead to low biased flux estimates. This can be attributed to the uneven spread of training samples across the concentration range, to changing illumination conditions throughout the year, an assumed sensor drift and to the varying water level, controlling bottom reflectance of the river. The findings indicate that future applications of autonomous above water reflectance should therefore focus on larger streams, ensuring optically deep conditions and a lower load flashiness. The system required low maintenance and low energy use, which makes it especially interesting for remote regions. Artificial lighting, an air cleaning system and inverse modelling for parameter retrieval should be investigated to limit the dependency on natural sunlight and training samples.

Despite their potential to monitor short-termed emission pulses, optical methods still have deficiencies with respect to their robustness towards environmental disturbances. If the expertise to process, interpret and link the great quantities of data is available, they can provide critical information to support management decisions as was exemplified here by the reservoir concentration gradients. Ultimately, these systems will to some degree always be dependent on laboratory analyses for calibrating them to the unique optical properties of the investigated water body. Hydrology-based sampling schemes can provide these data across a broad range of values and at the same time be used to establish accurate flux assessments as a reference. Replacing half of the low flow measurements in existing calendar schemes with flexibly obtained high-flow measurements may be the most cost-effective way of improving flux assessments around the world.

The technology to monitor substance concentrations at a high precision and high temporal resolution is available today. The more pressing issues are site prioritization, equipment maintenance and data interpretation. By deliberately combining traditional and novel monitoring concepts adapted to the

environmental behavior and impact of the target parameters, modern water resource management will be able to cope with todays and future challenges in the provision of an intact aquatic –and terrestrial– environment.

REFERENCES

Abbaspour, K. C., Rouholahnejad, E., Vaghefi, S., Srinivasan, R., Yang, H. and Kløve, B. 2015 A continental-scale hydrology and water quality model for Europe: Calibration and uncertainty of a high-resolution large-scale SWAT model. *Journal of Hydrology*, **524**, 733–752.

Abt. 9: Landesamt für Geologie, Rohstoffe und Bergbau 2015 *Organische Kohlenstoffvorräte der Böden in Baden-Württemberg.* LGRB-Fachbericht.

Ackerman, D., Stein, E. D. and Ritter, K. J. 2011 Evaluating performance of stormwater sampling approaches using a dynamic watershed model. *Environmental monitoring and assessment*, **180**(1-4), 283–302.

Amsler, M. L. and Drago, E. C. 2009 A review of the suspended sediment budget at the confluence of the Paraná and Paraguay Rivers. *Hydrol. Process.*, **23**(22), 3230–3235.

Apitz, S. and White, S. 2003 A conceptual framework for river-basin-scale sediment management. *J Soils & Sediments*, **3**(3), 132–138.

Arbeitsgruppe Oberflächengewässer (AG OW) der FGG Elbe; Bundesanstalt für Gewässerkunde (BfG) 2014 *Das Messprogramm Extremereignisse beim Junihochwasser der Elbe 2013*: *Schadstoffkonzentrationen und –frachten*. www.fgg-elbe.de/dokumente/fachberichte.html?file=tl_files/Download-Archive/Fachberichte/Hochwasser/Ergebnis_MP_Extreme_Okt_2014_Endversion.pdf.

Arnold, J. G., Moriasi, D. N., Gassman, P. W., Abbaspour, K. C., White, M. J., Srinivasan, R., Santhi, C., Harmel, D. D., Griensven, A. van, Liew, M. W. van, Kannan, N. and Jha, M. K. 2012 SWAT: Model Use, Calibration, and Validation. *Transactions of the ASABE*, **55**(4), 1491–1508.

Asselman, N.E.M. 2000 Fitting and interpretation of sediment rating curves. *Journal of Hydrology*, **234**(3-4), 228–248.

Audet, J., Martinsen, L., Hasler, B., Jonge, H. de, Karydi, E., Ovesen, N. B. and Kronvang, B. 2014 Comparison of sampling methodologies for nutrient monitoring in streams: uncertainties, costs and implications for mitigation. *Hydrol. Earth Syst. Sci.*, **18**(11), 4721–4731.

Baker, D. B., Richards, R. P., Loftus, T. T. and Kramer, J. W. 2004 A NEW FLASHINESS INDEX: CHARACTERISTICS AND APPLICATIONS TO MIDWESTERN RIVERS AND STREAMS. *J Am Water Resources Assoc*, **40**(2), 503–522.

Batalla, R. J. 2003 Sediment deficit in rivers caused by dams and instream gravel mining: A review with examples from NE Spain. *Rev. C. & G.*, **17**(3-4), 79–91.

Berechnung von Flächenanbindung, Anbindungswahrscheinlichkeit und Sedimenteintrag in die Gewässersysteme in Baden-Württemberg: *Endbericht* 2018, 41 pp. (accessed 17 September 2018).

Berlekamp, J., Lautenbach, S., Graf, N., Reimer, S. and Matthies, M. 2007 Integration of MONERIS and GREAT-ER in the decision support system for the German Elbe river basin. *Environmental Modelling & Software*, **22**(2), 239–247.

Beudert, G. 1997 *Gewässerbelastung und Stoffaustrag von befestigten Flächen in einem kleinen ländlichen Einzugsgebiet*. Dissertation, Institut für Siedlungswasserwirtschaft, Universität Karlsruhe, Karlsruhe.

Biemans, H., Haddeland, I., Kabat, P., Ludwig, F., Hutjes, R. W. A., Heinke, J., Bloh, W. von and Gerten, D. 2011 Impact of reservoirs on river discharge and irrigation water supply during the 20th century. *Water Resour. Res.*, **47**(3), 247.

Bilotta, G. S. and Brazier, R. E. 2008 Understanding the influence of suspended solids on water quality and aquatic biota. *Water research*, **42**(12), 2849–2861.

Blöschl, G. and Sivapalan, M. 1995 Scale issues in hydrological modelling: A review. *Hydrol. Process.*, **9**(3-4), 251–290.

Borah, D. K. and Bera, M. 2004 Watershed-Scale Hydrologic and Nonpoint-Source Pollution Models: Review of Applications. *American Society of Agricultural Engineers*, **47**(3), 789–803.

Bosworth, W. S. and Thibodeaux, L. J. 1990 Bioturbation: A facilitator of contaminant transport in bed sediment. *Environ. Prog.*, **9**(4), 211–217.

Bouraoui, F. and Grizzetti, B. 2014 Modelling mitigation options to reduce diffuse nitrogen water pollution from agriculture. *The Science of the total environment*, **468-469**, 1267–1277.

Bravard, J.-P. and Petts, G. E. 1996 Human impacts on fluvial hydrosystems. In: *Fluvial Hydrosystems*, G. E. Petts and C. Amoros (eds.), Springer Netherlands, Dordrecht, pp. 242–262.

Brils, J. 2008 Sediment monitoring and the European Water Framework Directive. *Annali dell'Istituto superiore di sanita*, **44**(3), 218–223.

Brombach, H., Weiss, G. and Fuchs, S. 2005 A new database on urban runoff pollution: comparison of separate and combined sewer systems. *Water science and technology : a journal of the International Association on Water Pollution Research*, **51**(2), 119–128.

Bundesministerium der Justiz und für Verbraucherschutz (BMJV) 2016 *Verordnung zum Schutz der Oberflächengewässer (Oberflächengewässerverordnung - OGewV)*: OGewV.

Burkart, A., Cogliati, S., Schickling, A. and Rascher, U. 2014 A Novel UAV-Based Ultra-Light Weight Spectrometer for Field Spectroscopy. *IEEE Sensors J.*, **14**(1), 62–67.

Carter, J., Owens, P. N., Walling D. and Leeks, G. J. L. 2003 Fingerprinting suspended sediment sources in a large urban river system. *Science of The Total Environment*, **314-316**, 513–534.

Chapman, D.V. (ed.) 1996 *Water quality assessments*: *A guide to the use of biota, sediments and water in environmental monitoring,* 2. ed. Spon, London.

Chen, J., Wang, F., Meybeck, M., He, D., Xia, X. and Zhang, L. 2005 Spatial and temporal analysis of water chemistry records (1958-2000) in the Huanghe (Yellow River) basin. *Global Biogeochem. Cycles*, **19**(3), 249.

Chipman, J. W., Olmanson, L. G. and Gitelson, A. A. 2009 *Remote sensing methods for lake management*: *A guide for resource managers and decision-makers.* North American Lake Management Society, [Madison, WI].

Ciszewski, D. and Grygar, T. M. 2016 A Review of Flood-Related Storage and Remobilization of Heavy Metal Pollutants in River Systems. *Water, air, and soil pollution*, **227**(7), 239.

Cogliati, S., Rossini, M., Julitta, T., Meroni, M., Schickling, A., Burkart, A., Pinto, F., Rascher, U. and Colombo, R. 2015 Continuous and long-term measurements of reflectance and sun-induced

chlorophyll fluorescence by using novel automated field spectroscopy systems. *Remote Sensing of Environment*, **164**, 270–281.

Correll, D. L. 1999 Phosphorus: a rate limiting nutrient in surface waters. *Poultry science*, **78**(5), 674–682.

Coynel, A., Schäfer, J., Hurtrez, J.-E., Dumas, J., Etcheber, H. and Blanc, G. 2004 Sampling frequency and accuracy of SPM flux estimates in two contrasted drainage basins. *Science of The Total Environment*, **330**(1-3), 233–247.

Davies-Colley, R. J. and Smith, D. G. 2001 TURBIDITY SUSPENI)ED SEDIMENT, AND WATER CLARITY: A REVIEW. *J Am Water Resources Assoc*, **37**(5), 1085–1101.

Deelstra, J. and Iital, A. 2008 The use of the flashiness index as a possible indicator for nutrient loss prediction in agricultural catchments. *Boreal environment research*(13), 209–221.

Deemer, B. R., Harrison, J. A., Li, S., Beaulieu, J. J., DelSontro, T., Barros, N., Bezerra-Neto, J. F., Powers, S. M., dos Santos, M. A. and Vonk, J. A. 2016 Greenhouse Gas Emissions from Reservoir Water Surfaces: A New Global Synthesis. *BioScience*, **66**(11), 949–964.

Dehghani, M. H.i. and Taleb Beydokhti, T. 2018 Investigating the quality and quantity of effluent in wastewater treatment plants of Iran: A case study of Tehran. *MethodsX*, **5**, 871–880.

Dekker, A. G., Brando, V. E., Anstee, J. M., Pinnel, N., Kutser, T., Hoogenboom, E. J., Peters, S., Pasterkamp, R., Vos, R., Olbert, C. and Malthus, T. J.M. 2001 Imaging Spectrometry of Water. In: *Imaging spectrometry*. *Basic principles and prospective applications*, F. van der Meer and S. M. d. Jong (eds.), vol. 4. Remote Sensing and Digital Image Processing v. 4, Kluwer Academic Publishers, Dordrecht, Boston, pp. 307–359.

Dekker, A. G. and Donze, M. 1994 Imaging Spectrometry as a Research Tool for Inland Water Resources Analysis. In: *Imaging Spectrometry — a Tool for Environmental Observations*, J. Hill and J. Mégier (eds.), vol. 4. Eurocourses: Remote Sensing, Springer Netherlands, Dordrecht, pp. 295–317.

Deutsche Vereinigung für Wasserwirtschaft, Abwasser und Abfall e.V. (DWA) 2009-2014 *Leistungsvergleich der kommunalen Kläranlagen Baden-Württemberg*.

Diaz, R. J. 2001 Overview of Hypoxia around the World. *Journal of Environment Quality*, **30**(2), 275.

Dierssen, H. M. and Randolph, K. 2012 Remote Sensing of Ocean Color. In: *Encyclopedia of Sustainability Science and Technology*, R. A. Meyers (ed.), Springer New York, New York, NY, pp. 8952–8975.

Djoukbala, O., Hasbaia, M., Benselama, O. and Mazour, M. 2019 Comparison of the erosion prediction models from USLE, MUSLE and RUSLE in a Mediterranean watershed, case of Wadi Gazouana (N-W of Algeria). *Model. Earth Syst. Environ.*, **5**(2), 725–743.

Dokulil, M. T. and Teubner, K. 2003 Eutrophication and restoration of shallow lakes – the concept of stable equilibria revisited. *Hydrobiologia*, **506**(1), 29–35.

Dudgeon, D., Arthington, A. H., Gessner, M. O., Kawabata, Z.-I., Knowler, D. J., Lévêque, C., Naiman, R. J., Prieur-Richard, A.-H., Soto, D., Stiassny, M. L. J. and Sullivan, C. A. 2006 Freshwater biodiversity: importance, threats, status and conservation challenges. *Biological reviews of the Cambridge Philosophical Society*, **81**(2), 163–182.

Dunne, T. and Reid, L. M. 2016 Sediment budgets as an organizing framework in fluvial geomorphology. In: *Tools in fluvial geomorphology*, G. M. Kondolf and H. Piégay (eds.), Second edition. Advancing river restoration and management, Wiley Blackwell, Chichester, West Sussex, UK, Hoboken, NJ, pp. 357–379.

European Parliament and Council 2000 *DIRECTIVE 2000/60/EC OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 23 Oktober 2000 establishing a framework for Community action in the field of water policy*: *Directive 2000/60/EC*.

Evrard, O., Navratil, O., Ayrault, S., Ahmadi, M., Némery, J., Legout, C., Lefèvre, I., Poirel, A., Bonté, P. and Esteves, M. 2011 Combining suspended sediment monitoring and fingerprinting to determine the spatial origin of fine sediment in a mountainous river catchment. *Earth Surf. Process. Landforms*, **36**(8), 1072–1089.

Facchi, A., Gandolfi, C. and Whelan, M. J. 2007 A comparison of river water quality sampling methodologies under highly variable load conditions. *Chemosphere*, **66**(4), 746–756.

Faria, F. A. M. de, Jaramillo, P., Sawakuchi, H. O., Richey, J. E. and Barros, N. 2015 Estimating greenhouse gas emissions from future Amazonian hydroelectric reservoirs. *Environ. Res. Lett.*, **10**(12), 124019.

Ferguson, R. I. 1982 Channel forms and channel changes. In: *British rivers. John Lewin (ed.). George Allen & Unwin, London, 1981.*, K. J. Gregory (ed.) 7, pp. 90–125.

Förstner, U. 2015 Sediments and the EU-Water Framework Directive: revisiting the Elbe 2015 River Basin Management Plan. *J Soils & Sediments*, **15**(9), 1863–1864.

Förstner, U. and Owens, P. 2007 Sediment Quantity and Quality Issues in River Basins. In: *Sediment Dynamics and Pollutant Mobility in Rivers*. *An Interdisciplinary Approach*, B. Westrich and U. Förstner (eds.). Environmental Science and Engineering, Environmental Science, Springer-Verlag Berlin Heidelberg, Berlin, Heidelberg, pp. 1–15.

Fraser, R. N. 1998 Hyperspectral remote sensing of turbidity and chlorophyll a among Nebraska Sand Hills lakes. *International Journal of Remote Sensing*, **19**(8), 1579–1589.

