



**University of
Zurich**^{UZH}

Performance Analysis of Different Flash Flood
Nowcasting Chains Based on Ensemble Precipitation
Nowcasting in the Emme and Verzasca Catchments
The Impact of Blending an Extrapolated Nowcast into
a Numerical Weather Prediction

GEO 511 Master's Thesis

Author

Severin Koller
14-722-680

Supervised by

Dr. Massimiliano Zappa (massimiliano.zappa@wsl.ch)

Faculty representative

Prof. Dr. Jan Seibert

27.06.2019

Department of Geography, University of Zurich



**University of
Zurich**^{UZH}

Performance Analysis of Different Flash Flood
Nowcasting Chains Based on Ensemble Precipitation
Nowcasting in the Emme and Verzasca Catchments
The Impact of Blending an Extrapolated Nowcast into
a Numerical Weather Prediction

GEO 511 Master's Thesis

Author

Severin Koller
14-722-680

Supervised by

Dr. Massimiliano Zappa (massimiliano.zappa@wsl.ch)

Faculty representative

Prof. Dr. Jan Seibert

27.06.2019

Department of Geography, University of Zurich



**University of
Zurich** ^{UZH}



Performance Analysis of Different Flash Flood Nowcasting Chains Based on Ensemble Precipitation Nowcasting in the Emme and Verzasca Catchments

The Impact of Blending an Extrapolated Nowcast into a Numerical
Weather Prediction

GEO 511.1 Master Thesis in Hydrology and Climate

Author:

Severin Koller

14-722-680

Supervision

Dr. Massimiliano Zappa

Swiss Federal Institute for Forest, Snow and Landscape Research (WSL),
Mountain Hydrology and Mass Movements Hydrological Forecasts

Member of faculty

Prof. Dr. Jan Seibert

University of Zurich,

Department of Geography, Hydrology and Climate

30th June 2019

Department of Geography, University of Zurich

Abstract

Suddenly evolving flash floods exhibit large hazard potential. Not only they do cause costly damages, but are also a threat to life. Very short-term precipitation nowcasting techniques in combination with a hydrological model are a means of predicting flash floods as early and reliably as possible.

This thesis analyses the benefit of blending a radar precipitation extrapolation into a numerical weather prediction (NWP) in comparison to raw NWPs. Accordingly, three nowcasting chains called NP1 NPC and NPE are set up. All of them incorporate (i) initial conditions in form of a quantitative precipitation estimation, (ii) a probabilistic precipitation nowcast extrapolation generating an ensemble of precipitation nowcasts, except for the NPC nowcasting chain which uses a deterministic nowcast extrapolation, (iii) a deterministic NWP continuing the NP1 and NPC nowcasting chain or a probabilistic NWP continuing the NPE nowcasting chain, (iv) a blending scheme, which merges the extrapolation nowcast into the NWP with regard to their spread in the ensemble, while the spread of deterministic predictions is taken from their probabilistic counterparts, and (v) a probabilistic hydrological model generating runoff predictions. The merged nowcasting chains are, then, compared to two nowcasting chains called nowcasting chain COSMO-1 (CO1) and nowcasting chain COSMO-E (COE), which generate their predictions without precipitation extrapolation and blending scheme, but only use the deterministic COSMO-1 and probabilistic COSMO-E NWP. This comparison is done in terms of the nowcasting chains skills in relation to a radar and rain gauge hindcast. A set of events is analysed and the best performing nowcasting chain is identified. For the evaluation, the proportion of best performing nowcasting chains is illustrated for each catchment. Hence, this thesis analyses whether an update cycle with frequently updated initial conditions leads to more reliable flash flood predictions in the used nowcasting chains.

The results show that the blended nowcasting chains outperform the raw NWP nowcasting chains in the main catchments of the Emme and Verzasca by over 70 %. In smaller subcatchments of the Emme and Verzasca, the superiority of the blended nowcasting chains is less pronounced with a proportion of best performing nowcasting chains of over 60 %. The used update cycle shows that with frequently updated initial conditions all nowcasting chains improve their skill.

