PhD Thesis

Reducing Bias and Increasing Transparency in the Designation Process of Critical Source Areas Across Scales

submitted in satisfaction of the requirements for the degree
Doctor of Science in Civil Engineering
of the TU Wien, Faculty of Civil and Environmental Engineering

---

Dissertation

Reduktion von Bias und Erhöhung der Transparenz im Zuge der maßstabsübergreifenden Ausweisung kritischer Herkunftsgebiete

ausgeführt zum Zwecke der Erlangung des akademischen Grads
Doktor der technischen Wissenschaften
eingereicht an der TU Wien, Fakultät für Bau- und Umweltingenieurwesen

Dipl.-Ing. **Gerold Hepp**

Matr.Nr.: 00140429

Betreuung: Univ. Prof. Dipl.-Ing. Dr. techn. **Matthias Zessner**
Assistant Prof. Dr.in techn. **Ottavia Zoboli,** MSc
Institut für Wassergüte und Ressourcenmanagement
Forschungsbereich Wassergütewirtschaft
Technische Universität Wien
Karlsplatz 13/226, 1040 Wien, Österreich

Begutachtung: Associate Prof. Dr.in **Adrienne Clement**
Faculty of Civil Engineering
Department of Sanitary and Environmental Engineering
Budapest University of Technology and Economics
Műegyetem rkp. 3, 1111 Budapest, Hungary

Begutachtung: Priv.-Doz. **Juraj Parajka,** PhD
Institut für Wasserbau und Ingenieurhydrologie
Forschungsbereich Ingenieurhydrologie und Wassermengenwirtschaft
Technische Universität Wien
Karlsplatz 13/222, 1040 Wien, Österreich

Wien, im Juni 2023

Kurzfassung

Sogenannte kritische Herkunftsgebiete sind ein etabliertes Konzept, wenn es um die Eindämmung diffuser Gewässerbelastungen geht. Es beruht auf der Idee, dass nur ein begrenzter Teil eines Flusseinzugsgebiets wesentlich zu einer stofflichen Belastung der zugehörigen Oberflächengewässer beiträgt. Kennzeichnend für kritische Herkunftsgebiete sind daher hohe Stoffemissionen und ein hoher Anbindungsgrad an Oberflächengewässer. Obwohl dieses Konzept grundsätzlich auf unterschiedlichste stoffliche Belastungen angewendet werden kann, ist es im Fall von Schwebstoffen und sedimentgebundenen Stoffen wie partikulärem Phosphor besonders relevant.

Die Hauptmotivation für seine Anwendung stellt die Einbeziehung ökonomischer Prinzipien in der Wassergütewirtschaft dar. Empfohlene Anpassungen der landwirtschaftlichen Praxis, die auf eine Verbesserung der Wasserqualität abzielen, gehen oft zu Lasten der Ernteerträge und werden daher von den zuständigen Behörden häufig mit finanziellen Anreizen begleitet. Selbst wenn solche Behörden bereit sind, das Konzept der kritischen Herkunftsgebiete anzuwenden, ist eine detaillierte Ausweisung auf beispielsweise der Feldebene schwierig, weshalb sich diese Behörden oft an die Flusseinzugsgebietsebene halten, die einfacher zu handhaben ist und in der Regel ausreichend Daten für die Anwendung der Wahrscheinlichkeitstheorie als Entscheidungsgrundlage aufweist.

Der hohe Bedarf an Daten und Rechenressourcen sowie fehlendes Prozessverständnis verhindern oft eine detaillierte Ausweisung kritischer Herkunftsgebiete. Ein gestaffelter Ansatz, der räumlich aggregierte mit räumlich verteilten Modellen kombiniert, ist jedoch eine mögliche Lösung für dieses Problem. Diese Arbeit konzentriert sich daher auf die Reduktion von Bias und Erhöhung der Transparenz im Zuge der maßstabsübergreifenden Ausweisung kritischer Herkunftsgebiete, indem sie (i) ein einfaches Bayessches hierarchisches Modell zur Vorhersage von Sedimentfrachten auf Flusseinzugsgebietsebene entwickelt, das sich für die Einbindung in räumlich aggregierte Flusseinzugsgebietsmodelle eignet, (ii) einen Weg aufzeigt, wie man die Verbesserung eines semi-empirischen, räumlich verteilten Phosphoremissions- und Transportmodells mit Hilfe einer Kartierungskampagne und eines weiteren Bayesschen hierarchischen Modells vorantreiben kann, um den Einfluss kulturtechnischer Maßnahmen auf diffuse partikuläre Phosphoremissionen in landwirtschaftlichen Einzugsgebieten zu bewerten, und (iii) eine Methode zur Ausweisung kritischer Herkunftsgebiete auf Feldebene für Flusseinzugsgebiete mit einer Größe von mehreren hundert Quadratkilometern entwickelt, die auf den Ergebnissen potenziell unterschiedlicher räumlich verteilter Modelle basiert und in der Möglichkeitstheorie verankert ist.

Da diese Arbeit hauptsächlich in einem Umfeld operiert, in dem bestimmte Daten nur äußerst spärlich vorhanden sind, wodurch insbesondere die frequentistische Wahrscheinlichkeitstheorie,

die sich auf Punktschätzungen konzentriert, nur eine sehr begrenzte Orientierungshilfe darstellt, macht sie in erheblichem Maß von der Bayesschen Wahrscheinlichkeitstheorie und der Möglichkeitstheorie Gebrauch. Insbesondere letztere passt perfekt zum Vorsorgeprinzip der EU-Wasserrahmenrichtlinie und kann in einem sich ständig wandelnden Umfeld an der Schnittstelle von Wirtschaft, menschlichem Verhalten, Umwelt und technischem Fortschritt Orientierung bieten.

Abstract

Critical source areas are a well-established concept in the field of diffuse water pollution control. It is based on the idea that only a limited share of a river catchment area significantly contributes to a pollution load of the corresponding surface waters. Characteristics of critical source areas are therefore high pollutant emissions and a high connectivity to surface waters. While this concept is basically applicable to many different pollutants, it is particularly relevant in the case of suspended solids and sediment-bound pollutants like particulate phosphorus.

The primary motivation for its application lies in the inclusion of economic principles in water quality management. Recommended changes in agricultural practices aimed at improving water quality often come at the expense of reduced crop yields and thus are commonly monetarily incentivised by responsible authorities. Even when such authorities are willing to adopt the concept of critical source areas, their detailed designation at, for example, the field level is difficult, hence why these authorities often stick to the catchment level, which is easier to handle and often holds enough data for the application of probability theory as a basis for decisions.

Demanding data and computational resource requirements as well as lacking process knowledge often prevent the designation of critical source areas at larger scales. However, a tiered approach combining lumped catchment models with spatially-distributed models is a possible solution to this. This thesis therefore focuses on reducing bias and increasing transparency in the designation process of critical source areas across scales by (i) developing a parsimonious Bayesian hierarchical model to predict river sediment yields at catchment scale suitable to incorporate into lumped catchments models, (ii) showing a way how to push the enhancement of a semi-empirical, spatially distributed phosphorus emission and transport model with the help of a small mapping campaign and another Bayesian hierarchical model in order to assess the impact of agricultural as well as civil engineering structures on diffuse particulate phosphorus emissions in agricultural catchments, and (iii) developing a method for the designation of critical source areas at field level within catchments of several hundred square kilometres based on the results of potentially diverse spatially distributed models and rooted in the theory of possibility.

As this thesis mainly operates in a sparse data environment where particularly frequentist probability theory focusing on point estimates only provides limited guidance, it makes considerable use of Bayesian probability theory and the theory of possibility. Especially the latter fits perfectly to the precautionary principle of the European Union Water Framework Directive and can provide guidance in an ever changing environment at the interface of economics, human behaviour, environment and technological progress.

Contents

Chapter 1

Introduction

Who should act and how? This is an important and common question when it comes to manage and mitigate environmental problems. A most prominent example is climate change these days. While its causes and underlying principles are well understood by now, the question who has to reduce emissions of climate noxious gases by what extent and in which time is still under heavy debate. However, the current debate would lack an important scientific basis without the various existing greenhouse gas emission inventories. An important aspect of such inventories is their scale. One can establish a single inventory for the whole world, one for each e. g. continent/economic area, national state, province or city. Even distinct inventories (so-called carbon footprints) for single products, services or human beings are available.

Whereas the calculation of inventories for large areas at small map scales (e. g. for economic areas like the European Economic Area) can make use of all kind of average values, inventories at large (map) scales (e. g. for human beings) require detailed data about the individual under "assessment", which in general are more difficult to acquire if available at all. Trying to deduce individual properties from properties of groups is therefore tempting, but poses a great threat in this context and is commonly known as ecological fallacy (Freedman, 1999; W. S. Robinson, 2009; Huddart Kennedy et al., 2015). On the other hand, primarily studies focusing on large areas may suffer from a problem known as (spatial) aggregation bias (Gehlke and Biehl, 1934; Clark and Avery, 1976; Tukker et al., 2020).

Besides such challenges, all scales have their distinct advantages and serve different purposes as well as stakeholders. While the focus of smaller scales often lies on the interested public and/or national/international authorities, larger scales frequently address authorities of smaller geographic units (e. g. state or city governments) and/or even single companies or citizens. When it comes to mitigating environmental problems, however, its always the acts and/or omissions of indivisible units, i. e. individuals, which count and add up.

From the point of view of managing and mitigating an environmental problem, important questions, among others, are therefore:

1. Which (geographical) units are affected?
2. Which (geographical) units cause the problem (source areas)?
3. Are the main sources evenly distributed among the source areas or is there maybe a small share responsible for the bulk of the problem?
4. Who are the responsible stakeholders in the source areas or maybe a critical subset thereof?
5. How can the responsible stakeholders be enabled and/or motivated or even forced to take appropriate actions and/or to cease harmful practices?

Turning to river sediment yield and particulate phosphorus inputs into Austrian surface waters, which are the main topics of this thesis, Kroiss, Lampert, et al. (2005) and Kroiss, Zessner, et al. (2006) state that phosphorus originating from the Danube basin is the main driver for phytoplankton growth and consequently algae blooms in the shallow north-western parts of

the Black Sea. In order to improve this situation, they recommend the application of a strong precautionary principle with respect to nutrient emissions; measures should focus on reducing inputs from waste water treatment plants (point sources) as well as losses from agricultural areas (diffuse sources). They further state that particularly Germany, Austria and the Czech Republic already made significant improvements regarding phosphorus emissions from waste water treatment plants in the past.

A more recent Austrian-wide nutrient emission modelling study (Gabriel et al., 2011) with the lumped catchment model MONERIS (Behrendt et al., 1999; Venohr et al., 2010) adapted to Austrian conditions (Zessner, Kovacs, et al., 2011) estimates that despite recent advancements in reducing phosphorus emissions from settlements and the industry, these sources still dominate phosphorus exports from Austria in the direction of the Black Sea. However, when looking at the water bodies of about 20 % of the assessed catchments at risk of not meeting their type-specific environmental quality standard for orthophosphate, this picture changes: The most prominent source of phosphorus emissions in those catchments is soil loss from agricultural areas. Sediment-bound, i. e. particulate phosphorus, is entering surface waters and becomes partly dissolved and therefore plant available again in this process.

Given these facts, it is evident that assessing particulate phosphorus emissions into surface waters can take place on different tiers with different scales. In the field of diffuse pollution of surface waters, the concept of critical source areas is a well established one. It refers to the principle that the majority of phosphorus inputs into the surface waters of a watershed usually originate from a limited share of its area. In principle, this concept can be applied to relations of arbitrary scales: larger catchments of various sizes (e. g. a multiple of 10^2 or 10^3 km^2) within basins (e. g. a multiple of 10^5 or 10^6 km^2), hillslopes (e. g. several ha) within smaller catchments of various sizes (e. g. a multiple of 10^2 or 10^3 ha) or even single fields within either smaller or larger catchments of various sizes or even within whole basins.

A primary motivation for applying the concept of critical source areas in water quality management is to address economic principles. The larger the first and the smaller the second scale in the above ratio, the more effective can mitigation measures be targeted. This is not only true when it comes to the economic costs of possible subsidies, but also when considering potential losses of crop yields due to changes in cultivation practices. So in fact, the application of the concept of critical source areas constitutes an optimisation problem primarily dealing with an optimal allocation of crops, the application of best management practices and the avoidance of direct as well as external costs. Furthermore, since action to reduce agricultural soil loss takes place on single fields and is carried out by individual farmers, targeting mitigation measures on catchment level may result in diffuse responsibilities.

Targeting single fields may pose a great challenge from an administrative point of view. The Integrated Administration and Control System of the Common Agricultural Policy of the European Union already collects various data about every field subunit (so-called “Schläge” in Austria) nonetheless. This includes geodata, which were available for fields (so-called “Feldstücke” in Austria) only up to the year 2014. After a period of maturing, these data can be reliably

used from the year 2019 on (Agrarmarkt Austria, 2018). Therefore, it should be reasonable to integrate the targeting of single fields designated as critical source areas into the existing system.

More severe problems with critical source areas especially at large scales are process knowledge, higher spatial as well as temporal variabilities, which require high-resolution data and generally lead to higher uncertainties, and the fact that they have to be designated with the help of spatially-distributed models, which are computationally expensive. Lumped catchment models, on the other hand, are in general computationally cheap. Consequently, a tiered approach is a possible solution to the problems of required vs. available computational resources in a specific amount of time as well as data requirements (A. L. Heathwaite, Dils, et al., 2005; Doody et al., 2012; Zessner, Gabriel, Kuderna, et al., 2014). In this approach, critical source areas are first designated at catchment scale with the help of a lumped catchment model (first tier) and a more demanding spatially-distributed model (second tier) is only applied to those catchments, which were designated as critical source areas in the first tier. However, the higher spatial as well as temporal variabilities at the larger scale of the second tier are not the only challenges. Lacking process knowledge and sparse or even missing data affect almost any scale when it comes to particulate phosphorus inputs into surface waters. Their degree of severity is generally higher for larger scales though.

1.1 Structure of the thesis

1.1.1 Chapter 2

The lumped catchment model MoRE (Fuchs et al., 2017), a derivative of the MONERIS model, considers eight pathways for nutrient emissions into surface waters, which can be grouped into direct and diffuse emissions:

- direct emissions from point sources
  - municipal waste water treatment plants
  - industrial direct dischargers
- diffuse emissions
  - atmospheric deposition on water surface areas
  - erosion
  - surface run-off
  - tile drainages
  - groundwater
  - sewer systems

As the fundamental pathway for river sediment yield is erosion, a more detailed look is taken on it. The calculation of river sediment yield originating from agricultural areas involves several processes in MoRE (Amann et al., 2019):

1. soil loss
2. overland sediment delivery
3. in-stream retention and/or mobilisation

In the case of river particulate phosphorus yield and orthophosphate concentrations there are even further processes involved: (i) the enrichment of particulate phosphorus during overland transport and (ii) the transformation of particulate phosphorus into orthophosphate within streams. The former describes the selective transport of soil particles. Smaller and lower weight particles are transported at a higher rate and at the same time possess a higher specific surface than bigger and heavier particles, which sediment more easily. Since a high specific surface allows for the adsorption of more sediment-bound phosphorus, i. e. particulate phosphorus, the prevalence of transporting these smaller and lower weight particles leads to a higher ratio of particulate phosphorus per unit weight of suspended sediments over time. MoRE considers this process with the help of the so-called enrichment ratio (Auerswald, 1989).

The main motivation for Chapter 2, which is based on the paper of Zoboli, Hepp, et al. (2020), is the fact that the individual processes of this chain are difficult to assess and evaluate on their own. For example, Parsons et al. (2006) consider the concept of the sediment delivery ratio, which is defined as the ratio between annual gross erosion and annual river sediment yield, a fallacy. They argue that it is merely a practical tool in special cases with inherent fundamental scaling problems on the spatial as well as temporal scale. So why should one consider it after all when the actual goal is the estimation of total fluxes at a certain point in a catchment for a given period in time?

MoRE's role in a tiered approach is one that usually spans the scales from catchments to basins. For this purpose, it is absolutely not necessary to look into the different processes at work within the modelled catchments as is the case with the existing process chain. Actually, when it comes to discharge, Amann et al. (2019) utilise the balance equation

$$Q_{GW} = Q_{net} - (Q_{WS} + Q_{SR} + Q_{TD} + Q_{US} + Q_{OR} + Q_{WWTP} + Q_{ID}),$$

where:

Q_{GW} is the contribution of groundwater discharge
Q_{net} is the net discharge
Q_{WS} is the balance of precipitation on and evaporation from water surface areas
Q_{SR} is the contribution of surface run-off
Q_{TD} is the contribution of tile drainages
Q_{US} is the contribution of urban sewers
Q_{OR} is the contribution of overland roads
Q_{WWTP} is the contribution of waste water treatment plants

Q_{ID} is the contribution of industrial direct dischargers

A distinct property of the groundwater discharge term here is that it is estimated as the difference of the net discharge and all other contributions, which makes it a catch-all contributor. While one could argue that groundwater discharge is a worse catch-all contributor than, for example, surface run-off, the principle of modelling and validating the sum of all discharge contributions, directly estimate those contributors, which can in principal be observed and therefore validated with the help of a representative sample, and put the rest into a catch-all term is an appealing one.

Instead of modelling suspended sediment inputs into surface waters from agricultural areas via the existing process chain, whose individual steps are practically impossible to validate, one could model and validate total river sediment yields and estimate the former or, going one step further, the erosive suspended sediment inputs from all land use types together as the remaining amount analogous to groundwater discharge. This, however, implies that the contributions of the different land use types to the total river sediment yield have to be assessed with the help of spatially distributed techniques, which in consequence should lead to sounder results. In case the delineation of the catchments is concerned about maximizing distinct characteristics, each catchment can be associated with a certain predefined class (e. g. glacial, alpine, agricultural and urban), which in turn allows to draw an appropriate conclusion.

In order to estimate the catchments' annual net discharges, Amann et al. (2019) utilised observed daily discharges where possible. Ungauged catchments, on the other hand, were estimated with the help of a geostatistical interpolation technique called top-kriging (Skøien et al., 2006). In comparison to ordinary kriging, this technique is not only based on spatial autocorrelation, but as well on topological relationships, which are characteristic for river networks. Since the Austrian network of discharge gauges is a dense one, kriging, which is also known as Gaussian process regression, generally performs well for this purpose, but top-kriging outperforms ordinary kriging. Considering the topology of river networks and their axiomatic direction of flow, it is not surprising that the information from neighbouring catchments is less informative than the one from downstream and particularly upstream catchments. This is in essence what top-kriging makes use of.

Taking up the fact that kriging can be considered as a generalisation of linear regression, the central hypothesis of Chapter 2 is that hierarchical linear models, which pose another generalisation of linear regression models and are also known as multilevel models or mixed-effect models, hold the theoretical power to considerably improve the performance and reliability of linear regression modelling when it comes to predict river sediment yields. In contrast to plain multiple linear regression models, which assume statistically independent samples, hierarchical linear models can deal with dependent observations resulting from nested structures as they are often found in environmental data. In our case, this nested structure consists of annual observations of river sediment yields at gauges, multiple gauges along rivers and multiple rivers within basins.

Hierarchical linear models work by estimating mean coefficients, the so-called fixed-effects, and at the same time individual coefficients for each of the defined levels, the so-called random-effects. This makes it a balanced approach between complete pooling of information in a grand mean, which tends to an underfitted model with high bias, and no pooling at all, i. e. individual coefficients for each observation, which usually leads to an overfitted model with high variance. They are therefore capable of capturing variations in space and time, but put less weight on levels with only few observations and ultimately lead to a more balanced bias-variance trade-off (cf. Geman et al., 1992). Furthermore, due to the possibility of considering multiple scales at once, they can be used to overcome the problem of aggregation bias (Raudenbush, 1988).

It is thus put forward that annual river sediment yields can be described and predicted more efficiently by using explicitly the information contained in the similarity within groups. To test this hypothesis, a novel Bayesian hierarchical model is developed, applied to a sample of heterogeneous catchments and its fixed- and mixed-effects performances incorporating different group levels, namely gauges, rivers, basins and clusters of catchments, are compared. With a parsimonious linear model consisting of only four variables (specific and extreme discharge, elevation and retention coefficient), good performance criteria in the calibration (NSE of 0.79 to 0.85 instead of 0.69) and cross-validations for temporal and spatial predictions (NSE of 0.71 and 0.72, respectively, instead of 0.65 and 0.64) are achieved. These results support the promising potential of this technique.

A major challenge, however, remain spatial predictions of river sediment yields in Austria. The number of gauges reporting on daily loads of suspended solids as of the time of writing Chapter 2 were less than 30. These are further located in mainly alpine catchments. Agricultural catchments and particularly catchments with a high share of arable land are almost none among them. Compared to the number of gauges reporting on daily discharges, which amounts to over five hundred, this number is very small. Since MoRE only requires annual river sediment yields and not daily loads as input data, several additional dozens of well placed gauges reporting on daily loads of suspended solids should improve spatial predictions already tremendously.

Furthermore, while the interpolation results of kriging are in principle based on spatial autocorrelation, but, among others, can be combined with a regression (regression-kriging) or make use of cross-correlations (co-kriging), hierarchical linear models are capable to incorporate all this and even more. Chagneau et al. (2011), for example, present a hierarchical Bayesian model for the simultaneous simulation of dependent Gaussian, count and ordinal spatial fields. In simpler cases, they are likewise merely enhanced with the help of Gaussian processes (e. g. over size, elevation, shares of land use types or space as in kriging).

In particular, incorporating one or more Gaussian processes over one or more shares of land use types may pose a promising future development for the improvement of spatial predictions, as one can assume that similar land use types result in similar sediment delivery processes. Chapter 3, however, indicates that this is only true to a certain extent, hence, one should probably consider additional appropriate interaction terms (see Section 1.1.2).

Although, due to the sparse data environment, one cannot expect highly reliable results from any model, neither the outlined integration of *BaHSYM* into MoRE nor its current approach is for vain. The importance of managing suspended sediment inputs and inputs of sediment-bound pollutants into surface waters is evident and taking appropriate action does not require overly reliable results. Nonetheless, as Chapter 4 shows, there exist methods to improve the accuracy and transparency of possible measures taken even in sparse data environments.

Results of multiple scenarios and models as well as expert judgement can be combined and objectified in a formalised way into a final result via fuzzy logic based on fuzzy sets (Zadeh, 1965). While Chapter 4 shows how to apply fuzzy logic to the second tier (field scale), it very well could be applied to the first tier (catchment scale) as well. Apart from that, integrating the results of statistical models like hierarchical linear models into MoRE could have the advantage of providing previously unavailable measures of uncertainty.

1.1.2 Chapter 3

The focus of this thesis now moves to the second tier of the described tiered approach for the designation of critical source areas at field scale. Chapter 3 is based on the paper of Hepp and Zessner (2019) and has another, yet not mentioned fallacy as its topic on the meta-level: cross-level fallacy (Clark and Avery, 1976). This fallacy can occur when one unconditionally assumes that insights from one group are also valid for other groups of the the same level/scale. In fact, this fallacy can easily manifest itself when making use of spatial or similar types of autocorrelation.

As the only way to not falling prey to this fallacy is to make observations of every existent group, knowledge about the number and nature of all possible groups is of major importance in the first place. An approach to tackle this problem is cluster analysis. While Luxburg et al. (2012) initially wonder whether clustering is actually a science or rather an art, they come to the conclusion that labelling it as engineering is more appropriate, if any, since its purpose lies in solving end-user problems. This implies, however, that there is as often no guarantee for finding the "correct" solution and, as Luxburg et al. (2012) further point out, is an "intrinsic psychological component" involved.

A maybe not "correct", but therefore optimal solution could be one that balances effort taken for the collection of data and acceptable risk of falling prey to cross-level fallacy. The state of this equilibrium has to be defined on a responsible team or even personal level. A serious problem in this context is, though, that those responsible for data collection and those for modelling usually are not the same. While such an unfortunate situation may lead to something comparable to omitted-variable bias in statistics when it comes to the designation of critical source areas at field level, i. e. the effect of a transport process missing in a certain modelling study is attributed to those included, there are means to avoid this or at least to point it out.

One such means is the simple mapping key presented in Chapter 3. It is designed for quick and systematic assessments of the types of agricultural and civil engineering structures present in a certain agricultural catchment as well as the impact they may have on the spatial distribution of

critical source areas at field scale. The main motivation for its creation was that the preparation and evaluation of several applications of the semi-empirical, spatially distributed PhosFate model (Kovacs, Honti, and Clement, 2008; Kovacs, Honti, Zessner, et al., 2012) in the Austrian federal state of Upper Austria (Kovacs, 2013; Zessner, Hepp, Kuderna, Weinberger, Gabriel, and Windhofer, 2014) suggest that storm drains may have a significant influence on sediment transport into surface waters.

An application of this mapping key to three small sub-catchments of a case study catchment with an area of several hundred square kilometres (one-stage cluster sampling) in the Innviertel region of Upper Austria clearly reveals that road embankments with subsurface drainage can exert a major influence on emissions and transport pathways of sediment-bound pollutants like particulate phosphorus.

Due to this, the PhosFate model is extended to separately model particulate phosphorus emissions into surface waters via storm drains and sewers along road embankments. Furthermore, the overall share of road embankments with subsurface drainage on all road embankments in the case study catchment is inferred with the help of a Bayesian hierarchical model. The combination of the results of these two models shows that the share of storm drains at road embankments on total particulate phosphorus emissions ranges from about one fifth to one third in the investigated area.

Here, a hierarchical model is once more used to connect scales. For the one-stage cluster sampling, the case study catchment was divided into approximately 150 small sub-catchments, whose fields were then systematically mapped. The scales connected by the hierarchical model, however, are the hillslope (zero-order catchment) and the case study catchment scale; field and sub-catchment (cluster) scales only acted as mediators in this process.

While three mapped clusters are definitely a sparse data environment and, from a frequentist statistical point of view focusing on point estimates, hold way too little power, the mode and highest posterior density interval of a Bayesian approach still provide some valuable information. In the present case, the mode of the share of road embankments with subsurface drainage on all road embankments is 77 % and the 90 % highest posterior density interval ranges from 54 to 100 %. Particularly in the light of the precautionary principle of the European Union Water Framework Directive (European Comission, 2000), this may already be enough information for a decision maker to take action though.

A more precise quantitative estimate for the size of the share of road embankments with subsurface drainage on all road embankments cannot be provided at the moment and has to be left to future research. Moreover, future research should deal with the question if there is a correlation between fields with high particulate phosphorus emissions and a corresponding downstream road embankment with subsurface drainage, as generally assuming mean conditions in this respect could pose another bias. Despite all this, a relevant and so far omitted transport pathway for particulate phosphorus inputs into surface waters could be revealed and its consideration successfully be incorporated into a selected spatially distributed phosphorus emission and transport model.

As initially stated, cross-level fallacy is a dangerous trap. A more recent application of PhosFate to various regions of Austria (Schmaltz et al., 2022) in turn suggests that these findings from the Innviertel region cannot be transferred unconditionally to other regions of Austria. In particular, the Weinviertel region seems to be different, since storm drains at road embankments appear to be less common. The presented mapping key should therefore be applied to at least this region as well, in order to assess if there exist different types of agricultural and civil engineering structures (e. g. roadside ditches without subsurface drainage and inter-field ditches), which perhaps require different model extensions. In this context, a future research question could read the following: Is there an interaction between the number of storm drains at road embankments and the amount of precipitation or discharge across regions?

1.1.3 Chapter 4

Poesen (2018) considers scaling up sediment yields from field to catchment scale as one of the main challenges in geomorphological research. The same is true for particulate phosphorus and other sediment-bound pollutants. Chapter 4, which is based on the paper of Hepp, Zoboli, et al. (2022), takes up this challenge. The importance of connecting these two scales unfolds from the fact that a subsidised mitigation of diffuse phosphorus emissions from agricultural soils into surface waters at catchment scale cannot be cost-effective. Such an inefficient allocation of money is neither desirable nor in line with the inclusion of economic principles in river basin and water quality management according to the European Union Water Framework Directive, hence the existence of the second tier in the described tiered approach.

In fact, the concept or critical source areas in itself poses a modifiable areal unit problem and is thus prone to aggregation bias. So while the designation of critical source areas at small scales is, flippantly said, a waste of resources, it becomes impossible to handle at very large scales. To illustrate this, let us imagine a raster-based fate and transport model like PhosFate applied to a spatial resolution of, for example, $1 \times 1\,\mathrm{cm}$. The number of cells with high inputs into surface waters compared to the total number of cells would probably be negligible. Pushing this to the extreme, a raster with an infinite small spatial resolution would result in the designation of infinite small critical source areas. This thesis therefore considers the field scale as the optimal scale for this purpose.

Designations at such a large scale, however, require great trust of all relevant stakeholders in order to prove useful. So how can they be told which fields are potential critical source areas in the context of the respective catchment? As there are complex processes at work and all fields have to be compared and ranked among each other, the application of spatially distributed models is a must to successfully tackle this problem. At the same time, the results of such models are difficult to comprehend for those not accustomed to them and have to be interpreted for as well as communicated to the relevant stakeholders and in particular to the individual farmers responsible for taking action in the end.

One of the main challenges in this connection is thus the coupling of complex, data-intensive fate and transport models with easy-to-use information on field level for management purposes at

the scale of large watersheds. To fill such a gap and create a bridge between these two tasks, this thesis puts forward the new Particulate PhozzyLogic Index based on the innovative combination of the results of a complex watershed model (in this case the PhosFate model) with fuzzy logic. Its main feature is the ability to transform the results of diverse scenarios or even models into a final map showing a catchment-wide ranking of the possibility of high particulate phosphorus emissions reaching surface waters for all agricultural fields.

For this purpose, the PhosFate model is enhanced with a new algorithm for the allocation of particulate phosphorus loads entering surface waters to their sources of origin. It is based on the idea that the ratio of the outflowing local particulate phosphorus emissions to the total particulate phosphorus turnover in a raster cell is related to the amount of particulate phosphorus originating from the respective raster cell. To put this in other words, raster cells with a high inflow and low outflow contribute little to the particulate phosphorus inputs into surface waters; the sources of origin are rather located upstream in such cases.

By means of a sensitivity analysis, this thesis further investigates the impacts of storm drains, discharge frequencies and flow directions on the designation of critical source areas with the help of present-day scenarios for a case study catchment with an area of several hundred square kilometres. The upfront model calibration exhibits a Nash-Sutcliffe efficiency (NSE) of about 0.95 and a modified Nash-Sutcliffe efficiency (mNSE) of around 0.83. A core result of the sensitivity analysis is that the scenarios at least partially disagree on the identified critical source areas and suggest that especially open furrows at field borders have the potential to lead to deviating outcomes.

All scenario results nevertheless support the 80:20 rule also known as Pareto principle, which states that about 80 % of the phosphorus inputs into the surface waters of a catchment originate from only about 20 % of its area (Sharpley, Kleinman, Jordan, et al., 2009). Vice versa, as this rule is in line with the scenario results, which were obtained at the field scale, the choice of the field scale for the designation of critical source areas from a practical point of view is reassured. From the same point of view the choice of the specific mitigation measure as well as its exact location on a critical source field yet should be left to the local expertise of the individual farmer.

Deviations between different scenarios or models can be taken into account by the methodology used to create the final map constituting the Particulate PhozzyLogic Index though. The range of the index lies between zero and one, where zero can be translated into “not possible” and one into “perfectly possible”. Zadeh (1978) states that “when our main concern is with the meaning of information—rather than with its measure—the proper framework for information analysis is possibilistic rather than probabilistic in nature”. This means that this index does not say anything about the probability of a field being a critical source area or not, but about the degree of it being possible.

A grassland field with a slope of almost zero, for example, will exhibit no or almost no particulate phosphorus emissions into surface waters in any serious modelling study. Its possibility of being a critical source area is therefore always close to zero, no matter if the field borders a surface water or not. On the other hand, an arable field with a high slope holds a high intrinsic possibility of

being a critical source area. This high intrinsic possibility is even higher when the field borders a surface water or a road embankment with a rather high known probability of having a subsurface drainage compared to a similar field, which does not. Further, as the possibility/probability consistency principle states that impossible events are improbable, but improbable events not impossible (Zadeh, 1978), the theory of possibility fits perfectly to the precautionary principle.