Frauendorf, J. 2002 *Entwicklung und Anwendung von Fernerkundungsmethoden zur Ableitung von Wasserqualitätsparametern verschiedener Restseen des Braunkohlentagebaus in Mitteldeutschland*. Dissertation, Martin-Luther-Universität Halle-Wittenberg, Halle.

Fryirs, K. 2013 (Dis)Connectivity in catchment sediment cascades: a fresh look at the sediment delivery problem. *Earth Surf. Process. Landforms*, **38**(1), 30–46.

Fryirs, K. and Brierley, G. J. 2001 Variability in sediment delivery and storage along river courses in Bega catchment, NSW, Australia: implications for geomorphic river recovery. *Geomorphology*, **38**(3-4), 237–265.

Fuchs, S., Butz, J. and Bechtel, A. 2004 *Stoffstromanalysen für kleine bis mittlere Flussgebiete als Grundlage für die Planung und Umsetzung von Gewässerschutzmaßnahmen*: *BWC 21003 Laufzeit: 01.10.2001 bis 30.09.2003*, 273 pp.

Fuchs, S., Haritopoulou, T. and Wilhelmi, M. 1996 Biofilms in freshwater ecosystems and their use as a pollutant monitor. *Water science and technology : a journal of the International Association on Water Pollution Research*, **34**(7-8), 137–140.

Fuchs, S., Kaiser, M., Kiemle, L., Kittlaus, S., Rothvoß, S., Toshovski, S., Wagner, A., Wander, R., Weber, T. and Ziegler, S. 2017 Modeling of Regionalized Emissions (MoRE) into Water Bodies: An Open-Source River Basin Management System. *Water*, **9**(4), 239.

Fuchs, S., Mayer, I., Haller, B. and Roth, H. 2014 Lamella settlers for storm water treatment - performance and design recommendations. *Water science and technology : a journal of the International Association on Water Pollution Research*, **69**(2), 278–285.

References

Fuchs, S. and Nickel, J. P. 2018 *Qualitative Untersuchung von Mischwasserentlastungen in Bayern*: *5. Zwischenbericht*. Unpublished, Karlsruhe.

Fuchs, S., Scherer, U., Wander, R., Behrendt, H., Venohr, M., Opitz, D., Hillenbrand, T., Marscheider-Weidemann, F. and Götz, T. 2010 *Calculation of Emissions into Rivers in Germany using the MONERIS Model*: *Nutrients, heavy metals and polycyclic aromatic hydrocarbons,* 1st edn, TEXTE, **46/2010.** Federal Environment Agency (Umweltbundesamt), Dessau-Roßlau.

Gao, M., Zhu, K., Bi, Y. and Hu, Z. 2016 Spatiotemporal patterns of surface-suspended particulate matter in the Three Gorges Reservoir. *Environmental Science and Pollution Research*, **23**(4), 3569–3577.

Gebel, M., Halbfaß, S., Bürger, S. and Uhlig, M. 2018 *Stoffbilanz*: *Modellerläuterung,* Gesellschaft für angewandte Landschaftsforschung bR (GALF). http://www.stoffbilanz.de/home/index.php?lang=de&m2=doku&m1= (accessed 02 July 2019).

Gellis, A. C. 2013 Factors influencing storm-generated suspended-sediment concentrations and loads in four basins of contrasting land use, humid-tropical Puerto Rico. *CATENA*, **104**, 39–57.

Gholizadeh, M. H., Melesse, A. M. and Reddi, L. 2016 A Comprehensive Review on Water Quality Parameters Estimation Using Remote Sensing Techniques. *Sensors (Basel, Switzerland)*, **16**(8), 1–43.

Gonzalez-Hidalgo, J. C., Batalla, R. J., Cerdà, A. and Luis, M. de 2010 Contribution of the largest events to suspended sediment transport across the USA. *Land Degrad. Dev.*, **21**(2), 83–91.

Gordon, H. R. and McCluney, W. R. 1975 Estimation of the depth of sunlight penetration in the sea for remote sensing. *Appl. Opt.*, **14**(2), 413–416.

Gordon, H. R. and Morel, A. Y. 1983 *Remote assessment of ocean color for interpretation of satellite visible imagery*: *A review,* Lecture notes on coastal and estuarine studies, **4.** Springer, New York.

Govender, M., Chetty, K. and Bulcock, H. 2007 A review of hyperspectral remote sensing and its application in vegetation and water resource studies. *Water SA*, **33**(2), 145–152.

Grill, G., Lehner, B., Thieme, M., Geenen, B., Tickner, D., Antonelli, F., Babu, S., Borrelli, P., Cheng, L., Crochetiere, H., Ehalt Macedo, H., Filgueiras, R., Goichot, M., Higgins, J., Hogan, Z., Lip, B., McClain, M. E., Meng, J., Mulligan, M., Nilsson, C., Olden, J. D., Opperman, J. J., Petry, P., Reidy Liermann, C., Sáenz, L., Salinas-Rodríguez, S., Schelle, P., Schmitt, R. J. P., Snider, J., Tan, F., Tockner, K., Valdujo, P. H., van Soesbergen, A. and Zarfl, C. 2019 Mapping the world's free-flowing rivers. *Nature*, **569**(7755), 215–221.

Gruber, G., Bertrand-Krajewski, J.-L., Beneditis, J. D., Hochedlinger, M. and Lettl, W. 2006 Practical aspects, experiences and strategies by using UV/VIS sensors for long-term sewer monitoring. *Water Practice and Technology*, **1**(1).

Harrington, S. T. and Harrington, J. R. 2013 An assessment of the suspended sediment rating curve approach for load estimation on the Rivers Bandon and Owenabue, Ireland. *Geomorphology*, **185**, 27–38.

Hauer, C., Leitner, P., Unfer, G., Pulg, U., Habersack, H. and Graf, W. 2018 The Role of Sediment and Sediment Dynamics in the Aquatic Environment. In: *Riverine Ecosystem Management: Science for Governing Towards a Sustainable Future*, S. Schmutz and J. Sendzimir (eds.), Springer International Publishing, Cham, pp. 151–169.

Hilden, M. 2003 *Ermittlung von Stoff-Frachten in Fliessgewässern*: *Probenahmestrategien und Berechnungsverfahren,* Oberirdische Gewässer : Konzepte und Strategien. Kulturbuch-Verl., Berlin.

Hjulström, F. 1935 Studies of the morphological activity of rivers as illustrated by the river Fyris. @Uppsala, Phil. Diss., 23. Mai 1935. *Bulletin of the Geological Institution of the University of Upsala ; 25*, **10**.

Hollert, H., Dürr, M., Haag, I., Wölz, J., Hilscherova, K., Blaha, L. and Gerbersdorf, S. 2007 Influence of Hydrodynamics on Sediment Ecotoxicity. In: *Sediment Dynamics and Pollutant Mobility in Rivers*. *An Interdisciplinary Approach*, B. Westrich and U. Förstner (eds.). Environmental Science and Engineering, Environmental Science, Springer-Verlag Berlin Heidelberg, Berlin, Heidelberg, pp. 401–416.

Honkavaara, E., Hakala, T., Kirjasniemi, J., Lindfors, A., Mäkynen, J., Nurminen, K., Ruokokoski, P., Saari, H. and Markelin, L. 2013 New Light-Weight Stereosopic Spectrometric Airborne Imaging Technology For High-Resolution Environmental Remote Sensing - Case Studies In Water Quality Mapping. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, **XL-1/W1**, 139–144.

Horowitz, A. J. 2003 An evaluation of sediment rating curves for estimating suspended sediment concentrations for subsequent flux calculations. *Hydrol. Process.*, **17**(17), 3387–3409.

Horowitz, A. J. 2008 Determining annual suspended sediment and sediment-associated trace element and nutrient fluxes. *Science of The Total Environment*, **400**(1-3), 315–343.

Horowitz, A. J. 2013 A review of selected inorganic surface water quality-monitoring practices: are we really measuring what we think, and if so, are we doing it right? *Environmental science & technology*, **47**(6), 2471–2486.

Horowitz, A. J., Clarke, R. T. and Merten, G. H. 2015 The effects of sample scheduling and sample numbers on estimates of the annual fluxes of suspended sediment in fluvial systems. *Hydrol. Process.*, **29**(4), 531–543.

Hudson, N. W. 1993 *Field measurement of soil erosion and runoff,* FAO soils bulletin, **68,** Rome.

Hupfer, M. and Lewandowski, J. 2008 Oxygen Controls the Phosphorus Release from Lake Sediments - a Long-Lasting Paradigm in Limnology. *Internat. Rev. Hydrobiol.*, **93**(4-5), 415–432.

Jain, S. K. and Singh, V. P. 2003 *Water resources systems planning and management,* 1st ed., Developments in water science, **51.** Elsevier, Amsterdam, Boston.

Jerlov, N. G. 1976 *Marine Optics,* 2nd edn, **14.** Elsevier Science, Amsterdam.

Johnes, P. J. 2007 Uncertainties in annual riverine phosphorus load estimation: Impact of load estimation methodology, sampling frequency, baseflow index and catchment population density. *Journal of Hydrology*, **332**(1-2), 241–258.

Karaska, M. A., Huguenin, R. L., Beacham, J. L., Wang, M.-H., Jensen, J. R. and Kaufmann, R. S. 2004 AVIRIS Measurements of Chlorophyll, Suspended Minerals, Dissolved Organic Carbon, and Turbidity in the Neuse River, North Carolina. *photogramm eng remote sensing*, **70**(1), 125–133.

Keller, P. A. 2001 *Imaging spectroscopy of lake water quality parameters*. Diss. phil. II Zürich (Austausch beschränkt). s.n, S.l.

Keller, S., Maier, P. M., Riese, F. M., Norra, S., Holbach, A., Börsig, N., Wilhelms, A., Moldaenke, C., Zaake, A. and Hinz, S. 2018 Hyperspectral Data and Machine Learning for Estimating CDOM, Chlorophyll a, Diatoms, Green Algae and Turbidity. *International journal of environmental research and public health*, **15**(9).

Kentroti, M. 2017 *Relation of water quality parameters with discharge Relation of water quality*

parameters with discharge dynamics for selected catchments in Baden–Württemberg. Master Thesis, Department of Aquatic Environmental Engineering, Karlsruhe Institute for Technology, Karlsruhe.

Khalid A. and Ortiz, J. D. 2016 Multivariate approach for chlorophyll-a and suspended matter retrievals in Case II type waters using hyperspectral data. *Hydrological Sciences Journal*, **61**(1), 200–213.

Kiemle, L., Wagner, A., Hüsener, J., Fuchs, S., Henning, K. and Haile, C. 2019 *Modellierung der Nährstoffeinträge in die Fließgewässer Baden-Württembergs für die Aktualisierung der Bewirtschaftungspläne nach WRRL*: *Modellbeschreibung und Ergebnisse der MONERIS-BW Version „August 2019",* Karlsruhe.

King, K. W. 2003 Considerations in Selecting a Water Quality Sampling Strategy. *TRANSACTIONS OF THE ASAE ONLINE*, **v. 46**(no. 1), pp. 63-73-2003 v.46 no.1.

Kjelland, M. E., Woodley, C. M., Swannack, T. M. and Smith, D. L. 2015 A review of the potential effects of suspended sediment on fishes: potential dredging-related physiological, behavioral, and transgenerational implications. *Environ Syst Decis*, **35**(3), 334–350.

Knaeps, E., Sterckx, S. and Raymaekers, D. 2010 A Seasonally Robust Empirical Algorithm to Retrieve Suspended Sediment Concentrations in the Scheldt River. *Remote Sensing*, **2**(9), 2040–2059.

Kobus und Partner (KUP) 2018 *SIMCOP-Interpolation für Phosphat im Grundwasser*: *Ergebnisse März 2018*.

Kondolf, G. M., Gao, Y., Annandale, G. W., Morris, G. L., Jiang, E., Zhang, J., Cao, Y., Carling, P., Fu, K., Guo, Q., Hotchkiss, R., Peteuil, C., Sumi, T., Wang, H.-W., Wang, Z., Wei, Z., Wu, B., Wu, C. and Yang, C. T. 2014 Sustainable sediment management in reservoirs and regulated rivers: Experiences from five continents. *Earth's Future*, **2**(5), 256–280.

Krasa, J., Dostal, T., Jachymova, B., Bauer, M. and Devaty, J. 2019 Soil erosion as a source of sediment and phosphorus in rivers and reservoirs – Watershed analyses using WaTEM/SEDEM. *Environmental Research*, **171**, 470–483.

Krysanova, V., Müller-Wohlfeil, D.-I. and Becker, A. 1998 Development and test of a spatially distributed hydrological/water quality model for mesoscale watersheds. *Ecological Modelling*, **106**(2-3), 261–289.

Kulasova, A., Smith, P. J., Beven, K. J., Blazkova, S. D. and Hlavacek, J. 2012 A method of computing uncertain nitrogen and phosphorus loads in a small stream from an agricultural catchment using continuous monitoring data. *Journal of Hydrology*, **458-459**, 1–8.

Laanen, M. L. 2007 *Yellow matters*: *Improving the remote sensing of coloured dissolved organic matter in inland freshwaters.* s.n.], [S.l.

Landesamt für Geoinformation und Landentwicklung (LGL) 2013 *ATKIS Basis-DLM NOrA-BW. Tatsächliche Nutzung*.

Landwirtschaftliches Technologiezentrum Augustenberg (LTZ) 2015 *P2O5-Oberbodengehalte der Jahre 2009-2014*.

Lee, H., Swamikannu, X., Radulescu, D., Kim, S.-j. and Stenstrom, M. K. 2007 Design of stormwater monitoring programs. *Water research*, **41**(18), 4186–4196.

Lehner, B., Liermann, C. R., Revenga, C., Vörösmarty, C., Fekete, B., Crouzet, P., Döll, P., Endejan, M., Frenken, K., Magome, J., Nilsson, C., Robertson, J. C., Rödel, R., Sindorf, N. and Wisser, D. 2011 High-resolution mapping of the world's reservoirs and dams for sustainable river-flow management.

Frontiers in Ecology and the Environment, **9**(9), 494–502.

Lempérière, F. and Lafitte, R. 2006 The role of dams in the XXI Century to achieve a sustainable development target 1065. In: *Dams and reservoirs, societies and environment in the 21st century. Proceedings of the International Symposium on Dams in the Societies of the 21st Century, ICOLD-SPANCOLD, 18 June 2006, Barcelona, Spain ; [… as part of the 22nd ICOLD International Congress*, L. Berga (ed.). Balkema-proceedings and monographs in engineering, water and earth sciences, Taylor & Francis, London, pp. 1065–1072.

Li, B., Morris, J. and Martin, E. B. 2002 Model selection for partial least squares regression. *Chemometrics and Intelligent Laboratory Systems*, **64**(1), 79–89.

Liu, Y., Villalba, G., Ayres, R. U. and Schroder, H. 2008 Global Phosphorus Flows and Environmental Impacts from a Consumption Perspective. *Journal of Industrial Ecology*, **12**(2), 229–247.