The credibility of this thesis's results is underpinned by the inclusion of other relevant and up-to-date nowcasting related studies. Nonetheless, the limitations of the thesis are shown and

discussed. Moreover, this thesis shows that current issues such as climate change pose major challenges to the discipline of flash flood nowcasting. However, the findings of this thesis and related studies encourage that the further development of extrapolated nowcasts and blending schemes will be beneficial in mitigating the hazard potential of flash floods in the future.

Acknowledgments

I would like to thank my supervisor Massimiliano, from whose advice and experience I was able to benefit uncomplicatedly and at any time. In this respect many thanks to the mountain hydrology and mass movements hydrological forecast group at the WSL, who involved me in their team as a fully-fledged member from the very first day and were available to answer all kind of questions.

Furthermore, I would like to thank Jan. As professor of the hydrology and climate group at the university of Zurich, he always promoted my fascination for hydrology.

Finally, the thesis would not have the same value if Daniele from MeteoSwiss had not supported me all the time with his unconditional commitment and I would like to thank him for that.

List of Abbreviations

Anom r	Anomaly correlation coefficient
CO1	Nowcasting chain COSMO-1
COE	Nowcasting chain COSMO-E
COSMO	Consortium for small-scale modelling
COSMO-1	COSMO 1.1 km
COSMO-7	COSMO 6.6 km
COSMO-E	COSMO Ensemble
COSMO-LEPS	COSMO limited-area ensemble prediction system
CPC	CombiPrecip
CV	Coefficient of variance
DRP	Dominant runoff process
EFAS	European flood awareness system
EPIC	European precipitation index based on simulated climatology
EPS	Ensemble predicting system
FAR	False alarm ratio
FF	Flash flood
FFG	Fash flood guidance
GOF	Goodness of fit
HPE	Heavy precipitation event
HRU	Hydrological response units
INCA	Integrated nowcasting through comprehensive analysis system
KGE	Kling-Gupta efficiency
MAE	Mean absolute error
MAP D-PHASE	Mesoscale alpine programme demonstration of probabilistic hydrological and atmospheric simulation of flood events
NowPAL	Nowcasting of precipitation accumulations
NP1	Nowcasting chain nowcast product COSMO-1
NPC	Nowcasting chain nowcast product control member
NPE	Nowcasting chain nowcast product COSMO-E
NSE	Nash-Sutcliffe efficiency
NWP	Numerical weather prediction
OBS	Observation

PDF	Probability density function
POD	Probability of detection
POFD	Probability of false detection
PQPF	Probabilistic quantitative precipitation forecast
QPN	Probabilistic quantitative precipitation nowcast
PREVAH	Precipitation-Runoff-Evapotranspiration HRU Model
QPE	Quantitative precipitation estimation
QPF	Quantitative precipitation forecast
QPN	Quantitative precipitation nowcast
RGM-PRO	Process-based runoff generation model
RMSE	Root mean square error
ROC	Relative operating characteristic
RT	Runoff type
SPPT	Stochastically perturbed parametrisation tendencies
S-PROG	Spectral prognosis model
STEPS	Short-term ensemble prediction system
UTC	Coordinated universal time

Contents

Abstract	I
Acknowledgments	III
List of Abbreviations	IV
1. Introduction	1
1.1. Motivation	1
1.2. Literature Review	3
1.2.1. Nowcast Extrapolation and NWP	3
1.2.2. Dealing with Uncertainty	4
1.2.3. Blending Scheme	7
1.2.4. Further Prediction Components	10
1.2.4.1. Alert System	11
1.2.4.2. Radar Technique	12
1.2.4.3. Hydrological Model	12
1.2.4.4. Communication	13
1.3. Objectives	14
2. Data and Methods	17
2.1. Study Catchments	17
2.2. Hydrological Model	20
2.3. Events Selection	20
2.4. Nowcast Extrapolation and Blending	23
2.5. Numerical Weather Prediction	24
2.6. Experimental Set-up	25
2.7. Verification Methods	27
2.7.1. Efficiency Scores	28
2.7.2. Taylor Diagram	29
2.7.3. Peak-box	29
3. Results	31
3.1. Comparison of Nowcasting Chains	31
3.1.1. Performance Overview	35
3.1.2. Ranked Events	40
3.1.3. Best Performing Nowcasting Chains	42
3.1.3.1. Deterministic and Probabilistic Nowcasting Chains	46
3.2. Impact of an Update Cycle	48
3.2.1. Progression Analysis	51
3.2.2. Flood Event Assessment	55