The deviating outcomes due to open furrows at field borders and other less important reasons are thus taken into account by combining the possibilities of all the modelled scenarios into a final possibility. This combination can be carried out in different ways. For example, by simply picking the lowest (fuzzy logic AND operator) or highest (fuzzy logic OR operator) possibility, where the former reflects a conservative and the latter a precautionary approach. In this thesis, however, the combination is carried out by the so-called convex combination operator (Dubois and Prade, 1985; Kandel, 1986), i. e. an optionally weighted mean operator, which neither constitutes a particularly conservative nor precautionary, but a balanced approach. In this regard, future research should assess the impacts of the different operators especially on the applicability of the 80:20 rule to the final index.

A practical problem with the application of the final index and, generally speaking, with the application of every kind of designation of critical source areas could be that focus on all other areas is lost. If particulate phosphorus emissions are decreased on critical source areas, but at the same time increased on other areas, it could happen that in total nothing is gained. In order to avoid this, additional precautions have to be taken. One such precaution could be the simultaneous introduction of a sediment or even particulate phosphorus loss inventory on farm level analogous to, for example, a farm's carbon footprint. This should lead to the application of best management practices and a better choice as well as allocation of the cultivated crops to the fields at hand in the long term.

The exploration of reasonable designs for such inventories and associated actions has to be left to future research though. It should, among others, assess if farms with a high footprint should be treated differently than farms with a low footprint and if the available budgets should be lowered over time so that environmental quality standards can eventually be reached. Furthermore, the question if the respective actual crop distributions and management practices or standardised conditions should be used for the initial determination of the available budgets should be pursued.

1.2 Overview of the research objectives and questions as well as their core answers

The following research objectives and questions are addressed by this thesis:

Objective 1 Does the hierarchical linear model approach considerably improve the performance and reliability of the otherwise unstable linear regression approach regarding the prediction of annual river sediment yields in Austria?

Yes, the consideration of gauges and rivers or gauges and basins as hierarchical levels increases the Nash-Sutcliffe efficiency from 0.69 to 0.85 for the best-fit model. Furthermore, cross-validation results for temporal prediction exhibit an increase from 0.65 to 0.71 when considering the same hierarchical levels and for spatial prediction an increase from 0.64 to 0.72 when considering catchment clusters as the only hierarchical level.

Objective 2 How can practitioners quickly and systemically assess the types of agricultural and civil engineering structures present in a certain catchment as well as the impact they may have on the spatial distribution of critical source areas?

This task can be achieved with the help of the simple and extendible mapping key presented in this thesis. It is supposed to be applied to mapping units like fields or zero-order catchments and its main features are the assessment of the following mapping unit properties:

- connected or not in a natural way
- explanatory details for mapping units not connected in a natural way
- explanatory details for mapping units connected in a natural way
- presence or absence of an artificial connection
- explanatory details for mapping units connected in an artificial way
- presence and absence of natural features as well as artificial point-shaped and linear structures

Objective 3 Are particulate phosphorus emissions entering surface waters via storm drains and sewers at road embankments a relevant factor in an Upper Austrian case study catchment with an area of several hundred square kilometres?

Yes, the share of storm drains at road embankments on total particulate phosphorus emissions ranges from about one fifth to one third in the investigated area. This estimate is obtained by combining the results of a Bayesian hierarchical model with the results of an extended semi-empirical, spatially distributed PhosFate model.

Objective 4 Can PhosFate's existing algorithm for allocating the particulate phosphorus emissions actually reaching surface waters to their respective source areas be redesigned so that conservation of mass is guaranteed under all circumstances?

Yes, whereas the original allocation algorithm based on a single top-down computation guarantees conservation of mass only on the level of zero-order catchments, the newly developed algorithm based on a top-down as well as bottom-up computation guarantees conservation of mass in every single cell.

Objective 5 What are the impacts of selected uncertain input data (flow directions) and model parameters (discharge frequencies) of PhosFate on the designation of critical source areas at field scale?

The conducted sensitivity analysis based on 18 scenarios shows that while the impacts of tillage directions and discharge frequencies are rather negligible, open furrows at field borders have the potential to cause deviating outcomes.

Objective 6 Can decision makers be provided with a transparent identification of critical source areas, which is as easy-to-use as the one resulting from the Phosphorus Index, but more accurate, by combining spatially distributed model results with fuzzy logic?

Yes, with the help of fuzzy membership functions results of complex watershed models can be transformed into well understandable maps ranking all agricultural fields of a potential large catchment with respect to their possibility of emitting high particulate phosphorus emissions actually reaching surface waters. Furthermore, transformed results of diverse scenarios or even models can be overlaid in order to create a final index map taking into account, among others, the uncertainty of input data. The fields designated as critical source areas then can easily be pointed out by colour-coding the displayed possibilities. This constitutes a major improvement compared to the semi-quantitative Phosphorus Index used for similar purposes in the past.

1.3 Data and software availability

The dataset and scripts of Chapter 2 are available on Zenodo (doi: 10.5281/zenodo.3514720) and the source code of the enhanced PhosFate model developed in Chapters 3 and 4 is available on GitHub in the form of an R package called `RPhosFate`.

1.4 Authorship

The author contributions to the papers on which the thesis is based on are made transparent via Contributor Role Taxonomy (CRediT) statements (Allen et al., 2019).

1.4.1 Chapter 2

Chapter 2 is based on the paper of Zoboli, Hepp, et al. (2020) with shared first authorship between Gerold Hepp and Ottavia Zoboli.

Gerold Hepp Conceptualisation, methodology, software, formal analysis, data curation, writing – original draft as well as visualisation.

Ottavia Zoboli Conceptualisation, methodology, validation, formal analysis, data curation, writing – original draft as well as visualisation.

Jörg Krampe Resources as well as writing – review & editing.

Matthias Zessner Writing – review & editing, supervision, project administration as well as funding acquisition.

1.4.2 Chapter 3

Chapter 3 is based on the paper of Hepp and Zessner (2019).

Gerold Hepp Conceptualisation, methodology, software, formal analysis, data curation, writing – original draft as well as visualisation.

Matthias Zessner Conceptualisation, writing – review & editing, supervision, project administration as well as funding acquisition.

1.4.3 Chapter 4

Chapter 4 is based on the paper of Hepp, Zoboli, et al. (2022).

Gerold Hepp Conceptualisation, methodology, software, formal analysis, data curation, writing – original draft as well as visualisation.

Ottavia Zoboli Writing – review & editing.

Eva Strenge Data curation as well as writing – review & editing.

Matthias Zessner Conceptualisation, writing – review & editing, supervision, project administration as well as funding acquisition.

Chapter 2

BaHSYM: parsimonious Bayesian hierarchical model to predict river sediment yield

2.1 Introduction

The delivery of sediments to surface water bodies as a result of soil erosion can exert a critical effect on flood risk (S. N. Lane, Tayefi, et al., 2007), on the lifetime of reservoirs (Kondolf et al., 2014) and on the health of benthic ecosystems (Greig et al., 2005). It can also be responsible for increased water treatment costs and for the decline of fisheries resources (Bilotta and Brazier, 2008). Further, the transport of sediments is mostly coupled with the transfer of organic carbon, phosphorus and a broad spectrum of particle-bound contaminants from soil into water, which further contribute to the degradation of aquatic environments (Long et al., 2006; Moran et al., 2017). Prediction and control of riverine sediment transport are thus fundamental goals for water managers worldwide. In this context, models are essential tools to estimate sediment yield (SY, e. g. $\mathrm{t\,yr^{-1}}$) at catchment outlets, to interpret spatial and temporal dynamics as well as to quantify and predict the consequences of climate and land use changes. However, the extreme complexity and variability of the processes linking soil erosion with river SY make these tasks very challenging.

The need for estimates of sediment yield and for the understanding of the major factors and processes controlling SY from watersheds across spatial scales is a field of research of long-standing nature (L. J. Lane et al., 1997). Yet, the reviews of Merritt et al. (2003) and de Vente, Poesen, Verstraeten, Govers, et al. (2013), which compared and critically discussed existing models designed to predict soil erosion and sediment yield at catchment scale, revealed a scattered and still unsatisfactory situation. Based on the classification system proposed by Wheater et al. (1993), Merritt et al. (2003) distinguished between empirical, conceptual and physics-based models. Similarly, according to de Vente, Poesen, Verstraeten, Govers, et al. (2013) models can be conceptually classified as follows: spatially lumped and spatially distributed, empirical, regression, physics-based, and factorial scoring models. LISEM (A. P. J. de Roo et al., 1996), PESERA (Kirkby et al., 2008) and SWAT (Arnold et al., 1998) are examples of widely-used physics-based models, which are based on the numerical solution of fundamental physical equations, such as transfer of mass, momentum and energy. Empirical and conceptual models, e. g. WaTEM/SEDEM (Van Oost et al., 2000; Verstraeten and Poesen, 2001), do not make explicit inference on detailed physical processes and rely instead on observed or stochastic relationships between causal variables and sediment yield. This distinction should not be interpreted in absolute terms, because there are also models which consist of a mix of physics-based and empirical components (e. g. AGNPS; Young et al., 1989). Factorial scoring models like FSM (Verstraeten, Poesen, de Vente, et al., 2003) are semi-quantitative methods which characterise catchments through factors coupled with scoring and which require in general an expert assessment in the field. Approaches for statistical modelling of SY, which are most typically spatially lumped, include multiple linear regression (e. g. de Vente, Verduyn, et al., 2011) and non-linear regression models (e. g. BQART; Syvitski and Milliman, 2007). The main outcome of Merritt et al. (2003) was that simpler empirical and conceptual approaches were more appropriate for the estimation of SY than physics-based or more complex conceptual models because these were limited by the lack of sufficient spatially distributed data, by the

over-dependency of the results on the experience of the modeller and by high computational requirements. The more recent review of de Vente, Poesen, Verstraeten, Govers, et al. (2013) similarly found that the elevated requirement of calibration parameters for most physics-based models often leads to equifinality and limits their applicability for spatial extrapolations and for scenario studies. Further outcomes of this review were, among others, that: (i) the accuracy of existing models differs across spatial and temporal scales and that different models should be selected according to the size of the catchments of interest, (ii) a drawback of many models lies in the fact that they depict only selected erosion and sediment transport processes, which limits their applicability to catchments where such processes are dominant and in turn requires an extensive prior knowledge of those (e. g. sheet, rill as well as ephemeral gully, hillslope and channel erosion), (iii) there is definitely need for further model development and for balancing between model complexity and quality of input data.

In this respect, there is a powerful technique that has been so far overlooked in this field. We refer to hierarchical linear models, also known as multilevel models or random coefficient models. In such an approach, data is grouped in a hierarchy of successively higher-level units and, instead of considering observations as independent from each other, it is assumed that groups within each level (e. g. annual SY of gauges, gauges of rivers and rivers of basins) share certain attributes and show similarities. The key idea is to explicitly use this information by considering both the within and between group variances, with the goal to improve model efficiency and estimate accuracy. From the statistical point of view this means that model parameters vary at more than one level and that inferences made about one group affect inferences about another. In other words, the model operates a partial pooling within levels, providing a balanced approach between complete pooling (same intercept and slopes for all data points, i. e. underfitting) and no pooling (individual intercepts and slopes for each data point, i. e. overfitting). Major advantages of hierarchical models are improved estimates for repeated and imbalanced samples, the explicit modelling of the variation across individuals within groups of the data, the fact that there is no need to perform averaging and consequently no associated loss of information as well as an optimal trade-off between overfitting and underfitting (McElreath, 2015). Hierarchical models are a well-established method in social and medical sciences to divide subjects into groups, and they are increasingly used in environmental and ecological sciences, because they enable incorporating cross-scales interactions and thus enhance the model effectiveness both in understanding causal effects and in prediction (Qian et al., 2010). Further, thanks to the exceptional progress of computational power achieved over the last decades, it is now technically feasible to develop hierarchical models within a Bayesian framework. This offers important possibilities, among which the explicit incorporation of prior knowledge into the model and the obtainment of probability distributions of both model parameters and estimates. The latter in turn allows for a thorough analysis of the significance and the uncertainty of the results (Gelman and Hill, 2006).

Based on these characteristics, we hypothesise that this technique holds a considerable potential to efficiently and reliably predict SY at catchment level and that it might help to overcome, at

least partially, the limitation of having to rely on different models depending on the scale and the properties of the catchments. The main consideration behind this hypothesis is that on the one hand there are common processes regulating soil erosion and transport of sediments in all catchments and that on the other hand their relative importance changes greatly in space and time. We can for instance expect that catchments with similar morphological traits, hydrological regimes and land use also show similarities in the dominant processes determining SY at their outlets. Further, individual catchments typically belong to larger groups, such as river systems or basins. This constitutes an ideal problem for Bayesian hierarchical models, which are designed to optimally use the information contained in the variability of the data across nested levels.

A widely-used technique in hydrology, top-kriging (Skøien et al., 2006), relies on a similar idea. It also makes use of the fact that information provided by gauges of the same river system helps to predict a stream flow related variable at an unobserved location better than information provided by gauges of other river systems. While top-kriging incorporates similarity of topological relation and geographical location only, *BaHSYM* is also capable of incorporating similarity of other factors.

The goal of this work is to test our hypothesis by developing a Bayesian hierarchical model able to describe and predict SY in heterogeneous river catchments in Austria. For the development and validation of the model we consider both temporal and spatial cross-validation. According to the outcomes of their review, de Vente, Poesen, Verstraeten, Govers, et al. (2013) discarded linear regression equations as suitable prediction models, since in a number of case studies they proved unstable and unsuitable to extrapolate sediment yield beyond the calibration datasets (de Vente, Verduyn, et al., 2011; Grauso et al., 2008; Haregeweyn et al., 2008; Verstraeten and Poesen, 2001). We hypothesise that the Bayesian hierarchical approach holds the theoretical power to considerably improve the performance and reliability of otherwise unstable linear regressions. This is why in this work we develop and test parsimonious linear regression models consisting of few explanatory variables.

2.2 Methods

2.2.1 Study area

The study area includes 30 Austrian river catchments for which data of suspended solids transported at the outlet is available for multiple years with high temporal resolution. This sample consists of catchments that are highly heterogeneous in their total area (135 to 10 660 km^2), average elevation (256 to 2495 m a. s. l.), mean slope (9 to 61 %), long-term mean discharge (1 to 305 $m^3 s^{-1}$) and land use. Most of the gauges are located in alpine or mountainous areas, whereas agricultural catchments in lowland are more sparse. Figure 2.1 depicts the geographical location of the gauges, whereas the basic properties of the corresponding catchments are reported in Table 2.1. It is important to observe the enormous temporal variability of sediment transport in the dataset, with annual SY varying in some cases up to an order of magnitude at the very same gauge.

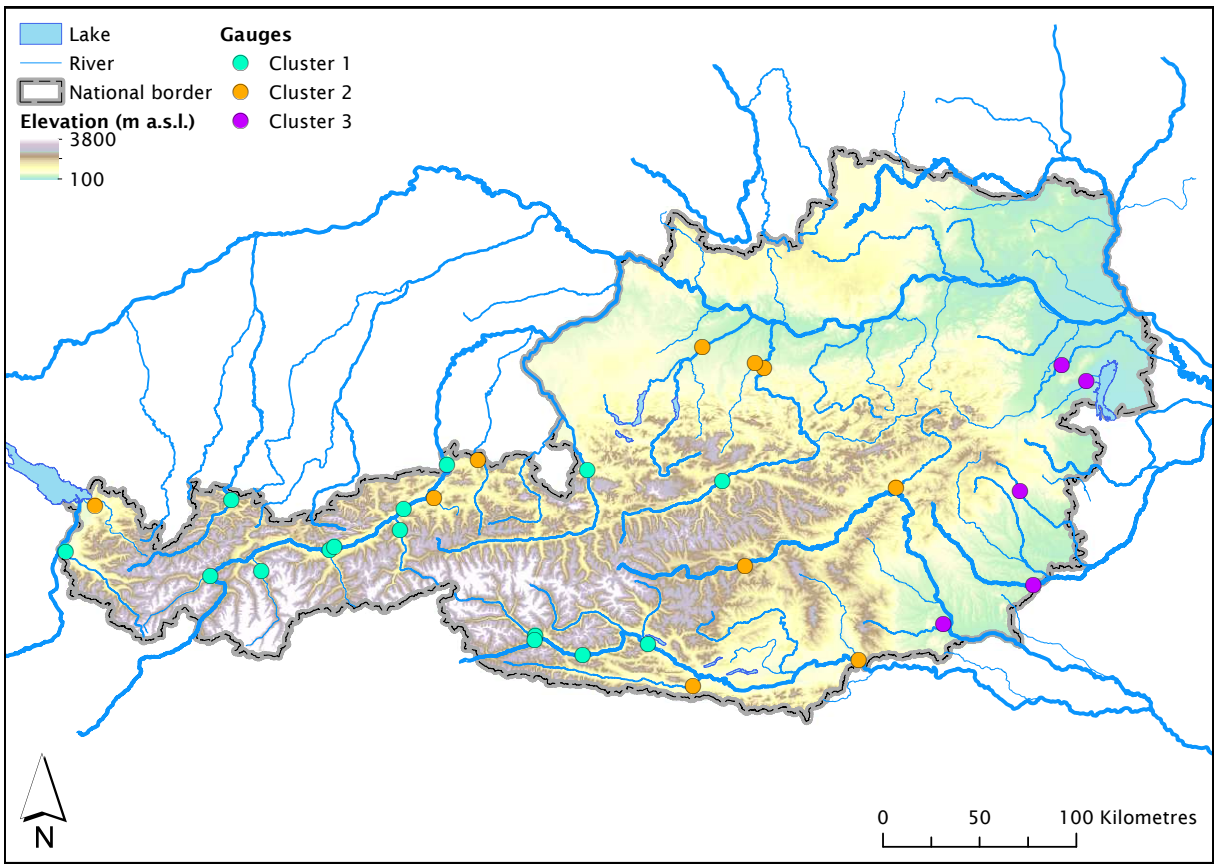

Fig. 2.1: Location of the catchment outlet gauges, from which the SY dataset stems. Gauges are reported with different colours depending on the cluster they belong to (please refer to Sections 2.2.5 and 2.3.1 for details regarding the clusters).

2.2.2 Bayesian Hierarchical Sediment Yield Model (*BaHSYM*)

BaHSYM consists of a linear regression model embedded within a Bayesian hierarchical approach. The basic level that we consider in the hierarchical model structure are individual gauges, i. e. we assume that at each gauge SY observations in different years are not fully independent and that dominant processes are to a certain degree similar. Further, we hypothesise that a specific gauge shares more similarities with other gauges within the same river system or within the same basin than with gauges located in other river systems or basins. Therefore, we created two model variants in which we nested the first level into a second higher level, namely the level of rivers or that of basins, respectively. To evaluate whether and how much this hierarchical structure improves the model performance, we also tested the model without any level, which would correspond to an ordinary linear regression model in a non-Bayesian context. We refer to the variant without levels as fixed-effects model and to the variants with hierarchical levels, i. e. random-effects, as mixed-effects models. In a mixed-effects model the fixed-effects describe the effect sizes of the overall mean and the random-effects the individual effect sizes of the hierarchical levels, which have to be either added to or subtracted from the fixed-effects.

Tab. 2.1: Basic properties of the catchments whose outlet gauge data is used to test the model: A (total area), E (average elevation), S (average slope), Q (long-term mean discharge), SY (range of annual sediment yields in the period 2009 to 2014).

Basin	River	Gauge	A (km^2)	E (m a. s. l.)	S (%)	Q ($m^3\,s^{-1}$)	SY ($1000\,t\,yr^{-1}$)
Bregenzerach	Bregenzerach	Kennelbach	838	1121	40	46	177–643
Drau	Drau	Lienz-Falkensteinsteg	658	1905	56	14	55–200
Drau	Drau	Dellach	2172	1981	56	63	286–622
Drau	Drau	Amlach	4779	1832	54	129	330–811
Drau	Drau	Lavamünd	10660	1379	43	256	70–498
Drau	Gail	Federaun	1297	1286	45	44	77–525
Drau	Isel	Lienz	1202	2145	58	39	276–769
Enns	Enns	Trautenfels	1518	1468	48	49	49–336
Enns	Enns	Jägerberg	5942	1162	49	166	173–1048
Ill	Ill	Gisingen	1230	1626	54	65	129–343
Inn	Brixentaler Ache	Bruckhäusl	326	1318	44	11	48–478
Inn	Großache	Kössen-Hütte	706	1201	43	27	151–834
Inn	Inn	Innsbruck	5786	2148	57	164	863–2310
Inn	Inn	Rattenberg	8483	2027	56	257	1113–3167
Inn	Inn	Oberaudorf	9683	1917	54	305	1327–4104
Inn	Salzach	Golling	3620	1577	52	141	181–1332
Inn	Sanna	Landeck-Bruggen	728	2139	58	20	51–379
Inn	Sill	Innsbruck-Reichenau	856	1911	56	25	83–379
Inn	Ziller	Hart im Zillertal	1135	1897	57	45	96–434
Inn	Ötztaler Ache	Tumpen	781	2495	57	26	338–917
Lech	Lech	Lechaschau	1012	1761	61	44	178–626
Leitha	Leitha	Deutsch Brodersdorf	1592	694	28	9	6–53
Mur	Mur	Gestüthof	1701	1637	44	36	23–134
Mürz	Mürz	Kapfenberg-Diemlach	1506	1071	44	22	16–83
Pinka	Pinka	Pinkafeld	135	743	25	1	4–22
Raab	Raab	Neumarkt	1009	492	20	7	20–33
Steyr	Steyr	Pergern	919	946	48	36	18–94
Sulm	Sulm	Leibnitz	1113	582	23	16	19–156
Traun	Traun	Wels-Lichtenegg	3379	847	31	129	56–294
Wulka	Wulka	Schützen am Gebirge	404	256	9	1	2–9

Annual SY was not modelled as such, but instead as the logarithmic transformation of the specific annual sediment yield (SSY, $t\,km^{-2}\,yr^{-1}$). Choosing SSY instead of SY enables to overcome the masking effect of different catchment sizes. The logarithmic transformation was necessary to meet the assumption of normality.

The set of Equations 2.1 presents the mathematical formulation of the mixed-effects model with two different group levels. It is formulated using non-centred parametrisation, i. e. with subtraction of the mean (fixed-effects) and factored out standard deviations of the random-effects:

$$\log(SSY_i) \sim \text{Normal}(\mu_i, \sigma)\,, \tag{2.1a}$$

$$\mu_i = X_{ij}\mathcal{B}_{j_i}\,, \tag{2.1b}$$

$$\mathcal{B}_{j_i} = \beta_j + \beta_{j,Level_{1_i}} \odot \sigma_{j,Level_1} + \beta_{j,Level_{2_i}} \odot \sigma_{j,Level_2}\,, \tag{2.1c}$$

$$\beta_j \sim \text{Normal}(0, 0.5)\,, \tag{2.1d}$$

$$\beta_{j,Level_{1_i}} \sim \text{MVNormal}(0_j, \text{P}_{jk,Level_1})\,, \tag{2.1e}$$

$$\beta_{j,Level_{2_i}} \sim \text{MVNormal}(0_j, \text{P}_{jk,Level_2})\,, \tag{2.1f}$$

$$\mathrm{P}_{jk,Level_1} \sim \mathrm{LKJcorr}(1)\,, \tag{2.1g}$$

$$\mathrm{P}_{jk,Level_2} \sim \mathrm{LKJcorr}(1)\,, \tag{2.1h}$$

$$\sigma_{j,Level_1} \sim \mathrm{Exponential}(1)\,, \tag{2.1i}$$

$$\sigma_{j,Level_2} \sim \mathrm{Exponential}(1)\,, \tag{2.1j}$$

$$\sigma \sim \mathrm{Exponential}(1)\,, \tag{2.1k}$$

where $\log(SSY_i)$ is assumed to be normally distributed with mean μ_i and standard deviation σ. X_{ij} represents the matrix of explanatory variables; $\mathcal{B}_{j_i}$ the combined fixed- and random-effects (i. e. mixed-effects); β_j the slopes of the fixed-effects; $\beta_{j,Level_{1_i}}$ and $\beta_{j,Level_{2_i}}$ the slopes of the random-effects of the two different group levels; $\sigma_{j,Level_1}$ and $\sigma_{j,Level_2}$ the factored out standard deviations of the slopes of the random-effects. The symbol $\odot$ represents the Hadamard product, which means variable-wise multiplication of the slopes of the random-effects with their corresponding standard deviation in the present case. Furthermore, $\mathrm{P}_{jk,Level_1}$ and $\mathrm{P}_{jk,Level_2}$ stand for the correlation matrices of the slopes of the random-effects. With respect to the matrix/vector indices, i indexes observations (rows of the matrix of explanatory variables), whereas j and k index explanatory variables (columns of the matrix of explanatory variables, vectors of the effect sizes of the explanatory variables, rows and columns of the correlation matrices of the explanatory variables).

Due to the fact that the response variable has been centred and scaled and thus has a mean of zero, the model is formulated without any intercept. Adding an intercept in such a case would not significantly improve model predictions.

Equations 2.1d to 2.1k define standard priors for centred and scaled variables according to McElreath (2020) (see Section 2.2.3 for details on centring and scaling). While these standard priors do not improve model predictions, they prevent implausible parameter values. Furthermore, having only centred and scaled explanatory variables allows us to use the same standard priors for all of them. Apart from the possibility to incorporate subjective or prior information into the model, in our opinion an underestimated advantage of Bayesian statistics is that it can make certain modelling steps simpler. For example, back transformation of logarithmic response variables, i. e. exponentiation, is straight forward and does not require a correction such as the one explained by Laurent (1963).

2.2.3 Explanatory variables

We considered a pool of potential explanatory variables for *BaHSYM*, which are reported in Table 2.2. Since this work focuses on the methodological approach and not so much on identifying the best performing variables, we constrained the selection to a set of relatively few attributes, which have the theoretical potential to explain the spatial and/or temporal variability of erosion and sediment transfer at catchment scale. We thus have chosen average and extreme hydrological variables, morphometric traits of the catchments as well as main land uses. Additionally, we

Tab. 2.2: Description of selected potential explanatory variables for the model.

Variable	Description (unit)
Hydrological and morphometric attributes	
A	Total area of the catchment (km^2)
E	Average elevation of the catchment (m a. s. l.)
S	Average slope of the catchment (%)
C	Compactness ratio: square root function of the ratio between A and the area of the circle having the same perimeter (-)
ξ	Sediment retention coefficient calculated according to Equation 2.2 (-)
l_p	Specific main channel length: total length of the principal waterway normalised by catchment area (km^{-1})
l_a	Specific tributary length: total length of secondary waterways normalised by catchment area (km^{-1})
q	Specific discharge: average annual river discharge normalised by catchment area ($m^3\,s^{-1}\,km^{-2}$)
Q_{p95}	Extreme discharge: maximum annual fraction of discharge above the 95th percentile of monthly discharge (%)
P_{p95}	Extreme precipitation: maximum annual fraction of precipitation above the 95th percentile of monthly precipitation (%; cf. Hanel et al., 2016)
Land use	
Glc	Percentage of total catchment area occupied by glaciers (%)
Agr	Percentage of total catchment area occupied by arable land (%)
Agr_4	Percentage of total catchment area occupied by arable land with slope $> 4\,\%$ (%)
Agr_8	Percentage of total catchment area occupied by arable land with slope $> 8\,\%$ (%)
Nat	Percentage of total catchment area with natural cover (mostly forests) (%)
Alp	Percentage of total catchment area occupied by bare alpine surfaces (%)
Grl	Percentage of total catchment area occupied by grassland (%)
Lakes	Percentage of total catchment area occupied by lakes (%)

have considered the sediment retention coefficient ξ, a variable conceived by Gavrilovic (1976) as sediment delivery component to be combined with an erosion rate in a semi-quantitative model to predict SY at basin scale (de Vente and Poesen, 2005). It is thus based on characteristics which mainly influence sediment transport and retention processes, such as morphometric traits and waterway length. It is calculated as follows:

$$\xi = \frac{\sqrt{Pe \cdot E}\,(L_p + L_a)}{(L_p + 10)\,A}\,, \tag{2.2}$$

where Pe is the perimeter of the catchment (km), E the average elevation in km a. s. l., L_p the length of the principal waterway (km) and L_a the cumulated length of secondary waterways (km).

The selection of these variables was driven by expert knowledge regarding the specific study area. In other regions, different variables might be more suitable and relevant. All variables, including the response variable, have been centred and scaled for modelling purposes. Centring refers to the practice of subtracting the sample mean from all values of a variable, whilst scaling describes the practice of dividing all values of a variable by its sample standard deviation. As a result, all variables have a mean of zero and a standard deviation of one and their effect sizes in the model can be compared independently of the scale of the original variables. The additional

centring and scaling of the response variable mean that increasing or decreasing an explanatory variable by one standard deviation, while keeping all other explanatory variables constant, causes the response variable to change by one standard deviation times the effect size of the adjusted explanatory variable. Increasing an explanatory variable with, for example, an effect size of plus 0.5 by 0.5 would therefore cause the response variable to change by 0.25 standard deviations. Thus, centring and scaling eases the interpretation of the modelling results.

When testing different combinations of explanatory variables, we selected for each run a maximum of three to four variables to keep the model parsimonious. Apart from additive effects, we also tested multiplicative interactions between the variables.

2.2.4 Data

For 27 out of 30 gauges, data on daily loads of transported suspended solids were provided for the years 2009 to 2014 by the Austrian Hydrographic Central Bureau (delivered in May 2017). For the rivers Pinka and Wulka, data on suspended solids concentrations corresponding to 48 h composite samples were provided for the same period of time by the State Government of Burgenland. As for the river Raab, data for the years 2009 to 2014 stems from a station equipped with devices for continuous monitoring of water quality parameters, which is operated by the Institute for Water Quality and Resource Management of the TU Wien (Camhy et al., 2013; Fuiko et al., 2016). For all gauges, daily or more detailed available loads were summed up to obtain annual sediment yields for each year, i. e. the modelled response variable in the proposed *BaHSYM* approach.

Daily precipitation data with 1×1 km spatial resolution were extracted from the SPARTACUS dataset of the Austrian Central Institution for Meteorology and Geodynamics (Hiebl and Frei, 2018). Daily discharges were obtained from the web GIS eHYD (BMLRT, 2020). Elevation and slope data stem from the official digital terrain model of Austria, which has a spatial resolution of 10×10 m (geoland.at, 2016a). Detailed land use data at catchment scale for the period 2009 to 2014, including river network length and lakes surface, was made available by Amann et al. (2019).

2.2.5 Model combined with catchment clustering

Should the model be applied for spatial prediction, i. e. to estimate annual SY for unmonitored locations, it would not be meaningful to use specific gauges as group level. For this purpose we employed and tested a variation of the mixed-effects model, in which the group level consists of clusters of similar catchments only. In this way, it shall be possible to predict SY for new gauges by assigning them to one of the identified clusters.

We formed clusters of catchments by following a two-step procedure. In the first place, we carried out a Principal Component Analysis (PCA; Jolliffe, 2002) based on a subset of the variables reported in Table 2.2, including topographic attributes, traits of the surface water network, land use and river discharge.

Tab. 2.3: Principal components obtained in the PCA: percentage and cumulative explained variance of SY.

Component	Explained variance (%)	Cumulative explained variance (%)
1	48.7	48.7
2	14.5	63.2
3	10.8	74.0
4	8.4	82.4
5	5.8	88.2
6–12	11.8	100.0

Thus, in a second step we used the first two principal components, which explain approximately 63 % of the total variance (Table 2.3), to identify clusters of catchments. To perform the cluster analysis, we applied the partitioning around medoids (PAM) algorithm (Kaufman and Rousseeuw, 1987; Park and Jun, 2009). The identified clusters are described and discussed in Section 2.3.1.