Lodhi, M., Rundquist, D., Han, L. and Kuzila, M. 1998 Estimation of suspended sediment concentration in water using integrated surface reflectance. *GeoCarto International*, **13**(2), 11–15.

Loizeau, J.-L. and Dominik, J. 2000 Evolution of the Upper Rhone River discharge and suspended sediment load during the last 80 years and some implications for Lake Geneva. *Aquat. sci.*, **62**(1), 54.

Lopez, E., Soto, B., Rubinos, D. and Diaz-Fierros, F. 2000 Flow-variation-paced sampling: a method for automatic sampling of streamflow during peak runoff periods. *Journal of Hydrology*, **229**(3-4), 255–264.

LUBW 2016 *Abfluss-BW. Regionalisierte Abfluss-Kennwerte Baden-Württemberg. Mittlere Abflüsse und mittlere Niedrigwasserabflüsse,* Institut für Wasser und Gewässerentwicklung (IWG) - Bereich Hydrologie, Karlsruher Institut für Technologie (KIT), Karlsruhe.

Mahmood, K. 1987 *Reservoir sedimentation*: *Impact, extent, and mitigation,* 1. publ, World Bank technical paper, **71.** World Bank, Washington DC.

Malthus, T. J., Hestir, E. L., Dekker, A. G. and Brando, V. E. 2012 The case for a global inland water quality product. In: *IGARSS 2012 - 2012 IEEE International Geoscience and Remote Sensing Symposium*, pp. 5234–5237.

Martínez-Carreras, N., Schwab, M. P., Klaus, J. and Hissler, C. 2016 In situ and high frequency monitoring of suspended sediment properties using a spectrophotometric sensor. *Hydrol. Process.*, **30**(19), 3533–3540.

Matthews, M. W. 2011 A current review of empirical procedures of remote sensing in inland and near-coastal transitional waters. *International Journal of Remote Sensing*, **32**(21), 6855–6899.

Mevik, B.-H. and Cederkvist, H. R. 2004 Mean squared error of prediction (MSEP) estimates for principal component regression (PCR) and partial least squares regression (PLSR). *J. Chemometrics*, **18**(9), 422–429.

Mevik, B.-H. and Wehrens, R. 2007 The pls Package: Principal Component and Partial Least Squares Regression in R. *Journal of Statistical Software*, **18**(2), 1–23.

Meybeck, M. and Helmer, R. 1989 The quality of rivers: From pristine stage to global pollution. *Palaeogeography, Palaeoclimatology, Palaeoecology*, **75**(4), 283–309.

Miedema, S. A. 2010 Constructing the shields curve, a new theoretical approach and its applications. *WODCON XIX, Beijing, China*.

Milliman, J. D. and Farnsworth, K. L. 2013 *River discharge to the coastal ocean*: *A global synthesis,* 1st

References

time paperback. Cambridge University Press, New York.

Minella, J.P.G., Merten, G. H., Walling, D. E. and Reichert, J. M. 2009 Changing sediment yield as an indicator of improved soil management practices in southern Brazil. *CATENA*, **79**(3), 228–236.

Mize, S. V., Murphy, J. C., Diehl, T. H. and Demcheck, D. K. 2018 Suspended-sediment concentrations and loads in the lower Mississippi and Atchafalaya rivers decreased by half between 1980 and 2015. *Journal of Hydrology*, **564**, 1–11.

Moatar, F., Floury, M., Renard, B., Meybeck, M., Piffady, J., Chandesris, A., Gold, A., Lowder, K.A. and Pinay, G. 2019 A novel framework to characterize solute and sediment export regime andoptimize their monitoring. In: *Geophysical Research Abstracts, EGU General Assembly 2019*, vol. 21.

Moriasi, D. N., Arnold, J. G., van Liew, M. W., Bingner, R. L., Harmel, R. D. and Veith, T. L. 2007 Model Evaluation Guidelines for Systematic Quantification of Accuracy in Watershed Simulations. *Transactions of the ASABE*, **50**(3), 885–900.

Mucha, Z. and Kułakowski, P. 2016 Turbidity measurements as a tool of monitoring and control of the SBR effluent at the small wastewater treatment plant – preliminary study. *Archives of Environmental Protection*, **42**(3), 33–36.

Multidisciplinary data acquisition as the key for a globally applicable water resource management (MuDak-WRM) 2019 *Data provided or used in the project.* Full list of partners and project description: https://www.mudak-wrm.kit.edu/.

Nash, J. E. and Sutcliffe, J. V. 1970 River flow forecasting through conceptual models part I — A discussion of principles. *Journal of Hydrology*, **10**(3), 282–290.

Ndomba, P., Mtalo, F. and Killingtveit, Å. 2008 A guided swat model application on sediment yield modeling in Pangani river basin: lessons learnt. *JUEE*, **2**(2), 53–62.

Nickel, J. P. and Fuchs, S. 2019 Micropollutant emissions from combined sewer overflows: In review. *Water science and technology*.

Olmanson, L. G., Brezonik, P. L. and Bauer, M. E. 2013 Airborne hyperspectral remote sensing to assess spatial distribution of water quality characteristics in large rivers: The Mississippi River and its tributaries in Minnesota. *Remote Sensing of Environment*, **130**, 254–265.

Owens, P. 2005 Conceptual Models and Budgets for Sediment Management at the River Basin Scale (12 pp). *J Soils & Sediments*, **5**(4), 201–212.

Owens, P. N., Batalla, R. J., Collins, A. J., Gomez, B., Hicks, D. M., Horowitz, A. J., Kondolf, G. M., Marden, M., Page, M. J., Peacock, D. H., Petticrew, E. L., Salomons, W. and Trustrum, N. A. 2005 Fine-grained sediment in river systems: environmental significance and management issues. *River Res. Applic.*, **21**(7), 693–717.

Padial, A. A. and Thomaz, S. M. 2008 Prediction of the light attenuation coefficient through the Secchi disk depth: Empirical modeling in two large Neotropical ecosystems. *Limnology*, **9**(2), 143–151.

Paerl, H. W. 2009 Controlling Eutrophication along the Freshwater–Marine Continuum: Dual Nutrient (N and P) Reductions are Essential. *Estuaries and Coasts*, **32**(4), 593–601.

Petts, G.E. and Amoros, C. (eds.) 1996 *Fluvial Hydrosystems.* Springer Netherlands, Dordrecht.

Pimentel, D. 2006 Soil Erosion: A Food and Environmental Threat. *Environ Dev Sustain*, **8**(1), 119–137.

Pohle, I., Glendell, M., Baggaley, N. and Stutter, M. 2019 A classification scheme for concentration-discharge relationships based on long-term low-frequency water quality data. In: *Geophysical*

Research Abstracts, EGU General Assembly 2019, vol. 21.

Qiu, L., Zheng, F. and Ying, R. 2012 SWAT-based runoff and sediment simulation in a small watershed, the loessial hilly-gullied region of China: capabilities and challenges. *International Journal of Sediment Research*, **27**(2), 226–234.

Rahmani, V., Kastens, J., deNoyelles, F., Jakubauskas, M., Martinko, E., Huggins, D., Gnau, C., Liechti, P., Campbell, S., Callihan, R. and Blackwood, A. 2018 Examining Storage Capacity Loss and Sedimentation Rate of Large Reservoirs in the Central U.S. Great Plains. *Water*, **10**(2), 190.

Randerson, J. T., Chapin, F. S., Harden, J. W., Neff, J. C. and Harmon, M. E. 2002 NET ECOSYSTEM PRODUCTION: A COMPREHENSIVE MEASURE OF NET CARBON ACCUMULATION BY ECOSYSTEMS. *Ecological Applications*, **12**(4), 937–947.

Rauen, W. B., Castro, C. O. de and da Silva, M. G. 2017 Caracterização Hidrossedimentológica Do Rio Passauna, PR, Brasil, A Partir De Dados Históricos. In: *XX Simpósio Brasileiro de Recursos Hídricos*, Associação Brasileira De Recursos Hídricos (ed.).

Regierungspräsidium Freiburg, Landesamt für Geologie, Rohstoffe und Bergbau (RP Freiburg, LGRB) 2015a *Bodenerosion durch Wasser in Baden-Württemberg. Ermittlung mit der Allgemeinen Bodenabtragsgleichung (ABAG)*.

Regierungspräsidium Freiburg, Landesamt für Geologie, Rohstoffe und Bergbau (RP Freiburg, LGRB) 2015b *Bodenkarte von Baden-Württemberg 1 : 50 000 (BK50)*: *GeoFachdaten BW - Boden (BK-BW)*, Freiburg im Breisgau.

Reyna, L., Dube, F., Barrera, J. A. and Zagal, E. 2017 Potential Model Overfitting in Predicting Soil Carbon Content by Visible and Near-Infrared Spectroscopy. *Applied Sciences*, **7**(7), 708.

Riechel, M., Matzinger, A., Pawlowsky-Reusing, E., Sonnenberg, H., Uldack, M., Heinzmann, B., Caradot, N., Seggern, D. von and Rouault, P. 2016 Impacts of combined sewer overflows on a large urban river - Understanding the effect of different management strategies. *Water research*, **105**, 264–273.

Rieger, L., Vanrolleghem, P. A., Langergraber, G., Kaelin, D. and Siegrist, H. 2008 Long-term evaluation of a spectral sensor for nitrite and nitrate. *Water science and technology : a journal of the International Association on Water Pollution Research*, **57**(10), 1563–1569.

Riggsbee, J. A., Orr, C. H., Leech, D. M., Doyle, M. W. and Wetzel, R. G. 2008 Suspended sediments in river ecosystems: Photochemical sources of dissolved organic carbon, dissolved organic nitrogen, and adsorptive removal of dissolved iron. *J. Geophys. Res.*, **113**(G3), 735.

Ritchie, J. C., Schiebe, F. R. and McHenry, J. R. 1976 Remote Sensing of Suspended Sediments in Surface Water. *Photogrammetric Engineering and Remote Sensing*, **42**(12), 1539–1545.

Ritchie, J. C., Zimba, P. V. and Everitt, J. H. 2003 Remote Sensing Techniques to Assess Water Quality. *photogramm eng remote sensing*, **69**(6), 695–704.

Rügner, H., Schwientek, M., Egner, M. and Grathwohl, P. 2014 Monitoring of event-based mobilization of hydrophobic pollutants in rivers: Calibration of turbidity as a proxy for particle facilitated transport in field and laboratory. *Science of The Total Environment*, **490**, 191–198.

Salomons, W. and Förstner, U. 1984 *Metals in the Hydrocycle.* Springer Berlin Heidelberg, Berlin, Heidelberg.

Salomons, W. and Förstner, U. 2010 Sediments and the "System": From site-specific to regional-scale research. *J Soils Sediments*, **10**(8), 1436–1439.

References

Savitzky, A. and Golay, M. 1964 Smoothing and Differentiation of Data by Simplified Least Squares Procedures. *Analytical Chemistry*, **36**(8), 1627–1638.

Schalles, J. F., Rundquist, D. and Schiebe, F. R. 2001 The influence of suspended clays on phytoplankton reflectance signatures and the remote estimation of chlorophyll. *Verhandlungen des Internationalen Verein Limnologie*(27), 3619–3625.

Schleiss, A. J., Franca, M. J., Juez, C. and Cesare, G. de 2016 Reservoir sedimentation. *Journal of Hydraulic Research*, **54**(6), 595–614.

Seitzinger, S. P., Mayorga, E., Bouwman, A. F., Kroeze, C., Beusen, A. H. W., Billen, G., van Drecht, G., Dumont, E., Fekete, B. M., Garnier, J. and Harrison, J. A. 2010 Global river nutrient export: A scenario analysis of past and future trends. *Global Biogeochem. Cycles*, **24**(4), n/a-n/a.

Shepard, D. 1968 A two-dimensional interpolation function for irregularly-spaced data. In: *Proceedings of the 1968 23rd ACM national conference on -*, R. B. Blue and A. M. Rosenberg (eds.), ACM Press, New York, New York, USA, pp. 517–524.

Silván-Cárdenas, J. L. and Wang, L. 2010 Retrieval of subpixel Tamarix canopy cover from Landsat data along the Forgotten River using linear and nonlinear spectral mixture models. *Remote Sensing of Environment*, **114**(8), 1777–1790.

Soonthornnonda, P. and Christensen, E. R. 2008 Source apportionment of pollutants and flows of combined sewer wastewater. *Water research*, **42**(8-9), 1989–1998.

Stelzer, R. S. and Likens, G. E. 2006 Effects of sampling frequency on estimates of dissolved silica export by streams: The role of hydrological variability and concentration-discharge relationships. *Water Resour. Res.*, **42**(7), 1973.

Stevens, R. J. and Smith, R. V. 1978 A comparison of discrete and intensive sampling for measuring the loads of nitrogen and phosphorus in the river main, County Antrim. *Water research*, **12**(10), 823–830.

Su, T.-C. and Chou, H.-T. 2015 Application of Multispectral Sensors Carried on Unmanned Aerial Vehicle (UAV) to Trophic State Mapping of Small Reservoirs: A Case Study of Tain-Pu Reservoir in Kinmen, Taiwan. *Remote Sensing*, **7**(8), 10078–10097.

Sultana, Q. 2017 Useful Life of a Reservoir and its Dependency on Watershed Activities. *ARTOAJ*, **8**(5).

Syvitski, J. P. M., Vörösmarty, C. J., Kettner, A. J. and Green, P. 2005 Impact of Humans on the Flux of Terrestrial Sediment to the Global Coastal Ocean. *Science*, **308**(5720), 376–380.

Teodoru, C. R., Bastien, J., Bonneville, M.-C., del Giorgio, P. A., Demarty, M., Garneau, M., Hélie, J.-F., Pelletier, L., Prairie, Y. T., Roulet, N. T., Strachan, I. B. and Tremblay, A. 2012 The net carbon footprint of a newly created boreal hydroelectric reservoir. *Global Biogeochem. Cycles*, **26**(2), n/a-n/a.

Tetzlaff, B. and Wendland, F. 2007 Modeling P-Fluxes from Diffuse and Point Sources in Heterogeneous Macroscale River Basins Using MEPhos. In: *Sediment Dynamics and Pollutant Mobility in Rivers*. *An Interdisciplinary Approach*, B. Westrich and U. Förstner (eds.). Environmental Science and Engineering, Environmental Science, Springer-Verlag Berlin Heidelberg, Berlin, Heidelberg, pp. 206–215.

Tolk, B. L., Han, L. and Rundquist, D. C. 2000 The impact of bottom brightness on spectral reflectance of suspended sediments. *International Journal of Remote Sensing*, **21**(11), 2259–2268.

Torres, A. and Bertrand-Krajewski, J.-L. 2008 Partial Least Squares local calibration of a UV-visible spectrometer used for in situ measurements of COD and TSS concentrations in urban drainage
124

systems. *Water science and technology*, **57**(4), 581–588.