4.	Discussion	59
4.1.	Blending an Extrapolation Nowcast into a NWP.....	59
4.1.1.	Blending Scheme and Weighting Function.....	59
4.1.2.	Extrapolation Nowcast and Ensemble Size.....	60
4.2.	Performance Improvement through Frequently Updated QPEs	61
4.3.	Nowcasting Chains Specific Characteristics	62
4.4.	Catchment Specific Characteristics	64
4.5.	The Benefit of Probabilistic Nowcasting Chains	65
4.6.	Limitations.....	66
5.	Summary and Conclusion	69
6.	References	72
7.	Appendix	78
7.1.	Event Selection NowPAL Output.....	78
7.2.	Performance Overview NSE and MAE.....	79
7.3.	Boxplot Overview KGE, NSE and MAE	82
7.4.	Ranked Events KGE, NSE and MAE.....	90
7.5.	Best Performing Nowcasting Chains.....	94
7.5.1.	All Nowcasting Chains NSE.....	94
7.5.2.	Deterministic Nowcasting Chains NSE and MAE.....	95
7.5.3.	Probabilistic Nowcasting Chains NSE and MAE	96
7.6.	Entire Update Cycle.....	97
7.6.1.	Accumulated Precipitation Prediction.....	97
7.6.2.	Runoff Simulation	101
7.6.3.	Median Progression NSE	106
7.6.4.	Boxplot Progression Subcatchments	107
7.6.5.	Taylor Diagram and Peak-box	108
	Personal Declaration	113

1. Introduction

The introduction first describes a short motivation, which underpins the importance of this thesis. This is followed by a longer literature review, which puts the thesis in the right scientific context and provides a basis to capture the topic entirely. The objectives in the end of the introduction formulate the research questions to answer.

1.1. Motivation

Very short-term precipitation forecasting (nowcasting) is fast becoming a key instrument in the field of runoff prediction. The term nowcasting differs from the term forecasting in this respect that only short lead times up to six hours are considered in the nowcast. Flash floods (FFs) caused by extreme runoff and initialised by heavy precipitation events (HPEs) pose a threat not only to infrastructure, but also to human lives. FFs and co-occurring debris flows are by far the most important natural hazards in Switzerland: 94 % of the total natural damages in the year 2017 were caused by FFs and debris flows. Other natural hazards like landslides or rockfall are almost negligible (Andres and Badoux, 2018). Between the years 1972 and 2007, costs with a total amount of 7'110 million Euros (exchange rate on 31 December 2007 and taking inflation into account) were caused only by floods and inundations. This vast amount of costs is mainly caused by a few extreme flood events. This can be shown by six extreme events during the years from 1978 to 2005, which are responsible for more than half of the total costs in the period from 1972 to 2007. The ongoing climate change has the consequence that the number of extreme precipitation events will increase the hazard potential of FFs in the future (Hilker et al., 2009). Therefore, a reliable event-based prediction system is of undeniable importance.

The name flash flood, however, indicates that these floods evolve rapidly and are therefore difficult to predict in both magnitude and timing. Especially, in small and mountainous catchments, which are prone to FF, an early and accurate forecast is challenging. In addition, most catchments are divided into several nested subcatchments. It is difficult to accurately capture how the subcatchments interact with each other, which consequently brings another source of uncertainty into the runoff prediction. Finally, small and mountainous catchments are often ungauged. The only possible method to collect precipitation data and give an estimation about the precipitation in these catchments is by radar measurements. As a result, no observation data from rain gauges are available for verification (Liechti et al., 2013a).