2.2.6 Model evaluation

As described in the previous sections, we tested different combinations of group levels, which bring into being the *BaHSYM* variants depicted by the set of Equations 2.3. These equations correspond to variants of Equation 2.1c, whereas the rest of the general modelling approach is the same for all. For the reasons discussed previously, we tested Equations 2.3a to 2.3c for temporal prediction (e. g. to fill yearly gaps), whereas Equation 2.3d was employed for spatial prediction (e. g. to extrapolate SY for catchments without monitoring of sediment transport):

$$\mathcal{B}_{j_i} = \beta_j + \beta_{j,Gauge_i} \odot \sigma_{j,Gauge} \tag{2.3a}$$

$$\mathcal{B}_{j_i} = \beta_j + \beta_{j,Gauge_i} \odot \sigma_{j,Gauge} + \beta_{j,River_i} \odot \sigma_{j,River} \tag{2.3b}$$

$$\mathcal{B}_{j_i} = \beta_j + \beta_{j,Gauge_i} \odot \sigma_{j,Gauge} + \beta_{j,Basin_i} \odot \sigma_{j,Basin} \tag{2.3c}$$

$$\mathcal{B}_{j_i} = \beta_j + \beta_{j,Cluster_i} \odot \sigma_{j,Cluster} \tag{2.3d}$$

The *BaHSYM* variants were tested through a k-fold cross-validation procedure. To test the use of the model for spatial prediction (to estimate SY for out-of-sample gauges), we applied a 10-fold cross-validation leaving out gauges stratified by cluster. The available data was split in ten training and test sets. Given the 30 available gauges, each training set consists of approximately 27 sites and each test set of approximately three different sites each time. It is important to note that all annual observations at one site stay together each time, i. e. it cannot happen that, for example, the years 2009, 2010, 2012 and 2014 of one site are in the training set and the years 2011 and 2013 of the same site are in the test set. This type of cross-validation is solely used to test the model's capability to predict SY at new sites. Furthermore, stratifying by clusters makes sure that each test set contains approximately one site of each cluster. The goodness of fit metrics are then calculated from the collected predictions of all test sets. Since the available

dataset of SY comprises six years, to test the performance of the models for temporal prediction (to estimate SY for out-of-sample years), we applied a 6-fold leave-one-year-out cross-validation. The six years of available data was split into six training and test sets. Each training set consists of the data corresponding to all sites for five years, whereas each test set contains the data corresponding to all sites for the remaining year (in each fold a different year). This type of cross-validation is solely used to test the model's capability to predict SY for new years. The goodness of fit metrics are then likewise calculated from the collected predictions of all test sets.

Following criteria were selected to quantify the performance of the models in estimating annual SY: Nash-Sutcliffe efficiency (NSE), modified Nash-Sutcliffe efficiency (mNSE) and coefficient of determination (R^2). NSE measures the goodness of fit of the model, with values ranging from $-\infty$ for poor predictive power to one for perfect match between modelled values and data (Nash and Sutcliffe, 1970). mNSE is a modification of NSE which is less sensitive to extreme values and is more influenced by low values (Krause et al., 2005). We applied it to better evaluate the model performance for catchments with relatively smaller SSY. In addition, we selected the root mean square error (RMSE) and percent bias (PBIAS) (Yapo et al., 1996). Whereas RMSE calculates the standard deviation of the model prediction error, PBIAS indicates the average tendency of modelled values to be larger or smaller than the observed ones. In addition, *BaHSYM* was applied to the whole dataset to identify the best-fit model. The general best-fit model is presented and discussed in Section 2.3.2, while the results for temporal and spatial prediction are described in Section 2.3.3.

2.2.7 Software

The model was developed and tested with R version 3.6.1 (R Core Team, 2019). In particular, we made use of the brms package version 2.10.0, which is specifically designed to implement Bayesian multilevel models in R using the probabilistic programming language Stan (Bürkner, 2017; Bürkner, 2018).

2.3 Results and discussion

2.3.1 Clusters of catchments

The three groups of catchments obtained through the cluster analysis are depicted in Figure 2.1. Before analysing their distinctive traits, it is important to observe that they largely differ in size. Clusters No. 1, 2 and 3 are composed by 15, ten and five catchments respectively. This unbalance is mainly caused by the fact that available SY data stems in the majority from gauges located in alpine and mountainous regions, whereas lowland areas are rather under-represented. Figure 2.2 shows for each cluster the range of variation of the variables selected for the PCA.

The first cluster (No. 1) is characterised by a median elevation of 1911 m and a median slope of 56 %. Other distinctive attributes are the presence of glaciers and large shares of alpine bare areas and alpine grassland. The second cluster (No. 2) is mainly composed of mountainous catchments,

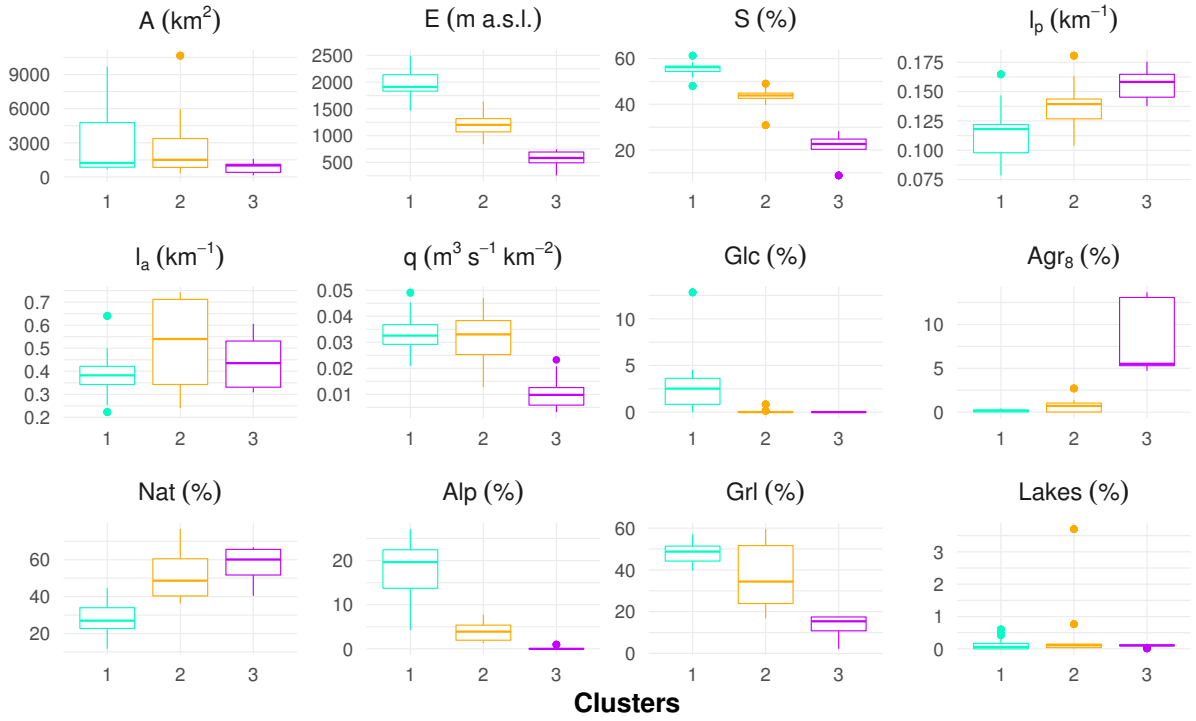

Fig. 2.2: Ranges of variation of the variables selected for the cluster analysis within the three resulting clusters.

although with lower median elevation (1182 m) and slope (44 %). In this cluster, glaciers and bare areas are present as well, though they occupy a much smaller share of land, whereas forests and grassland are largely prevalent. The third and smallest cluster (No. 3) has a median elevation of 582 m and a median slope of 23 %. There are obviously no glaciers, and bare areas and grassland are far less important. This cluster has a high average share of forest cover, but its most distinct trait is the significantly larger presence of steep arable land. With respect to specific discharge, the first two clusters are quite similar with a median value of $0.03\,\mathrm{m^3\,s^{-1}\,km^{-2}}$, whereas the third cluster presents a significantly lower value of $0.01\,\mathrm{m^3\,s^{-1}\,km^{-2}}$. Total catchment area is not a distinctive factor for these clusters. We can only observe that whereas cluster No. 3 presents a very narrow range of variation around small areas, the other two are characterised by broad ranges.

2.3.2 Best-fit model

The best model and cross-validation results were obtained with an additive model comprising four variables, namely average elevation (E), specific discharge (q), extreme discharge (Q_{p95}) and sediment retention coefficient (ξ). It performs best in its variant with two group levels: gauges and rivers. This model is described in the set of Equations 2.4, where priors consist of

Tab. 2.4: Estimate and significance of the slope parameters (β_j) for the input explanatory variables of *BaHSYM* (best variant with two group levels: gauges and rivers) applied to the whole dataset (SE: standard error; CI: confidence interval).

Variable	Estimate	SE	99 % CI	90 % CI
E	0.46	0.14	0.04–0.83	0.22–0.69
q	0.64	0.09	0.42–0.88	0.50–0.78
Q_{p95}	0.27	0.07	0.10–0.46	0.16–0.38
ξ	0.25	0.13	−0.13–0.59	0.03–0.46

standard priors for centred and scaled variables according to McElreath (2020) and are described by Equations 2.1d to 2.1k:

$$\log(SSY_i) \sim \text{Normal}(\mu_i, \sigma) \tag{2.4a}$$

$$\mu_i = X_{ij}\mathcal{B}_{j_i} \tag{2.4b}$$

$$X_{ij} = \begin{bmatrix} E_i & q_i & Q_{\text{p95}_i} & \xi_i \end{bmatrix} \tag{2.4c}$$

$$\mathcal{B}_{j_i} = \beta_j + \beta_{j,Gauge_i} \odot \sigma_{j,Gauge} + \beta_{j,River_i} \odot \sigma_{j,River} \tag{2.4d}$$

As reported in Table 2.4, the slope parameters (β_j) of the first three variables are statistically significant with a 99 % confidence interval, whereas this is true for the sediment retention coefficient ξ with a 90 % confidence interval. An investigation of the residuals furthermore confirmed that the assumption of normality was met and that the residuals do not show any signs of non-linearity or heteroscedasticity. The residuals plot is reported in Figure A.1 in Appendix A.

Although catchment area has been considered in the past to be a relevant predictor for SSY, the critical review of de Vente, Poesen, Arabkhedri, et al. (2007) concluded that, depending on scales and regional specificities, the relation between A and SSY can vary from positive to negative and is often non-linear. It is thus in general a poor predictor of SSY. This is in line with our outcomes. In fact, although A is indirectly considered in this work as component of the sediment retention coefficient, its inclusion as separate variable does not improve the model performance. Slope is almost interchangeable with elevation, although E performs slightly better. The fact that they hold a similar explanatory power can be explained through their very high correlation coefficient of 0.92. Likewise, high correlation coefficients between most land use variables and E explain to a large extent why these do not bring almost any improvement to the performance of the parsimonious model indicated above. The study of Gericke and Venohr (2012) on erosion in German mountainous catchments similarly found a strong correlation between SY and average elevation. For a complete overview of the correlation coefficients between the variables reported in Table 2.2, please refer to Figure A.2 in Appendix A.

Figure 2.3 graphically shows the comparison between observed and modelled SSY obtained with the best *BaHSYM* variant with two group levels (gauges and rivers), whereas Table 2.5 reports the performance criteria for all *BaHSYM* variants applied to the whole dataset.

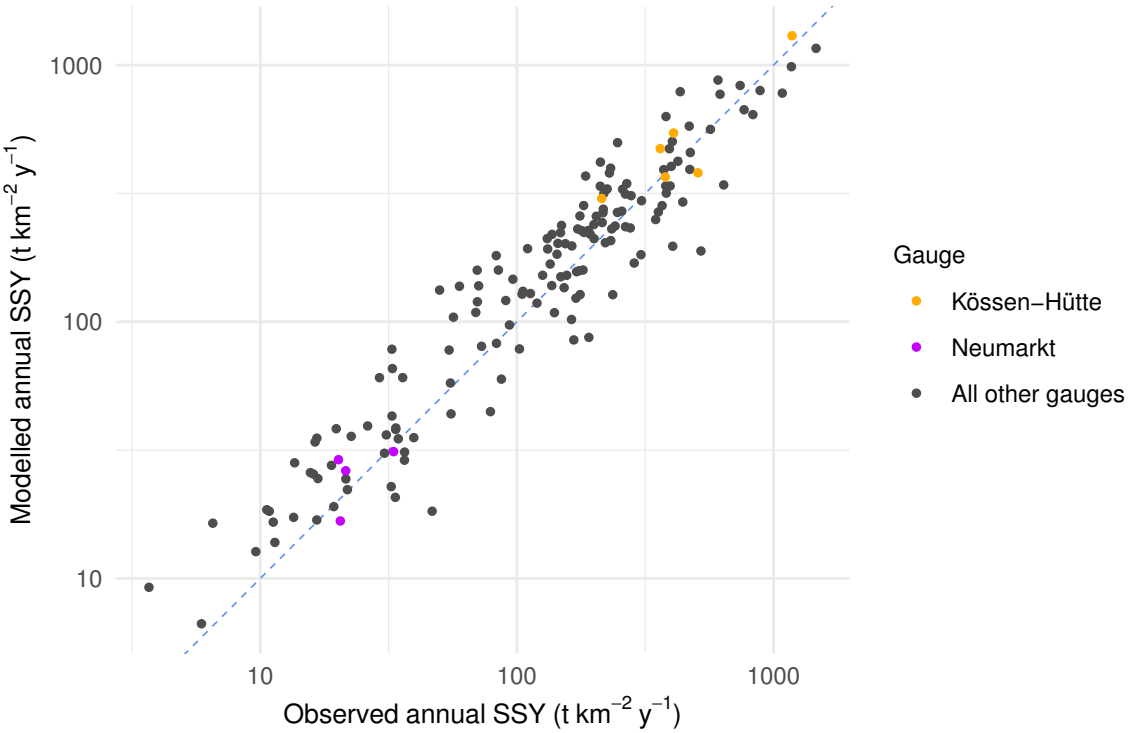

Fig. 2.3: Observed annual SSY vs. annual SSY modelled with *BaHSYM* (best variant with two group levels: gauges and rivers) applied to the whole dataset. The blue dashed line indicates the perfect match between modelled and observed values. The results for two gauges are highlighted in colour, as examples of different performance of the model for specific gauges.

Tab. 2.5: Results of the different *BaHSYM* variants applied to the whole dataset for the estimation of annual SSY.

Model	R^2 (-)	NSE (-)	mNSE (-)	RMSE (t km^{-2})	PBIAS (%)
Fixed-effects model	0.70	0.69	0.43	133	6.9
Mixed-effects model utilising Equation 2.3a					
One group level: gauges	0.84	0.83	0.62	96	5.5
Mixed-effects model utilising Equation 2.3b					
Two group levels: gauges, rivers	0.85	0.85	0.63	92	6.1
Mixed-effects model utilising Equation 2.3c					
Two group levels: gauges, basins	0.85	0.85	0.63	93	5.5
Mixed-effects model utilising Equation 2.3d					
One group level: catchment clusters	0.80	0.79	0.54	108	11.3

In support of our initial hypothesis, we achieve a notable improvement of the model performance through the technique of partial pooling. This is true for all variants of *BaHSYM* with distinct group levels, although they lead to partially different outcomes. The variant based on catchment clusters as group level brings a notable enhancement compared to the fixed-effects model, with R^2 raised from 0.70 to 0.80, NSE from 0.69 to 0.79 and mNSE from 0.43 to 0.54 respectively, whilst RMSE was lowered from 1.33 t ha^{-1} to 1.08 t ha^{-1}. This improvement comes however at the expense of a bias increase from 6.9 % to 11.3 %. Nevertheless, it was by partially pooling over gauges and over the combination of gauges with rivers or basins that we achieved the real breakthrough in model performance. R^2 increased to 0.84 to 0.85, NSE to 0.83 to 0.85, mNSE to 0.62 to 0.63, whereas RMSE was decreased below 1 t ha^{-1}. Further, these three variants even reduced the bias.

Despite the improvements, we can observe that mNSE is consistently lower than NSE in all model variants. In other words, the model's estimates are more reliable for catchments with greater SSY, independently from their size. This is exemplified by the performance of the best *BaHSYM* variant with two group levels (gauges and rivers) for two gauges highlighted in colour in Figure 2.3. The model performed very well for the gauge Kössen-Hütte, located in the mountainous river Großache and included in cluster No. 2 (R^2: 0.93, NSE: 0.88, mNSE: 0.56, RMSE: 106 and PBIAS: 10.3). The performance was however rather poor for the gauge Neumarkt, located in the rather lowland river Raab and included in cluster No. 3 (R^2: 0.32, NSE: −0.11, mNSE: −0.07, RMSE: 6 and PBIAS: 9.7). While their performance visually might appear comparable, the goodness of fit metrics clearly show the difference. Especially the NSE is a sensitive metric for the relative relationship of the magnitude of modelled residual variance and measured data variance. In this respect, viewed in isolation, the model captures the measured data variance of Kössen-Hütte way better than of Neumarkt, which is also clearly reflected by the R^2 metric.

2.3.3 Model for temporal and spatial prediction

Taking into account the huge temporal variability of SY observed in our dataset, the fixed-effects model does not perform badly for temporal predictions, with a NSE of 0.65 and a RMSE of 1.41 t ha^{-1} (Table 2.6). This means that the selected combination of variables has a high explanatory power with respect to the temporal variability of SSY. Two of the four variables, namely specific and extreme discharge, are time-dependent and especially the latter largely varies from one year to the next.

Table 2.6 shows the considerable improvement of the model performance for temporal prediction achieved through partial pooling. The three *BaHSYM* variants achieve the same NSE of 0.71. The combination of gauges with rivers or with basins in a mixed-effects model with two group levels slightly improves RMSE from 1.41 t ha^{-1} to 1.27 t ha^{-1}, at the expense of a small bias increase. Taking individual gauges as single group level instead than the combination of two group levels leads to a better fit for individual catchments, also for the ones with smaller SSY, which is reflected in a slightly higher mNSE of 0.50. Nevertheless, this improvement comes at a

Tab. 2.6: Results of the 6-fold leave-one-year-out cross-validation for different *BaHSYM* variants tested for temporal prediction of annual SSY.

Model	R^2 (-)	NSE (-)	mNSE (-)	RMSE ($t\,km^{-2}$)	PBIAS (%)
Fixed-effects model	0.66	0.65	0.39	141	6.9
Mixed-effects model utilising Equation 2.3a					
One group level: gauges	0.72	0.71	0.50	128	7.2
Mixed-effects model utilising Equation 2.3b					
Two group levels: gauges, rivers	0.73	0.71	0.49	127	8.6
Mixed-effects model utilising Equation 2.3c					
Two group levels: gauges, basins	0.73	0.71	0.49	127	8.4

small expense of the general performance, since R^2 and RMSE reflect to a greater extent the dominance of catchments with high SSY. The ability of the fixed-effects model for temporal prediction is thus enhanced by the structure of *BaHSYM*. Including random-effects allows the effect sizes of the explanatory variables to be correlated. Making use of at least one explanatory variable that varies in time can cause the effect sizes of the other explanatory variables to depend on the effect size of that time-dependent variable. For example, in a year with high extreme discharge, the effect size of the retention coefficient (which does not vary in time) can be different from the one it has in a year with low extreme discharge. While such correlations are often not statistically significant, they still affect the predictive power of models making use of random-effects. In other words, random-effects have the potential to add sometimes complicated "it-depends-structures" to a linear regression model.

de Vente, Poesen, Verstraeten, Govers, et al. (2013) state that extrapolating SY for different years for catchments that were used for calibration generally leads to better validation results than extrapolating SY for out-of-sample catchments, since differences in e. g. land use and dominant erosion processes between calibration and validation datasets are relatively small. In our case, the model performs almost equally well for both purposes. The two hydrological variables in the model are fundamental to describe temporal variability, but their combination with elevation and the morphometric variables contained in the sediment retention coefficient also allow capturing to a great extent the spatial variability.

This model application further supports our hypothesis regarding the potential of Bayesian hierarchical models in this field. As reported in Table 2.7, adding partial pooling over clusters of catchments notably improves all performance criteria. Although lowland catchments with generally lower SSY are under-represented in the sample available for model training, partial pooling over the clusters significantly improves mNSE from 0.39 to 0.48. This implies that even though the cluster of this type of catchments is relatively small, it conveys a significant amount of information on the different erosion and sediment transport processes that distinguish these catchments from the mountainous and alpine ones. That is an exemplary benefit of this technique in case of unbalanced datasets. Nevertheless, it is clear from the difference between NSE (0.72) and mNSE (0.48) that the good performance of the model is dominated by mountainous and alpine catchments with greater erosion and larger transfer of sediments. Our outcomes show that

Tab. 2.7: Results of the 10-fold cross-validation leaving out gauges stratified by cluster for different *BaHSYM* variants tested for spatial prediction of annual SSY.

Model	R^2 (-)	NSE (-)	mNSE (-)	RMSE ($t\,km^{-2}$)	PBIAS (%)
Fixed-effects model	0.66	0.64	0.39	142	6.1
Mixed-effects model utilising Equation 2.3d					
One group level: catchment clusters	0.74	0.72	0.48	128	10.5

the idea of combining *BaHSYM* with clusters of catchments as group level holds a great potential for spatial extrapolation, but they also reveal its limitations. In order to apply the model to make robust predictions, it is essential that new catchments share fundamental similarities with the available clusters. In our case, given the largely heterogeneous sample available, elevation, slope, land use and discharge were all important factors to determine the clustering. However, if a more homogeneous sample was available, more specific and targeted criteria could be used.

The outcomes of *BaHSYM* are very promising when compared to those of ordinary linear regressions. For example, de Vente, Verduyn, et al. (2011) attempted to spatially extrapolate SY and SSY based on linear regression and on a wide number of variables for a sample of 61 catchments with areas comprised between 30 and 13 000 km^2 in Spain. Although they could achieve quite good results for calibration (NSE: 0.58), the model performed very poorly in the validation step, with a NSE of −0.10. For the same dataset, they did achieve better validation results (NSE of 0.35 to 0.67 with spatially distributed models and 0.72 with the factorial score model), but such models are more complex and require considerably more data and expert assessments than the variants of *BaHSYM* presented in this paper (de Vente and Poesen, 2005; de Vente, Poesen, Verstraeten, Van Rompaey, et al., 2008). Our benchmark is not the model by de Vente, Verduyn, et al. (2011) per se, but rather the use of ordinary linear regressions of which their study is an example. Nevertheless, we tested the performance of their model for our study area. To do that, we reproduced the *BaHSYM* model with the variables selected by de Vente, Verduyn, et al. (2011) to model SY and SSY: (i) two topographic variables, namely average slope (%), also termed "mean slope gradient", based on a 25 × 25 m digital terrain model and "relief ratio" ($m\,km^{-2}$) calculated as

$$\frac{E_{max} - E_{min}}{A};$$

(ii) the climatic variable "precipitation concentration index" (%) calculated as

$$\frac{\sum_{i=1}^{12} p_i^2}{P^2} \cdot 100,$$

where p_i is the average monthly precipitation (mm) and P is the average annual precipitation (mm); (iii) the land use variable "Matorral" and (iv) three lithological and soil texture type variables, namely percentage of "acid metamorphic rock", "limestone" and "Fluvisols" (%;

Tab. 2.8: Results of *BaHSYM* variants tested for the estimation of annual SSY with the variables of the ordinary regression model developed by de Vente, Verduyn, et al. (2011). Results are reported for the model with sparsely vegetated area as land use variable.

Model	R^2 (-)	NSE (-)	mNSE (-)	RMSE ($t\,km^{-2}$)	PBIAS (%)
Model fit to whole dataset					
Fixed-effects model	0.21	0.10	0.08	226	15.3
Mixed-effects model utilising Equation 2.3a					
One group level: gauges	0.61	0.60	0.47	150	6.1
Mixed-effects model utilising Equation 2.3b					
Two group levels: gauges, rivers	0.61	0.60	0.47	150	6.1
Mixed-effects model utilising Equation 2.3c					
Two group levels: gauges, basins	0.61	0.60	0.47	150	6.1
Mixed-effects model utilising Equation 2.3d					
One group level: catchment clusters	0.36	0.29	0.21	201	15.6
Cross-validation for temporal prediction					
Fixed-effects model	0.13	−0.11	−0.05	250	19.3
Mixed-effects model utilising Equation 2.3a					
One group level: gauges	0.43	0.38	0.29	187	9.7
Mixed-effects model utilising Equation 2.3b					
Two group levels: gauges, rivers	0.43	0.37	0.28	188	10.7
Mixed-effects model utilising Equation 2.3c					
Two group levels: gauges, basins	0.41	0.36	0.27	190	8.9
Cross-validation for spatial prediction					
Fixed-effects model	0.001	−18.7	−1.2	1055	107.0
Mixed-effects model utilising Equation 2.3d					
One group level: catchment clusters	0.007	$< -10^3$	$< -10^3$	$> 10^6$	$> 10^6$

European Commission and European Soil Bureau Network, 2004). With respect to the land use variable, there is no perfect match for "Matorral" in Austria, which corresponds to a Mediterranean and sub-Mediterranean evergreen bush and scrub land use type. Instead, we tested three variables closely resembling this land use type in the Austrian landscape, namely percentage of sparsely vegetated area, transitional woodland-shrub and natural grassland (%, Corine Land Cover codes 333, 324 and 321, respectively). The results of this comparative exercise are reported in Table 2.8. The original fixed-effects model overall did not perform well and an important reason might lie in the variables, which were selected for catchments very different from the ones included in our study area. Nevertheless, it is interesting to observe that adding group levels, i. e. turning it into a mixed-effects model, did improve the best-fit model (NSE of 0.29 to 0.60 instead of 0.10) as well as the model for temporal prediction (NSE of 0.36 to 0.38 instead of −0.11) considerably. However, adding group levels failed for the use of the model for spatial prediction (NSE $< -10^3$ instead of −18.7). The comparison of these two models is thus useful to show that more complex structures, such as the one of *BaHSYM*, can indeed be very beneficial, but only in combination with suitable variables for each application. Otherwise, they might even worsen the model performance.

2.4 Conclusions and outlook

The outcomes of this study support the hypothesis that Bayesian hierarchical models hold a great potential to improve the prediction of sediment yield in rivers. We have shown that through the implementation of this technique even parsimonious linear regression models can provide relatively robust temporal and spatial extrapolations. This means that with a reduced amount of data availability for few variables, this technique enables filling annual gaps, performing predictions for future scenarios and extrapolating SY for catchments without monitoring of sediment transport. We have also shown that through *BaHSYM* the limitation of having unbalanced datasets for the model training is partially compensated for. Nevertheless, the power of this technique can overcome the lack of information only to a certain extent. The robustness and reliability of the predictions remain constrained by the availability of sediment transport data. For the Austrian case study, for example, it is evident that at present an enhanced monitoring network would be required in lowland catchments with dominant erosion on arable land.

What we put forward is the use of this technique for an enhanced extrapolation of sediment yield across scales, but the model per se will likely need to be adapted for each case study. *BaHSYM* shall be thus seen as a methodological approach. Specific purpose, data availability and required temporal and spatial scales shall determine in each application the most adequate variables and group levels to be used.

Future lines of research include upgrading *BaHSYM* via advanced correlation structures (e. g. Gaussian processes), formulating informative priors as well as extending the application of this technique to the investigation of the selected transfer of particulate-bound contaminants in river catchments.

Chapter 3

Assessing the impact of storm drains at road embankments on diffuse particulate phosphorus emissions in agricultural catchments

3.1 Introduction

Many efforts to control diffuse pollution of surface waters in agricultural catchments are based on the concept of critical source areas. This concept commonly refers to relevant source areas connected to surface waters in a way that allows significant amounts of pollutants to enter surface waters (e. g. A. L. Heathwaite, Quinn, et al., 2005; Pionke et al., 2000; Strauss et al., 2007). While the concept of critical source areas is basically applicable to many different pollutants, it is particularly relevant in the case of particulate phosphorus (PP) and other sediment-bound emissions.

PP emissions in agricultural catchments are in fact often dominated by soil erosion processes and PP transport into surface waters by overland flow. In this context, a common approach to identify critical source areas is the use of spatially distributed models. de Vente, Poesen, Verstraeten, Govers, et al. (2013) compared different models developed to predict soil erosion rates, sediment yield and in case of spatially distributed models also sediment sources and sinks. They conclude that while empirical/conceptual, spatially distributed models show good model performances in comparison to data requirements, the application of this type of models is only appropriate in cases where the considered erosion (e. g. sheet, rill and ephemeral gully erosion) and sediment transport processes (e. g. overland flow, soil creep and debris flow) effectively dominate the catchment of interest.

This implies that in order to successfully apply such a model to a certain catchment, a priori knowledge on its dominant erosion and sediment transport processes as well as pathways is required. Jetten, Govers, et al. (2003) even advocate including observed patterns in modelling studies, since this can greatly improve model results. Such knowledge, however, is only scarcely available.

A number of studies report that agricultural and civil engineering structures (e. g. roads, ditches and culverts) may exert a decisive influence on erosion and sediment transport within catchments. Roads are found to "accelerate water flows and sediment transport" (Forman and Alexander, 1998) and are responsible for altering flow routing patterns as well as slope lengths (Aurousseau et al., 2009; Van Oost et al., 2000). Particularly in combination with roadside ditches and culverts, they can directly connect sections of catchment areas to stream networks, which otherwise would only be indirectly connected or not connected at all (Alder et al., 2015; Croke et al., 2005; Doppler et al., 2012; Hösl et al., 2012; La Marche and Lettenmaier, 2001; Wemple et al., 1996).

Moussa et al. (2002) furthermore point out that inter-field ditch networks increase the portions of catchment areas directly contributing to the discharge of surface waters. In short, linear agricultural and civil engineering structures hold a large potential to boost direct pollutant emissions into surface waters even from remote and at first glance disconnected fields.

This a priori knowledge is very valuable when it comes to applying a model to a certain catchment. In case of catchments where roads play a major role, they are capable of extending the natural stream network considerably. Or to say it with the words of Hösl et al. (2012): "Ditches may occur near almost every road or track to drain road systems."

Despite all these observations, linear structures are rarely considered in models and modelling studies (Fiener et al., 2011). A notable exception is the WaTEM/SEDEM model (Van Oost et al., 2000; Van Rompaey et al., 2001; Verstraeten, Van Oost, et al., 2002), which optionally takes a roads data layer as additional input. Instead of integrating linear structures into models, there is, however, also the alternative of digital elevation model (DEM) pre-processing. The RIDEM model (Duke et al., 2003; Duke et al., 2006) takes this approach. It utilises cross-sectional templates of different configurations of roads and roadside ditches to enforce their effect on flow directions. Although DEM pre-processing has the advantage of not necessarily being dependent on a certain model, it does not seem to be widely applied in the context of linear structures.

While the role of roads and ditches in altering overland flow patterns is generally rather well-known, still vague is the role of another common agricultural and civil engineering structure: storm drains (Gramlich et al., 2018). This point-shaped structure is similarly capable of directly connecting remote fields to surface waters. The main difference to ditches lies in its subsurface transport pathway via sewers. In direct comparison with ditches, this means there is less retention, given that roadside ditches are often at least partly vegetated with grass and occasionally cleaned.

The preparation and evaluation of several applications of the semi-empirical, spatially distributed PhosFate model (Kovacs, Honti, and Clement, 2008; Kovacs, Honti, Zessner, et al., 2012) in the Austrian federal state of Upper Austria (Kovacs, 2013; Zessner, Hepp, Kuderna, Weinberger, Gabriel, and Windhofer, 2014) suggest that storm drains may have a significant influence on sediment transport into surface waters. Verstraeten, Poesen, Gillijns, et al. (2006) state that the effect of riparian vegetated buffer strips at catchment scale may be rather low, since a considerable amount of sediment bypasses them through sewers, among others. Another study found that artificial zones like farm courtyards are responsible for about half of the fields connected to the stream network (Gascuel-Odoux et al., 2011). While they do not decidedly mention storm drains or sewers, it can be reasoned, that they are a major cause of this. In addition, Djodjic and Villa (2015) express that surface water inlets are common in Sweden and due attention should be paid to them.

As a result of a field mapping campaign in a catchment in north-western Switzerland, Bug (2011) reports that storm drains occur mainly at asphalt roads leading across a hillside and in areas with flow convergence. However, for another catchment in the German federal state of Lower Saxony, he depicts a dense network of ditches but no storm drains at all. This indicates that there exist different traditions in implementing agricultural and civil engineering measures.

Since detailed datasets on the existence of ditches and storm drains are rarely available but a prerequisite for modelling studies aiming at the identification of critical source areas, there is a need to map them. Although several studies have carried out related field mapping campaigns in the past (e. g. Bug, 2011; Croke et al., 2005; Hösl et al., 2012), no ready to use mapping key exists for this task. This study therefore intends to develop, present and discuss a simple mapping key suitable for practitioners interested in a quick and systematic assessment of the types of agricultural and civil engineering structures present in a certain catchment as well as the impact they may have on the spatial distribution of critical source areas.

Tab. 3.1: Additional properties of the case study catchment.

	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
Field size in ha	0.01	0.36	0.90	1.60	2.02	28.16
Slope[a] in %	0.00	4.59	8.41	10.22	13.50	241.23
Discharge[b] in $m^3 s^{-1}$	0.94	1.87	2.68	4.58	4.23	117.00

[a] Based on a DEM with 10×10 m resolution.
[b] Period 2008 to 2013 of the gauge 204867, Pramerdorf/Pram close to the outlet with a catchment area of approx. 340 km^2 (BMLFUW, 2015).