Tranvik, L. J., Downing, J. A., Cotner, J. B., Loiselle, S. A., Striegl, R. G., Ballatore, T. J., Dillon, P., Finlay, K., Fortino, K., Knoll, L. B., Kortelainen, P. L., Kutser, T., Larsen, S., Laurion, I., Leech, D. M., McCallister, S. L., McKnight, D. M., Melack, J. M., Overholt, E., Porter, J. A., Prairie, Y., Renwick, W. H., Roland, F., Sherman, B. S., Schindler, D. W., Sobek, S., Tremblay, A., Vanni, M. J., Verschoor, A. M., Wachenfeldt, E. von and Weyhenmeyer, G. A. 2009 Lakes and reservoirs as regulators of carbon cycling and climate. *Limnol. Oceangr.*, **54**(6part2), 2298–2314.

Tyler, A. N., Svab, E., Preston, T., Présing, M. and Kovács, W. A. 2006 Remote sensing of the water quality of shallow lakes: A mixture modelling approach to quantifying phytoplankton in water characterized by high-suspended sediment. *International Journal of Remote Sensing*, **27**(8), 1521–1537.

Uhlenbrook, S., Roser, S. and Tilch, N. 2004 Hydrological process representation at the meso-scale: the potential of a distributed, conceptual catchment model. *Journal of Hydrology*, **291**(3-4), 278–296.

VDLUFA 2018 *Phosphordüngung nach Bodenuntersuchung und Pflanzenbedarf,* Speyer.

Vercruysse, K., Grabowski, R. C. and Rickson, R. J. 2017 Suspended sediment transport dynamics in rivers: Multi-scale drivers of temporal variation. *Earth-Science Reviews*, **166**, 38–52.

Verstraeten, G., Rommens, T., Peeters, I., Poesen, J., Govers, G. and Lang, A. 2009 A temporarily changing Holocene sediment budget for a loess-covered catchment (central Belgium). *Geomorphology*, **108**(1-2), 24–34.

Vörösmarty, C. J., Meybeck, M., Fekete, B., Sharma, K., Green, P. and Syvitski, J. P. M 2003 Anthropogenic sediment retention: major global impact from registered river impoundments. *The supply of flux of sediment along hydrological pathways: Anthropogenic influences at the global scale*, **39**(1–2), 169–190.

Voulvoulis, N., Arpon, K. D. and Giakoumis, T. 2017 The EU Water Framework Directive: From great expectations to problems with implementation. *The Science of the total environment*, **575**, 358–366.

Wagner, A., Hilgert, S., Kattenborn, T. and Fuchs, S. 2018 Proximal VIS-NIR spectrometry to retrieve substance concentrations in surface waters using partial least squares modelling. *Water Science and Technology: Water Supply*.

Wagner, A., Hilgert, S., Kishi, R. T., Drummond, S., Kiemle, L., Nickel, J. P., Sotiri, K. and Fuchs, S. 2019 Flow-proportional large volume composite sampling to assess substance fluxes: EGU2019-18649. In: *Geophysical Research Abstracts, EGU General Assembly 2019*, vol. 21.

Walling, D. E. 1977 Assessing the accuracy of suspended sediment rating curves for a small basin. *Water Resour. Res.*, **13**(3), 531–538.

Walling, D. E. 1983 The sediment delivery problem. *Journal of Hydrology*, **65**(1-3), 209–237.

Walling, D. E. 2006 Human impact on land–ocean sediment transfer by the world's rivers. *Geomorphology*, **79**(3-4), 192–216.

Walling, D. E., Owens, P. N. and Leeks, G. J. L. 1999 Fingerprinting suspended sediment sources in the catchment of the River Ouse, Yorkshire, UK. *Hydrol. Process.*, **13**(7), 955–975.

Walling, E. (ed.) 1996 *Erosion and sediment yield*: *Global and regional perspectives : proceedings of an international symposium held at Exeter, UK, from 15 to 19 July 1996,* IAHS publication, **236.** IAHS Press, Wallingford.

Wang, G., Wu, B. and Wang, Z.-Y. 2005 Sedimentation problems and management strategies of

Sanmenxia Reservoir, Yellow River, China. *Water Resour. Res.*, **41**(9), 3.

Wang, H., Yang, Z., Saito, Y., Liu, J. P., Sun, X. and Wang, Y. 2007 Stepwise decreases of the Huanghe (Yellow River) sediment load (1950–2005): Impacts of climate change and human activities. *The supply of flux of sediment along hydrological pathways: Anthropogenic influences at the global scale*, **57**(3-4), 331–354.

Wang, Z.-Y. and Hu, C. 2009 Strategies for managing reservoir sedimentation. *International Journal of Sediment Research*, **24**(4), 369–384.

Ward, J. V. and Stanford, J. A. 1995 Ecological connectivity in alluvial river ecosystems and its disruption by flow regulation. *Regul. Rivers: Res. Mgmt.*, **11**(1), 105–119.

Wilber, D. H. and Clarke, D. G. 2001 Biological Effects of Suspended Sediments: A Review of Suspended Sediment Impacts on Fish and Shellfish with Relation to Dredging Activities in Estuaries. *North American Journal of Fisheries Management*, **21**(4), 855–875.

Wisser, D., Frolking, S., Hagen, S. and Bierkens, M. F. P. 2013 Beyond peak reservoir storage? A global estimate of declining water storage capacity in large reservoirs. *Water Resour. Res.*, **49**(9), 5732–5739.

Wold, S., Sjöström, M. and Eriksson, L. 2001 PLS-regression: A basic tool of chemometrics. *Chemometrics and Intelligent Laboratory Systems*, **58**(2), 109–130.

Wu, J.-L., Ho, C.-R., Huang, C.-C., Srivastav, A. L., Tzeng, J.-H. and Lin, Y.-T. 2014 Hyperspectral Sensing for Turbid Water Quality Monitoring in Freshwater Rivers: Empirical Relationship between Reflectance and Turbidity and Total Solids. *Sensors (Basel, Switzerland)*, **14**(12), 22670–22688.

Yang, D., Kanae, S., Oki, T., Koike, T. and Musiake, K. 2003 Global potential soil erosion with reference to land use and climate changes. *Hydrol. Process.*, **17**(14), 2913–2928.

Zang, W., Lin, J., Wang, Y. and Tao, H. 2012 Investigating small-scale water pollution with UAV Remote Sensing Technology. In: *World Automation Congress (WAC), 2012. 24-28 June 2012, Puerto Vallarta, Mexico*, IEEE, Piscataway, NJ, pp. 1–4.

Zarfl, C., Lumsdon, A. E., Berlekamp, J., Tydecks, L. and Tockner, K. 2015 A global boom in hydropower dam construction. *Aquat. sci.*, **77**(1), 161–170.

Zehe, E., Maurer, T., Ihringer, J. and Plate, E. 2001 Modeling water flow and mass transport in a loess catchment. *Physics and Chemistry of the Earth, Part B: Hydrology, Oceans and Atmosphere*, **26**(7-8), 487–507.

Zheng, M. 2018 *Assessment of sediment and nutrient fluxes from discharge measurements by means of Large Volume Sampling and other quantification approaches for the Kraichbach catchment*. Master Thesis, Institute for Water and River Basin Management, Karlsruhe Institute for Technology, Karlsruhe.

APPENDIX
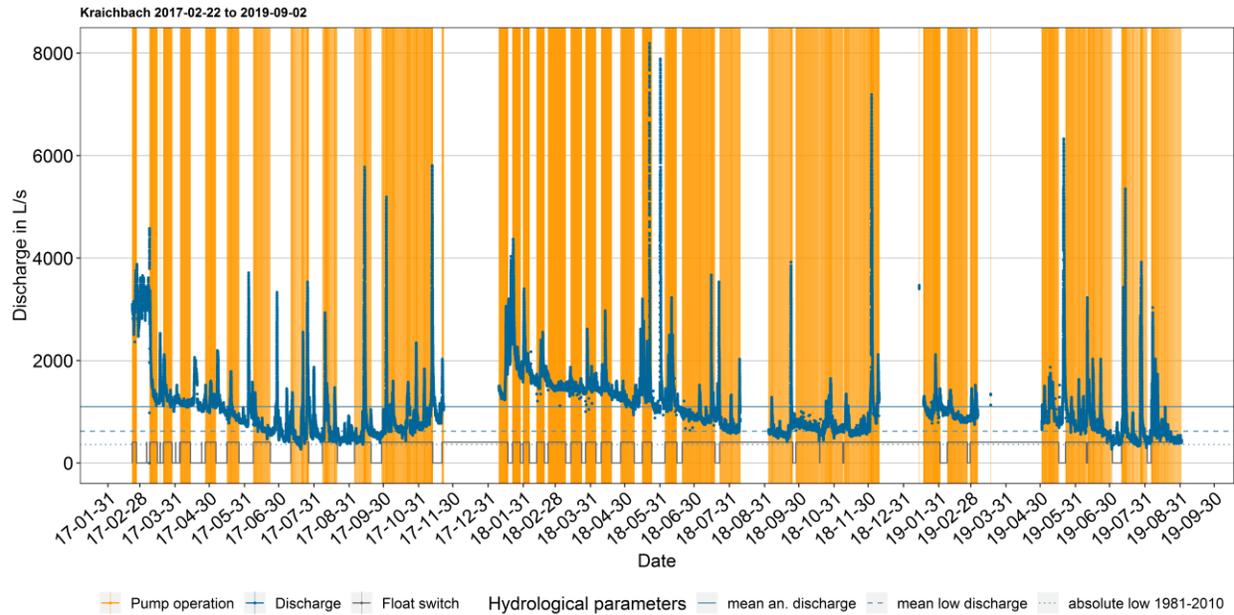
Figure 63: Kraichbach hydrograph (dark blue) during the large volume sampling period. Orange vertical lines indicate times when the pump was active. The float switch (grey line) terminated the sampling when the tank was full. Mean annual discharge, mean low discharge and the lowest discharge recorded between 1981 and 2010 are indicated.
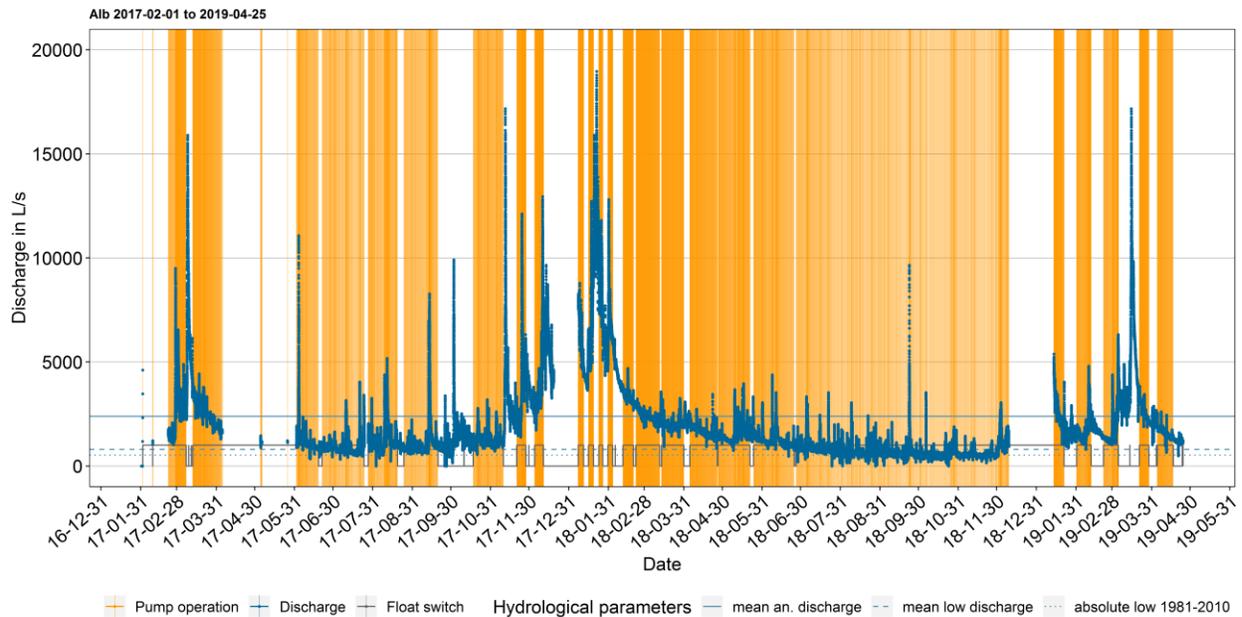
Figure 64: Alb hydrograph (blue) during the large volume sampling period. Orange vertical lines indicate times when the pump was active. The float switch (grey line) terminated the sampling when the tank was full. Mean annual discharge, mean low discharge and the lowest discharge recorded between 1981 and 2010 are indicated.
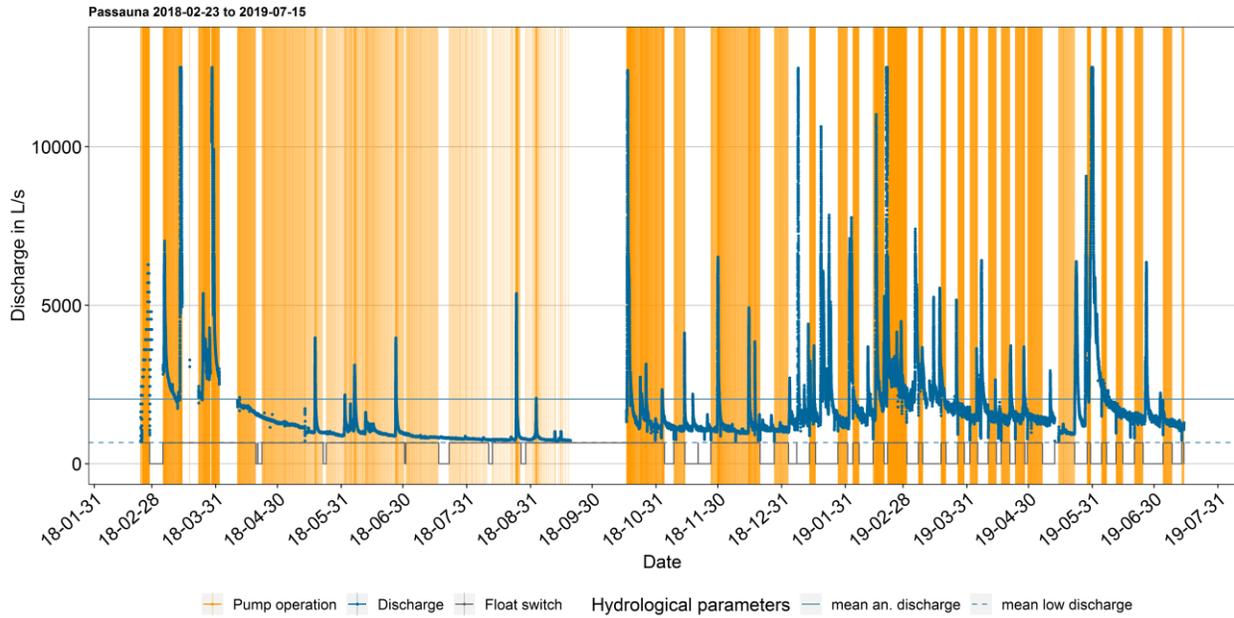
Figure 65: Passauna hydrograph (blue) during the large volume sampling period. Orange vertical lines indicate times when the pump was active. The float switch (grey line) terminated the sampling when the tank was full. Mean annual discharge and mean low discharge are indicated.
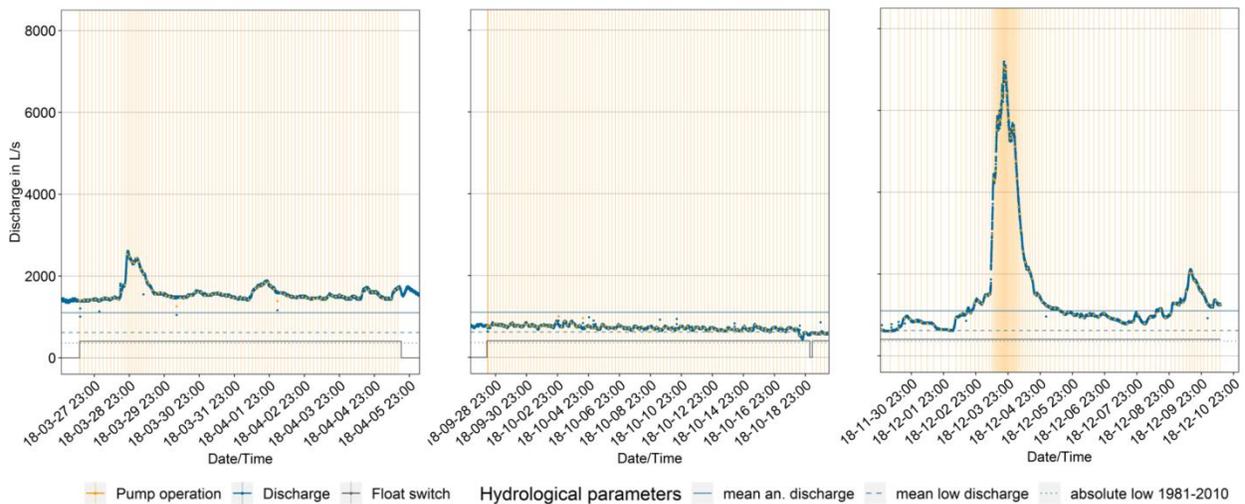
Figure 66: Detailed view of three selected large volume sampling periods at Kraichbach. Left: Wet pre-conditions in spring with flows above mean flow. Middle: Low flow conditions after prolonged drought. Right: Flood event. Orange vertical lines indicate times when the pump was active. The float switch (grey line) terminated the sampling when the tank was full. Mean annual discharge, mean low discharge and the lowest discharge recorded between 1981 and 2010 are indicated.
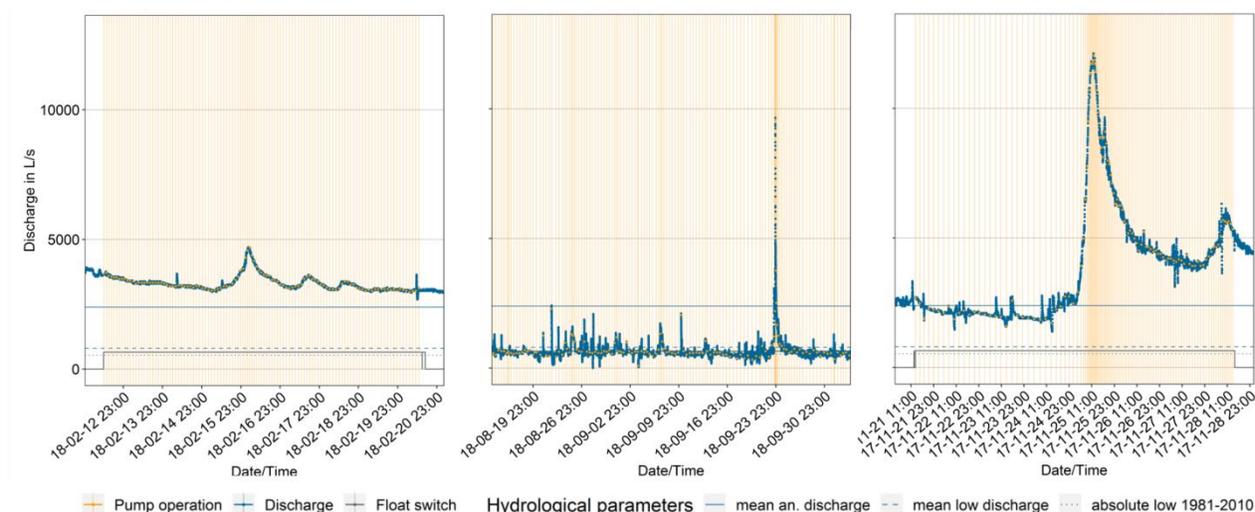
Figure 67: Detailed view of three selected large volume sampling periods at Alb. Left: Wet pre-conditions in spring with flows above mean flow. Middle: Low flow conditions during dry summer with sharp increases due to isolated thunderstorms. Right: Flood event. Orange vertical lines indicate times when the pump was active. The float switch (grey line) terminated the sampling when the tank was full. Mean annual discharge, mean low discharge and the lowest discharge recorded between 1981 and 2010 are indicated.
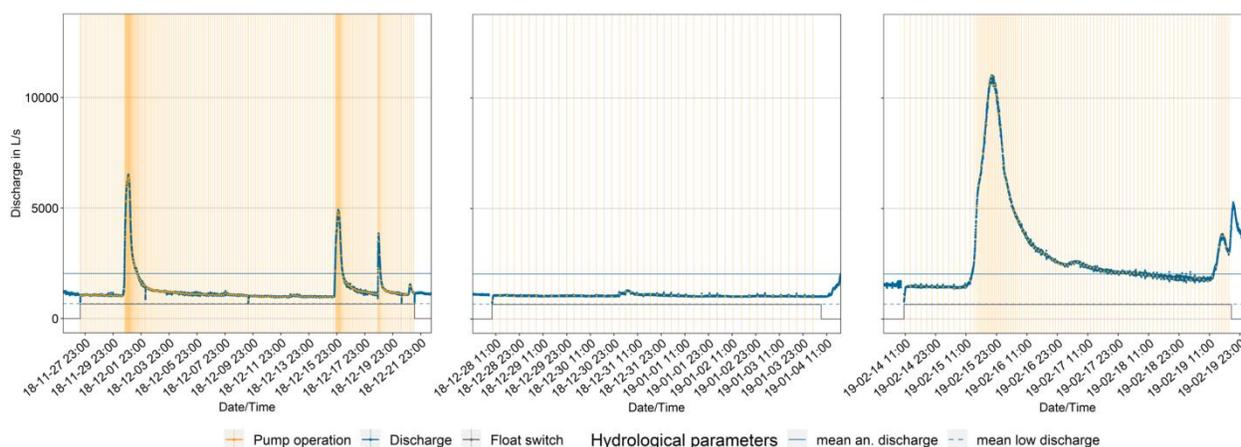
Figure 68: Detailed view of three selected large volume sampling periods at Passauna. Left: Several small peaks during a dry period. Middle: Low flow conditions close to the mean low discharge. Right: Single peak flood event. Orange vertical lines indicate times when the pump was active. The float switch (grey line) terminated the sampling when the tank was full. Mean annual discharge and mean low discharge are indicated.