Very short-term forecasting with recently updated precipitation estimation is the most promising method to detect FFs in advance. Vague long-term predictions with a long lead struggle to capture suddenly occurring FFs. Neither for scientists nor for policy-makers are long-term predictions, which try to predict a FF far in time, an option to assess the threat of a FF (Antonetti et al., 2019). A meteorological nowcast, which describes the state of the atmosphere in only the next few hours, offers a more reliable estimation of the amount of precipitation. Consequently, nowcasting with a suitable hydrological model and the very short-term precipitation prediction is the most sophisticated way to predict FFs, and therefore to mitigate hazards (Mass, 2011). Consequently, progress in the application of nowcasting systems needs to be pushed further ahead in order to maintain civil defence in the future (Thorndahl et al., 2017).

Important for the entire discipline of detecting FFs is a reliable quantitative precipitation estimation (QPE), which could be radar-based and/or rain gauge-based. The QPE is used as the initial condition for the FF prediction. Since radar and rain gauge data are available in real time, a nowcast product in the form of an optimal extrapolation algorithm, which models the precipitation fields into the future, gives a nearly immediate estimation of the precipitation progression of the QPE (Germann and Zawadzki, 2002). However, a pure extrapolation of radar images does not consider crucial growth and decay processes in the atmosphere and quickly becomes inaccurate. Nonetheless, a switch from the QPE directly into a numerical weather prediction (NWP) can be subject to great uncertainties since the prediction needs some computation time, and hence uses outdated observation data (Nerini et al., 2019). In this thesis, a new nowcasting approach is used to achieve a more reliable prediction of precipitation, and as a result a better basis for the hydrological nowcast. This nowcasting approach seamlessly blends the extrapolated nowcast of the rain fields with the NWP gradually over time depending on their ensemble spreads. Seamless blending makes it possible to use the strengths of both the nowcast product and the NWP by adjusting their weights in inverse proportion to their uncertainties (Foresti et al., 2016). The term seamless indicates a consistent prediction regardless of location, lead time or forecasting procedure (World Meteorological Organisation 2015). The outcome is a more realistic simulation of short-term heavy precipitation events, and therefore a more realistic simulation of the runoff. This more sophisticated approach is able to assess the hazard potential of FFs more accurately and will have a trendsetting impact in the future.

1.2. Literature Review

There is a growing body of literature that recognises the importance of the implementation of a seamless blending from the QPE into the NWP. A large number of contributors have all been part of the establishment of various forecasting systems, which blend an extrapolation nowcast into a NWP (Bowler et al. 2006, Foresti and Seed 2015, Foresti et al. 2016, Haiden et al. 2011 and Nerini et al. 2019). In this section, an overview of the different components used for FFs nowcasting is given and their developments brought into a scientific context.

1.2.1. Nowcast Extrapolation and NWP

Regarding the extrapolation nowcast, one of the first assessment of it is attributed to Wilson et al., (1979). In the study, rain fields were produced by a rainfall forecasting model, which is able to extrapolate precipitation into the future by making some assumptions. This approach can be seen as one of the first nowcast extrapolation approaches ever realized in rainfall-runoff modeling, although the term extrapolation or nowcast never appears in the study. The used assumptions about velocity and direction of rain fields or their correlation with time and space were quite simple, but still used in today's extrapolation nowcasts. The study could show that a useful precipitation extrapolation with a proper rainfall spatial distribution has a large impact on accuracy of the runoff prediction (Wilson et al., 1979).

Nowadays, it is a fact that the spatial distribution of rainfall fields plays a decisive role in predicting runoff (e.g. Cristiano et al., 2017; Ochoa-Rodriguez et al., 2015). Especially small and mountainous catchments, as they occur in Switzerland, suffer from large errors in the runoff prediction if the spatial distribution of the rainfall fields cannot be represented correctly (Liechti et al., 2013a). Further improvements took place in the type of initial data used for the extrapolation nowcast. The extrapolation nowcast uses initial rain gauge data as a starting point for the extrapolation. With today's radar techniques, it is possible to use up-to-date initial rain data provided by radar measurements. The use of raing gauge measurements, as initial data, is therefore not absolutely necessary. For the first time, this enables the estimation of precipitation in ungauged basins (Liechti et al., 2012).