A further objective is to estimate the share of PP emissions entering surface waters via storm drains and sewers at road embankments. The relevance of these structures is assessed for a case study catchment with an area of several hundred square kilometres. This is seen as essential information for policy makers facing the task to meet water quality targets in catchments of similar size.

3.2 Material and methods

3.2.1 Case study catchment

This study was performed on the catchment of the river Pram in the Austrian federal state of Upper Austria, which is located in the northern part of the country. Its area is approximately 380 km^2 in size and it belongs to the geologic formation Molasse basin with small parts in the north and northeast belonging to the crystalline Bohemian Massif. The prevalent soil texture is silt loam (nearly 50 % of the soil surface) followed by other types of loam. Only 5 % of the soil surface are dominated by clay. Soil textures dominated by sand are not present in the catchment area.

Annual precipitation ranges from roughly 900 mm in the north with an elevation of about 300 m a. s. l. to around 1200 mm in the south with an elevation of about 800 m a. s. l. Agricultural land (approx. 45 % arable land and 25 % grassland) is prevailing in the predominantly hilly terrain. Forests account for about 20 % and built-up areas for almost 10 % of the catchment area. The most important crops are winter grains and maize covering approximately 40 and 30 % of the arable land respectively. Additional catchment properties are listed in Table 3.1. Diffuse PP emissions are totally dominated by erosion from, especially, agricultural land. Other types of sources are negligible (Zessner, Gabriel, Kovacs, et al., 2011). Altogether, anthropogenic influence can be considered as high in this catchment.

3.2.2 Mapping key

The mapping key was developed based on the hydrological connectivity definitions related to landscape features and hillslope scale as collected by Ali and Roy (2009, Table 1). Among these definitions, the one from Croke et al. (2005) distinguishes “two types of connectivity: direct

connectivity via new channels or gullies, and diffuse connectivity as surface run-off reaches the stream network via overland flow pathways."

For this mapping key, the direct connectivity type was extended to include not only channels and gullies but also ditches, culverts, storm drains and sewers. Direct connectivity was furthermore split into known and unknown direct connectivity in order to differentiate between areas connected to an existing (e. g. governmental mapped) stream network dataset and areas connected to the actual stream network via features unknown to such a dataset (cf. Mosimann et al., 2007). The term diffuse connectivity was kept as it is.

A major issue with all connectivity mapping methods is that they merely record "snapshots" (Bracken, Wainwright, et al., 2013). Within the concept of hydrological connectivity one can distinguish static/structural and dynamic/functional elements of connectivity (Bracken and Croke, 2007; Turnbull et al., 2008). A simplified view on these elements comprehends space as structure and space-time as function (Wainwright et al., 2011). Omitting the functional element leads to a purely structural approach, which is solely capable of pointing out potential connectivity (Bracken, Wainwright, et al., 2013).

de Vente, Poesen, Verstraeten, Govers, et al. (2013) note that empirical/conceptual, spatially distributed models are frequently built on the RUSLE (Renard et al., 1997). Since this erosion model as well as its precursor USLE (Schwertmann et al., 1987; Wischmeier and Smith, 1978) estimate long term yearly erosion rates under average climatic conditions only, which can likewise be considered a potential (erosion potential), they match comparatively well with purely structural connectivity approaches.

Ultimately, the development of this mapping key aimed at aiding practitioners in preparing modelling studies with empirical/conceptual, spatially distributed models for water quality management at catchment scale. It therefore was based on a purely structural approach focusing on discrete mapping units like fields or (parts of) zero-order catchments. With the term zero-order catchments we refer to the catchment areas of all zones in a landscape exhibiting an overland-channel transition (Figure 3.1).

For each mapping unit, this mapping key fundamentally records if it is (i) connected or (ii) not connected in a natural way and/or if an (additional) artificial connection is (i) present, (ii) not present or (iii) present in a downstream mapping unit. It is thus required to map from downstream to upstream. Both of these types of connectivity are further supported by explanatory details, which include the cause of a mapping unit being connected or not in a natural way and, in case one is present, the type of the artificial connection.

Additionally, this mapping key allows for mapping artificial point-shaped and linear structures as well as natural features like channels missing in existing datasets. This can help with retracing flow pathways in case of questions once the mapping has been completed. The full ready to use and basically adaptable mapping key tables can be found in Appendix B.

Since connectivity is linked, among others, to discharge frequency, especially the categorisation of mapping units regarding connected in a natural way or not is somewhat arbitrary and in danger of subjectivity of the person(s) carrying out a mapping campaign. This issue can be

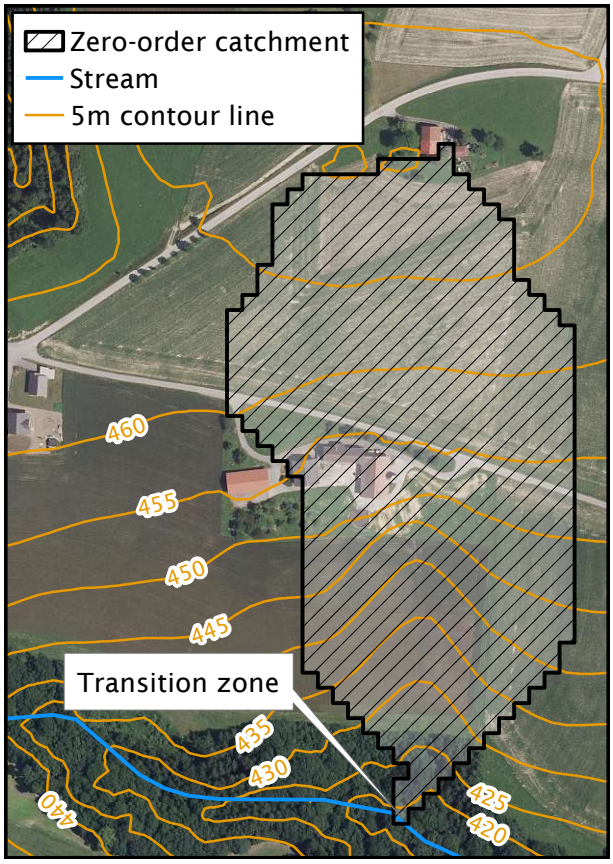

Fig. 3.1: Example of a zero-order catchment (catchment area belonging to a single overland-channel transition zone) based on a DEM with 10×10 m resolution and the D8 single flow direction algorithm (O'Callaghan and Mark, 1984).

addressed by defining a threshold obstacle height. Mosimann et al. (2007) acknowledge obstacles higher than 50 cm as effective barriers for surface run-off, preventing connectivity, and so do we.

In order to efficiently apply this mapping key, a mobile geographic information system (GIS) supporting a global navigation satellite system (GNSS) with at least two datasets is required: (i) an existing stream network and (ii) borders of the desired mapping units as well as the ability to alter them. Additional digital orthophotos help furthermore with orientation and locating certain landscape features.

3.2.3 Application of the mapping key

For the application of the mapping key, agricultural fields were chosen as mapping units. The statistical population of the case study catchment is therefore the total amount of its fields. Survey sampling was carried out with the help of one-stage cluster sampling. Cluster sampling is particularly useful when the statistical population under consideration can be grouped geographically. In the present case, it was grouped into small sub-catchments, which allowed for minimising travel times.

Tab. 3.2: Mapped agricultural land and number of fields per catchment and agricultural land use type.

		Cluster A	Cluster B	Cluster C	Total
Mapped agricultural land in ha	Arable land	128	206	149	483
	Grassland	23	69	76	168
	Total	151	275	225	651
No. of fields	Arable land	37	67	73	177
	Grassland	21	49	76	146
	Total	58	116	149	323
Average field size in ha	Arable land	3.5	3.1	2.0	2.7
	Grassland	1.1	1.4	1.0	1.2
	Total	2.6	2.4	1.5	2.0

Generally speaking, one-stage cluster sampling refers to the random selection of a subset of all clusters and the systematic assessment of all elements within the selected clusters. On the contrary, two-stage cluster sampling refers to not only the random selection of a subset of all clusters but also to the random selection of a subset of the elements to be assessed within the selected clusters. In this context, three small sub-catchments/clusters (Figure 3.2) were randomly selected from the total of 154 small sub-catchments/clusters of the case study catchment. The one-stage cluster sampling was then carried out by systematically mapping all of their fields.

For this task, a tablet computer running Microsoft Windows 10 with Esri ArcPad version 10.2.1 employing an external Global Positioning System (GPS) receiver was used. The field borders could be taken from a (geo-)database related to the Integrated Administration and Control System (IACS) of the Common Agricultural Policy (CAP) of the European Union (EU) (Hofer et al., 2014). A high-quality governmental mapped stream network was available from the Federal Office of Metrology and Surveying (2015) and digital orthophotos were contributed by the State Government of Upper Austria. Moreover, for referencing, a picture geodatabase was created containing overview pictures and pictures showing the relevant details of each field.

About 650 ha of agricultural land corresponding to 323 fields were mapped in total (Table 3.2). These are approximately 2.5 % of the overall agricultural land within the case study catchment. Each mapped cluster has its own characteristics: cluster A is elongated with tendentiously steep slopes and roads mainly on the ridges, cluster B has a tree-like stream network, tendentiously gentle slopes and roads mainly on the ridges as well and cluster C has roads primarily leading across hillsides and partially cased streams. The fairly smaller average field size of cluster C may be due to its roads leading across hillsides, which split more fields apart. While cluster A's diffuse PP emissions reaching surface waters are with roughly 3.4 $kg\,ha^{-1}$ rather high due to its steep slopes, those of cluster B and C are lower and amount to roughly 1.4 and 1.1 $kg\,ha^{-1}$ respectively.

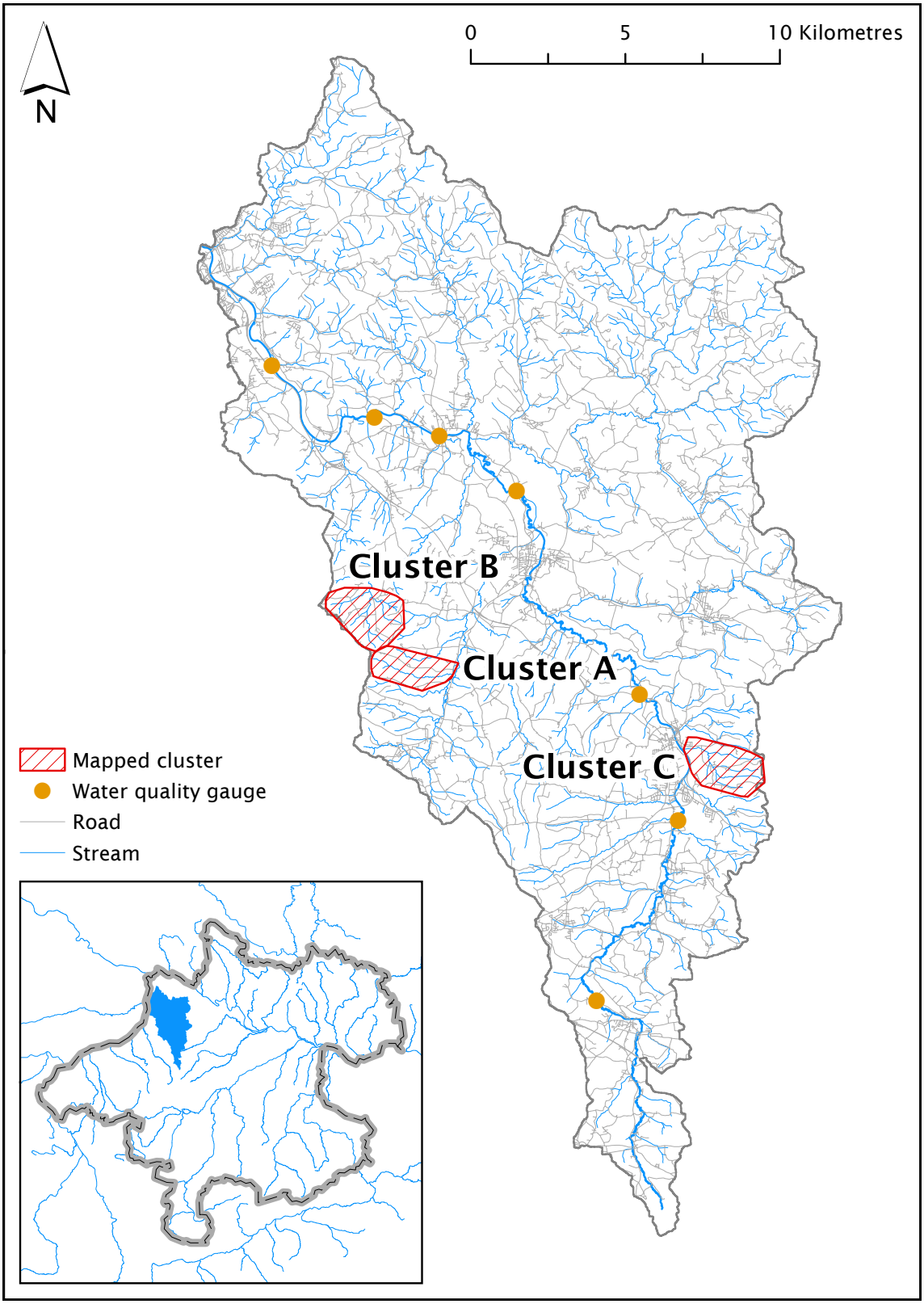

Fig. 3.2: The case study catchment with the three randomly selected and systematically mapped clusters as well as the seven water quality gauges along the river Pram used for calibrating PhosFate.

Fig. 3.3: Roadside ditch with multiple storm drains at varying distances. The left picture was taken upstream of the right.

3.2.4 Modelling the impact of storm drains at road embankments on PP transport

3.2.4.1 The PhosFate model

PhosFate is a semi-empirical, spatially distributed phosphorus emission and transport model created for the identification of critical source areas in a management context at catchment scale (Kovacs, Honti, and Clement, 2008; Kovacs, Honti, Zessner, et al., 2012). It is based on raster GIS data and utilises the (R)USLE (Renard et al., 1997; Schwertmann et al., 1987; Wischmeier and Smith, 1978) incorporating a single flow algorithm version of the slope length factor of Desmet and Govers (1996). PP emissions are then calculated from the erosion and PP content of each raster cell considering a local enrichment ratio (Kovacs, 2013).

The PP transport part of PhosFate consists of the computation of the PP retention via an exponential function of the cell residence time and a mass balance equation considering the inflowing PP load, the local PP emission and the local as well as the transfer PP retention. Computing the cell residence time requires the calculation of the hydraulic radius among others. This variable in turn involves model parameters related to discharge frequency (Kovacs, 2013). So again, a potential (transport potential) is calculated. PP transport as calculated by PhosFate thus reflects maximum potential PP transport for a chosen discharge frequency.

3.2.4.2 Storm drains model extension

The current version of PhosFate does not consider any linear structures potentially influencing PP transport. As a result of the field mapping campaign (see Section 3.3.1), it was deemed necessary to take storm drains at varying distances along road embankments often in combination with roadside ditches (Figure 3.3) into account. An algorithm incorporating known locations of storm water infrastructure is the (W)ASI algorithm (Choi et al., 2011; Choi, 2012). At catchment scale, however, particularly the locations of rural storm drains are hardly known.

It was therefore decided to regard every raster cell bordering and flowing in the direction of a road as a storm drain with its outlet at the nearest stream cell (cf. Alder et al., 2015). Since the

flow lengths in roadside ditches between the actual locations of surface run-off leaving a field and entering a storm drain remain unknown choosing such an approach, the retention potential in roadside ditches is likewise unknown. Retention in roadside ditches was thus considered globally by means of a transfer coefficient. This transfer coefficient represents the share of PP emissions reaching a storm drain, which is further routed to a stream cell. In this way, it emulates retention in roadside ditches.

In order not to account retention in roadside ditches to the bottom cells of fields bordering roads, it was necessary to introduce an additional raster layer taking care of this. The purpose of this layer is to store the amount of PP emissions further routed to a stream cell, which allows for calculating the amount of PP retained in roadside ditches among others.

3.2.4.3 Modelling case study

The storm drains model extension was tested at a spatial resolution of 10×10 m on several scenarios within the case study catchment: two road and three transfer coefficient scenarios. Both types of scenarios were combined one by one adding up to six scenarios in total. The first road scenario takes all asphalt roads (termed all roads) and the second just major roads into account. Road data was taken from an up-to-date governmental reference routing dataset (geoland.at, 2016b) and while the major roads scenario encompasses about 100 urban and 200 km non-urban roads, the all roads scenario encompasses about 500 urban and 850 km non-urban roads.

Compared to the natural stream network, which is about 700 km long, this means that the non-urban road network is even longer than the natural stream network. In this comparison, it has to be considered, though, that roads usually have one uphill side only, whereas streams are fed from both of their sides.

Regarding the transfer coefficient scenarios, values of 0.4, 0.6 and 0.8 were chosen and kept constant during calibration. The channel deposition rate was furthermore calibrated on a fixed total in-channel PP retention at catchment outlet of approximately 20 %. This value was adopted from a modelling study (Zessner, Hepp, Kuderna, Weinberger, Gabriel, and Windhofer, 2014) with the lumped catchment model MONERIS (Behrendt et al., 1999; Venohr et al., 2010; Zessner, Kovacs, et al., 2011). The period of the modelling case study ranged from the year 2008 to the year 2013 and all model parameters related to discharge frequency for the calculation of the hydraulic radius were set accordingly, i. e. to a recurrence interval of six years.

As most important input data for the model served the already mentioned (geo-)database related to the IACS of the CAP of the EU (Hofer et al., 2014). It not only contains field borders of agricultural land but also detailed information on cultivated crops as well as the different factors of the USLE. Other important input data were a DEM with 10×10 m resolution utilised for flow routing as well as slope calculation, a dataset based on the digital cadastre map providing non-agricultural land use and another dataset derived from the digital soil map of Austria encompassing top soil characteristics. The State Government of Upper Austria contributed the latter three datasets. In addition, Manning's roughness coefficients were taken from Engman

(1986) and data on PP accumulation in top soil from Zessner, Gabriel, Kovacs, et al. (2011) and Zessner, Zoboli, et al. (2016). Zessner, Hepp, Zoboli, et al. (2016) and Zessner, Hepp, Kuderna, Weinberger, and Gabriel (2017) give more detailed information on the input data.

Calibration quality was assessed with the help of Nash-Sutcliffe efficiency (NSE), modified Nash-Sutcliffe efficiency (mNSE), percent bias (PBIAS) and ratio of the root mean square error to the standard deviation of measured data (RSR) (Krause et al., 2005; Moriasi et al., 2007) by comparing observed mean annual PP loads of seven water quality gauges along the river Pram (Figure 3.2) with modelled mean annual PP loads at the same locations (Table 3.4).

Water quality data was available from the surface water monitoring programme of the State Government of Upper Austria (Kapfer, 2014) for the full period of the modelling case study (2008 to 2013). It is a routine sampling programme with a sampling interval of two weeks and measures PP concentration according to EN ISO 15681-2 and EN ISO 6878 with a minimum limit of quantification (LOQ) of $0.005\,\mathrm{mg\,L^{-1}}$. This data was then combined with stream discharge from the web GIS eHYD (BMLFUW, 2015) and mean annual PP river loads were calculated according to the flow intervals method of Zoboli, Viglione, et al. (2015).

This method facilitates the calculation of river loads for different flow conditions. Since the PP loads modelled by PhosFate represent solely rainfall induced loads stemming from erosion, they cannot be directly compared to total PP river loads. For this, the total PP river loads first have to be reduced by the amount of PP emanated during base flow conditions.

One of the results of the extended PhosFate model is the ratio of emissions reaching surface waters via cells defined as storm drains on total emissions. In order to check the chosen transfer coefficients, which strongly affect this ratio, for plausibility, it was possible to obtain data concerning roadside ditch cleaning from the State Government of Upper Austria (personal communication). Although this data is from the year 2015 only, it integrates over a period of approximately 10 to 15 years as this is the regular interval of cleaning campaigns. It was not possible to get additional years, since data on these operations is not part of general record keeping.

In total, data on the amount of material removed from a cleaned length of about 80 km was collected. This data was then used to estimate average minimum and maximum values of PP retention per metre roadside ditch (3.0 to $11.3\,\mathrm{g\,m^{-1}}$). Despite the consideration of different soil particle densities and material types as well as cleaning intervals to account for uncertainty, this best to worst case range has to be considered as indicative only, as it originates from little data. Nonetheless, the minimum, mean and maximum values of this estimated range could be used to check if the modelled retention rate due to the chosen transfer coefficients in roadside ditches is plausible.

3.2.5 Statistical inference of the overall impact of storm drains at road embankments on diffuse PP emissions

While fields appear to be the “natural” mapping unit from a management point of view, they are problematic from a statistical point of view. Due to potential flow routing from one field to

another, not all of them are statistically independent. This problem was solved, however, by allocating all mapped fields influenced by road embankments to delineated zero-order catchments.

Subsequently, we applied a Bayesian hierarchical model on the reallocated data. The purpose of this model was to infer the mode as well as an equal-tailed credible and highest posterior density interval (HPDI) for the share of road embankments with subsurface drainage on all road embankments in the case study catchment. In order to make a best estimate of the overall impact of storm drains at road embankments on diffuse PP emissions, the results of the Bayesian model in turn were combined with the results of the extended PhosFate model.

The Bayesian model considered all 154 clusters of the case study catchment of which the data of three was known. First, the number of zero-order catchments influenced by road embankments was drawn from a Poisson distribution with a weakly informative gamma prior for the unmapped clusters. Then the number of influenced zero-order catchments with subsurface drainage was estimated from the total number by means of a binomial distribution. Its success probability was allowed to vary on the logit-scale for each cluster. For this, a weakly informative normal prior on its mean and a weakly informative half-Cauchy prior as a conservative choice on its standard deviation (Gelman, 2006; Gelman, Jakulin, et al., 2008) were applied. Since the parameter of interest is the share of zero-order catchments influenced by road embankments with subsurface drainage on all zero-order catchments influenced by road embankments, a possibly occurring overdispersion in the lower level of the model is not relevant. The model was written in R (R Core Team, 2016) using JAGS (Just Another Gibbs Sampler; Plummer, 2003).

3.3 Results

3.3.1 Field mapping campaign

Among the mapped clusters and agricultural land use types, the minimum share of fields connected in a natural way is about 50 and the maximum about 75 %, whereas fields connected in an artificial way constitute approximately 10 to 30 % (Figure 3.4). The proportion of fields showing both attributes lies between roughly 5 and 15 % and fields not connected at all exhibit a minimum share of 0 and a maximum of about 10 %.

Furthermore, a fifth category was identified and termed “vegetated buffer strips” (VBS). These are grassland fields connected in a natural way and located between arable land and streams showing some of the properties (i. e. dimensions) of “normal” VBS. They are, however, not arable land with VBS but leftovers, which for some reason were not turned into arable land and are not subject to fertiliser restrictions. Their share varies between approximately 10 and 15 % among the grassland fields and constitutes roughly 5 % of all mapped fields. Moreover, a general comparison of all these results—especially when accounting the mapped “VBS” to grassland fields connected in a natural way—indicates a rather low overall variability.

Due to the assessment of the explanatory details, it was possible to analyse the causes of fields not being connected in a natural way and the different types of artificial connections present in the mapped clusters (Figure 3.5). Fields categorised as “VBS” were excluded from this analysis.

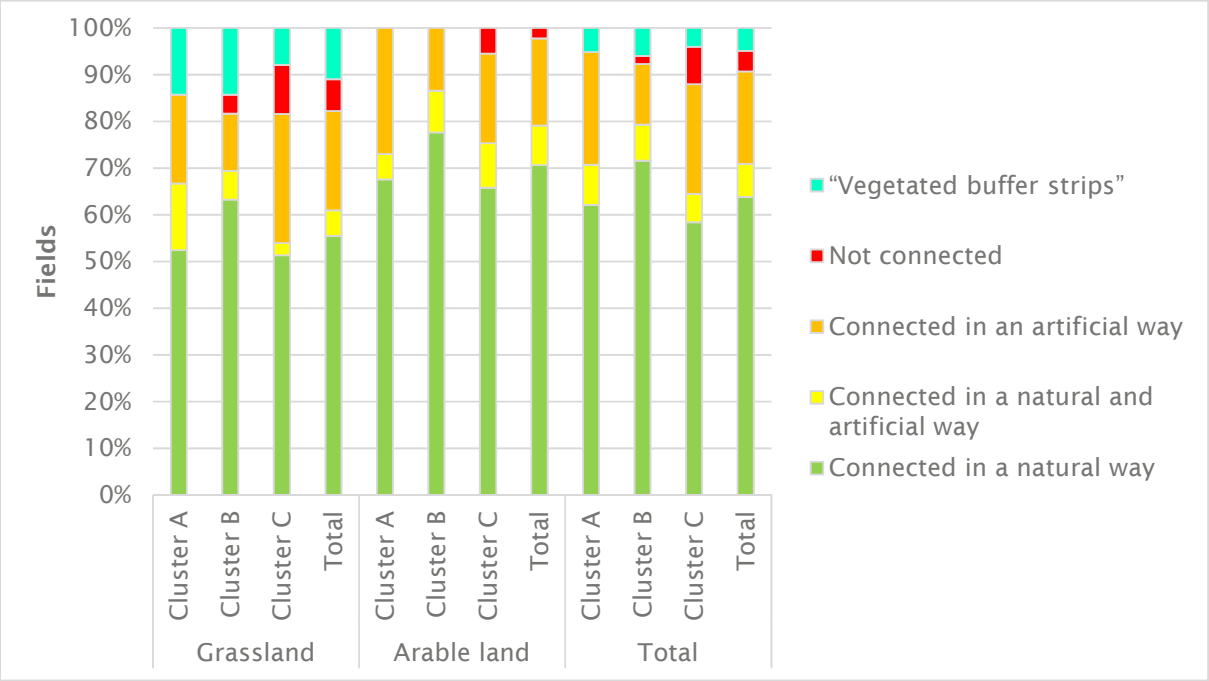

Fig. 3.4: Proportions of the assessed connectivity types among the mapped clusters and agricultural land use types. The results indicate a rather low overall variability ("Vegetated buffer strips" can be considered as grassland fields connected in a natural way).

Table 3.3 shows the causes and their respective shares as identified in the field mapping campaign. These data clearly reveal that road embankments are the main cause of influence on emissions and transport pathways (approx. 65 %). They are followed by single storm drains unrelated to roads or roadside ditches and influences stemming from differences between the actual and the mapped stream network (e. g. as a result of cased streams or unknown channels) (both approx. 15 %). The remaining almost 10 % are composed of a variety of causes including sinks and settlements.

Considering all mapped fields, road embankments with subsurface drainage (i. e. storm drains at varying distances along road embankments often in combination with roadside ditches) are responsible for influencing emissions and transport pathways of almost a quarter of the fields. Given that about 45 % of the mapped fields are subject to an influence, road embankments with subsurface drainage account for approximately half of all influences present in the mapped clusters. They furthermore constitute about 80 % of the fields influenced by road embankments.

3.3.2 Impact of storm drains at road embankments on PP transport

The calibration quality of the six modelled scenarios (two road and three transfer coefficient scenarios) indicates good model performance (cf. Moriasi et al., 2007) and is consistent across all scenarios (Table 3.4). Unfortunately, a validation with independent river loads could not be

5m contour line
Previously mapped stream

Point-shaped and linear structures

Storm drain
Roadside ditch
Sewer
Culvert

Connectivity types

Connected in a natural way
Connected in an artificial way

Explanatory details

Directly connected to a previously mapped stream
Direct connectivity via an inter-field ditch
Direct connectivity via a storm drain
Direct connectivity via a (roadside) ditch with (a) storm drain(s)
Diffuse connectivity via a storm drain
Other (diffuse connectivity via a roadside ditch, culvert and storm drain)
Other (diffuse connectivity via roadside ditches with storm drains and a culvert)

Fig. 3.5: Exemplified visualisation of some of the results of the field mapping campaign with connectivity types shown on the left and explanatory details on the right side of the figure. The part above connectivity types on the left side of the legend is common to both sides. Please note that the top- and rightmost, bright looking field behind the north arrow does not belong to the mapped cluster.

Tab. 3.3: Causes of influence on emissions and transport pathways expressed as percentage of all influenced fields as well as all mapped fields.

Cause	Arable land	Grassland	Total
	% of influenced fields		
Road embankments with subsurface drainage[a]	57	45	51
Road embankments with overland drainage[b]	10	5	8
Road embankments without artificial drainage	4	6	5
Single storm drains[c]	10	15	13
Stream network differences[d]	12	17	14
Other	7	12	9
	% of mapped fields		
Road embankments with subsurface drainage	24	21	23
Road embankments with overland drainage	5	2	3
Road embankments without artificial drainage	2	3	2
Single storm drains	5	7	6
Stream network differences	5	8	6
Other	3	5	4
Total	43	45	44

[a] Storm drains with or without roadside ditches. [b] Roadside ditches and/or culverts.
[c] Storm drains unrelated to roads or roadside ditches.
[d] Differences between the actual and the mapped stream network (e. g. due to cased streams).

performed, since not all of the input data are available at the same spatial resolution in another period of time.

The plausibility check of the chosen PP transfer coefficients in roadside ditches utilising the estimated PP retention per metre roadside ditch range demonstrates that a transfer coefficient of 0.6 fits best with both road scenarios. Whereas a transfer coefficient of 0.8 scrapes at the lower limit of the estimated PP retention per metre roadside ditch range in both instances, a transfer coefficient of 0.4 only scrapes at the upper limit of the major road scenario and indicates that the best fit actually seems to be somewhere between a transfer coefficient of 0.5 and 0.6 in the other case.

A comparison of the modelled shares of storm drains at road embankments on total PP emissions between the six modelled scenarios results in a small range of 8 to 14 % for the major roads scenarios and a big range of 27 to 44 % for the scenarios considering all roads. Overall, the range encompasses a share of as little as approximately one tenth and as much as almost one half.

3.3.3 Best estimate of the overall impact of storm drains at road embankments on diffuse PP emissions

Overall, the retention rate as calculated by the extended PhosFate model fits well to real world data. By comparing the values of the transfer coefficients and the resulting mean estimated PP transfer ratios in roadside ditches, it can be concluded that a transfer coefficient of 0.6 fits best

Tab. 3.4: Modelled shares of storm drains at road embankments on total PP emissions including the calibration quality of the six modelled scenarios and plausibility check of the chosen PP transfer coefficients in roadside ditches utilising the estimated PP retention per metre roadside ditch range. The closer the values of the modelled mean PP retention in roadside ditches and the mean estimated PP retention per metre roadside ditch range for a given transfer coefficient scenario are, the more plausible we consider the coefficient.

	all roads			major roads		
PP transfer coefficient	0.4	0.6	0.8	0.4	0.6	0.8
	Calibration quality					
NSE	0.95	0.94	0.94	0.95	0.95	0.95
mNSE	0.83	0.83	0.82	0.84	0.84	0.84
PBIAS	−2.3	−1.6	−1.5	−2.6	−2.6	−2.6
RSR	0.22	0.22	0.22	0.21	0.21	0.21
	Plausibility check utilising the estimated PP retention per metre roadside ditch range: 3.0 to 11.3 g m^{-1} with a mean of 7.1 g m^{-1}					
Modelled mean PP retention in roadside ditches in g m^{-1}	9.5	5.7	2.6	11.4	7.4	3.6
Share of storm drains at road embankments on total PP emissions	0.27	0.36	0.44	0.08	0.11	0.14

and is therefore the most likely coefficient. However, neither the major roads nor the all roads scenario seems to appropriately reflect the situation in the case study catchment.

According to the results of the Bayesian model, the mode of the share of road embankments with subsurface drainage on all road embankments is 77 % and the equal-tailed 95 % credible interval ranges from 20 to 98 %. While this is quite a big range, the 90 % HPDI ranges from 54 to 100 % only. Taking into account that the case study catchment appears all in all fairly homogeneous, our best estimate of the share of storm drains at road embankments on total PP emissions consequently ranges from about one fifth to one third. This estimate was obtained by combining the 90 % HPDI with the PhosFate results of the all roads scenario with a transfer coefficient of 0.6.

3.4 Discussion

By applying the mapping key presented here, it could be shown that it is capable of collecting valuable information on agricultural and civil engineering structures potentially influencing emissions and transport pathways in agricultural catchments. It clearly revealed that the main influencing factor in the mapped clusters are road embankments with subsurface drainage and

that the most important question in this case study catchment is whether a road embankment is equipped with one or more storm drains or not.