	SSC in mg/L	TP in mg/L	o-PO$_4$-P in mg/L
2017-03-09 — 2017-03-16	54	0.19	0.09
2017-03-21 — 2017-03-29	26	0.18	0.13
2017-04-05 — 2017-04-14	33	0.15	0.1
2017-04-27 — 2017-05-06	27	0.22	0.17
2017-05-16 — 2017-05-27	56	0.27	0.17
2017-06-08 — 2017-06-23	76	0.26	0.14
2017-07-11 — 2017-07-27	100	0.45	0.24
2017-08-08 — 2017-08-21	58	0.3	0.19
2017-09-05 — 2017-09-20	240	0.47	0.15
2017-09-29 — 2017-10-07	200	0.37	0.11
2017-11-03 — 2017-11-13	160	0.46	0.19
2017-11-21 — 2017-11-22			
2018-01-10 — 2018-01-18	44		0.15
2018-01-22 — 2018-01-29	52	0.2	0.12
2018-01-31 — 2018-02-06	47	0.23	0.13
2018-02-12 — 2018-02-19	30		
2018-02-22 — 2018-03-10	26	0.15	0.11
2018-03-14 — 2018-03-24	18	0.17	0.14
2018-03-27 — 2018-04-05	32	0.16	0.11
2018-04-10 — 2018-04-19	41		
2018-04-27 — 2018-05-09	29		0.16
2018-05-16 — 2018-05-24	830	1.42	0.19
2018-06-05 — 2018-06-15	190	0.53	0.21
2018-06-20 — 2018-06-27	42	0.22	0.17
2018-06-27 — 2018-07-04	39	0.22	0.16
2018-07-04 — 2018-07-19	72	0.29	0.18
2018-07-23 — 2018-08-10	55	0.2	0.12
2018-09-04 — 2018-09-28	99	0.2	0.07
2018-09-28 — 2018-10-19	22	0.09	0.07
2018-10-19 — 2018-11-08			
2018-11-09 — 2018-11-28	12	0.18	0.17
2018-11-30 — 2018-12-10	320	0.55	0.14
2019-01-18 — 2019-02-02	26	0.16	0.12
2019-02-08 — 2019-02-26	16	0.11	0.09
2019-05-02 — 2019-05-17	23	0.17	0.13
2019-05-23 — 2019-06-10	23	0.19	0.15
2019-06-11 — 2019-07-03	31	0.25	0.2
2019-07-11 — 2019-08-03	230	0.53	0.21
2019-08-06 — 2019-09-02	13		0.2

discharge in m$^3 \cdot$s^{-1}

--- two-year flood   — mean annual discharge   – – mean low water discharge   ⋯⋯ absolute low water discharge 1981-2010

Figure 69: Boxplot of hourly discharges for each of the 42 large volume sampling periods at Kraichbach (The 32 in 2017 and 2018 used for constructing rating curves) with event mean substance concentrations of suspended sediment (SSC), total phosphorous (TP) and orthophosphate (o-PO$_4$-P) indicated. Missing values are due to handling or analysis errors. Two-year flood, mean annual discharge, mean low discharge and the lowest discharge recorded between 1981 and 2010 are indicated.