The most important reason why nowcast extrapolation is extremely promising is that measured precipitation values (i.e. QPEs) used as initial conditions, from radar measurements and/or rain gauges, can be incorporated into the nowcast of precipitation in near real time. These initial conditions can be frequently updated, which enables the detection of potentially dangerous convective processes in the atmosphere, which in turn can lead to HPEs and therefore FFs. A major disadvantage of nowcast extrapolation is that it is not calculated by numeri-

cal models or dynamic equations as it is done by NWP. As a result, the nowcast extrapolation loses rapidly predictive skill with increasing lead time. On the other hand, a NWP with observation data, which can be renewed less frequently than those of the extrapolation nowcast, struggles to incorporate condensation processes into the prediction, especially if they are small-scaled. The forecast skill of the NWP is, therefore, mainly depending on the time past the last update of the initial data and the degree by which the atmosphere has changed during this time (Haiden et al., 2011). The question can therefore be asked whether it makes sense to use numerical models at all to predict single extreme precipitation events, which mainly trigger FFs. The study of Klasa et al. (2018) demonstrates that there have been far-reaching improvements in recent decades regarding the development of NWP. These improvements made the NWP systems suitable also for FFs predictions. One of the most important improvements was that current NWP are able to incorporate convection processes in the simulations. It has been demonstrated that these convection-permitting predicting systems provide more reliable forecasts in all cases investigated (Klasa et al., 2018).

1.2.2. Dealing with Uncertainty

When dealing with forecasts or nowcasts, it is always essential to incorporate uncertainties into the simulation. A nowcast which models only one possible runoff does not provide any estimation about uncertainty. However, with increasing computer power, it became possible to generate multiple runoff simulations in a meaningful computation time. Each runoff simulation, provided from the hydrological model, is different to the other because the meteorological initial conditions and/or the parameter set of the hydrological model are slightly perturbed for each run. The resulting outcome is known as an ensemble or probabilistic forecast, which is used in an ensemble predicting system (EPS). The numerous generated runoff hydrographs provide a good estimation of uncertainty through their different predictions, with each one of them representing one possible scenario of the runoff. It is, therefore, not surprising that the use of EPSs has increased rapidly in the past (Rossa et al., 2011).

Nevertheless, not only the hydrological model, but also the meteorological predictions of precipitation estimation, done by the nowcast extrapolation and the NWP, uses the ensemble approach and therefore the EPS approach. Regarding the quantification of uncertainty in the nowcast extrapolation, Bowler et al., (2006) presented in their study a trend-setting methodology which is still used today. The ensemble predicting system from Bowler et al. (2006) is based on the findings of Seed (2003), which in turn introduces a spectral prognosis model (S-PROG) based on the circumstance that the persistence in time of observed precipitation features behave proportional to their size. Differently spoken, this means that smaller features of

the precipitation field change faster in time than larger ones. The term feature in this context is associated to smaller or larger precipitation patterns within the precipitation field. It is assumed that these features occur in any precipitation field. Since large-scale features behave more persistently in the nowcast extrapolation over time, their simulation in the extrapolation is more reliable than the simulation of the small-scale features, which are in many cases entirely unpredictable. Thus, Bowler et al. (2006) used these findings to extract the small-scale features and to replace them by stochastic perturbation. By adding stochastic perturbation to the nowcast, it is possible to generate different simulations, and as a result an ensemble of nowcast simulations (members). The different simulations of the precipitation field make it possible to quantify the uncertainty in the nowcast. It therefore applies that the larger the spread of the different members is, the more uncertainty there is in the model (Nerini et al., 2019).

Regarding the NWP, the amount of error sources, which leads to uncertainty in the model prediction, is nearly endless. A numerical model tries to represent the dynamic processes in the atmosphere by spatial and temporal discretisation. However, the amount of processes in the atmosphere is uncountably large and their variability is difficult to capture. For this reason, many atmospheric processes are simplified to make their implementation into the model possible, or to simply make them more understandable. Equations like the Navier-Stokes flow equation, the mass continuity equation, and the first law of thermodynamics are mathematical formulations which try to capture the complex processes in the atmosphere. It is not possible to solve these equations analytically, but it is numerically. This means that the numerical solution of the equation is just an approximation. Furthermore, for some physical processes in the atmosphere even their scale of motion is unknown. As a consequence, these atmospheric processes have to be parameterized, even if these processes are just a simplified representation of the reality (Bauer et al., 2015).