A limitation of the mapping key is that it is adequate to assess structural but not functional elements of connectivity. It thus only points out potential connectivity (Bracken, Wainwright, et al., 2013). This, however, fits well to spatially distributed models like PhosFate, which incorporate the (R)USLE, as this equation only estimates an (erosion) potential as well. Beyond that, due to the involvement of parameters for the calculation of the hydraulic radius related to discharge frequency, PhosFate again calculates a (transport) potential only. This calculation of mere potentials can be considered a limitation of the chosen modelling approach, which otherwise shows a good ratio of model performance to data requirements (cf. de Vente, Poesen, Verstraeten, Govers, et al., 2013).

Calibrating PhosFate on several water quality gauges along the main river of the case study catchment ensured that the utilised input data reflect the spatial distribution of the relevant transport pathways and that the assumption of global calibration parameters was justified. The fact that all modelled scenarios show an equally good calibration quality further supports this.

When interpreting the channel and in particular the field deposition rate resulting from calibration, one has to bear in mind, though, that they include a temporal component. This means that apart from integrating catchment properties influencing transport (e. g. density and average width of hedges as well as reins not covered by the land use input data), they also relate the local modelled mean annual transport loads corresponding to the chosen discharge frequency of the hydraulic radius to the global observed mean annual loads at water quality gauges. In other words, they at least partly control whether a certain area contributes to the overall river load (i. e. becomes connected) or not (cf. S. N. Lane, Reaney, et al., 2009).

Another limitation of the mapping key is its use of discrete mapping units. For one thing, they are of different sizes, for another, due to potential flow routing from one unit to another, not all of them may be statistically independent. The latter can be solved by delineating zero-order catchments and using these as mapping units.

Addressing the issue of different sizes is more challenging. In fact, assuming an equal-sized sampling grid at a spatial resolution smaller than the average field size would increase the issue of dependent samples, whereas a sampling grid at a spatial resolution larger than the average field size would lead to ambiguity. This already is a problem with fields (e. g. same field on both sides of a ridge) and means that a certain property is sometimes only valid for the majority but not all of the assessed area.

Getting back to zero-order catchments, though, and assuming that the sizes of those (parts) showing a certain property and those not showing a certain property average out in a sampling cluster (i. e. small sub-catchment in this case), this issue can probably be neglected, as the weights stay the same in both cases. This would even allow for a geostatistical approach, interpolating ratios between clusters (Krivoruchko et al., 2011). The total number of affected units and their associated areas or pollutant loads could then be estimated with the help of auxiliary data.

Ali and Roy (2009) compare different spatial sampling schemes and conclude that cyclic sampling is the most adequate sampling design for geostatistical analysis. It should therefore be investigated how and with how much effort it can be successfully applied to cluster sampling within catchment areas of several hundred square kilometres and if the above assumption holds true.

So while the use of fields as mapping units worked well, there may be much precision to gain by switching to zero-order catchments in the future. This, nonetheless, has the complication of many possible statistical populations (cf. Poeppl and Parsons, 2018), each strongly depending on the type and resolution of elevation data as well as flow routing algorithm used for their delineation. Again, future research should investigate if mapping results are sensitive to these and if it may be advantageous to only consider zero-order catchments with a certain minimum size or showing a certain shape for statistical inference.

Furthermore, another limitation of the chosen modelling approach results from using a global transfer coefficient. As storm drains are usually built at locations of high overland flow concentration, it is not unlikely that their locations also exhibit high erosion and in turn high diffuse PP emissions along with steep slopes and short or even no passages through roadside ditches where retention can take place. A global transfer coefficient simply considering average conditions may therefore overestimate retention under such conditions. Thus, the overall impact of storm drains at road embankments on diffuse PP emissions may be even higher. This consideration should likewise be subject of future research.

Finally, if one wanted to know the exact locations of, for example, storm drains or culverts, a conceivable solution could be to focus on a narrow strip along roads only. In this case, the question could be formulated as machine learning classification problem for which several existing model types could be utilised (e. g. logistic regression, decision trees, support vector machines and artificial neural networks).

A good start for this might be to use high resolution orthophotos in combination with high resolution airborne topographic Light Detection And Ranging (LiDAR) data as input to learn from. Another possibility would be to equip a car with appropriate sensors and collect topographic LiDAR data including pictures and videos from the roadsides by merely driving along all of them or at least along those of interest (Holland et al., 2016). In the light of rather recent advancements in image recognition with deep neural networks (e. g. He et al., 2015) and their emerging ability to produce a measure of uncertainty even in the context of large-scale image classification problems (Heek and Kalchbrenner, 2019), such a task becomes less and less utopian. To get an idea of what is going on in a catchment area, such detail, however, is usually not required.

3.5 Conclusions

The influence of road embankments with subsurface drainage can be of importance in catchment areas with a strong tradition towards this drainage type. In case there is indication of this in

a certain catchment, a field mapping campaign assessing the degree of their influence should be performed. This is particularly important when trying to identify critical source areas and finding suitable spots for the implementation of measures reducing pollutant emissions into surface waters (cf. Blackwell et al., 1999).

Field observations capable of collecting valuable information on agricultural and civil engineering structures potentially influencing emissions and transport pathways can already be accomplished with very simple means. Carried out with a mobile GIS application and a simple mapping key as the one presented in this study, they can easily be integrated into the preparation of modelling studies aiming at the identification of critical source areas at catchment scale. In the end, the combination and comparison of field observations with simplifying hypothesis (Ludwig et al., 1995) and model output (Jetten, A. de Roo, et al., 1999; van Dijk et al., 2005) proved to be of great help in assessing the impact of agricultural and civil engineering structures in this case study catchment.

Chapter 4

Particulate PhozzyLogic Index for policy makers—an index for a more accurate and transparent identification of critical source areas

4.1 Introduction

An innovative aspect of the European Union Water Framework Directive (European Comission, 2000) is the inclusion of economic principles in river basin and water quality management (Martin-Ortega and Balana, 2012). In particular, conservation measures must be evaluated with respect to how cost-effectively they contribute to the achievement of the "good status" of water bodies. The effectiveness and thus the cost-effectiveness of measures, however, is not uniform and strongly depends on the location of their implementation. This non-uniformity exists on different scales and ranges from the national to the local scale and beyond.

In the field of diffuse pollution, one can find the concept of critical source areas (CSAs) on one of the largest scales, i. e. often the field scale. CSAs are commonly defined as those areas within a watershed, which contribute disproportionately to the pollution load in surface waters via one of many possible transport pathways (e. g. A. L. Heathwaite, Quinn, et al., 2005; Pionke et al., 2000; Strauss et al., 2007). This concept is especially applicable to particulate-bound pollutants, whose emission into water bodies is primarily driven by soil erosion. Phosphorus is a classic example of such substances. Upon excessive fertilisation, the surplus of phosphorus not taken up by crops becomes almost "fixed" in the soil through processes of sorption, precipitation, immobilisation and mineralisation (Tiessen, 2008).

The transfer of phosphorus strongly bound to particles (particulate phosphorus: PP) via soil erosion and surface run-off into surface waters represents a major environmental concern, due to its eutrophication and water quality impairment potential (Sims and Sharpley, 2005). Differences in e. g. slope, crop cover, agricultural practices and connectivity to streams lead to considerable differences in the extent of contributions to the total PP load reaching water bodies. Usually, only a small portion of a watershed is responsible for the majority of phosphorus inputs into its surface waters (e. g. Kovacs, Honti, Zessner, et al., 2012; Strauss et al., 2007; White, Storm, et al., 2009). Sharpley, Kleinman, Jordan, et al. (2009) state as a rule of thumb that about 80 % of the phosphorus inputs into the surface waters of a catchment originate from only about 20 % of its area. This relationship is also known as the 80:20 rule or the Pareto principle.

An early approach to identify such hotspot areas is the Phosphorus Index (Lemunyon and Gilbert, 1993), which was originally developed in the USA but has found its way to Europe since then (L. Heathwaite et al., 2003). By now, there exist several modifications and/or localisations of the Phosphorus Index (Buczko and Kuchenbuch, 2007). A general advantage of all of them is that they are simple, easy to communicate and capable of identifying high-risk regions for a cost-effective implementation of emission mitigation measures (Cherry et al., 2008).

The separation of source and transport factors and the incorporation of hydrologic return periods by Gburek et al. (2000) comprises an important modification. Although the Phosphorus Index is still a semi-quantitative tool, especially the Iowa Phosphorus Index tries to approximate biologically available phosphorus loads entering surface waters (Mallarino et al., 2002). Beyond that, with its stronger physical basis, the Swedish Phosphorus Index requires more input data and is not as easy to calculate as other Phosphorus Indices (Buczko and Kuchenbuch, 2007).

This leads us to another approach to identify CSAs: semi-empirical/conceptual, spatially distributed soil erosion and phosphorus transport modelling. While this approach usually requires even more input data and higher calculation efforts, it can be expected to provide more accurate results than the Phosphorus Index approach, due to its improved quantitative basis. It generally represents a good compromise between solely empirical and process based models (Cherry et al., 2008). Examples are the WaTEM/SEDEM (Onnen et al., 2019; Van Oost et al., 2000; Van Rompaey et al., 2001; Verstraeten, Van Oost, et al., 2002) and PhosFate (Kovacs, Honti, and Clement, 2008; Kovacs, Honti, Zessner, et al., 2012) models.

However, as it is well pointed out by Ghebremichael et al. (2013), one major drawback of this approach is that it is not outright clear how to translate model results into catchment-wide risk assessments at field level, i. e. at the lowest spatial unit managed by farmers and thus the level at which conservation measures can actually be applied. One problem with linking results of hydrological models to prioritisation at field scale is that field and farm boundaries do not usually coincide with hydrological boundaries. This gap makes the outcome of complex catchment-scale models of limited application and use for policy makers, farmers and practitioners involved in agri-environmental programmes. We put forward that fuzzy logic could be the key to both overcome this shortage and to reduce vagueness in identifying CSAs.

Fuzzy logic was developed in the mid-sixties of the twentieth century and is based on fuzzy sets (Zadeh, 1965). Later it was developed further into a theory of possibility (Zadeh, 1978). As possibility theory and Phosphorus Index surprisingly share some very basic ideas, we hypothesise that the potential of phosphorus loss or movement into surface waters from a certain site as it is used by the Phosphorus Index is conceptually related to possibility theory. It is yet lacking its theoretical foundation in mathematics. Hence, why it is called an index.

While fuzzy logic has been previously applied successfully to, for example, river quality (e. g. Lermontov et al., 2009; Liou et al., 2003), groundwater quality (e. g. Rebolledo et al., 2016; Vadiati et al., 2016) and landslide susceptibility (e. g. Champati ray et al., 2007; Pourghasemi et al., 2012; Tien Bui et al., 2012) analyses, to our knowledge it has not been used in the context of the identification of CSAs so far.

The main research goal of this study is to develop a novel semi-empirical “index” for PP based on the combination of spatially distributed models, i. e. PhosFate in our case, with fuzzy logic. As a prerequisite to this, we aim to completely redesign and enhance the model’s existing algorithm for allocating the PP emissions actually reaching surface waters to their respective source areas so that conservation of mass is guaranteed under all circumstances. Furthermore, the study conducts a sensitivity analysis to determine the impacts of selected uncertain input data (flow directions) and model parameters (discharge frequencies) on the designation of CSAs in the form of present-day scenarios for a case study catchment with an area of several hundred square kilometres.

Tab. 4.1: Additional properties of the case study catchment.

	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
Slope[a] in %	0.00	4.59	8.41	10.22	13.50	241.23
Discharge[b] in $m^3 s^{-1}$	0.94	1.87	2.68	4.58	4.23	117.00
Field size in $10^4 m^2$	0.01	0.36	0.90	1.60	2.02	28.16

[a] Based on a DEM with 10×10 m resolution covering the whole catchment area.
[b] Period 2008 to 2013 of the gauge 204867, Pramerdorf/Pram close to the outlet with a catchment area of approx. 340 km^2 (BMLFUW, 2015).

4.2 Material and methods

4.2.1 Case study catchment

The present study was conducted on the catchment of the river Pram in the north-western part of the Austrian federal state of Upper Austria. According to Zessner, Gabriel, Kovacs, et al. (2011), erosion from agricultural land is its only relevant source of diffuse PP emissions. The catchment predominantly belongs to the geologic formation Molasse basin and is approximately 380 km^2 in size. Only some small parts in the north and north-east belong to the crystalline Bohemian Massif. Various types of loam with silt loam being the most prominent (nearly 50 % of the catchment area) dominate its soil surface. Clay is only dominant in about 5 % of the area and sandy soils are not present at all.

The elevation of about 300 m a. s. l. at the mouth of the river in the north-west gradually rises to about 800 m a. s. l. in the south. Annual precipitation shows a similar gradient ranging from roughly 900 mm to around 1200 mm. The catchment's major land use type is agricultural land covering approximately 70 % of the area and can be further divided into about 45 % arable land and 25 % grassland. Other important land use types are forests and settlements, which sum up to about 20 % and almost 10 % of the area respectively. Winter grains are cultivated on approximately 40 % and maize on approximately 30 % of the mostly hilly arable land. Table 4.1 lists a few additional catchment properties.

4.2.2 PP emission and transport modelling

4.2.2.1 The PhosFate model

Originally created by Kovacs, Honti, and Clement (2008) and Kovacs, Honti, Zessner, et al. (2012), the semi-empirical, spatially distributed phosphorus emission and transport model PhosFate was recently extended by Hepp and Zessner (2019) with a module capable of taking into account storm drains at road embankments. It models erosion with the help of the (R)USLE (Renard et al., 1997; Schwertmann et al., 1987; Wischmeier and Smith, 1978) making use of a raster GIS data based single (D8) flow algorithm version of the slope length factor (Desmet and Govers, 1996). The PP emission part of the model then combines the erosion with the PP content of the top soil and a local enrichment ratio of each raster cell in order to calculate local gross PP emissions (Kovacs, 2013).

PP retention in turn is computed via an exponential function of the cell residence time and a mass balance equation including terms for the inflowing PP load, the local gross PP emission and the local as well as the transfer PP retention. The hydraulic radius, among others, is a requirement for the calculation of the cell residence time and itself involves model parameters related to discharge frequency (Kovacs, 2013). With the channel as well as overland PP deposition rates and the PP transfer coefficient of the storm drains model extension, PhosFate features three potential calibration parameters.

4.2.2.2 Revised allocation algorithm

One major output of PhosFate is the calculation of PP cell loads via an allocation algorithm. The PP cell loads describe the amount of local PP emissions actually reaching surface waters, i. e. the PP cell load of a single cell represents its local PP emission minus the cumulated amounts of retention taking place in all of its downstream cells (Kovacs, 2013). Since the original allocation algorithm of the PhosFate model based on transmission coefficients was developed for larger cell sizes (e. g. 100×100 m resolution) than used in this study (10×10 m resolution), it was necessary to revise it. A major drawback of the original algorithm is that conservation of mass is not guaranteed in every single cell, but only on the level of zero-order catchments.

The original algorithm requires a single top-down computation starting with the lowest and finishing with the highest flow accumulations. Its results are the PP retentions as well as transports. The revised algorithm builds on this computation and adds an additional bottom-up computation starting with the highest and finishing with the lowest flow accumulations to it. Its results are the PP cell loads for which the local net PP emissions calculated as local PP emissions minus PP retentions constitute the upper PP cell load limits.

This latter computation in turn involves the calculation of PP cell transfers. These are the cumulated amounts of PP entering a cell from upstream cells, which are transferred through the cell and actually reach a surface water (again minus the cumulated amounts of retention taking place in all of its downstream cells). The complete bottom-up computation thus consists of cycles of the following steps, which, however, are executed for overland cells only:

1. Initialising an intermediate cell load either as the transported amount of PP entering a surface water (riparian cells – first cycle) or as the apportioned cell transfer in the last step of the previous cycle (all other cells – subsequent cycles).

2. Initialising an intermediate cell transfer by setting it equal to the previously initialised intermediate cell load.

3. Updating the intermediate cell load by multiplying it with the ratio of the local and the local plus the sum of the inflowing PP transports.

4. Setting the final cell load utilising the local net PP emission as upper limit (minimum of the previously updated intermediate cell load and local net PP emission).

5. Calculating the cell load carry-over as the maximum of the intermediate cell transfer minus the final cell load and zero.

6. Setting the final cell transfer to the previously calculated cell load carry-over.

7. Weighted apportioning of the cell load carry-over to the inflowing cells utilising their PP transports as weights.

4.2.2.3 Coupling of the PhosFate with the STREAM model

The STREAM model was designed to simulate overland flow and erosion in agricultural catchments (Cerdan, Bissonnais, et al., 2002; Cerdan, Souchère, et al., 2002; Le Bissonnais et al., 2005). Yet it does lack the ability to simulate the emission and transport of chemical substances like phosphorus. What makes it appealing from the point of view of phosphorus emission and transport modelling, though, is its capability to simulate a flow network, which accounts for tillage directions among others (Couturier et al., 2013). The effect of tillage direction on flow direction has been thoroughly studied by Souchere et al. (1998) and Takken, Govers, Jetten, et al. (2001), Takken, Govers, Steegen, et al. (2001), and Takken, Jetten, et al. (2001) and depends mainly on the angle between tillage direction and topographic slope aspect, topographic slope gradient and surface roughness.

In order to let PhosFate benefit from the last mentioned feature, we coupled both models, i. e. we used version 3.7.1 of the STREAM model to pre-process the flow direction data handed over to PhosFate. For this, we applied a stepwise approach with (i) topographic flow directions only (TOPO scenario), (ii) the sole enforcement of tillage directions on topographic flow directions for arable land (TILL scenario) and (iii) the combined enforcement of tillage directions and open furrows at all field borders (sides as well as headlands, which demonstrates an extreme case) on topographic flow directions for arable land (FURR scenario). Fortunately, field border data was available from a (geo-)database related to the Integrated Administration and Control System (IACS) of the Common Agricultural Policy (CAP) of the European Union (EU) (Hofer et al., 2014).

As tillage directions are an input parameter to the STREAM model, they first had to be derived. This was accomplished by means of minimum bounding rectangles with smallest areas enclosing each field and the assumption that tillage directions are parallel to the longer sides of those rectangles. A sample of 176 fields was then used to compare such derived tillage directions to visually determined tillage directions from orthophotos. In only about 16 % of the cases (29 fields), they deviated by more than ±22.5°. Two more required input parameters are the average surface roughness in tillage direction and the average surface roughness perpendicular to tillage direction. These were globally set to 1 to 2 cm and 2 to 5 cm, respectively. Slopes were generally calculated in the direction of the topographic or enforced D8 flow directions.

4.2.2.4 Present-day case study scenarios

Drained roads and storm drains can be a relevant emission pathway into surface waters in Switzerland (Alder et al., 2015; Bug, 2011; Doppler et al., 2012; Prasuhn, 2011; Remund et al., 2021). Hepp and Zessner (2019) came to the same conclusion for Austria and estimated that in the same catchment as examined in this study, about 77 % of all fields upstream to roads are artificially drained by one or more storm drains with a 90 % highest posterior density interval (credible interval, sometimes also called the Bayesian confidence interval) ranging from approximately 54 % to almost 100 %. They further performed a plausibility check of the transfer coefficient of the storm drains model extension accounting for retention in roadside ditches, which delivered a result of roughly 0.60, i. e. 60 % of all PP emissions that enter roadside ditches and subsequently storm drains are reaching surface waters via such bypasses.

With these introductory words said, the three factors used to define the case study scenarios are as follows: (i) transfer coefficients of 0.32, 0.46 and 0.60 corresponding to 54 %, 77 % and 100 % of 0.60 (TC0.32, TC0.46, TC0.60), (ii) discharge frequencies of one (T1) and six years (T6) (utilised for the calculation of the hydraulic radius/residence time and in turn PP retention) and (iii) the three described STREAM model flow direction scenarios, namely, TOPO (topographic flow directions), TILL (enforcement of tillage directions) and FURR (tillage directions combined with open furrows at all field borders). So all in all 18 scenarios (three transfer coefficients times two discharge frequencies times three flow direction data sets) were modelled and assessed.

Storm drains at road embankments of almost all asphalt roads were taken into account with the help of the storm drains model extension. A governmental reference routing dataset provided the necessary road data for this (geoland.at, 2016b). The period of the modelled scenarios ranged from the year 2008 to the year 2013. Water quality data of seven water quality gauges (additionally shown in Figure 4.6) could be used to calculate mean annual PP river loads using the same methodology as of Hepp and Zessner (2019).

A detailed overview on all input data is provided by Zessner, Hepp, Zoboli, et al. (2016) and Zessner, Hepp, Kuderna, Weinberger, and Gabriel (2017). Some other notable datasets nevertheless are (i) the already mentioned (geo-)database related to the IACS of the CAP of the EU contributing detailed information on the cultivated crops, the different factors of the USLE as well as the field borders (Hofer et al., 2014), (ii) a DEM with 10×10 m resolution, which served as the main input data for the STREAM model, (iii) non-agricultural land use based on the digital cadastre map and (iv) top soil characteristics derived from the digital soil map of Austria. The latter three datasets were supplied by the State Government of Upper Austria. Lastly, data on PP accumulation in top soil was extracted from Zessner, Gabriel, Kovacs, et al. (2011) and Zessner, Zoboli, et al. (2016) and Manning's roughness coefficients from Engman (1986).

Each of the nine T6 scenarios was then calibrated individually under a channel deposition rate of zero utilising the previously calculated mean annual PP river loads as targets and the overland deposition rate as the only calibration parameter. This more or less reflects long-term conditions including in-stream phosphorus stock depletion effects caused by major flood events (Zoboli, Viglione, et al., 2015). It can also be considered a rather conservative approach for identifying

CSAs, as it generally leads to lower PP loads from cells farther away from surface waters due to higher overland deposition rates. The overland deposition rates of the nine T1 scenarios were subsequently adopted from the corresponding T6 scenarios thus simulating conditions as if no respective flood/transport event took place in the modelled period of six years. In order to achieve this, the discharge frequency related parameters of the hydraulic radius calculation are altered accordingly (Liu and De Smedt, 2004; Molnár and Ramírez, 1998).

4.2.3 Particulate PhozzyLogic Index (PPLI)

While providing interesting details, a map displaying the PP cell loads with, for example, $10 \times 10\,\mathrm{m}$ resolution is not outright helpful to policy makers, who have to decide where to cost-effectively implement an emission mitigation measure and where not. This information therefore has to be somehow transformed. For this purpose, we first calculate the total absolute PP contributions to surface waters for each field, i. e. the sum of all PP cell loads of each field with the help of zonal statistics and then apply a fuzzy membership function to those results.

4.2.3.1 Fuzzy membership function

A short introduction to fuzzy sets and fuzzy membership functions is given by V. B. Robinson (2003). Basically, fuzzy membership functions are functions turning data sets of arbitrary ranges and scales into fuzzy sets consisting solely of values between zero and one, where zero can be translated into "not possible" or not a member of a given set and one into "perfectly possible" or a member of a given set. All other values between zero and one represent a varying degree of possibility or membership in a given set. A value of, for instance, 0.7 thus stands for a higher degree of possibility or membership in a given set than a value of e. g. 0.4.

Possibility must not be confused with probability, even though they share the same value range. While probabilities are estimated based on data as well as certain assumptions and provide confidence intervals, possibility, especially due to the subjective choice of fuzzy membership functions, is less objective. Nonetheless, once one or more fuzzy membership functions have been chosen, possibility is perfectly objective and can even be used to objectively compare expert judgements. This approach is particularly interesting and helpful in a sparse data environment such as in the case of a cost-effective implementation of mitigation measures against PP inputs into surface waters.

The fuzzy membership function we have chosen is the so called fuzzy large membership function. It has the form

$$\mu(x) = \frac{1}{1 + \left(\frac{x}{p_2}\right)^{-p_1}},$$

where p_1 is the spread and p_2 is the midpoint. For the spread, we generally put in one, which makes the shape of the function somewhat similar to that of a logarithmic function, and for the midpoint, we put in a different value for each scenario. These values were calculated so that 80 % of the respective agricultural land received a value of less than 0.5 and 20 % a value of 0.5 or

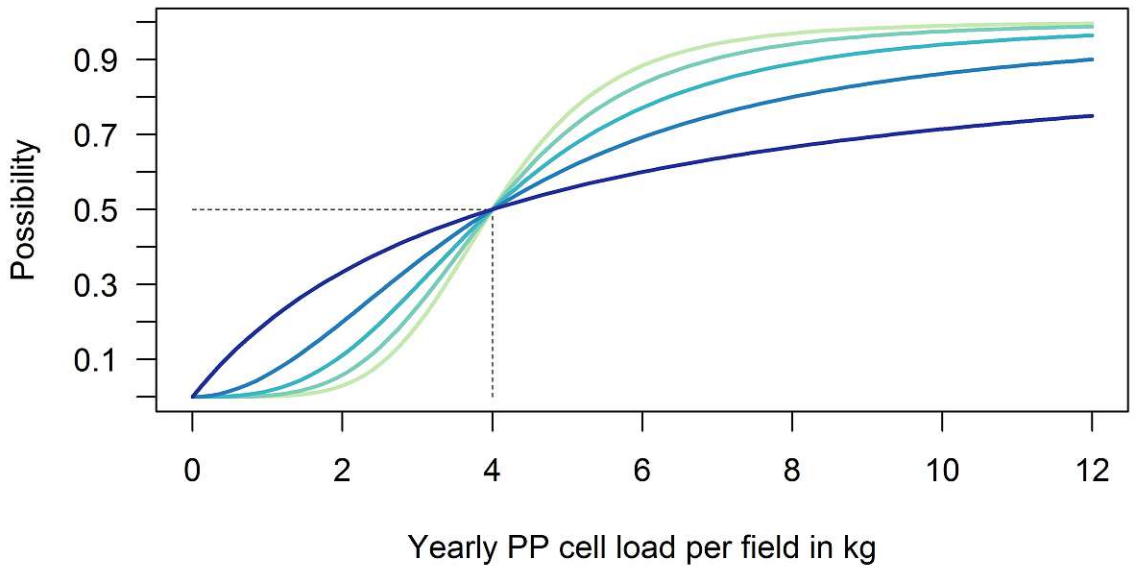

Fig. 4.1: Examples of fuzzy large membership functions with a midpoint of 4 kg yr^{-1} PP cell load per field and spreads ranging from one (dark blue curve) to five (yellow-green curve). With increasing spread, the function gradually transforms itself from a function with a shape similar to that of a logarithmic function to one with a shape similar to that of a sigmoid function. The black dashed lines shall help illustrate the meaning of the midpoint (possibility of 0.5).

more (Pareto principle; cf. Sharpley, Kleinman, Jordan, et al., 2009). In this way, the 20 % of the total area of agricultural land belonging to the fields with the highest absolute PP inputs into surface waters received a value of 0.5 or more and are therefore what we consider CSAs. Figure 4.1 provides examples of fuzzy large membership functions with a midpoint of 4 kg yr^{-1} PP cell load per field and spreads ranging from one to five.

4.2.3.2 Final Particulate PhozzyLogic Index creation

While the application of the described fuzzy large membership function to the model results of a single scenario could already be referred to as a fuzzy logic based PP index, this is only the first step, since multiple fuzzy sets can further be overlaid in order to take into account, among others, the uncertainty of input data. A simple technique for this is the so-called convex combination (Dubois and Prade, 1985; Kandel, 1986). In the process of convex combination, each involved fuzzy set can be assigned a weight, which makes it a weighted mean operator with weights adding up to one. Charnpratheep et al. (1997) also propose a modified version of this operator, which sets the overall result to zero in case one or more of the combined fuzzy sets are zero.

As no indication exists that one of our 18 scenarios is more possible than another as far as it comes to a single field and we neither favour one discharge frequency over the other, we used equal weights to overlay the fuzzy sets of all our scenarios with the unmodified convex combination operator in order to create the final PPLI. Using the modified operator does not contribute to the quality of the result in the absence of a knock-out criterion.

4.3 Results and discussion

4.3.1 STREAM model flow directions

All in all, the flow direction of about 28 % (approx. 440 000 cells) of the arable land were altered by enforcing tillage directions on topographic flow directions (TOPO vs. TILL). While comparing TOPO and FURR results in a somewhat higher share of about 31 % (approx. 500 000 cells), comparing TILL and FURR only shows a rather small share of about 8 % (approx. 130 000 cells). This is reasonable, since the field borders merely comprise a small portion of the cells representing the arable land of the catchment. Additionally, these shares reveal that about 5 % of the cells were altered two times, first by the TILL and then by the FURR scenario.

Of the roughly 7000 arable fields, nearly 6200 (approx. 88 %) are affected by at least one cell with altered flow directions due to tillage directions and obviously all fields are affected by altered flow directions due to open furrows. An example of the resulting flow directions of the three scenarios is given in Figure 4.2.

4.3.2 Calibrated overland deposition rates

Calibration quality of the nine independently calibrated T6 scenarios is altogether very similar. They all exhibit a Nash-Sutcliffe efficiency (NSE) of about 0.95, a modified Nash-Sutcliffe efficiency (mNSE) of around 0.83, a percent bias (PBIAS) of almost zero and a ratio of the root mean square error to the standard deviation of measured data (RSR) of roughly 0.21 (Krause et al., 2005; Moriasi et al., 2007; Nash and Sutcliffe, 1970). Figure 4.3 shows the observed and modelled yearly PP loads of the representative TC0.46-T6-TOPO scenario at the seven water quality gauges along the river Pram. Deviations are bigger for smaller loads, which is also why the mNSE indicates a poorer performance. Poesen (2018) considers scaling up sediment yields from field to catchment scale one of the main challenges in geomorphological research. Particularly seen in this light, these results are thus very promising.

Calibrated overland deposition rates range from 0.91×10^{-3} to $1.46 \times 10^{-3}\,\mathrm{s}^{-1}$ and decrease with lower transfer coefficients (1.16×10^{-3} to $1.46 \times 10^{-3}\,\mathrm{s}^{-1}$ for 0.60 and 0.91×10^{-3} to $1.21 \times 10^{-3}\,\mathrm{s}^{-1}$ for 0.32), which is reasonable, as lower inputs into surface waters from fields upstream of roads have to be compensated with higher inputs from fields upstream of surface waters. Generally, the deposition rates of the FURR scenario group are approximately by $0.30 \times 10^{-3}\,\mathrm{s}^{-1}$ lower than those of the other two STREAM model scenario groups.

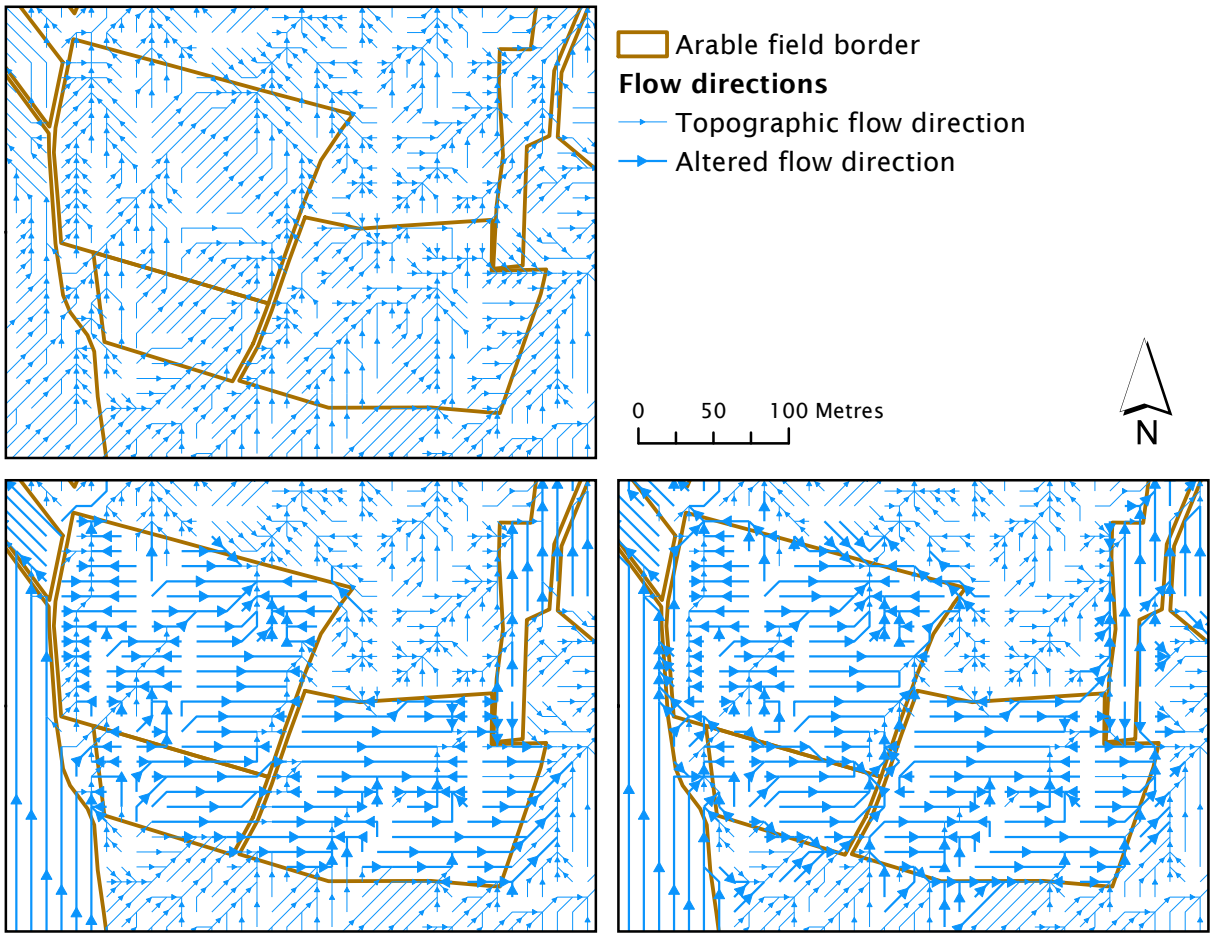

Fig. 4.2: Example of the resulting flow directions of the TOPO (top left), TILL (bottom left) and FURR scenario (bottom right). Especially the FURR scenario leads to partially more extreme flow accumulations, as it concentrates flow at field borders and limits spillovers to neighbouring fields to cells with a single possible downstream path.