|	SSC in mg/L	TP in mg/L	o-PO₄-P in mg/L
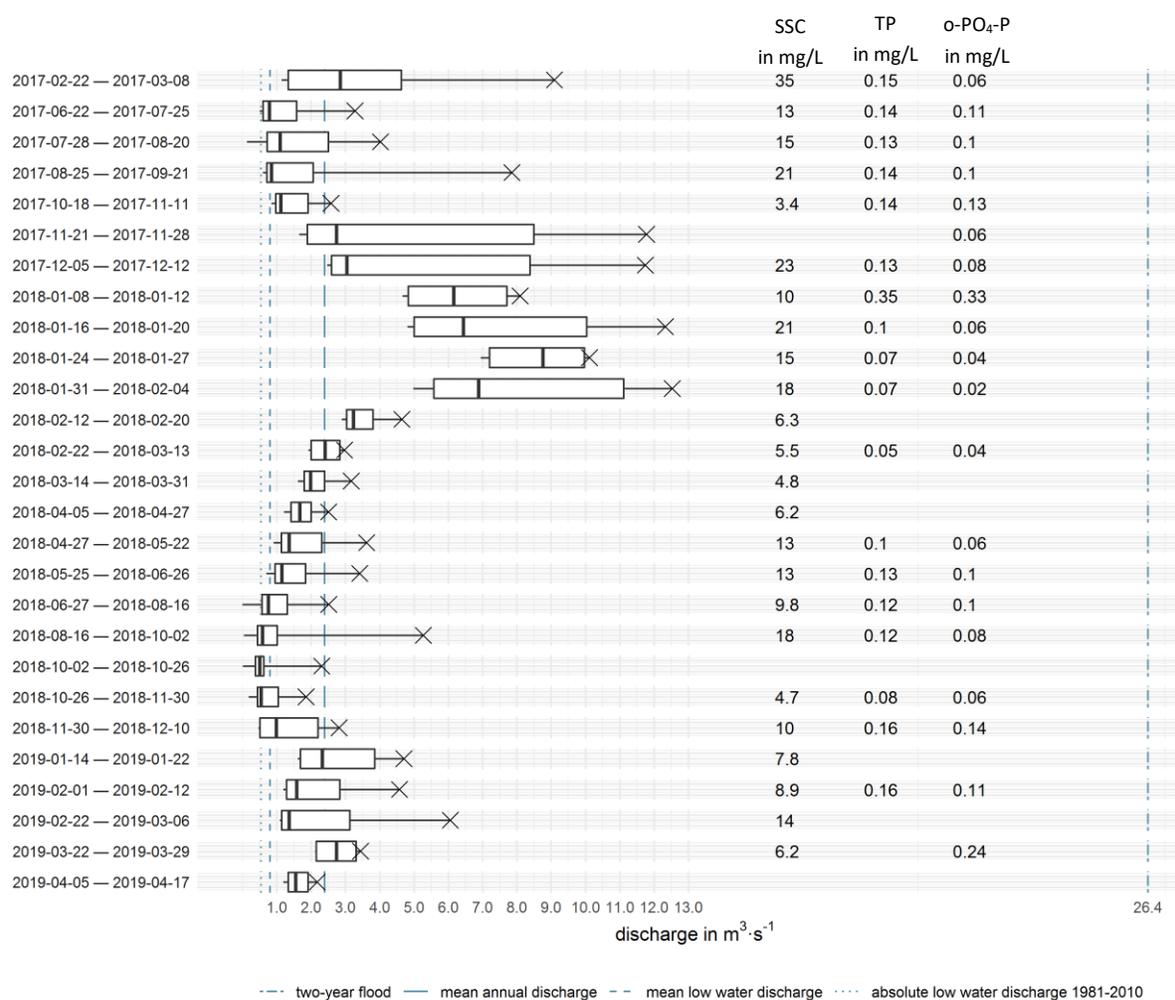
Figure 70: Boxplot of hourly discharges for each of the 32 large volume sampling periods at Alb with event mean substance concentrations of suspended sediment (SSC), total phosphorous (TP) and orthophosphate (o-PO₄-P) indicated. Missing values are due to handling or analysis errors. Two-year flood, mean annual discharge, mean low discharge and the lowest discharge recorded between 1981 and 2010 are indicated.
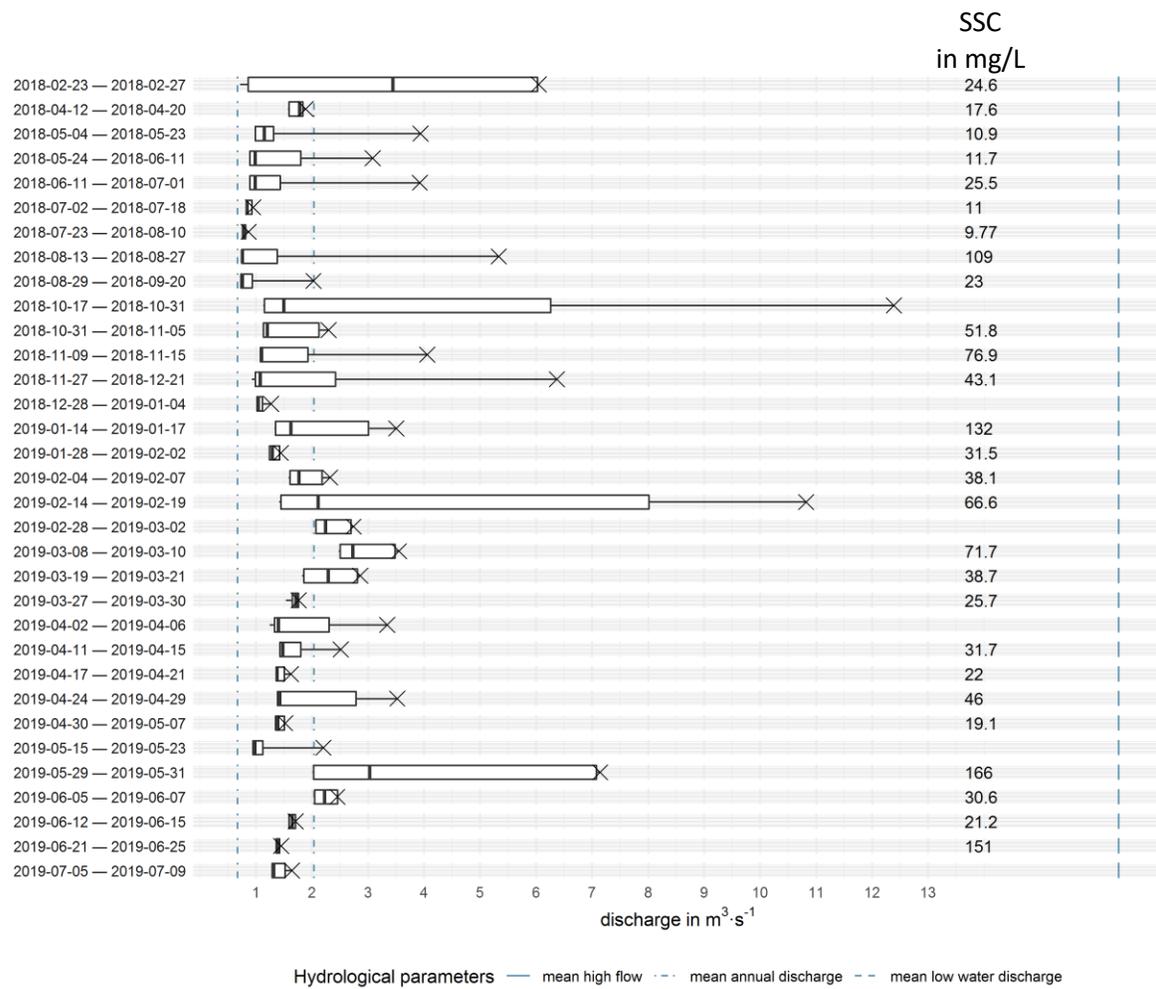
Figure 71: Boxplot of hourly discharges for each of the 33 large volume sampling periods at Passauna River with mean event suspended sediment concentration (SSC) indicated. Missing values are due to handling or analysis errors. Mean high flow, ,mean annual discharge and mean low discharge are indicated.
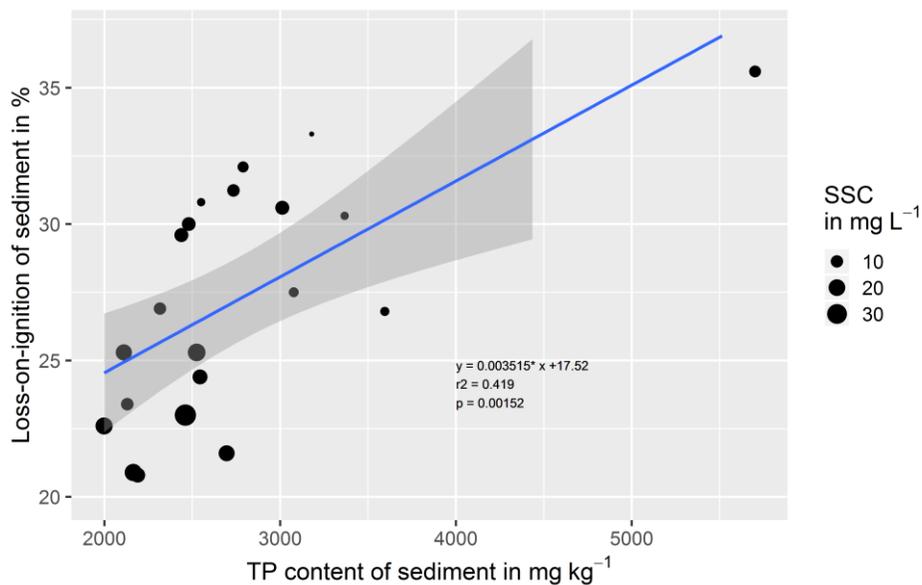
Figure 72: Loss-on-ignition against total phosphorous (TP) content of sediment from large volume composite samples at the Alb River. Linear regression line (blue), equation with R² and p-value indicated. Grey area: 95% confidence level interval for predictions. The mean suspended sediment concentration (SSC) is represented as point size.
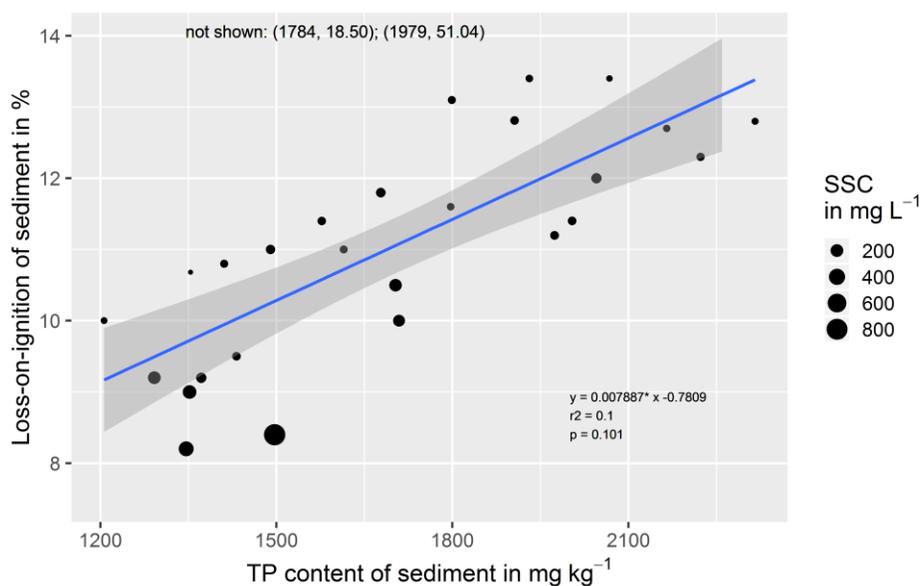
Figure 73: Loss-on-ignition versus total phosphorous (TP) content of sediment from large volume composite samples at the Kraichbach River. Linear regression line (blue), equation with R² and p-value indicated. Grey area: 95% confidence level interval for predictions. The mean suspended sediment concentration (SSC) is represented as point size.
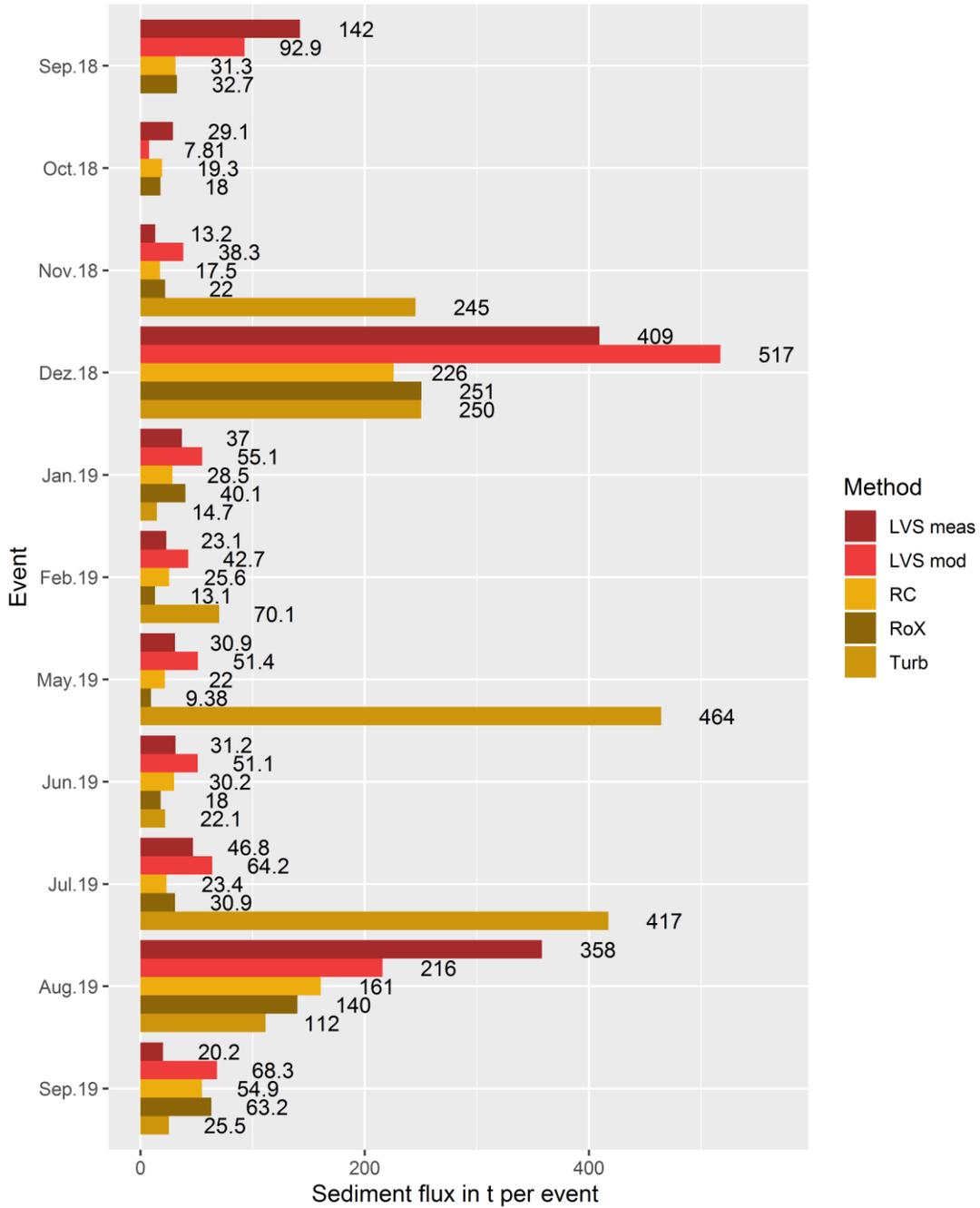
Figure 74: Event fluxes of suspended sediment based on different methods for eleven LVS monitoring periods. LVS: measured LVS flux; RC: Flux from uncapped grab sampling rating curve; RoX: Reflectance-based flux; Turb: Turbidity-based flux. The event names are derived from the end date of the LVS sampling.
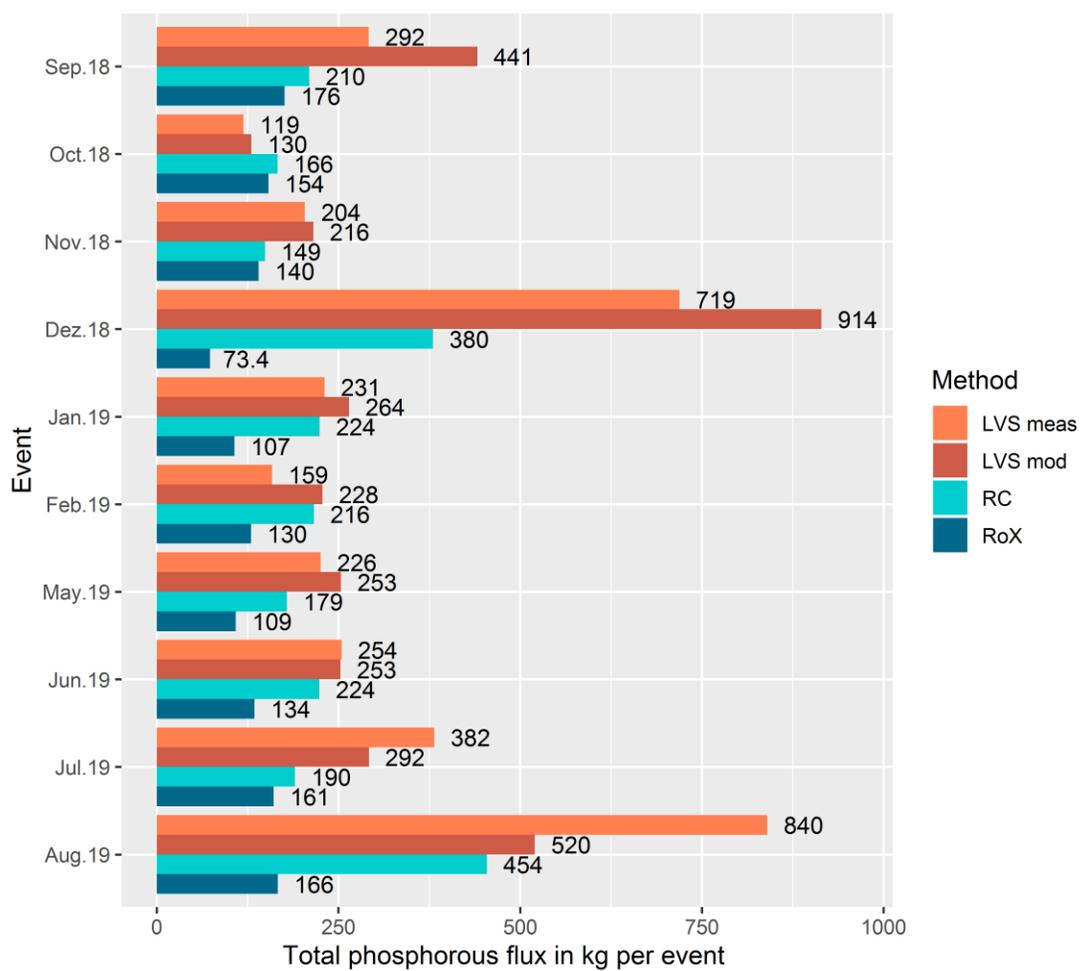
Figure 75: Event fluxes of total phosphorous based on different methods for ten LVS monitoring periods. LVS: measured LVS flux; RC: Flux from uncapped grab sampling rating curve; RoX: Reflectance-based flux. The event names are derived from the end date of the LVS sampling.
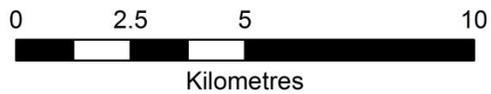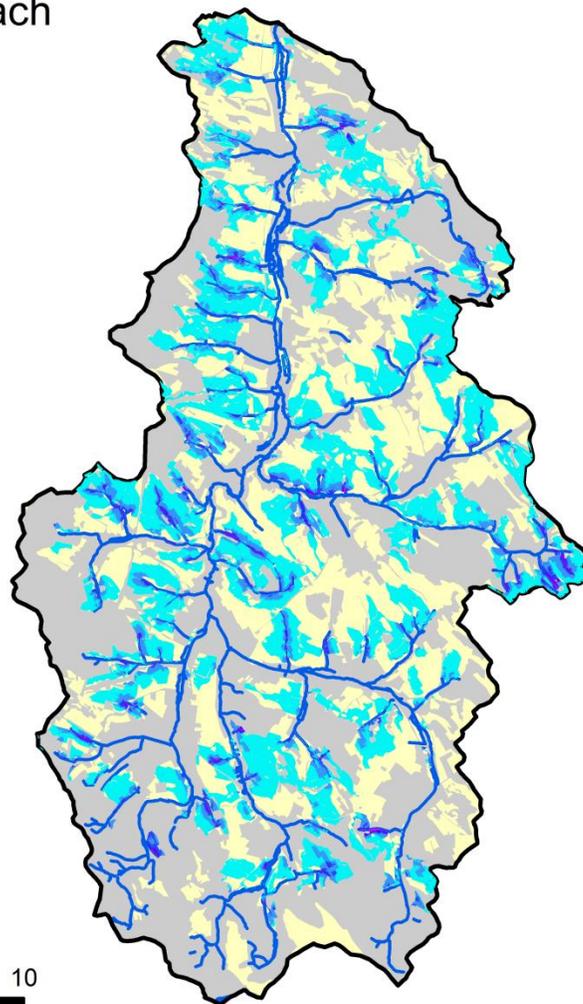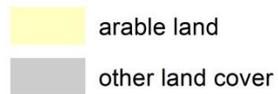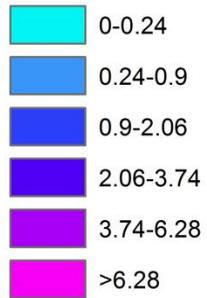
Figure 76: Sediment input from arable land in the Kraichbach catchment, based on soil loss and catchment connectivity (GALF 2018).