To find an optimal set of parameters is always a difficult challenge and adds further uncertainty to the prediction. Various verification methods with different strategies to get the goodness of fit (GOF) of the prediction were elaborated and used in the past. A perfect parameter set can not be found at all because it is the task of an experienced modeler to decide how to verify the prediction and which GOF should be used. This decision depends above all on how much and which experiences the modeler has already gained, and can therefore be challenged. Finally, the phenomenon and pitfall of equifinality must be mentioned when dealing with the uncertainties of parametrisation. Equifinality describes the problem when there are several optimal parameter sets which all have the same best GOF. This circumstance is contradictory,

since in reality there should be only one correct value per parameter. Accordingly, equifinality represents another source of uncertainty in prediction (Beven, 2012).

In order to deal with parameter uncertainties, there are various ways to optimally incorporate them into a model. The uncertainties in the parameters go along with the ensemble approach, as different ways have been developed to vary the parameters in a model, and thus get different simulations in an EPS. The strategy on how to choose the variation in the parameter set for each member of the EPS should improve the reliability of the prediction. One approach introduced by Houtekamer et al. (1996) is called multiphysics approach. This approach gives each simulation in the ensemble a specific set of parameters for each physical process. Another approach called multiparameter approach developed by Murphy et al., (2004) varies each individual parameter in a given range randomly around the default value. This approach is particularly valuable in a changing climate as the range of parameters can be adapted according to the knowledge of experts. Many other approaches to parametrise a model exist, and they all serve a distinct purpose. Furthermore, the approaches can be combined with each other to produce even better EPSs (Klasa et al., 2018). Nevertheless, it is not the intention to account for all of them when dealing with uncertainty, but to give an impression of how important the field of appropriate parametrisation of a model is in EPSs.

In the past it could be shown that the mean value of an EPS is superior to a single forecast. As a result, it became recommendable to use EPSs also in the prediction of FFs (Klasa et al., 2018). Additionally, the different simulations represent the range of uncertainty, and thus provide information about the reliability of the forecast (Bauer et al., 2015). It has to be considered, that the uncertainty is artificially generated in the meteorological input, which forces the hydrological model. However, the uncertainties stemming from the hydrological model are often not taken into account (Zappa et al., 2008). The reason for this is that the uncertainties of the hydrological model are only weakly pronounced in comparison to other uncertainties. The uncertainties originating from the meteorological input are far more essential, especially in mountainous and small catchments, where orographic uplifts contribute substantially to convection processes, and the precipitation associated with them.

The question arises as to why not consider and combine all type of uncertainties in the EPS. The answer to this question can be found in the total uncertainty when the different partial uncertainty sources superimpose. In this case, the total uncertainty does not correspond to the cumulated uncertainty of the individual partial uncertainties, but to a considerably higher value. This non-linear behaviour in uncertainty superposition is an important finding when dealing with FFs predictions (Zappa et al., 2011). It leads to the consequence that only the

meteorological uncertainties are considered in this master thesis.

The mentioned findings in the past can also give an impression of how the development in dealing with uncertainties in EPSs could continue. The steadily increasing computer power of the past made it possible to move the development in EPSs forward. This will continue to be the case in the future. Spatially high-resolution models with multiple members and incorporated processes with higher degrees of complexity provide more advanced predictions, but also require new developments in computer power. Moreover, also the processes of the various environmental spheres must continually be researched in depth and their complexity understood to deal with uncertainty in EPSs (Bauer et al., 2015).

1.2.3. Blending Scheme

The approach of blending an extrapolation nowcast into a NWP is relatively new. One of the reasons why it was likely to stay away from the extrapolation nowcast and rely more on the NWP was that there was a too large gap of complexity between the nowcast extrapolation and the NWP. While the nowcast extrapolation was avoided in many studies as a simple and rather rudimentary extrapolation of precipitation, the NWP was appreciated as a highly complex EPS. The NWP was, therefore, to be assumed to deliver highly sophisticated results instead of the nowcast extrapolation, which in its simplicity is incapable of grasping the complex processes of the atmosphere (Golding, 1998).