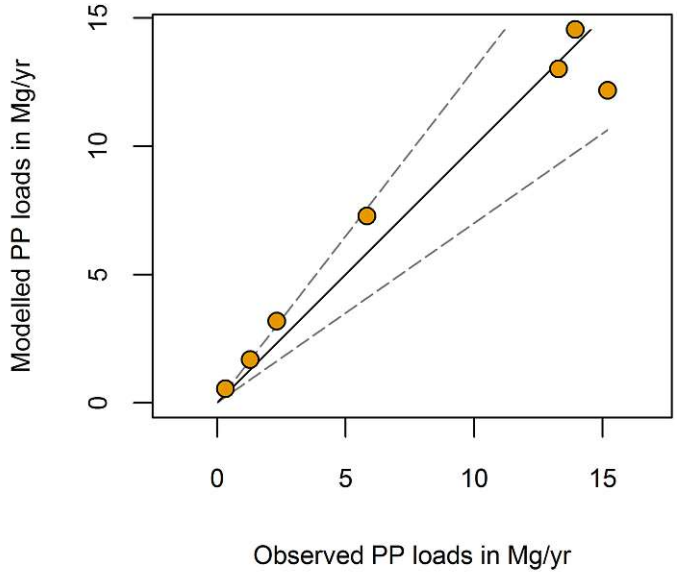

Fig. 4.3: Observed and modelled yearly PP loads of the representative TC0.46-T6-TOPO scenario at the seven water quality gauges along the river Pram. The solid black line represents the 1:1 line and the grey dashed lines 30 % deviations from the 1:1 line.

The lower deposition rates of the FURR scenarios may appear counter-intuitive. They are, however, a product of partially lower flow accumulations in the interior of the fields and higher flow accumulations along the field borders. Apparently, the higher transport potential along field borders, which affects rather few cells, does not fully compensate the lower transport potential, which affects the comparatively higher number of cells of field interiors. Thus, the process of calibration has to result in the overall lower deposition rates of the FURR scenario group in order to make up for the gap between the modelled and calculated PP river loads.

4.3.3 PP inputs into surface waters

Catchment-wide PP inputs into surface waters range from about 8.0 to 8.7 $\mathrm{Mg\,yr^{-1}}$ for the T1 scenario group and 15.9 to 16.2 $\mathrm{Mg\,yr^{-1}}$ for the T6 scenario group. Storm drains at road embankments account for approximately one quarter to almost one half of the PP inputs. The respective shares range from 24 to 29 % for the TC0.32 scenarios, 32 to 37 % for the TC0.46 scenarios and 39 to 44 % for the TC0.60 scenarios. This is in line with findings of Prasuhn (2011) whose ten year long field survey resulted in a share of 22 % of the total eroded soil. He also states that about half of the eroded soil was already retained within the borders of the source field though. This share therefore at least has to be doubled in order to account for the input into surface waters alone.

Tab. 4.2: Selected percentiles of the sums of the PP cell loads per field of arable land in $kg\,yr^{-1}$ for two selected scenarios.

Scenario/percentile	0.30	0.50	0.70	0.90	0.92	0.94	0.96	0.98
TC0.46-T1-TOPO	0.01	0.08	0.46	2.27	2.79	3.65	5.04	8.09
TC0.46-T6-TOPO	0.03	0.24	1.09	4.75	5.98	7.71	10.15	16.28

Fields of arable land are predominantly responsible for the catchment-wide PP inputs into surface waters (96.6 to 98.0 %). Grassland fields (1.1 to 1.8 %) and other land use types (0.9 to 1.6 %) contribute insignificant amounts only; hence, our focus lies on arable fields. Table 4.2 shows selected percentiles of the sums of yearly PP cell loads per field of arable land for two selected scenarios. The distributions are as expected very skewed. In the case of the TC0.46-T1-TOPO scenario, 70 % of the fields contribute less than $0.5\,kg\,yr^{-1}$ PP to the total PP emissions into surface waters and even in the case of the TC0.46-T6-TOPO scenario the same percentile amounts to as little as nearly $1.1\,kg\,yr^{-1}$ PP. The amounts of the higher percentiles are increasing rapidly from there on.

Individual yearly PP cell loads and transfers of a small area within the case study catchment are displayed in Figure 4.4 for the TC0.46-T6-TOPO scenario. The figure clearly depicts that forests hardly contribute to PP emissions into surface waters, but act as important transfer zones for upstream PP loads mainly via preferential flow pathways. This finds confirmation in the outcomes of tracer experiments conducted e. g. in France (van der Heijden et al., 2013) and in Germany (Julich et al., 2017), which have shown that phosphorus transport in forests occurs along preferential flow pathways, largely bypassing the nutrient-poor soil matrix of forests.

4.3.4 Identified critical source areas

The number of fields classified as a CSA, i. e. fields with a possibility of 0.5 or more, ranges from 1151 to 1233 for all the modelled scenarios. 14 538 fields, on the other hand, are not classified as a CSA in any of the scenarios. This denotes an average of 1196 fields, which in relation to the total number of fields of 16 320 constitutes a share of about 7.3 %. The differences in the number of CSAs can be regarded as negligible among the scenarios.

Furthermore, the median area of fields classified as a CSA in at least one scenario is $2.94 \times 10^4\,m^2$ with an interquartile range of $3.30 \times 10^4\,m^2$. The median area of fields never classified as a CSA is $0.74 \times 10^4\,m^2$ with an interquartile range of $1.31 \times 10^4\,m^2$. This considerable difference in the median field size and interquartile range explains the relative small share of fields regarded as CSAs compared to the approximately 20 % of the total area of agricultural land under consideration.

A major reason for this percentage mismatch is the choice of the sum of all PP cell loads per field as input to the fuzzy large membership function, which favours the selection of comparatively large fields. The purpose of this choice is the maximisation of the anticipated decrease in PP emissions into surface waters per field due to the implementation of suitable mitigation measures

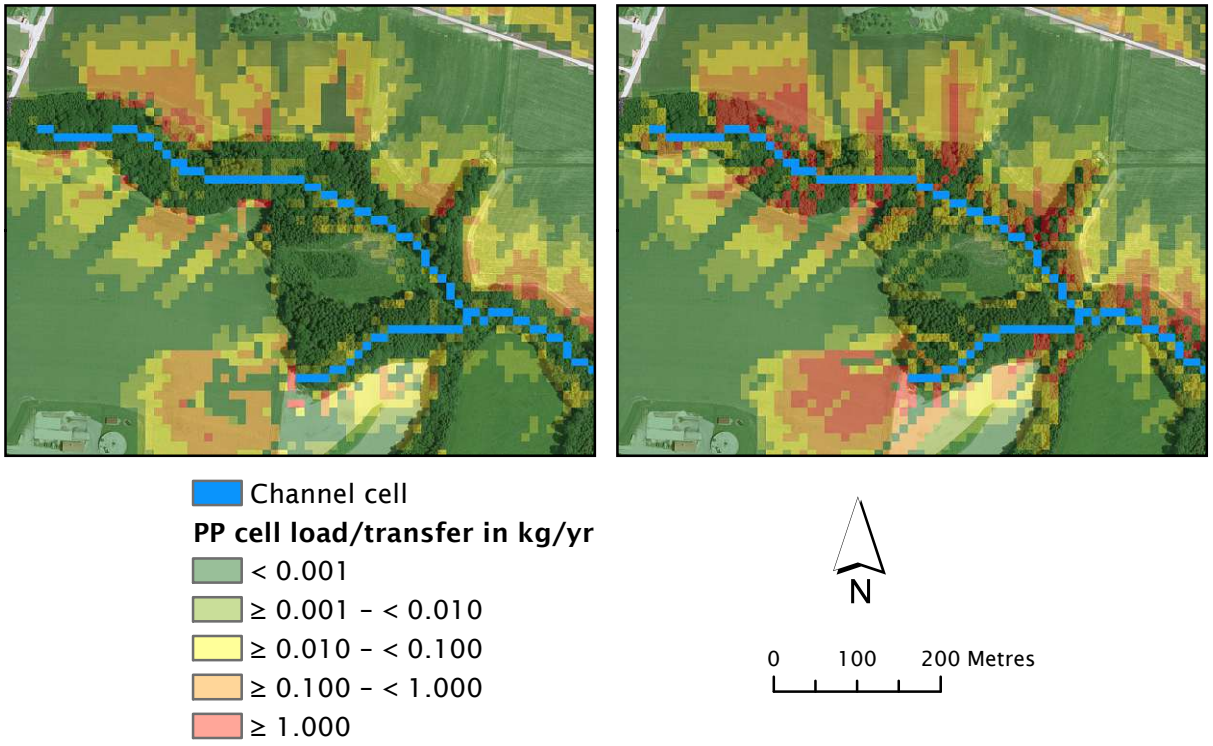

Fig. 4.4: Individual yearly PP cell loads (left) and transfers (right) of a small area within the case study catchment for the TC0.46-T6-TOPO scenario.

and, at the same time, the minimisation of the number of farmers involved. While other choices like the average of all PP cell loads per field may offer a better overall potential cost-effectiveness ratio, from a pragmatic perspective it would be likely more beneficial to have to convince fewer farmers to participate in a catchment-wide water quality protection programme. With potential cost-effectiveness ratio, we here refer to the area of arable land taken out of production and thus associated with possible monetary compensations in order to reduce PP inputs into surface waters by a given amount.

The applied fuzzy large membership functions are designed to classify 20 % of the total area of agricultural land as CSAs. Depending on the scenario, this share is responsible for 79 to 83 % of the total PP inputs into surface waters, in line with the general 80:20 rule (Sharpley, Kleinman, Jordan, et al., 2009). Furthermore, apart from two outliers, all fields classified as a CSA in at least one of the scenarios (1782 fields) are arable fields. As a result, the 20 % share of the total agricultural land classified as a CSA actually represents roughly 30 % of the total arable land. This nonetheless underlines the general principle at work.

Table 4.3 shows the number of times as well as the cumulative share in which the same field is classified as a CSA. For the TC and T scenario groups, the shares of fields classified as a CSA in all of their respective scenarios lie, despite their different group sizes, around 50 %. For the flow direction scenario groups (TOPO, TILL and FURR), the same shares account for somewhat more than 60 % of the fields. These differing shares can be explained by the fact that all flow

Tab. 4.3: Number of times as well as the cumulative share in which the same field is classified as a CSA shown separately for each of the assessed scenario groups. Please note that the scenarios of the TC and flow direction scenario groups each occur only six times (the three scenarios of the respective other group times the two scenarios of the T scenario group), while the scenarios of the T scenario group each occur nine times (the three scenarios of the TC scenario group times the three scenarios of the flow direction scenario group).

Scenarios	9	8	7	6	5	4	3	2	1	0
	No. of fields									
TC0.32				857	80	210	100	212	157	14 704
TC0.46				840	94	183	118	194	174	14 717
TC0.60				825	97	192	110	205	156	14 735
T1	815	92	75	141	72	75	133	104	145	14 668
T6	816	63	80	149	65	87	119	108	118	14 715
TOPO				928	114	113	96	116	130	14 823
TILL				927	121	108	100	108	139	14 817
FURR				924	103	92	99	94	114	14 894
	Cumulative %									
TC0.32				53	58	71	77	90	100	
TC0.46				52	58	70	77	89	100	
TC0.60				52	58	70	77	90	100	
T1	49	55	59	68	72	77	85	91	100	
T6	51	55	60	69	73	79	86	93	100	
TOPO				62	70	77	84	91	100	
TILL				62	70	77	84	91	100	
FURR				65	72	78	85	92	100	

directions and in turn flow accumulations are homogeneous within the flow direction scenario groups, but heterogeneous within the other scenario groups.

So-called UpSet plots allow for a rather easy to interpret visualisation of multiple intersecting sets (Conway et al., 2017; Lex et al., 2014). Figure 4.5 shows such a plot for a subset of the intersections of the CSAs of all T6 scenarios. All nine scenarios agree on 816 (46 %) of the in total 1782 fields, which are classified as a CSA at least once. Furthermore, there are 153 (96 plus 57; 9 %) CSA fields where the FURR scenario group makes a clear difference. This discrepancy emphasises the importance of knowing the relevant preferential flow pathways in a catchment under consideration. Another recognisable pattern is that the scenarios with lower transfer coefficients (especially TC0.32) differ slightly from those with higher transfer coefficients (TC0.46 and especially TC0.60). The T1 scenarios exhibit a similar pattern, albeit not as distinct as the one of the T6 scenarios.

Evenson et al. (2021) compared the use of five different watershed-scale models for the identification of CSAs of phosphorus emissions in the 17 000 km^2 Maumee River watershed. Their study reveals that on average only 16 to 46 % of sub-watersheds were identified as CSAs by more than one model. This large disagreement is attributable to differences in input data, parametrisation and model structure. Based on these outcomes, Evenson et al. (2021) conclude

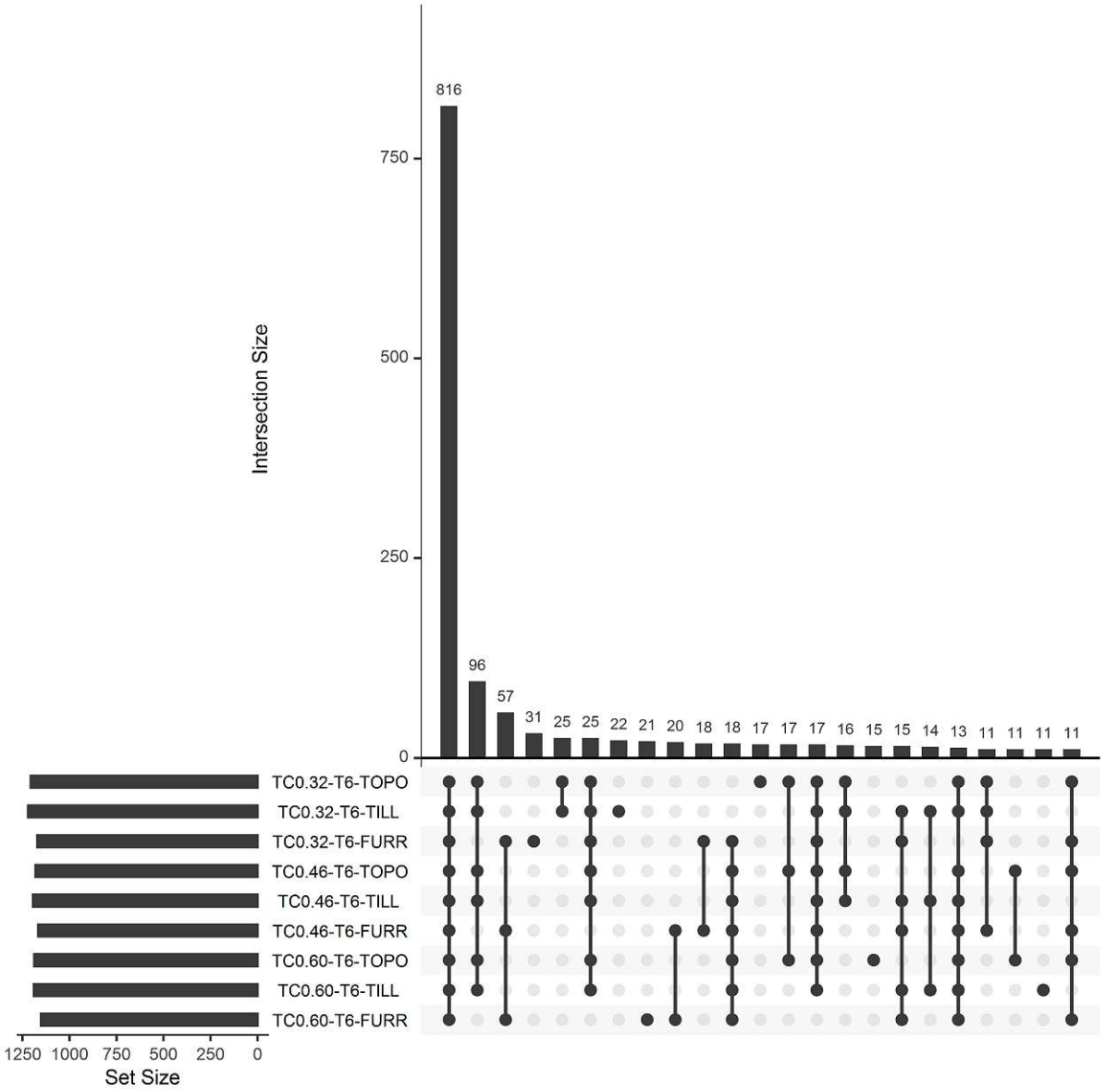

Fig. 4.5: UpSet plot for a subset of the intersections of the CSAs of all T6 scenarios. While the set sizes of the nine sets are presented in the bottom left and their intersection sizes in the top right area, the bottom right area provides information about the sets involved (black dots connected with a black line or single black dots) in each of the above displayed intersections.

that: (i) the share of watershed identified as CSAs by the different models can be selected for conservation measures with a high level of confidence; (ii) a comprehensive uncertainty analysis is essential in this field of research to enhance understanding and acceptance of modelling results by planners, decision makers and farmers.

All in all, our 18 scenarios perfectly agree on 720 of the 1782 fields classified as a CSA. This is a share of about 40 %. Therefore, similarly to the conclusion of Evenson et al. (2021), emission mitigation measures could concentrate on these areas with a high level of confidence. The remaining 60 % of fields allocated as CSAs are diversely distributed among the various possible degrees of scenario intersections. Through the set-up of scenarios accounting for different sources of uncertainty, this study provides a transparent basis, which allows to understand which reasons lead to the different identification of CSAs. By merging the information of this screening as well as the uncertainty reasons with local knowledge of farmers and practitioners active on the territory, an optimal selection of fields could be achieved for the implementation of measures.

Including even more scenarios could push these two numbers even further apart, particularly, since each of the assessed scenario groups comes with its own limitations. A limitation of the transfer coefficient scenario group is for sure the choice of a global transfer coefficient. This assumption may not hold, as areas with higher flow accumulations may also exhibit higher transfer rates. Another limitation is that the presence or absence as well as the exact location of individual storm drains remains uncertain (Hepp and Zessner, 2019).

While the erosion part of PhosFate based on the (R)USLE reflects average conditions, its transport part has to be adjusted to a specific discharge frequency. With longer periods both parts eventually represent average conditions. Especially the T1 scenarios exhibit a mismatch in this regard. Increasing the discharge frequency (e. g. from six to one year) yet has a similar effect as increasing the overland deposition rate. Since the latter acts as calibration and therefore as catch-all parameter for, among others, the effects of catchment elements, which cannot be represented by the chosen spatial resolution of 10×10 m, but where retention can take place (e. g. unploughed strips between fields and hedges), the former can be used as a proxy for studying the effects of the presence or absence of such elements. The global nature of such an increase in discharge frequency once more poses a severe limitation, as it mainly influences long-distance transport.

Finally, the limitations of the flow direction scenarios stem from limited knowledge on the actual tillage directions, surface roughness in and perpendicular to tillage directions and exact locations of open furrows capable of concentrating flow. Some or even all of these shortcomings may be remedied by better available data or improved methods to derive them from existing data in the future. Tillage directions, for example, could be more accurately derived for L-shaped fields from field border data encompassing every single cultivated crop and not only the outer border of all crops in direct vicinity, which are cultivated by the same farmer. The latter, unfortunately, applies in this case with the field borders currently available.

4.3.5 The final Particulate PhozzyLogic Index map

Figure 4.6 presents the final PPLI map of the case study catchment. The fields coloured violet can be considered the most possible CSAs. Most CSA fields are located near the catchment boundary, where the slopes are comparably higher, and in the northwest of the catchment. This map is the result of only a single option of overlaying multiple fuzzy sets. Further options, which should be tested in future research, are, for example, the application of different weights or other fuzzy overlay operators like the fuzzy gamma operator (Zimmermann and Zysno, 1980).

No matter how the different fuzzy sets are combined in the end, such a PPLI map can be a significant step forward in practice, since it allows each farmer to compare his or her fields to all the other fields within a certain catchment. This can be of great help when it comes to evaluating if – in the context of the entire catchment – a certain field substantially contributes to the overall PP emissions into surface waters or not. Furthermore, such a map provides a clear ranking for policy makers and could be used as a starting point for developing a state aid programme, which promotes the implementation of mitigation measures in a cost-effective way. For this purpose, the height of subsidies could even be linked to the degree of possibility of a field being a CSA or not.

Such a map can thus foremost provide a detailed screening for the selection of the farmers to be involved in a programme of measures. Yet we consider it important that the final decision on the kind and exact location of the implemented mitigation measure stays with the farmer, as he or she has the best knowledge with respect to local details, which are not available for modelling. For example, in many cases only a small portion of vegetated buffer strips receives the majority of overland flow (Djodjic and Villa, 2015; White and Arnold, 2009), hence, the concept of CSAs can even be translated to a single field. Allowing farmers to choose from a variety of mitigation measures and combine them in a smart way may be capable of significantly improve the implementation quality and effectiveness.

Sharpley, Kleinman, Flaten, et al. (2011), while praising easy-to-use and well understandable colour-coded maps derived from P Indices, criticise their reliability and address the need of considering the spatial complexity of watersheds, the heterogeneous response time as well as delayed release of legacy PP, which eventually determine the expected effectiveness of measures on the long-term. The PPLI map presented here has the same traits praised by Sharpley, Kleinman, Flaten, et al. (2011) and conveys information in a straightforward and easily interpretable way, but at the same time it builds on a much more solid basis, namely on the results of a complex fate and transport model.

More recently, Wang et al. (2020) highlight in their review of modelling of phosphorus loss from field to watershed the strong need for integrating comprehensive uncertainty analyses in the studies to provide a more transparent and robust support to decision makers. Despite the above-mentioned limitations, the PPLI map based on the one hand on the outcomes of the PhosFate model and on the other hand on the overlay of several present-day scenario fuzzy sets, which take into account multiple sources of uncertainty, clearly meets the criteria of the advocated way forward in this field of research.

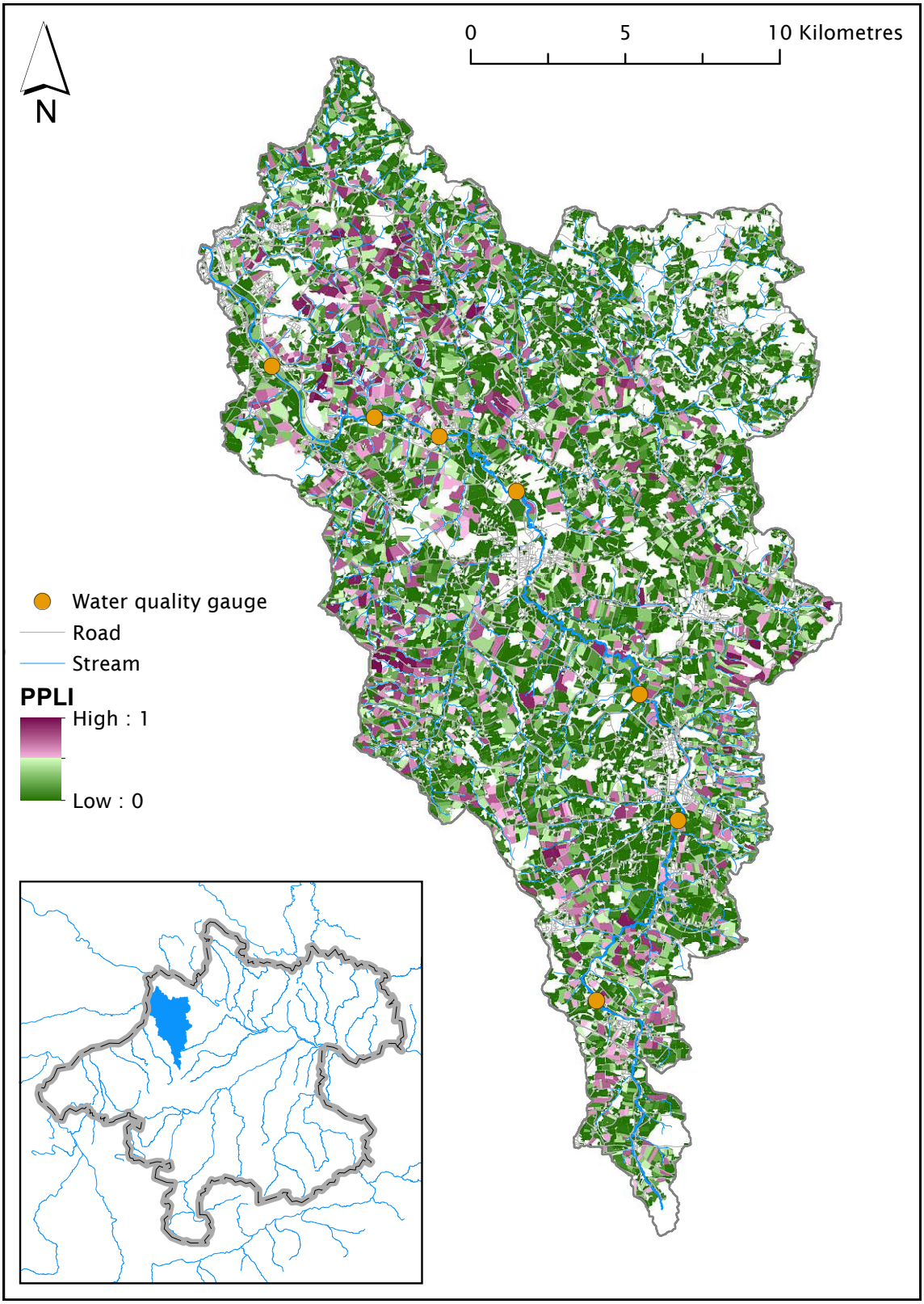

Fig. 4.6: Final PPLI map of the case study catchment: fields with an overall possibility of 0.5 or more, which can be considered the most possible CSAs, are coloured violet. A darker tone thereby indicates a higher possibility.

4.4 Conclusions

This study successfully developed a new algorithm for the allocation of PP emissions entering surface waters to their respective source areas. The algorithm was implemented into the existing semi-empirical, spatially distributed phosphorus emission and transport model called PhosFate. An innovative aspect of this algorithm is that in comparison to the existing one it does guarantee conservation of mass in every single cell and not only on the level of zero-order catchments.

With the help of fuzzy logic, it was then possible to translate these model results into a novel Phosphorus Index with a sound theoretical foundation in mathematics and a great value for management purposes. The novel Particulate PhozzyLogic Index (PPLI) ranks all agricultural fields of a potential large catchment with respect to their possibility of emitting high PP emissions actually reaching surface waters. Its range of possibility lies between zero ("not possible") and one ("perfectly possible"). Possibility values of 0.5 or higher imply that a field belongs to the 20 % of agricultural land responsible for 80 % of the PP inputs into surface waters, which is also our definition of CSAs.

The sensitivity analysis based on 18 scenarios shows that especially open furrows at field borders have the potential to cause deviating CSAs. A thorough validation of the identified CSAs within a catchment of several hundred square kilometres poses a major challenge and future research has to be carried out concerning the development of adequate validation strategies. While a validation strategy based on a representative sample of fields may be a conceivable option, an indirect strategy based on long-term measurements of PP concentrations in a representative sample of small sub-catchments may be a more viable yet also more approximative choice.

Another topic of future research should be the advancement of methods able to derive the positions of: (i) open furrows and other features capable of concentrating flow and (ii) agricultural and civil engineering structures (e. g. storm drains) acting as potential short cuts for surface run-off on its way to surface waters. In addition, a multi-flow version of the presented allocation algorithm, which is not only capable to model flow concentration, but also flow divergence may be a relevant topic for future research.

Appendix A

Residuals and correlation matrix

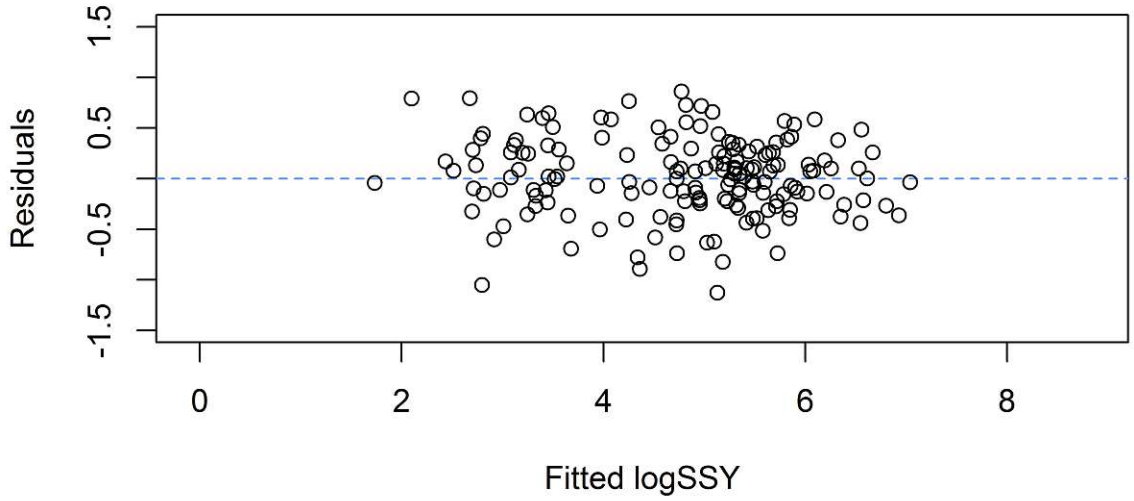

Fig. A.1: Plot of residuals for the best fit model.