Figure 77: Impressions of the Kraichbach catchment shortly after heavy rainfalls in December 2018. Sediment was apparently washed across roads and into drainage systems.

Figure 78: Top: Excavation of sediment deposited during and after the July/August high flow event 2019 in the Kraichbach/Kriegbach flood relief channel, ~7km downstream of the Ubstadt gauge. Bottom: Excavated material and yet to be removed in-channel deposits. Photo by the author, 1st of Oct 2019.

Table 6: Water quality data and discharges from Kraichbach River at/near Ubstadt gauge.

Date	Time (CET)	SSC in mg L^{-1}	TP in mg L-1	PO$_4$-P in mg L-1	Q in m³ s^{-1}
2017-01-17	10:53	12.2			0.71
2017-07-20	16:40	17.2	0.22	0.16	0.49
2017-07-27	15:17	39.0	0.30	0.18	0.69
2017-08-18	10:45	26.0	0.18	0.12	0.45
2017-10-06	14:20	43.2	0.31		0.74
2018-02-02	14:28	31.3			1.58
2018-02-14	11:21	13.5			1.14
2018-03-06	12:55	11.2			1.08
2018-03-27	14:13	8.1	0.12	0.09	1.01
2018-04-10	11:30	14.7	0.09	0.09	1.01
2018-04-27	13:30	13.8			0.89
2018-05-16	12:40	74.5			2.16
2018-06-05	12:40	43.4	0.22	0.13	0.83
2018-06-14	16:00	100.2	0.31		0.83
2018-06-20	14:30	29.8	0.24	0.16	0.71
2018-06-27	13:20	19.3	0.20	0.15	0.71
2018-07-04	16:20	17.9	0.23	0.16	0.66
2018-07-20	12:55	12.7	0.17	0.16	0.66
2018-07-20	13:12	32.4	0.23	0.18	0.66
2018-07-23	13:00	107.4	0.33	0.14	0.71
2018-07-27	15:00	29.9			0.53
2018-07-30	15:15	14.0			0.54
2018-08-03	09:55	26.7	0.27	0.19	0.56
2018-08-03	11:43	20.3	0.24	0.17	0.56
2018-08-10	11:40	42.2	0.27	0.16	0.61
2018-08-20	13:10	21.7	0.22	0.16	0.48
2018-08-20	14:00	21.4	0.20	0.15	0.48
2018-08-14	11:20	19.1	0.15	0.12	0.56
2018-09-04	12:40	16.3	0.18	0.12	0.52
2018-09-04	13:10	23.8	0.23	0.15	0.52
2018-09-13	11:40	13.7	0.14	0.07	0.48
2018-09-13	12:15	22.5	0.11	0.07	0.48
2018-09-20	16:00	17.7	0.10	0.05	0.45
2018-09-20	16:30	13.6	0.10	0.10	0.45
2018-09-28	13:30	7.0	0.10	0.10	0.52
2018-10-15	14:50	11.3	0.14	0.09	0.53
2018-10-15	16:00	6.9	0.16	0.11	0.56
2018-10-19	12:55	8.2	0.15	0.11	0.48
2018-10-31	12:50	7.2			0.48
2018-10-31	12:15	2.9			0.48
2018-11-07	14:15	4.9			0.52
2018-11-07	16:20	6.9			0.52

2018-11-12	10:30	7.2	0.26	0.21	0.56
2018-11-12	11:35	8.2	0.27	0.22	0.56
2018-11-16	10:20	3.3	0.15	0.14	0.48
2018-11-27	16:30	6.4	0.16	0.12	0.61
2018-11-30	15:00	5.9	0.19	0.14	0.53
2018-11-30	13:45	2.4	0.15	0.11	0.52
2018-12-04	14:00	121.0	0.35	0.10	1.50
2018-12-04	13:42	127.0	0.43	0.11	1.50
2018-12-10	14:00	14.8	0.21	0.10	0.89
2018-12-10	14:30	17.2	0.19	0.11	0.89
2018-12-13	13:45	12.1	0.11	0.08	0.56
2019-01-14	13:15	296.0	0.58	0.09	3.39
2019-01-17	12:00	20.5	0.13	0.09	1.01
2019-01-18	13:05	16.2	0.14	0.10	0.95
2019-01-18	13:45	17.3	0.14	0.01	0.95
2019-01-23	17:00	14.3	0.14	0.09	0.77
2019-02-08	10:45	11.3	0.12	0.08	0.77
2019-02-08	12:00	9.7	0.13	0.09	0.77
2019-02-14	10:55	11.0	0.12	0.08	0.71
2019-02-14	12:00	6.1	0.11	0.08	0.71
2019-02-28	11:40	3.3	0.15	0.11	0.66
2019-02-28	12:50	7.2	0.15	0.10	0.66
2019-03-19	11:15		0.09	0.06	1.08
2019-03-22	12:50	5.9	0.08	0.06	0.83
2019-03-22	13:30	4.2	0.10	0.08	0.83
2019-05-21	12:25	459.0	0.87	0.15	3.72
2019-05-23	11:20	19.2	0.17	0.10	1.01
2019-05-23	09:50	33.2	0.15	0.09	1.01

Table 7: Water quality data and discharges from Alb River at/near Ettlingen gauge. Downloaded from http://jdkfg.lubw.baden-wuerttemberg.de on 2019-09-01.

Date	Time (CET)	SSC in mg L^{-1}	PO$_4$-P in mg L-1	Q in m³ s^{-1}
2006-01-23	08:30		0.13	1.18
2006-02-21	09:25		0.12	2.62
2006-03-23	08:45		0.10	3.10
2006-04-19	08:30		0.05	2.85
2006-05-15	13:00		0.12	1.33
2006-06-14	09:40		0.08	1.58
2006-07-12	10:00		0.12	0.91
2006-08-09	09:40		0.14	0.85
2006-09-06	08:00		0.10	1.18
2006-10-04	12:00		0.07	6.85
2006-10-30	10:00		0.07	1.58
2006-11-29	10:00		0.11	1.67
2007-01-15	08:45	1	0.08	1.67
2007-02-12	09:00	11	0.08	8.58
2007-03-13	08:45	3	0.04	3.91
2007-04-11	08:10	4	0.05	2.85
2007-05-10	09:00	7	0.09	2.39
2007-06-04	07:45	4	0.06	2.39
2007-07-04	09:30	3	0.09	2.39
2007-07-30	09:10	30	0.11	2.17
2007-08-29	10:00	3	0.05	1.26
2007-09-26	10:45	1	0.08	1.26
2007-10-22	08:30	1	0.10	0.97
2007-11-21	08:15	1	0.10	1.33
2007-12-17	08:30	4	0.05	4.20
2008-01-15	08:45	3	0.07	2.17
2008-02-14	10:15	1	0.06	2.62
2008-03-13	10:00	8	0.05	6.85
2008-04-08	17:10	3	0.04	5.15
2008-05-06	08:45	11	0.04	2.85
2008-06-03	09:30	14	0.16	1.96
2008-07-01	08:15	1	0.08	1.11
2008-07-28	09:00	3	0.10	0.97
2008-08-26	10:00	1	0.09	0.97
2008-09-24	10:30	1	0.09	0.85
2008-10-23	08:15	5	0.10	3.23
2008-11-17	09:15	9	0.23	2.17
2008-12-16	08:30	1	0.06	2.07
2009-01-13	08:30	1	0.16	1.58
2009-02-09	10:15	3	0.08	1.50
2009-03-12	08:20	8	0.06	7.27
2009-04-07	09:15	3	0.06	2.85

2009-05-05	13:40	5	0.08	1.58
2009-06-04	10:30	5	0.06	1.67
2009-06-30	08:45	4	0.08	1.26
2009-07-29	09:15	1	0.07	1.33
2009-08-26	10:00	5	0.20	1.41
2009-09-22	09:00	1	0.09	0.85
2009-10-19	09:30	1	0.09	1.11
2009-11-18	10:00	4	0.13	2.28
2009-12-15	09:15	8	0.08	2.98
2010-01-18	09:45	39	0.05	4.82
2010-02-15	08:45	1	0.05	2.62
2010-03-15	10:15	3	0.08	2.62
2010-04-13	14:45	4	0.07	1.86
2010-05-10	08:30	8	0.08	1.77
2010-06-07	09:30	24	0.16	2.85
2010-07-07	08:45	3	0.09	1.11
2010-08-03	09:00	220	0.07	2.98
2010-08-30	09:15	7	0.12	2.73
2010-09-29	09:15	4	0.10	1.26
2010-10-25	09:30	3	0.13	1.86
2010-11-23	08:45	6	0.09	3.91
2010-12-20	09:15	9	0.06	6.85
2011-01-17	10:30		1.21	7.70
2011-02-17	08:55		0.01	2.39
2011-03-14	09:45		0.07	1.67
2011-04-13	09:00		0.10	1.50
2011-05-09	10:25		0.08	0.97
2011-06-06	09:45		0.12	1.41
2011-07-04	11:00		0.10	0.85
2011-08-04	13:35		0.21	0.97
2011-08-29	09:00		0.09	0.91
2011-09-27	08:45		0.11	0.73
2011-10-26	09:00		0.10	0.85
2011-11-21	09:00		0.10	0.73
2011-12-19	09:45		0.09	3.36
2012-01-16	08:45		0.06	4.20
2012-02-16	08:45		0.09	2.07
2012-03-13	08:15		0.05	1.58
2012-04-10	09:15		0.06	1.18
2012-05-10	09:00		0.07	1.11
2012-06-06	08:00		0.13	1.11
2012-07-02	12:30		0.15	1.41
2012-08-09	09:05		0.08	0.85

2012-08-27	08:30		0.18	0.91
2012-09-25	08:15		0.10	0.85
2012-10-23	10:10		0.07	0.91
2012-11-22	09:00		0.06	1.26
2012-12-18	09:00		0.05	9.44
2013-01-16	09:00	6	0.04	2.85
2013-02-13	09:00	5	0.05	4.35
2013-03-12	09:05	4	0.05	2.50
2013-04-10	08:45	5	0.41	2.50
2013-05-07	08:45	18	0.09	3.91
2013-06-03	09:30	11	0.04	15.39
2013-07-03	14:50	7	0.17	3.36
2013-08-01	09:10	26	0.11	1.11
2013-08-29	09:25	3	0.09	0.97
2013-09-24	09:30	1	0.08	1.26
2013-10-23	09:30	3	0.06	2.85
2013-11-18	09:30	3	0.11	2.73
2013-12-19	15:30	6	0.18	2.98
2014-01-15	08:25		0.06	3.10
2014-02-12	08:15		0.05	2.39
2014-03-12	08:30		0.05	1.96
2014-04-09	08:45		0.08	1.26
2014-05-07	08:40		0.09	1.26
2014-06-04	08:05		0.11	1.04
2014-07-02	08:35		0.09	0.79
2014-07-30	10:00		0.10	2.73
2014-08-27	08:50		0.06	7.06
2014-09-24	14:35		0.06	2.39
2014-10-22	09:45		0.12	1.77
2014-11-19	08:45		0.08	5.50
2014-12-17	09:00		0.06	1.96
2015-01-13	08:25		0.08	5.32
2015-02-11	08:40		0.06	2.28
2015-03-11	08:35		0.08	2.85
2015-04-09	08:50		0.02	2.73
2015-05-06	08:30		0.06	1.77
2015-06-01	15:00		0.10	1.67
2015-07-01	08:20		0.09	1.04
2015-07-29	08:15		0.30	0.85
2015-08-26	08:30		0.08	0.79
2015-09-23	08:45		0.12	0.91
2015-10-21	08:45		0.10	0.68
2015-11-18	08:35		0.09	0.62