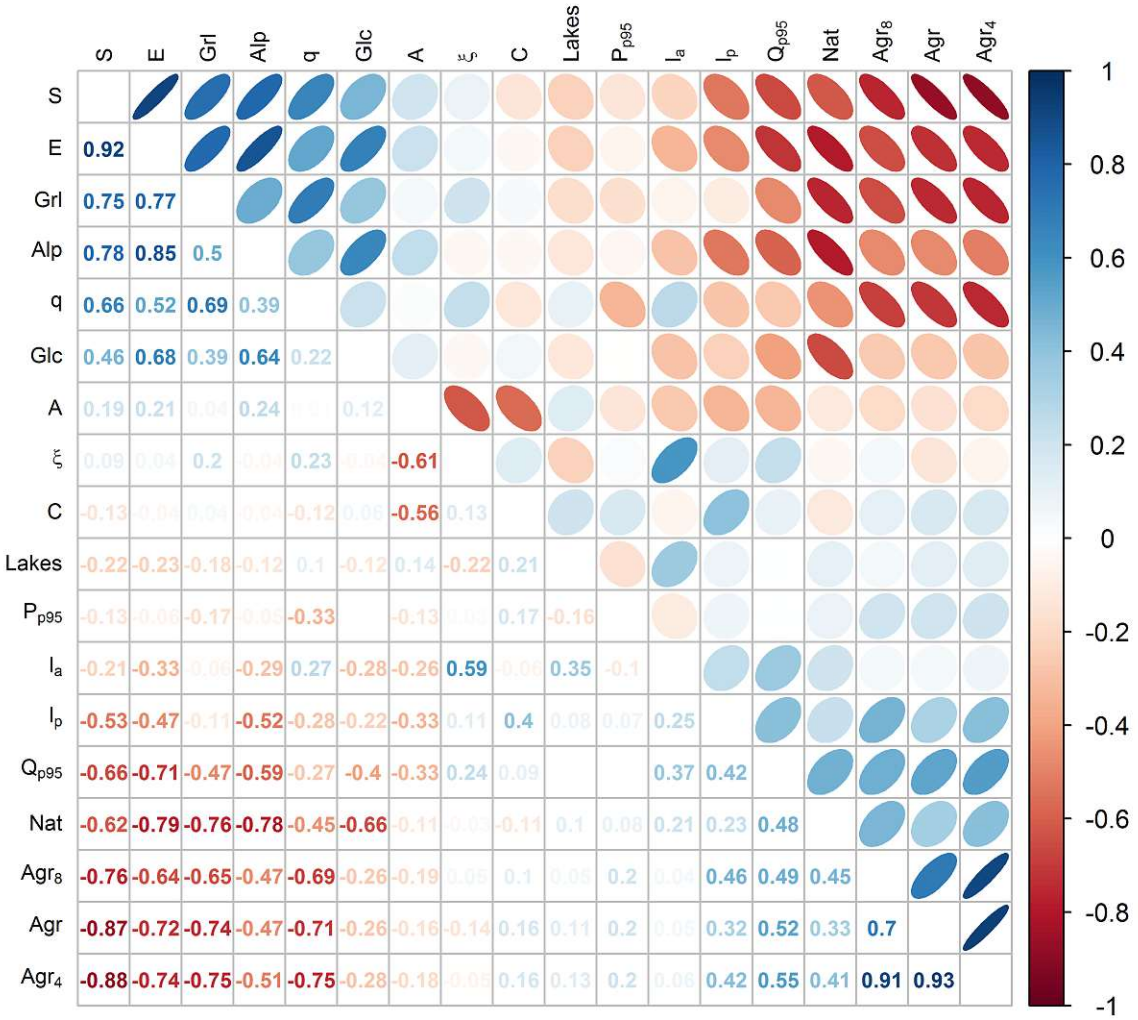

Fig. A.2: Correlation matrix for the explanatory variables tested in the model.

Appendix B

Mapping key tables

Tab. B.1: Connected or not in a natural way. Mosimann et al. (2007) acknowledge obstacles higher than 50 cm as effective barriers for surface runoff, preventing connectivity, and so do we.

Code	Description
0	Not connected in a natural way
1	Connected in a natural way

Tab. B.2: Explanatory details for mapping units not connected in a natural way. Particularly code 4 and 5 in combination with the codes of Tab. B.3 can also be used to indicate a change in connectivity (e. g. a change from direct towards diffuse connectivity).

Code	Description
1	Road embankment
2	Levee or other linear structure
3	Depression
4	Previously mapped stream is missing
5	Previously mapped stream is cased
99	Other

Tab. B.3: Explanatory details for mapping units connected in a natural way.

Code	Description
1	Directly connected to a previously mapped stream
2	Direct connectivity via a previously unknown channel
3	Direct connectivity via a gully
4	Diffuse connectivity
99	Other

Tab. B.4: Artificial connection.

Code	Description
0	Not present
1	Present
2	Diffusely present in a downstream mapping unit

Tab. B.5: Explanatory details for mapping units connected in an artificial way. Diffuse connectivity in this context means that an artificial connection is either diffusely present in a downstream mapping unit (see code 2 of Tab. B.4) or that an artificial structure is not directly connected to a surface water (e. g. a ditch ending in a forest).

Code	Description
1	Direct connectivity via an inter-field ditch
2	Direct connectivity via a roadside ditch
3	Direct connectivity via a storm drain
4	Direct connectivity via a culvert
5	Direct connectivity via a (roadside) ditch with (a) storm drain(s)
6	Direct connectivity via a (roadside) ditch with (a) culvert(s)
7	Diffuse connectivity via an inter-field ditch
8	Diffuse connectivity via a roadside ditch
9	Diffuse connectivity via a storm drain
10	Diffuse connectivity via a culvert
11	Diffuse connectivity via a (roadside) ditch with (a) storm drain(s)
12	Diffuse connectivity via a (roadside) ditch with (a) culvert(s)
99	Other

Tab. B.6: Presence and absence of natural features as well as artificial point-shaped and linear structures.

Code	Description
1	Channel
2	Gully
3	Inter-field ditch
4	Roadside ditch
5	Storm drain
6	Sewer
7	Culvert
8	Stream casing
9	Missing, previously mapped stream
99	Other

Bibliography

Agrarmarkt Austria (2018). *INVEKOS Schläge Österreich 2018 - Datensatz - data.gv.at*. https://www.data.gv.at/katalog/dataset/f7691988-e57c-4ee9-bbd0-e361d3811641. (Visited on 01/14/2023).

Alder, S., V. Prasuhn, H. Liniger, K. Herweg, H. Hurni, A. Candinas, and H. U. Gujer (2015). "A High-Resolution Map of Direct and Indirect Connectivity of Erosion Risk Areas to Surface Waters in Switzerland—a Risk Assessment Tool for Planning and Policy-Making". In: *Land Use Policy* 48.Supplement C, pp. 236–249. DOI: 10.1016/j.landusepol.2015.06.001. (Visited on 12/04/2017).

Ali, G. A. and A. G. Roy (2009). "Revisiting Hydrologic Sampling Strategies for an Accurate Assessment of Hydrologic Connectivity in Humid Temperate Systems". In: *Geogr. Compass* 3.1, pp. 350–374. DOI: 10.1111/j.1749-8198.2008.00180.x. (Visited on 02/25/2017).

Allen, L., A. O'Connell, and V. Kiermer (2019). "How Can We Ensure Visibility and Diversity in Research Contributions? How the Contributor Role Taxonomy (CRediT) Is Helping the Shift from Authorship to Contributorship". In: *Learn. Publ.* 32.1, pp. 71–74. DOI: 10.1002/leap.1210. (Visited on 01/08/2023).

Amann, A., M. Clara, O. Gabriel, G. Hochedlinger, F. Humer, M. Humer, S. Kittllaus, S. Kulcsar, C. Scheffknecht, H. Trautvetter, M. Zessner, and O. Zoboli (2019). *STOBIMO Spurenstoffe. Stoffbilanzmodellierung für Spurenstoffe auf Einzugsgebietsebene*. Tech. rep. Wien: Bundesministerium für Nachhaltigkeit und Tourismus.

Arnold, J. G., R. Srinivasan, R. S. Muttiah, and J. R. Williams (1998). "Large Area Hydrologic Modeling and Assessment Part I: Model Development". In: *JAWRA J. Am. Water Resour. Assoc.* 34.1, pp. 73–89. DOI: 10.1111/j.1752-1688.1998.tb05961.x. (Visited on 12/25/2022).

Auerswald, K. (1989). "Predicting Nutrient Enrichment from Long-Term Average Soil Loss". In: *Soil Technology* 2.3, pp. 271–277. DOI: 10.1016/0933-3630(89)90011-1. (Visited on 02/04/2023).

Aurousseau, P., C. Gascuel-Odoux, H. Squividant, R. Trepos, F. Tortrat, and M. O. Cordier (2009). "A Plot Drainage Network as a Conceptual Tool for the Spatial Representation of Surface Flow Pathways in Agricultural Catchments". In: *Computers & Geosciences* 35.2, pp. 276–288. DOI: 10.1016/j.cageo.2008.09.003. (Visited on 02/25/2017).

Behrendt, H., P. Huber, D. Opitz, O. Schmoll, G. Scholz, and R. Uebe (1999). *Nährstoffbilanzierung der Flußgebiete Deutschlands*. Tech. rep. 75/99. Berlin: Umweltbundesamt.

Bilotta, G. S. and R. E. Brazier (2008). "Understanding the Influence of Suspended Solids on Water Quality and Aquatic Biota". In: *Water Research* 42.12, pp. 2849–2861. DOI: 10.1016/j.watres.2008.03.018. (Visited on 12/25/2022).

Blackwell, M. S. A., D. V. Hogan, and E. Maltby (1999). "The Use of Conventionally and Alternatively Located Buffer Zones for the Removal of Nitrate from Diffuse Agricultural Run-Off". In: *Water Science and Technology* 39.12, pp. 157–164. DOI: 10.1016/S0273-1223(99)00331-5. (Visited on 02/25/2017).

BMLFUW (2015). *Hydrografisches Jahrbuch von Österreich 2013. Daten und Auswertungen*. Tech. rep. 121. Wien: Bundesministerium für Land- und Forstwirtschaft, Umwelt und Wasserwirtschaft. (Visited on 02/24/2017).

BMLRT (2020). *Hydrografisches Jahrbuch von Österreich 2017*. Tech. rep. 125. Wien: Bundesministerium für Landwirtschaft, Regionen und Tourismus.

Bracken, L. J., J. Wainwright, G. A. Ali, D. Tetzlaff, M. W. Smith, S. M. Reaney, and A. G. Roy (2013). "Concepts of Hydrological Connectivity: Research Approaches, Pathways and Future Agendas". In: *Earth-Science Reviews* 119, pp. 17–34. DOI: 10.1016/j.earscirev.2013.02.001. (Visited on 02/25/2017).

Bracken, L. J. and J. Croke (2007). "The Concept of Hydrological Connectivity and Its Contribution to Understanding Runoff-Dominated Geomorphic Systems". In: *Hydrol. Process.* 21.13, pp. 1749–1763. DOI: 10.1002/hyp.6313. (Visited on 02/25/2017).

Buczko, U. and R. O. Kuchenbuch (2007). "Phosphorus Indices as Risk-Assessment Tools in the U.S.A. and Europe—a Review". In: *J. Plant Nutr. Soil Sci.* 170.4, pp. 445–460. DOI: 10.1002/jpln.200725134. (Visited on 01/17/2018).

Bug, J. F. (2011). "Modellierung der linearen Erosion und des Risikos von Partikeleinträgen in Gewässer". PhD thesis. Hannover: Leibniz Universität Hannover. (Visited on 12/14/2016).

Bürkner, P.-C. (2017). "Brms: An R Package for Bayesian Multilevel Models Using Stan". In: *J. Stat. Softw.* 80, pp. 1–28. DOI: 10.18637/jss.v080.i01. (Visited on 12/25/2022).

— (2018). "Advanced Bayesian Multilevel Modeling with the R Package Brms". In: *R J.* 10.1, pp. 395–411. DOI: 10.32614/RJ-2018-017. (Visited on 12/27/2022).

Camhy, D., T. Ertl, R. Fuiko, V. Gamerith, G. Gruber, T. Hofer, M. Höller, C. Kinzel, M. Kleidorfer, L. Kornfeind, G. Leonhardt, D. Muschalla, A. Pressl, W. Rauch, D. Steffelbauer, B. Steger, K. Svardal, C. Urich, A. Winkelbauer, and S. Winkler (2013). *Integrierte Betrachtung eines Gewässerabschnitts auf Basis kontinuierlicher und validierter Langzeitmessreihen*. Tech. rep. Wien: Bundesministerium für Land- und Forstwirtschaft, Umwelt und Wasserwirtschaft.

Cerdan, O., V. Souchère, V. Lecomte, A. Couturier, and Y. Le Bissonnais (2002). "Incorporating Soil Surface Crusting Processes in an Expert-Based Runoff Model: Sealing and Transfer by Runoff and Erosion Related to Agricultural Management". In: *CATENA* 46.2, pp. 189–205. DOI: 10.1016/S0341-8162(01)00166-7. (Visited on 12/29/2020).

Cerdan, O., Y. L. Bissonnais, V. Souchère, P. Martin, and V. Lecomte (2002). "Sediment Concentration in Interrill Flow: Interactions between Soil Surface Conditions, Vegetation

and Rainfall". In: *Earth Surf. Process. Landf.* 27.2, pp. 193–205. DOI: 10.1002/esp.314. (Visited on 12/29/2020).

Chagneau, P., F. Mortier, N. Picard, and J.-N. Bacro (2011). "A Hierarchical Bayesian Model for Spatial Prediction of Multivariate Non-Gaussian Random Fields". In: *Biometrics* 67.1, pp. 97–105. DOI: 10.1111/j.1541-0420.2010.01415.x. (Visited on 02/04/2023).

Champati ray, P. K., S. Dimri, R. C. Lakhera, and S. Sati (2007). "Fuzzy-Based Method for Landslide Hazard Assessment in Active Seismic Zone of Himalaya". In: *Landslides* 4.2, p. 101. DOI: 10.1007/s10346-006-0068-6. (Visited on 01/23/2018).

Charnpratheep, K., Q. Zhou, and B. Garner (1997). "Preliminary Landfill Site Screening Using Fuzzy Geographical Information Systems". In: *Waste Management & Research* 15.2, pp. 197–215. DOI: 10.1006/wmre.1996.0076. (Visited on 02/02/2018).

Cherry, K. A., M. Shepherd, P. J. A. Withers, and S. J. Mooney (2008). "Assessing the Effectiveness of Actions to Mitigate Nutrient Loss from Agriculture: A Review of Methods". In: *Science of The Total Environment* 406.1, pp. 1–23. DOI: 10.1016/j.scitotenv.2008.07.015. (Visited on 01/17/2018).

Choi, Y. (2012). "A New Algorithm to Calculate Weighted Flow-Accumulation from a DEM by Considering Surface and Underground Stormwater Infrastructure". In: *Environmental Modelling & Software* 30, pp. 81–91. DOI: 10.1016/j.envsoft.2011.10.013. (Visited on 02/25/2017).

Choi, Y., H. Yi, and H.-D. Park (2011). "A New Algorithm for Grid-Based Hydrologic Analysis by Incorporating Stormwater Infrastructure". In: *Computers & Geosciences* 37.8, pp. 1035–1044. DOI: 10.1016/j.cageo.2010.07.008. (Visited on 04/21/2017).

Clark, W. A. V. and K. L. Avery (1976). "The Effects of Data Aggregation in Statistical Analysis". In: *Geogr. Anal.* 8.4, pp. 428–438. DOI: 10.1111/j.1538-4632.1976.tb00549.x. (Visited on 01/01/2023).

Conway, J. R., A. Lex, and N. Gehlenborg (2017). "UpSetR: An R Package for the Visualization of Intersecting Sets and Their Properties". In: *Bioinformatics* 33.18, pp. 2938–2940. DOI: 10.1093/bioinformatics/btx364. (Visited on 01/08/2021).

Couturier, A., J. Daroussin, F. Darboux, V. Souchère, Y. Le Bissonnais, O. Cerdan, and D. King (2013). "Improvement of Surface Flow Network Prediction for the Modeling of Erosion Processes in Agricultural Landscapes". In: *Geomorphology.* Studying Water-Erosion Processes with Geoinformatics 183.Supplement C, pp. 120–129. DOI: 10.1016/j.geomorph.2012.07.025. (Visited on 09/26/2017).

Croke, J., S. Mockler, P. Fogarty, and I. Takken (2005). "Sediment Concentration Changes in Runoff Pathways from a Forest Road Network and the Resultant Spatial Pattern of Catchment Connectivity". In: *Geomorphology* 68.3–4, pp. 257–268. DOI: 10.1016/j.geomorph.2004.11.020. (Visited on 02/25/2017).

de Roo, A. P. J., C. G. Wesseling, and C. J. Ritsema (1996). "LISEM: A Single-Event Physically Based Hydrological and Soil Erosion Model for Drainage Basins. I: Theory, Input and Output".

In: *Hydrol. Process.* 10.8, pp. 1107–1117. DOI: 10.1002/(SICI)1099-1085(199608)10:8<1107::AID-HYP415>3.0.CO;2-4. (Visited on 12/25/2022).

de Vente, J. and J. Poesen (2005). "Predicting Soil Erosion and Sediment Yield at the Basin Scale: Scale Issues and Semi-Quantitative Models". In: *Earth-Science Reviews* 71.1–2, pp. 95–125. DOI: 10.1016/j.earscirev.2005.02.002. (Visited on 02/25/2017).

de Vente, J., J. Poesen, M. Arabkhedri, and G. Verstraeten (2007). "The Sediment Delivery Problem Revisited". In: *Prog. Phys. Geogr. Earth Environ.* 31.2, pp. 155–178. DOI: 10.1177/0309133307076485. (Visited on 12/25/2022).

de Vente, J., J. Poesen, G. Verstraeten, G. Govers, M. Vanmaercke, A. Van Rompaey, M. Arabkhedri, and C. Boix-Fayos (2013). "Predicting Soil Erosion and Sediment Yield at Regional Scales: Where Do We Stand?" In: *Earth-Science Reviews* 127, pp. 16–29. DOI: 10.1016/j.earscirev.2013.08.014. (Visited on 02/25/2017).

de Vente, J., J. Poesen, G. Verstraeten, A. Van Rompaey, and G. Govers (2008). "Spatially Distributed Modelling of Soil Erosion and Sediment Yield at Regional Scales in Spain". In: *Global and Planetary Change* 60.3, pp. 393–415. DOI: 10.1016/j.gloplacha.2007.05.002. (Visited on 12/25/2022).

de Vente, J., R. Verduyn, G. Verstraeten, M. Vanmaercke, and J. Poesen (2011). "Factors Controlling Sediment Yield at the Catchment Scale in NW Mediterranean Geoecosystems". In: *J Soils Sediments* 11.4, pp. 690–707. DOI: 10.1007/s11368-011-0346-3. (Visited on 12/25/2022).

Desmet, P. J. J. and G. Govers (1996). "A GIS Procedure for Automatically Calculating the USLE LS Factor on Topographically Complex Landscape Units". In: *Journal of Soil and Water Conservation* 51.5, pp. 427–433. (Visited on 02/28/2017).

Djodjic, F. and A. Villa (2015). "Distributed, High-Resolution Modelling of Critical Source Areas for Erosion and Phosphorus Losses". In: *AMBIO* 44.2, pp. 241–251. DOI: 10.1007/s13280-014-0618-4. (Visited on 08/01/2017).

Doody, D. G., M. Archbold, R. H. Foy, and R. Flynn (2012). "Approaches to the Implementation of the Water Framework Directive: Targeting Mitigation Measures at Critical Source Areas of Diffuse Phosphorus in Irish Catchments". In: *Journal of Environmental Management* 93.1, pp. 225–234. DOI: 10.1016/j.jenvman.2011.09.002. (Visited on 02/25/2017).

Doppler, T., L. Camenzuli, G. Hirzel, M. Krauss, A. Lück, and C. Stamm (2012). "Spatial Variability of Herbicide Mobilisation and Transport at Catchment Scale: Insights from a Field Experiment". In: *Hydrol. Earth Syst. Sci.* 16.7, pp. 1947–1967. DOI: 10.5194/hess-16-1947-2012.

Dubois, D. and H. Prade (1985). "A Review of Fuzzy Set Aggregation Connectives". In: *Information Sciences* 36.1, pp. 85–121. DOI: 10.1016/0020-0255(85)90027-1. (Visited on 02/02/2018).

Duke, G. D., S. W. Kienzle, D. L. Johnson, and J. M. Byrne (2003). "Improving Overland Flow Routing by Incorporating Ancillary Road Data into Digital Elevation Models". In: *J. Spat. Hydrol.* 3.2. (Visited on 02/25/2017).

— (2006). "Incorporating Ancillary Data to Refine Anthropogenically Modified Overland Flow Paths". In: *Hydrol. Process.* 20.8, pp. 1827–1843. DOI: 10.1002/hyp.5964. (Visited on 03/29/2017).

Engman, E. T. (1986). "Roughness Coefficients for Routing Surface Runoff". In: *J. Irrig. Drain. Eng.* 112.1, pp. 39–53.

European Comission (2000). *Directive 2000/60/EC of the European Parliament and of the Council of 23 October 2000 Establishing a Framework for Community Action in the Field of Water Policy*. (Visited on 02/14/2018).

European Commission and European Soil Bureau Network (2004). *The European Soil Database Distribution Version 2.0*. Brussels.

Evenson, G. R., M. Kalcic, Y.-C. Wang, D. Robertson, D. Scavia, J. Martin, N. Aloysius, A. Apostel, C. Boles, M. Brooker, R. Confesor, A. T. Dagnew, T. Guo, J. Kast, H. Kujawa, R. L. Muenich, A. Murumkar, and T. Redder (2021). "Uncertainty in Critical Source Area Predictions from Watershed-Scale Hydrologic Models". In: *Journal of Environmental Management* 279, p. 111506. DOI: 10.1016/j.jenvman.2020.111506. (Visited on 12/08/2021).

Federal Office of Metrology and Surveying (2015). *Digitales Landschaftsmodell. Fließende Gewässer*. Wien.

Fiener, P., K. Auerswald, and K. Van Oost (2011). "Spatio-Temporal Patterns in Land Use and Management Affecting Surface Runoff Response of Agricultural Catchments—a Review". In: *Earth-Science Reviews* 106.1–2, pp. 92–104. DOI: 10.1016/j.earscirev.2011.01.004. (Visited on 02/25/2017).

Forman, R. T. T. and L. E. Alexander (1998). "Roads and Their Major Ecological Effects". In: *Annu. Rev. Ecol. Syst.* 29.1, pp. 207–231. DOI: 10.1146/annurev.ecolsys.29.1.207. (Visited on 02/14/2017).

Freedman, D. A. (1999). *Ecological Inference and the Ecological Fallacy*. Tech. rep. 549. Berkeley: University of California. (Visited on 01/01/2023).

Fuchs, S., M. Kaiser, L. Kiemle, S. Kittlaus, S. Rothvoß, S. Toshovski, A. Wagner, R. Wander, T. Weber, and S. Ziegler (2017). "Modeling of Regionalized Emissions (MoRE) into Water Bodies: An Open-Source River Basin Management System". In: *Water* 9.4, p. 239. DOI: 10.3390/w9040239. (Visited on 01/15/2023).

Fuiko, R., L. Kornfeind, K. Schilling, S. Weilguni, and A. Winkelbauer (2016). *Nachhaltige Wassergütewirtschaft Raab. Online-Monitoring*. Tech. rep. Wien: Bundesministerium für Land- und Forstwirtschaft, Umwelt und Wasserwirtschaft.

Gabriel, O., A. Kovacs, S. Thaler, M. Zessner, G. Hochedlinger, C. Schilling, and G. Windhofer (2011). *Stoffbilanzmodellierung für Nährstoffe auf Einzugsgebietsebene (STOBIMO-Nährstoffe) als Grundlage für Bewirtschaftungspläne und Maßnahmenprogramme*. Tech. rep. Wien: Bundesministerium für Land- und Forstwirtschaft, Umwelt und Wasserwirtschaft.

Gascuel-Odoux, C., P. Aurousseau, T. Doray, H. Squividant, F. Macary, D. Uny, and C. Grimaldi (2011). "Incorporating Landscape Features to Obtain an Object-Oriented Landscape Drainage

Network Representing the Connectivity of Surface Flow Pathways over Rural Catchments". In: *Hydrol. Process.* 25.23, pp. 3625–3636. DOI: 10.1002/hyp.8089. (Visited on 03/30/2017).

Gavrilovic, S. (1976). "Bujicni tokovi i erozija (Torrents and erosion)". In: *Gradevinski kalendar*.

Gburek, W. J., A. N. Sharpley, L. Heathwaite, and G. J. Folmar (2000). "Phosphorus Management at the Watershed Scale: A Modification of the Phosphorus Index". In: *J. Environ. Qual.* 29.1, pp. 130–144. DOI: 10.2134/jeq2000.00472425002900010017x. (Visited on 01/17/2018).

Gehlke, C. E. and K. Biehl (1934). "Certain Effects of Grouping upon the Size of the Correlation Coefficient in Census Tract Material". In: *J. Am. Stat. Assoc.* 29.185A, pp. 169–170. DOI: 10.1080/01621459.1934.10506247. (Visited on 03/05/2023).

Gelman, A. (2006). "Prior Distributions for Variance Parameters in Hierarchical Models (Comment on Article by Browne and Draper)". In: *Bayesian Anal.* 1.3, pp. 515–534. DOI: 10.1214/06-BA117A. (Visited on 05/29/2017).

Gelman, A. and J. Hill (2006). *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Analytical Methods for Social Research. Cambridge: Cambridge University Press. (Visited on 12/27/2022).

Gelman, A., A. Jakulin, M. G. Pittau, and Y.-S. Su (2008). "A Weakly Informative Default Prior Distribution for Logistic and Other Regression Models". In: *Ann. Appl. Stat.* 2.4, pp. 1360–1383. DOI: 10.1214/08-AOAS191. (Visited on 05/29/2017).

Geman, S., E. Bienenstock, and R. Doursat (1992). "Neural Networks and the Bias/Variance Dilemma". In: *Neural Computation* 4.1, pp. 1–58. DOI: 10.1162/neco.1992.4.1.1. (Visited on 03/11/2023).

geoland.at (2016a). *Digitales Geländemodell (DGM) Österreich*. Wien.

— (2016b). *Intermodales Verkehrsreferenzsystem Österreich (GIP.at)*. Tech. rep. Wien: Geodatenverbund der Länder.

Gericke, A. and M. Venohr (2012). "Improving the Estimation of Erosion-Related Suspended Solid Yields in Mountainous, Non-Alpine River Catchments". In: *Environmental Modelling & Software* 37, pp. 30–40. DOI: 10.1016/j.envsoft.2012.04.008. (Visited on 12/25/2022).

Ghebremichael, L. T., T. L. Veith, and J. M. Hamlett (2013). "Integrated Watershed- and Farm-Scale Modeling Framework for Targeting Critical Source Areas While Maintaining Farm Economic Viability". In: *Journal of Environmental Management* 114, pp. 381–394. DOI: 10.1016/j.jenvman.2012.10.034. (Visited on 12/08/2021).

Gramlich, A., S. Stoll, C. Stamm, T. Walter, and V. Prasuhn (2018). "Effects of Artificial Land Drainage on Hydrology, Nutrient and Pesticide Fluxes from Agricultural Fields – a Review". In: *Agriculture, Ecosystems & Environment* 266, pp. 84–99. DOI: 10.1016/j.agee.2018.04.005. (Visited on 10/09/2019).

Grauso, S., A. Pagano, G. Fattoruso, P. De Bonis, F. Onori, P. Regina, and C. Tebano (2008). "Relations between Climatic–Geomorphological Parameters and Sediment Yield in a Mediterranean Semi-Arid Area (Sicily, Southern Italy)". In: *Environ Geol* 54.2, pp. 219–234. DOI: 10.1007/s00254-007-0809-4. (Visited on 12/25/2022).

Greig, S. M., D. A. Sear, and P. A. Carling (2005). "The Impact of Fine Sediment Accumulation on the Survival of Incubating Salmon Progeny: Implications for Sediment Management". In: *Science of The Total Environment*. Linking Landscape Sources of Phosphorus and Sediment to Ecological Impacts in Surface Waters 344.1, pp. 241–258. DOI: 10.1016/j.scitotenv.2005.02.010. (Visited on 12/25/2022).

Hanel, M., P. Máca, P. Bašta, R. Vlnas, and P. Pech (2016). "The Rainfall Erosivity Factor in the Czech Republic and Its Uncertainty". In: *Hydrol. Earth Syst. Sci.* 20.10, pp. 4307–4322. DOI: 10.5194/hess-20-4307-2016. (Visited on 12/25/2022).

Haregeweyn, N., J. Poesen, J. Nyssen, G. Govers, G. Verstraeten, J. de Vente, J. Deckers, J. Moeyersons, and M. Haile (2008). "Sediment Yield Variability in Northern Ethiopia: A Quantitative Analysis of Its Controlling Factors". In: *CATENA*. Environmental Change, Geomorphic Processes and Land Degradation in Tropical Highlands 75.1, pp. 65–76. DOI: 10.1016/j.catena.2008.04.011. (Visited on 12/25/2022).

He, K., X. Zhang, S. Ren, and J. Sun (2015). "Deep Residual Learning for Image Recognition". In: *ArXiv151203385 Cs*. arXiv: 1512.03385 [cs]. (Visited on 10/08/2019).

Heathwaite, A. L., R. M. Dils, S. Liu, L. Carvalho, R. E. Brazier, L. Pope, M. Hughes, G. Phillips, and L. May (2005). "A Tiered Risk-Based Approach for Predicting Diffuse and Point Source Phosphorus Losses in Agricultural Areas". In: *Science of The Total Environment*. Linking Landscape Sources of Phosphorus and Sediment to Ecological Impacts in Surface Waters 344.1, pp. 225–239. DOI: 10.1016/j.scitotenv.2005.02.034. (Visited on 01/19/2018).

Heathwaite, A. L., P. F. Quinn, and C. J. M. Hewett (2005). "Modelling and Managing Critical Source Areas of Diffuse Pollution from Agricultural Land Using Flow Connectivity Simulation". In: *Journal of Hydrology*. Nutrient Mobility within River Basins: A European Perspective 304.1–4, pp. 446–461. DOI: 10.1016/j.jhydrol.2004.07.043. (Visited on 02/25/2017).

Heathwaite, L., A. Sharpley, and M. Bechmann (2003). "The Conceptual Basis for a Decision Support Framework to Assess the Risk of Phosphorus Loss at the Field Scale across Europe". In: *J. Plant Nutr. Soil Sci.* 166.4, pp. 447–458. DOI: 10.1002/jpln.200321154. (Visited on 01/17/2018).

Heek, J. and N. Kalchbrenner (2019). "Bayesian Inference for Large Scale Image Classification". In: *ArXiv190803491 Cs Stat*. arXiv: 1908.03491 [cs, stat]. (Visited on 10/08/2019).

Hepp, G. and M. Zessner (2019). "Assessing the Impact of Storm Drains at Road Embankments on Diffuse Particulate Phosphorus Emissions in Agricultural Catchments". In: *Water* 11.10, p. 2161. DOI: 10.3390/w11102161. (Visited on 12/07/2020).

Hepp, G., O. Zoboli, E. Strenge, and M. Zessner (2022). "Particulate PhozzyLogic Index for Policy Makers—an Index for a More Accurate and Transparent Identification of Critical Source Areas". In: *Journal of Environmental Management* 307, p. 114514. DOI: 10.1016/j.jenvman.2022.114514. (Visited on 12/27/2022).

Hiebl, J. and C. Frei (2018). "Daily Precipitation Grids for Austria since 1961—Development and Evaluation of a Spatial Dataset for Hydroclimatic Monitoring and Modelling". In: *Theor Appl Climatol* 132.1, pp. 327–345. DOI: 10.1007/s00704-017-2093-x. (Visited on 12/27/2022).

Hofer, O., W. Fahrner, G. Pavlis-Fronaschitz, S. Linder, and P. Gmeiner (2014). *INVEKOS-Datenpool 2014 des BMLFUW. Übersicht über alle im Ordner „Invekosdaten" enthaltenen Datenbanken mit ausführlicher Tabellenbeschreibung sowie Informationen zu sonstigen verfügbaren Datenbanken*. Tech. rep. Wien: Bundesministerium für Land- und Forstwirtschaft, Umwelt und Wasserwirtschaft.

Holland, D. A., C. Pook, D. Capstick, and A. Hemmings (2016). "The Topographic Data Deluge – Collecting and Maintaining Data in a 21st Century Mapping Agency". In: *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* Vol. XLI-B4. Prague, pp. 727–731. DOI: 10.5194/isprs-archives-XLI-B4-727-2016. (Visited on 06/08/2017).

Hösl, R., P. Strauss, and T. Glade (2012). "Man-Made Linear Flow Paths at Catchment Scale: Identification, Factors and Consequences for the Efficiency of Vegetated Filter Strips". In: *Landscape and Urban Planning* 104.2, pp. 245–252. DOI: 10.1016/j.landurbplan.2011.10.017. (Visited on 02/25/2017).

Huddart Kennedy, E., H. Krahn, and N. T. Krogman (2015). "Are We Counting What Counts? A Closer Look at Environmental Concern, pro-Environmental Behaviour, and Carbon Footprint". In: *Local Environ.* 20.2, pp. 220–236. DOI: 10.1080/13549839.2013.837039. (Visited on 01/01/2023).

Jetten, V., A. de Roo, and D. Favis-Mortlock (1999). "Evaluation of Field-Scale and Catchment-Scale Soil Erosion Models". In: *CATENA* 37.3–4, pp. 521–541. DOI: 10.1016/S0341-8162(99)00037-5. (Visited on 02/23/2017).

Jetten, V., G. Govers, and R. Hessel (2003). "Erosion Models: Quality of Spatial Predictions". In: *Hydrol. Process.* 17.5, pp. 887–900. DOI: 10.1002/hyp.1168. (Visited on 02/23/2017).

Jolliffe, I. T. (2002). *Principal Component Analysis*. Second. Springer Series in Statistics. New York: Springer. (Visited on 12/25/2022).

Julich, D., S. Julich, and K.-H. Feger (2017). "Phosphorus in Preferential Flow Pathways of Forest Soils in Germany". In: *Forests* 8.1, p. 19. DOI: 10.3390/f8010019. (Visited on 12/08/2021).

Kandel, A. (1986). *Fuzzy Mathematical Techniques with Applications*. Reading: Addison-Wesley.

Kapfer, S. (2014). *Jahresbericht 2013 – Korrektur. Zustand der Oö. Fließgewässer gem. WRRL*. Tech. rep. Linz: Amt der Oberösterreichischen Landesregierung.

Kaufman, L. and P. J. Rousseeuw (1987). "Clustering by Means of Medoids". In: *Stat. Data Anal. Based L1 Norm Relat. Methods*. Ed. by Y. Dodge. Amsterdam, New York: North-Holland, pp. 405–416.

Kirkby, M. J., B. J. Irvine, R. J. A. Jones, G. Govers, and Pesera Team (2008). "The PESERA Coarse Scale Erosion Model for Europe. I. – Model Rationale and Implementation". In: *Eur. J. Soil Sci.* 59.6, pp. 1293–1306. DOI: 10.1111/j.1365-2389.2008.01072.x. (Visited on 12/25/2022).

Kondolf, G. M., Y. Gao, G. W. Annandale, G. L. Morris, E. Jiang, J. Zhang, Y. Cao, P. Carling, K. Fu, Q. Guo, R. Hotchkiss, C. Peteuil, T. Sumi, H.-W. Wang, Z. Wang, Z. Wei, B. Wu, C. Wu, and C. T. Yang (2014). "Sustainable Sediment Management in Reservoirs and Regulated Rivers: Experiences from Five Continents". In: *Earths Future* 2.5, pp. 256–280. DOI: 10.1002/2013EF000184. (Visited on 12/25/2022).

Kovacs, A., M. Honti, and A. Clement (2008). "Design of Best Management Practice Applications for Diffuse Phosphorus Pollution Using Interactive GIS". In: *Water Sci. Technol.* 57.11, pp. 1727–1733. DOI: 10.2166/wst.2008.264. (Visited on 02/25/2017).

Kovacs, A. (2013). "Quantification of Diffuse Phosphorous Inputs into Surface Water Systems". PhD thesis. Vienna: Technische Universität Wien.

Kovacs, A., M. Honti, M. Zessner, A. Eder, A. Clement, and G. Blöschl (2012). "Identification of Phosphorus Emission Hotspots in Agricultural Catchments". In: *Science of The Total Environment* 433, pp. 74–88. DOI: 10.1016/j.scitotenv.2012.06.024. (Visited on 02/25/2017).

Krause, P., D. P. Boyle, and F. Bäse (2005). "Comparison of Different Efficiency Criteria for Hydrological Model Assessment". In: *Adv. Geosci.* 5, pp. 89–97. DOI: 10.5194/adgeo-5-89-2005. (Visited on 12/25/2022).

Krivoruchko, K., A. Gribov, and E. Krause (2011). "Multivariate Areal Interpolation for Continuous and Count Data". In: *Procedia Environmental Sciences*. 1st Conference on Spatial Statistics 2011 – Mapping Global Change 3, pp. 14–19. DOI: 10.1016/j.proenv.2011.02.004. (Visited on 06/06/2017).

Kroiss, H., M. Zessner, and C. Lampert (2006). "daNUbs: Lessons Learned for Nutrient Management in the Danube Basin and Its Relation to Black Sea Euthrophication". In: *Chem. Ecol.* 22.5, pp. 347–357. DOI: 10.1080/02757540600917518. (Visited on 01/08/2023).

Kroiss, H., C. Lampert, M. Zessner, C. Schilling, and O. Gabriel (2005). *Nutrient Management in the Danube Basin and Its Impact on the Black Sea*. Tech. rep. Vienna: European Commission.

La Marche, J. L. and D. P. Lettenmaier (2001). "Effects of Forest Roads on Flood Flows in the Deschutes River, Washington". In: *Earth Surf. Process. Landforms* 26.2, pp. 115–134. DOI: 10.1002/1096-9837(200102)26:2<115::AID-ESP166>3.0.CO;2-0. (Visited on 02/14/2017).

Lane, L. J., M. Hernandez, and M. Nichols (1997). "Processes Controlling Sediment Yield from Watersheds as Functions of Spatial Scale". In: *Environmental Modelling & Software* 12.4, pp. 355–369. DOI: 10.1016/S1364-8152(97)00027-3. (Visited on 12/27/2022).

Lane, S. N., S. M. Reaney, and A. L. Heathwaite (2009). "Representation of Landscape Hydrological Connectivity Using a Topographically Driven Surface Flow Index". In: *Water Resour. Res.* 45.8, W08423. DOI: 10.1029/2008WR007336. (Visited on 02/23/2017).

Lane, S. N., V. Tayefi, S. C. Reid, D. Yu, and R. J. Hardy (2007). "Interactions between Sediment Delivery, Channel Change, Climate Change and Flood Risk in a Temperate Upland Environment". In: *Earth Surf. Process. Landf.* 32.3, pp. 429–446. DOI: 10.1002/esp.1404. (Visited on 12/25/2022).

Laurent, A. G. (1963). "The Lognormal Distribution and the Translation Method: Description and Estimation Problems". In: *J. Am. Stat. Assoc.* 58.301, pp. 231–235. DOI: 10.1080/01621459.1963.10500844.

Le Bissonnais, Y., O. Cerdan, V. Lecomte, H. Benkhadra, V. Souchère, and P. Martin (2005). "Variability of Soil Surface Characteristics Influencing Runoff and Interrill Erosion". In: *CATENA*. Surface Characterisation for Soil Erosion Forecasting 62.2, pp. 111–124. DOI: 10.1016/j.catena.2005.05.001. (Visited on 12/29/2020).

Lemunyon, J. L. and R. G. Gilbert (1993). "The Concept and Need for a Phosphorus Assessment Tool". In: *J. Prod. Agric.* 6.4, pp. 483–486. DOI: 10.2134/jpa1993.0483. (Visited on 02/20/2018).

Lermontov, A., L. Yokoyama, M. Lermontov, and M. A. S. Machado (2009). "River Quality Analysis Using Fuzzy Water Quality Index: Ribeira Do Iguape River Watershed, Brazil". In: *Ecological Indicators* 9.6, pp. 1188–1197. DOI: 10.1016/j.ecolind.2009.02.006. (Visited on 01/26/2018).

Lex, A., N. Gehlenborg, H. Strobelt, R. Vuillemot, and H. Pfister (2014). "UpSet: Visualization of Intersecting Sets". In: *IEEE Trans. Vis. Comput. Graph.* 20.12, pp. 1983–1992. DOI: 10.1109/TVCG.2014.2346248.

Liou, S.-M., S.-L. Lo, and C.-Y. Hu (2003). "Application of Two-Stage Fuzzy Set Theory to River Quality Evaluation in Taiwan". In: *Water Research* 37.6, pp. 1406–1416. DOI: 10.1016/S0043-1354(02)00479-7. (Visited on 01/26/2018).

Liu, Y. B. and F. De Smedt (2004). *WetSpa Extension, a GIS-based Hydrologic Model for Flood Prediction and Watershed Management*. Tech. rep. Brussels: Department of Hydrology and Hydraulic Engineering of the Vrije Universiteit Brussel.

Long, E. R., C. G. Ingersoll, and D. D. MacDonald (2006). "Calculation and Uses of Mean Sediment Quality Guideline Quotients: A Critical Review". In: *Environ Sci Technol* 40.6, pp. 1726–1736. DOI: 10.1021/es058012d.

Ludwig, B., J. Boiffin, J. Chadœuf, and A.-V. Auzet (1995). "Hydrological Structure and Erosion Damage Caused by Concentrated Flow in Cultivated Catchments". In: *CATENA*. Experimental Geomorphology and Landscape Ecosystem Changes 25.1, pp. 227–252. DOI: 10.1016/0341-8162(95)00012-H. (Visited on 02/25/2017).

Luxburg, U. von, R. C. Williamson, and I. Guyon (2012). "Clustering: Science or Art?" In: *JMLR Workshop Conf. Proc.* Ed. by I. Guyon, G. Dror, V. Lemaire, G. Taylor, and D. Silver. Vol. 27. ML Research Press, pp. 65–79. (Visited on 02/19/2023).

Mallarino, A. P., B. M. Stewart, J. L. Baker, J. D. Downing, and J. E. Sawyer (2002). "Phosphorus Indexing for Cropland: Overview and Basic Concepts of the Iowa Phosphorus Index". In: *Journal of Soil and Water Conservation* 57.6, pp. 440–447. (Visited on 01/19/2018).

Martin-Ortega, J. and B. B. Balana (2012). "Cost-Effectiveness Analysis in the Implementation of the Water Framework Directive: A Comparative Analysis of the United Kingdom and Spain". In: *Eur. Water* 37, pp. 15–25. (Visited on 02/14/2018).

McElreath, R. (2015). *Statistical Rethinking: A Bayesian Course with Examples in R and Stan*. New York: Chapman and Hall/CRC.

— (2020). *Statistical Rethinking: A Bayesian Course with Examples in R and Stan*. Second. New York: Chapman and Hall/CRC.

Merritt, W. S., R. A. Letcher, and A. J. Jakeman (2003). "A Review of Erosion and Sediment Transport Models". In: *Environmental Modelling & Software*. The Modelling of Hydrologic Systems 18.8, pp. 761–799. DOI: 10.1016/S1364-8152(03)00078-1. (Visited on 12/27/2022).

Molnár, P. and J. A. Ramírez (1998). "Energy Dissipation Theories and Optimal Channel Characteristics of River Networks". In: *Water Resour. Res.* 34.7, pp. 1809–1818. DOI: 10.1029/98WR00983. (Visited on 04/03/2021).

Moran, P. W., L. H. Nowell, N. E. Kemble, B. J. Mahler, I. R. Waite, and P. C. Van Metre (2017). "Influence of Sediment Chemistry and Sediment Toxicity on Macroinvertebrate Communities across 99 Wadable Streams of the Midwestern USA". In: *Science of The Total Environment* 599–600, pp. 1469–1478. DOI: 10.1016/j.scitotenv.2017.05.035. (Visited on 12/25/2022).

Moriasi, D. N., J. G. Arnold, M. W. Van Liew, R. L. Bingner, R. D. Harmel, and T. L. Veith (2007). "Model Evaluation Guidelines for Systematic Quantification of Accuracy in Watershed Simulations". In: *Trans. ASABE* 50.3, pp. 885–900.

Mosimann, T., J. Backhaus, and H. Westphal (2007). *Gewässeranschluss von Ackerflächen. Ein Schlüssel für Betriebsleiter und Berater in Niedersachsen*. Hannover: Landwirtschaftskammer Niedersachsen.

Moussa, R., M. Voltz, and P. Andrieux (2002). "Effects of the Spatial Organization of Agricultural Management on the Hydrological Behaviour of a Farmed Catchment during Flood Events". In: *Hydrol. Process.* 16.2, pp. 393–412. DOI: 10.1002/hyp.333. (Visited on 02/23/2017).

Nash, J. E. and J. V. Sutcliffe (1970). "River Flow Forecasting through Conceptual Models Part I – a Discussion of Principles". In: *Journal of Hydrology* 10.3, pp. 282–290. DOI: 10.1016/0022-1694(70)90255-6. (Visited on 12/25/2022).

O'Callaghan, J. F. and D. M. Mark (1984). "The Extraction of Drainage Networks from Digital Elevation Data". In: *Computer Vision, Graphics, and Image Processing* 28.3, pp. 323–344. DOI: 10.1016/S0734-189X(84)80011-0. (Visited on 07/17/2017).

Onnen, N., G. Heckrath, A. Stevens, P. Olsen, M. B. Greve, J. W. M. Pullens, B. Kronvang, and K. Van Oost (2019). "Distributed Water Erosion Modelling at Fine Spatial Resolution across Denmark". In: *Geomorphology* 342, pp. 150–162. DOI: 10.1016/j.geomorph.2019.06.011. (Visited on 04/06/2021).

Park, H.-S. and C.-H. Jun (2009). "A Simple and Fast Algorithm for K-medoids Clustering". In: *Expert Systems with Applications* 36.2, Part 2, pp. 3336–3341. DOI: 10.1016/j.eswa.2008.01.039. (Visited on 12/25/2022).

Parsons, A. J., J. Wainwright, R. E. Brazier, and D. M. Powell (2006). "Is Sediment Delivery a Fallacy?" In: *Earth Surf. Process. Landf.* 31.10, pp. 1325–1328. DOI: 10.1002/esp.1395. (Visited on 12/25/2022).

Pionke, H. B., W. J. Gburek, and A. N. Sharpley (2000). "Critical Source Area Controls on Water Quality in an Agricultural Watershed Located in the Chesapeake Basin". In: *Ecological Engineering* 14.4, pp. 325–335. DOI: 10.1016/S0925-8574(99)00059-2. (Visited on 02/14/2017).

Plummer, M. (2003). "JAGS: A Program for Analysis of Bayesian Graphical Models Using Gibbs Sampling". In: *Proc. 3rd Int. Workshop Distrib. Stat. Comput.* Vienna.

Poeppl, R. E. and A. J. Parsons (2018). "The Geomorphic Cell: A Basis for Studying Connectivity". In: *Earth Surf. Process. Landf.* 43.5, pp. 1155–1159. DOI: 10.1002/esp.4300.

Poesen, J. (2018). "Soil Erosion in the Anthropocene: Research Needs". In: *Earth Surf. Process. Landf.* 43.1, pp. 64–84. DOI: 10.1002/esp.4250. (Visited on 04/06/2021).

Pourghasemi, H. R., B. Pradhan, and C. Gokceoglu (2012). "Application of Fuzzy Logic and Analytical Hierarchy Process (AHP) to Landslide Susceptibility Mapping at Haraz Watershed, Iran". In: *Nat Hazards* 63.2, pp. 965–996. DOI: 10.1007/s11069-012-0217-2. (Visited on 01/18/2018).

Prasuhn, V. (2011). "Soil Erosion in the Swiss Midlands: Results of a 10-Year Field Survey". In: *Geomorphology* 126.1, pp. 32–41. DOI: 10.1016/j.geomorph.2010.10.023. (Visited on 12/12/2017).

Qian, S. S., T. F. Cuffney, I. Alameddine, G. McMahon, and K. H. Reckhow (2010). "On the Application of Multilevel Modeling in Environmental and Ecological Studies". In: *Ecology* 91.2, pp. 355–361. DOI: 10.1890/09-1043.1. (Visited on 12/27/2022).

R Core Team (2016). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.

— (2019). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.

Raudenbush, S. W. (1988). "Educational Applications of Hierarchical Linear Models: A Review". In: *J. Educ. Stat.* 13.2, pp. 85–116. DOI: 10.3102/10769986013002085. (Visited on 01/03/2023).

Rebolledo, B., A. Gil, X. Flotats, and J. Á. Sánchez (2016). "Assessment of Groundwater Vulnerability to Nitrates from Agricultural Sources Using a GIS-compatible Logic Multicriteria Model". In: *Journal of Environmental Management* 171, pp. 70–80. DOI: 10.1016/j.jenvman.2016.01.041. (Visited on 01/25/2018).

Remund, D., F. Liebisch, H. P. Liniger, A. Heinimann, and V. Prasuhn (2021). "The Origin of Sediment and Particulate Phosphorus Inputs into Water Bodies in the Swiss Midlands – a Twenty-Year Field Study of Soil Erosion". In: *CATENA* 203, p. 105290. DOI: 10.1016/j.catena.2021.105290. (Visited on 04/06/2021).

Renard, K. G., G. R. Foster, G. A. Weesies, D. K. McCool, and D. C. Yoder (1997). *Predicting Soil Erosion by Water: A Guide to Conservation Planning with the Revised Universal Soil Loss Equation (RUSLE)*. Agriculture Handbook 703. Washington, DC: U.S. Government Printing Office.

Robinson, V. B. (2003). "A Perspective on the Fundamentals of Fuzzy Sets and Their Use in Geographic Information Systems". In: *Trans. GIS* 7.1, pp. 3–30. DOI: 10.1111/1467-9671.00127. (Visited on 01/23/2018).

Robinson, W. S. (2009). "Ecological Correlations and the Behavior of Individuals". In: *International Journal of Epidemiology* 38.2, pp. 337–341. DOI: 10.1093/ije/dyn357. (Visited on 01/01/2023).

Schmaltz, E., T. Brunner, E. Strenge, C. Weinberger, M. Kuderna, P. Strauss, and M. Zessner (2022). *Begrünte Fließwege AT. Identifizierung von Maßnahmenflächen für begrünte Fließwege und Pufferstreifen für ÖPUL*. Tech. rep. Wien: Bundesministerium für Landwirtschaft, Regionen und Tourismus.

Schwertmann, U., W. Vogl, and M. Kainz (1987). *Bodenerosion durch Wasser. Vorhersage des Abtrags und Bewertung von Gegenmaßnahmen*. 2nd ed. Stuttgart: Ulmer.

Sharpley, A. N., P. J. A. Kleinman, D. N. Flaten, and A. R. Buda (2011). "Critical Source Area Management of Agricultural Phosphorus: Experiences, Challenges and Opportunities". In: *Water Science and Technology* 64.4, pp. 945–952. DOI: 10.2166/wst.2011.712. (Visited on 12/08/2021).

Sharpley, A. N., P. J. A. Kleinman, P. Jordan, L. Bergström, and A. L. Allen (2009). "Evaluating the Success of Phosphorus Management from Field to Watershed". In: *J. Environ. Qual.* 38.5, pp. 1981–1988. DOI: 10.2134/jeq2008.0056. (Visited on 08/01/2017).

Sims, J. T. and A. N. Sharpley, eds. (2005). *Phosphorus: Agriculture and the Environment*. Agronomy Monographs 46. American Society of Agronomy. (Visited on 12/08/2021).

Skøien, J. O., R. Merz, and G. Blöschl (2006). "Top-Kriging – Geostatistics on Stream Networks". In: *Hydrol. Earth Syst. Sci.* 10.2, pp. 277–287. DOI: 10.5194/hess-10-277-2006. (Visited on 02/28/2017).

Souchere, V., D. King, J. Daroussin, F. Papy, and A. Capillon (1998). "Effects of Tillage on Runoff Directions: Consequences on Runoff Contributing Area within Agricultural Catchments". In: *Journal of Hydrology* 206.3, pp. 256–267. DOI: 10.1016/S0022-1694(98)00103-6. (Visited on 09/25/2017).

Strauss, P., A. Leone, M. N. Ripa, N. Turpin, J.-M. Lescot, and R. Laplana (2007). "Using Critical Source Areas for Targeting Cost-Effective Best Management Practices to Mitigate Phosphorus and Sediment Transfer at the Watershed Scale". In: *Soil Use Manag.* 23, pp. 144–153. DOI: 10.1111/j.1475-2743.2007.00118.x. (Visited on 02/25/2017).

Syvitski, J. P. M. and J. D. Milliman (2007). "Geology, Geography, and Humans Battle for Dominance over the Delivery of Fluvial Sediment to the Coastal Ocean". In: *The Journal of Geology* 115.1, pp. 1–19. DOI: 10.1086/509246. (Visited on 03/11/2020).

Takken, I., G. Govers, V. Jetten, J. Nachtergaele, A. Steegen, and J. Poesen (2001). "Effects of Tillage on Runoff and Erosion Patterns". In: *Soil and Tillage Research*. XVth ISTRO Conference on Tillage at the Threshold of the 21st Century: Looking Ahead 61.1, pp. 55–60. DOI: 10.1016/S0167-1987(01)00178-7. (Visited on 10/12/2017).

Takken, I., G. Govers, A. Steegen, J. Nachtergaele, and J. Guérif (2001). "The Prediction of Runoff Flow Directions on Tilled Fields". In: *Journal of Hydrology* 248.1, pp. 1–13. DOI: 10.1016/S0022-1694(01)00360-2. (Visited on 10/12/2017).

Takken, I., V. Jetten, G. Govers, J. Nachtergaele, and A. Steegen (2001). "The Effect of Tillage-Induced Roughness on Runoff and Erosion Patterns". In: *Geomorphology* 37.1, pp. 1–14. DOI: 10.1016/S0169-555X(00)00059-3. (Visited on 10/12/2017).

Tien Bui, D., B. Pradhan, O. Lofman, I. Revhaug, and O. B. Dick (2012). "Spatial Prediction of Landslide Hazards in Hoa Binh Province (Vietnam): A Comparative Assessment of the Efficacy of Evidential Belief Functions and Fuzzy Logic Models". In: *CATENA* 96, pp. 28–40. DOI: 10.1016/j.catena.2012.04.001. (Visited on 01/18/2018).

Tiessen, H. (2008). "Phosphorus in the Global Environment". In: *The Ecophysiology of Plant-Phosphorus Interactions*. Ed. by P. J. White and J. P. Hammond. Plant Ecophysiology. Dordrecht: Springer, pp. 1–7. (Visited on 12/08/2021).

Tukker, A., R. Wood, and S. Schmidt (2020). "Towards Accepted Procedures for Calculating International Consumption-Based Carbon Accounts". In: *Clim. Policy* 20.sup1, S90–S106. DOI: 10.1080/14693062.2020.1722605. (Visited on 01/01/2023).

Turnbull, L., J. Wainwright, and R. E. Brazier (2008). "A Conceptual Framework for Understanding Semi-Arid Land Degradation: Ecohydrological Interactions across Multiple-Space and Time Scales". In: *Ecohydrol.* 1.1, pp. 23–34. DOI: 10.1002/eco.4. (Visited on 02/21/2017).

Vadiati, M., A. Asghari-Moghaddam, M. Nakhaei, J. Adamowski, and A. H. Akbarzadeh (2016). "A Fuzzy-Logic Based Decision-Making Approach for Identification of Groundwater Quality Based on Groundwater Quality Indices". In: *Journal of Environmental Management* 184, pp. 255–270. DOI: 10.1016/j.jenvman.2016.09.082. (Visited on 01/25/2018).

van der Heijden, G., A. Legout, B. Pollier, C. Bréchet, J. Ranger, and E. Dambrine (2013). "Tracing and Modeling Preferential Flow in a Forest Soil—Potential Impact on Nutrient Leaching". In: *Geoderma* 195–196, pp. 12–22. DOI: 10.1016/j.geoderma.2012.11.004. (Visited on 12/08/2021).

van Dijk, P. M., A.-V. Auzet, and M. Lemmel (2005). "Rapid Assessment of Field Erosion and Sediment Transport Pathways in Cultivated Catchments after Heavy Rainfall Events". In: *Earth Surf. Process. Landforms* 30.2, pp. 169–182. DOI: 10.1002/esp.1182. (Visited on 02/25/2017).

Van Oost, K., G. Govers, and P. Desmet (2000). "Evaluating the Effects of Changes in Landscape Structure on Soil Erosion by Water and Tillage". In: *Landscape Ecology* 15.6, pp. 577–589. DOI: 10.1023/A:1008198215674. (Visited on 12/25/2022).

Van Rompaey, A., G. Verstraeten, K. Van Oost, G. Govers, and J. Poesen (2001). "Modelling Mean Annual Sediment Yield Using a Distributed Approach". In: *Earth Surf. Process. Landf.* 26.11, pp. 1221–1236. DOI: 10.1002/esp.275. (Visited on 12/25/2022).

Venohr, M., U. Hirt, J. Hofmann, D. Opitz, A. Gericke, A. Wetzig, K. Ortelbach, S. Natho, F. Neumann, and J. Hürdler (2010). *The Model System MONERIS Version 2.14.1vba*. Berlin: Leibniz-Institute of Freshwater Ecology and Inland Fisheries.

Verstraeten, G., K. Van Oost, A. Van Rompaey, J. Poesen, and G. Govers (2002). "Evaluating an Integrated Approach to Catchment Management to Reduce Soil Loss and Sediment Pollution through Modelling". In: *Soil Use Manag.* 18.4, pp. 386–394. DOI: 10.1111/j.1475-2743.2002.tb00257.x. (Visited on 02/25/2017).

Verstraeten, G. and J. Poesen (2001). "Factors Controlling Sediment Yield from Small Intensively Cultivated Catchments in a Temperate Humid Climate". In: *Geomorphology* 40.1, pp. 123–144. DOI: 10.1016/S0169-555X(01)00040-X. (Visited on 12/25/2022).

Verstraeten, G., J. Poesen, J. de Vente, and X. Koninckx (2003). "Sediment Yield Variability in Spain: A Quantitative and Semiqualitative Analysis Using Reservoir Sedimentation Rates". In: *Geomorphology* 50.4, pp. 327–348. DOI: 10.1016/S0169-555X(02)00220-9. (Visited on 12/25/2022).

Verstraeten, G., J. Poesen, K. Gillijns, and G. Govers (2006). "The Use of Riparian Vegetated Filter Strips to Reduce River Sediment Loads: An Overestimated Control Measure?" In: *Hydrol. Process.* 20.20, pp. 4259–4267. DOI: 10.1002/hyp.6155. (Visited on 02/23/2017).

Wainwright, J., L. Turnbull, T. G. Ibrahim, I. Lexartza-Artza, S. F. Thornton, and R. E. Brazier (2011). "Linking Environmental Régimes, Space and Time: Interpretations of Structural and Functional Connectivity". In: *Geomorphology.* Geomorphology on Multiscale Feedbacks in Ecogeomorphology 126.3–4, pp. 387–404. DOI: 10.1016/j.geomorph.2010.07.027. (Visited on 04/05/2017).

Wang, Z., T. Zhang, C. S. Tan, and Z. Qi (2020). "Modeling of Phosphorus Loss from Field to Watershed: A Review". In: *J. Environ. Qual.* 49.5, pp. 1203–1224. DOI: 10.1002/jeq2.20109. (Visited on 12/08/2021).

Wemple, B. C., J. A. Jones, and G. E. Grant (1996). "Channel Network Extension by Logging Roads in Two Basins, Western Cascades, Oregon". In: *JAWRA J. Am. Water Resour. Assoc.* 32.6, pp. 1195–1207. DOI: 10.1111/j.1752-1688.1996.tb03490.x. (Visited on 03/17/2017).

Wheater, H. S., A. J. Jakeman, and K. J. Beven (1993). "Progress and Directions in Rainfall-Runoff Modelling". In: *Modelling Change in Environmental Systems.* Ed. by A. J. Jakeman, M. B. Beck, and M. J. McAleer. Chichester: Wiley, pp. 101–132. (Visited on 12/27/2022).

White, M. J. and J. G. Arnold (2009). "Development of a Simplistic Vegetative Filter Strip Model for Sediment and Nutrient Retention at the Field Scale". In: *Hydrol. Process.* 23.11, pp. 1602–1616. DOI: 10.1002/hyp.7291. (Visited on 02/25/2017).

White, M. J., D. E. Storm, P. R. Busteed, S. H. Stoodley, and S. J. Phillips (2009). "Evaluating Nonpoint Source Critical Source Area Contributions at the Watershed Scale". In: *J. Environ. Qual.* 38.4, pp. 1654–1663. DOI: 10.2134/jeq2008.0375. (Visited on 01/19/2018).

Wischmeier, W. H. and D. D. Smith (1978). *Predicting Rainfall Erosion Losses. A Guide to Conservation Planning.* Agriculture Handbook 537. Washington, DC: U.S. Government Printing Office.

Yapo, P. O., H. V. Gupta, and S. Sorooshian (1996). "Automatic Calibration of Conceptual Rainfall-Runoff Models: Sensitivity to Calibration Data". In: *Journal of Hydrology* 181.1, pp. 23–48. DOI: 10.1016/0022-1694(95)02918-4. (Visited on 12/25/2022).

Young, R. A., C. A. Onstad, D. D. Bosch, and W. P. Anderson (1989). "AGNPS: A Nonpoint-Source Pollution Model for Evaluating Agricultural Watersheds". In: *J. Soil Water Conserv.* 44.2, pp. 168–173. (Visited on 12/27/2022).

Zadeh, L. A. (1965). "Fuzzy Sets". In: *Information and Control* 8.3, pp. 338–353. DOI: 10.1016/S0019-9958(65)90241-X. (Visited on 01/23/2018).

— (1978). "Fuzzy Sets as a Basis for a Theory of Possibility". In: *Fuzzy Sets and Systems* 1, pp. 3–28.

Zessner, M., O. Gabriel, A. Kovacs, M. Kuderna, C. Schilling, G. Hochedlinger, and G. Windhofer (2011). *Analyse der Nährstoffströme in oberösterreichischen Einzugsgebieten nach unterschiedlichen Eintragspfaden für strategische Planungen (Nährstoffströme Oberösterreich).* Tech. rep. Wien: Amt der Oberösterreichischen Landesregierung.

Zessner, M., O. Gabriel, M. Kuderna, C. Weinberger, G. Hepp, A. Kovacs, and G. Windhofer (2014). "Effektivität von Maßnahmen zur Reduktion der Phosphorbelastung der oberösterreichischen Fließgewässer". In: *Österr Wasser- und Abfallw* 66.1-2, pp. 51–58. DOI: 10.1007/s00506-013-0130-2. (Visited on 03/01/2017).

Zessner, M., G. Hepp, M. Kuderna, C. Weinberger, and O. Gabriel (2017). *Zustandserfassung, Nährstoffentwicklung und Quantifizierung der Maßnahmenwirksamkeiten von ÖPUL 2007 in oberösterreichischen Einzugsgebieten.* Tech. rep. Wien: Amt der Oberösterreichischen Landesregierung.

Zessner, M., G. Hepp, M. Kuderna, C. Weinberger, O. Gabriel, and G. Windhofer (2014). *Konzipierung und Ausrichtung übergeordneter strategischer Maßnahmen zur Reduktion von Nährstoffeinträgen in oberösterreichische Fließgewässer.* Tech. rep. Wien: Amt der Oberösterreichischen Landesregierung.

Zessner, M., G. Hepp, O. Zoboli, O. Mollo Manonelles, M. Kuderna, C. Weinberger, and O. Gabriel (2016). *Erstellung und Evaluierung eines Prognosetools zur Quantifizierung von Maßnahmenwirksamkeiten im Bereich der Nährstoffeinträge in oberösterreichische Oberflächengewässer.* Tech. rep. Wien: Amt der Oberösterreichischen Landesregierung.

Zessner, M., A. Kovacs, C. Schilling, G. Hochedlinger, O. Gabriel, S. Natho, S. Thaler, and G. Windhofer (2011). "Enhancement of the MONERIS Model for Application in Alpine Catchments in Austria". In: *International Review of Hydrobiology* 96.5, pp. 541–560. DOI: 10.1002/iroh.201111278. (Visited on 04/24/2017).

Zessner, M., O. Zoboli, G. Hepp, M. Kuderna, C. Weinberger, and O. Gabriel (2016). "Shedding Light on Increasing Trends of Phosphorus Concentration in Upper Austrian Rivers". In: *Water* 8.9, p. 404. DOI: 10.3390/w8090404. (Visited on 02/28/2017).

Zimmermann, H.-J. and P. Zysno (1980). "Latent Connectives in Human Decision Making". In: *Fuzzy Sets and Systems* 4.1, pp. 37–51. DOI: 10.1016/0165-0114(80)90062-7. (Visited on 01/25/2018).

Zoboli, O., G. Hepp, J. Krampe, and M. Zessner (2020). "BaHSYM: Parsimonious Bayesian Hierarchical Model to Predict River Sediment Yield". In: *Environmental Modelling & Software* 131, p. 104738. DOI: 10.1016/j.envsoft.2020.104738. (Visited on 01/04/2021).

Zoboli, O., A. Viglione, H. Rechberger, and M. Zessner (2015). "Impact of Reduced Anthropogenic Emissions and Century Flood on the Phosphorus Stock, Concentrations and Loads in the Upper Danube". In: *Science of The Total Environment* 518–519, pp. 117–129. DOI: 10.1016/j.scitotenv.2015.02.087. (Visited on 02/28/2017).