2015-12-14	07:50	0.20	0.91
2016-01-20	15:00	0.04	0.91
2016-02-17	08:35	0.05	0.91
2016-03-16	09:05	0.03	0.91
2016-04-13	08:30	0.03	0.91
2016-05-11	09:45	0.05	0.91
2016-06-08	09:00	0.09	0.91
2016-07-06	08:15	0.03	0.91
2017-02-15	08:30	0.04	1.16
2017-03-15	08:35	0.03	3.12
2017-04-12	16:50	0.03	1.65
2017-05-10	14:35	0.03	1.15
2017-06-08	15:05	0.06	1.05
2017-07-04	14:20	0.07	0.82
2017-08-02	08:25	0.11	1.20
2017-08-30	08:20	0.07	0.73
2017-09-27	08:30	0.05	0.84
2017-10-25	08:20	0.05	1.03
2017-11-22	08:45	0.04	2.14
2017-12-20	08:25	0.04	4.17
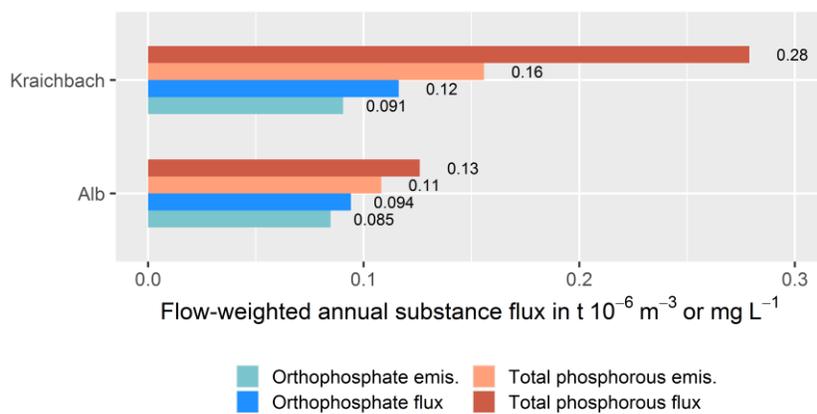
Figure 79: Flow-weighted annual total phosphorous and orthophosphate flux based on the best available estimate and annual flow.
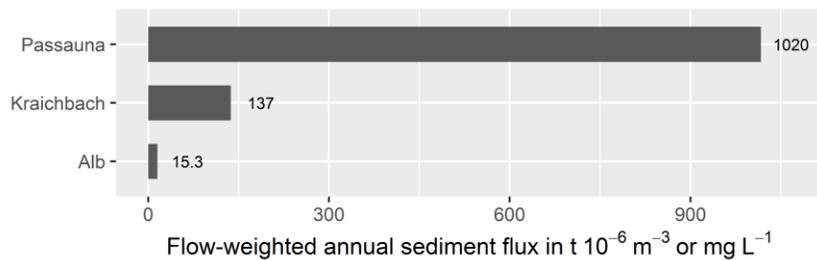

Figure 80: Flow-weighted annual sediment flux based on the best available estimate and annual flow.
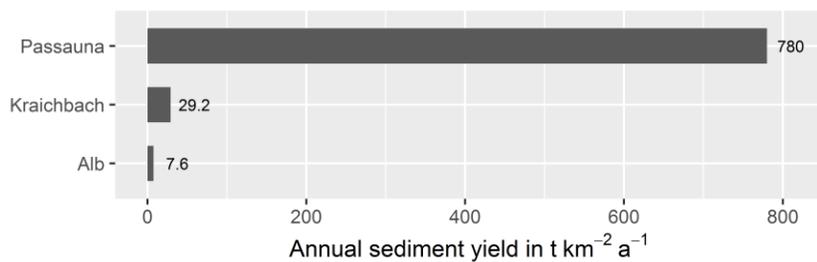

Figure 81: Annual sediment yield based on the best available flux estimate and catchment size.
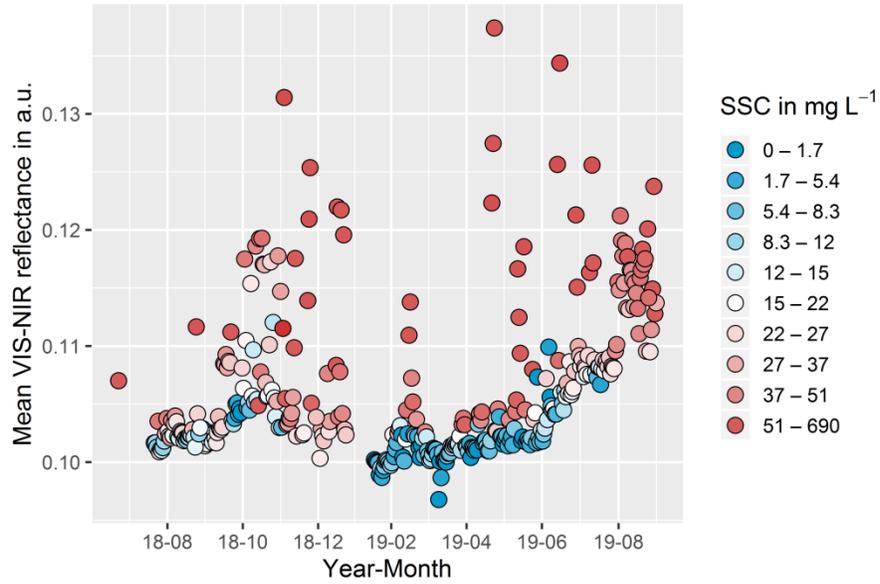
Figure 82: Mean visible (VIS) to near infrared (NIR) reflectance of 280 measurements of the Kraichbach monitoring station during low flow (discharge < 0.8 m³ s⁻¹). Colors indicate predicted suspended sediment concentration (SSC). For low flows, SSC would be expected to be around 1020 mg L, suggesting sensor drift. a.u.: Arbitrary units.
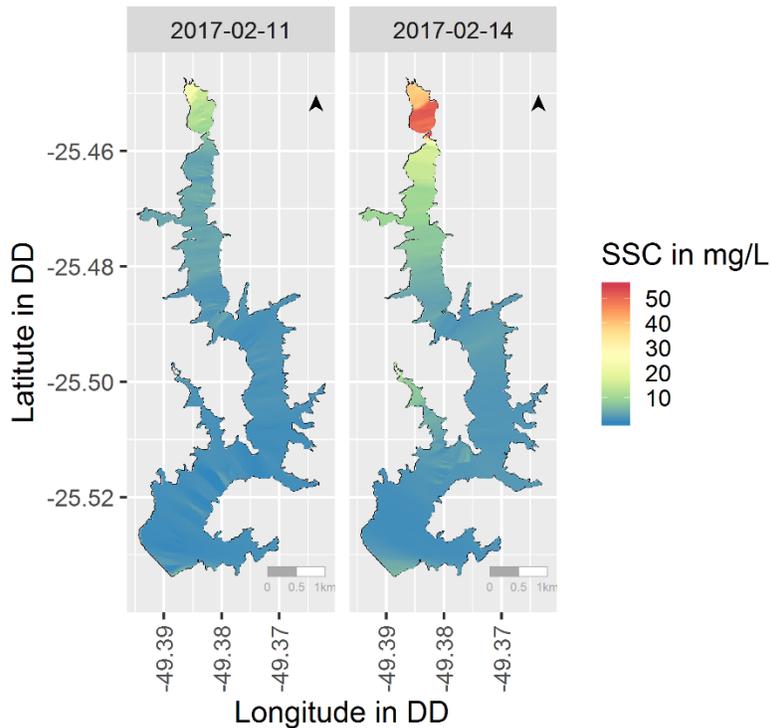

Figure 83: Inverse distance weighted interpolation of surface suspended sediment concentration (SSC) estimates from reflectance measrements at Passauna Reservoir in February 2017. Between the two dates, there were severe rainfalls triggering an SSC increase near the inflow. DD: Decimal degrees.
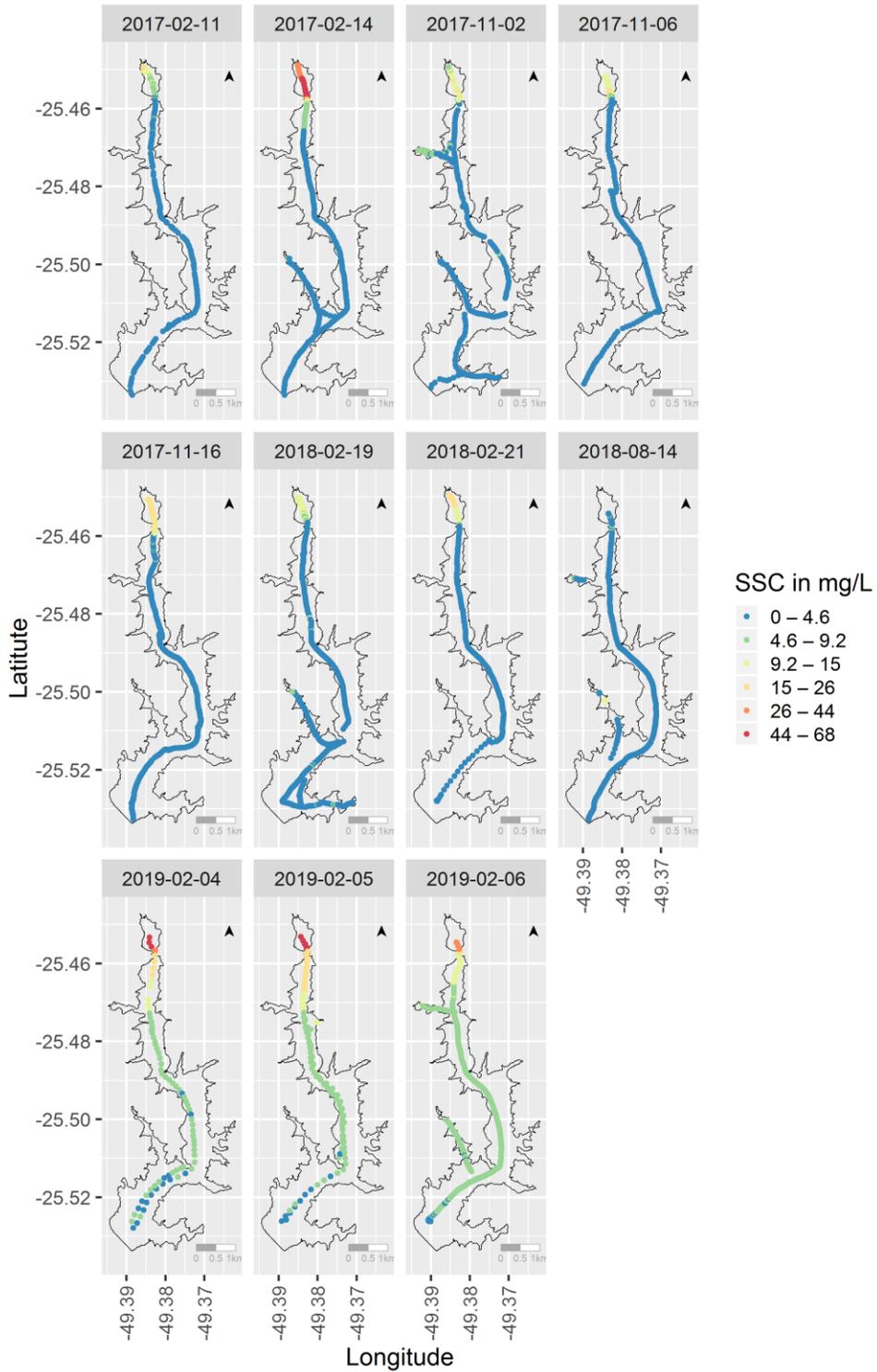
Figure 84: Eleven suspended sediment concentration (SSC) transects as determined from boat-based reflectance measurements in Passauna Reservoir.
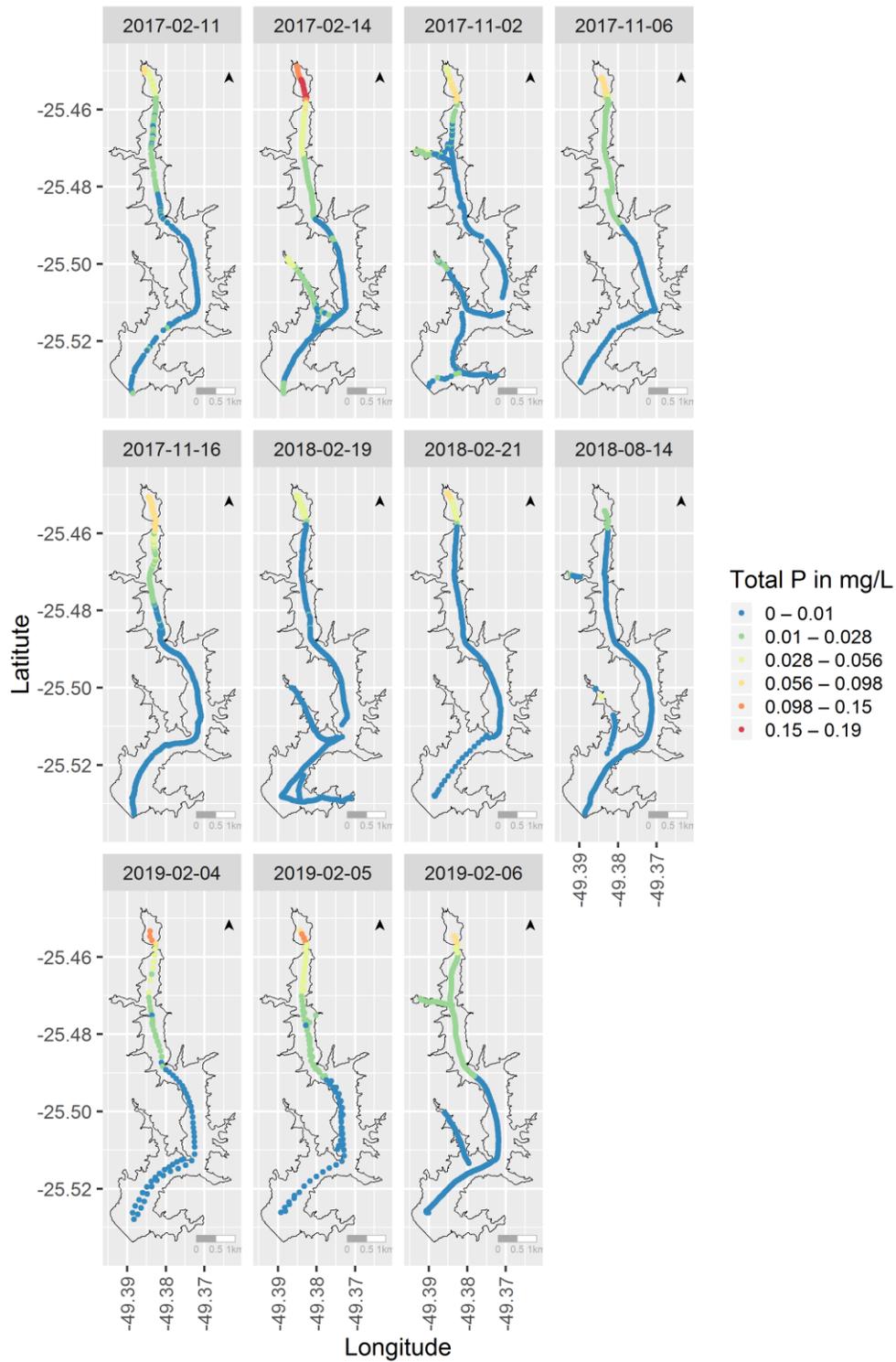
Figure 85 Eleven total phosphorus (P) transects as determined from boat-based reflectance measurements in Passauna Reservoir.