One of the first attempt to blend the extrapolation nowcast with the NWP was made by Golding (1998). This study recognised the advantages of both prediction systems and combined them. The advantage of the extrapolation nowcast is that the up-to-date initial conditions from the radar and/or rain gauges observations can be used in almost real time. Consequently, the extrapolation nowcast, starts from perfect initial conditions, if the uncertainties in the radar and/or rain gauges are neglected. Even with simple extrapolation algorithms, it is possible to obtain a high quality prediction of the QPE for short lead times, which is called quantitative precipitation nowcast (QPN) (Foresti and Seed, 2015). Furthermore, the computational costs can be kept low, because only one variable, the precipitation, has to be extrapolated and the complex processes of the atmosphere do not have to be calculated. The QPN is, as a result, available almost immediately. On the downside the nowcast extrapolation quickly loses skill, due to his rather simple approach, compared to the NWP approach.

The NWP on the other hand is very unsuitable for very short lead time predictions, and hence for the QPN. The reason for this is the high computational costs, which are based on the holistic computation of the atmospheric processes. In addition, the NWP uses low spatial

resolution initial conditions, which brings a certain degree of fuzziness into the prediction. The low spatial resolution in the NWP is used because otherwise the computation time would be strongly increased, and the initial condition would be even more outdated. Furthermore, the uncertainties in the parametrisation and the simplification of physical processes are reasons for an imperfect simulation. The main advantage of the NWP, however, is that the predominant large-scale events in the atmosphere can be well captured, which enables a reliable prediction of the QPE for long lead times called quantitative precipitation forecast (QPF). Only after a considerable long lead time, the NWP loses much of its skill due to the effect that subsequent errors of the simplified physical representations add up, and the development of small-scale features cannot be recorded. Summarizing these findings, figure 1 shows a schematic illustration of the skill behaviour with increasing lead time of an extrapolation nowcast and a NWP. The figure additionally shows the theoretically best possible prediction skill. The best possible prediction skill decreases with increasing lead time, since the chaotic system in the atmosphere inevitably leads to uncertainties in the prediction (Golding, 1998; Jenkner et al., 2008; Nerini et al., 2019).

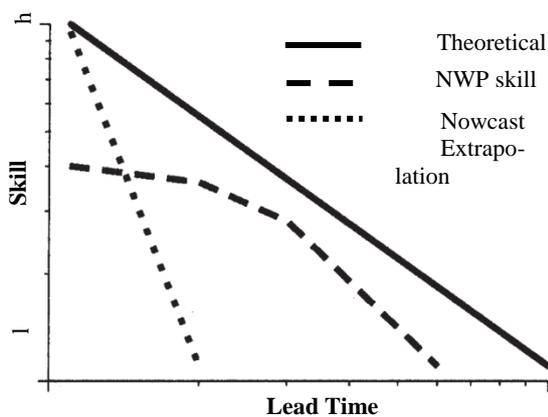


Figure 1: Schematic illustration of the skill loss with increasing lead time. According to Golding (1998).

The study by Golding (1998) could show that blending the extrapolation nowcast with the NWP does indeed produce significantly improved predictive results than a raw NWP or a raw extrapolation nowcast. Based on the findings of Golding (1998) and, not to be neglected, through increasing computing power in the past, it was to be expected that more sophisticated blending schemes and enhanced extrapolation nowcasting system would emerge.

One of these more recent QPF system is called integrated nowcasting through comprehensive analysis system (INCA), which is fully described in Haiden et al. (2011). The main reason why INCA was developed was that the conventional simple extrapolation nowcast provided insufficient predictive skills in mountainous areas. Orographic cloud formation leading to precipitation could not be captured with the existing nowcast extrapolation. This is because the simple extrapolation of a single variable like precipitation is quite challenging in mountainous areas. Thus, the lead time of a meaningful prediction is dramatically reduced. INCA is able to integrate variables like temperature, humidity, global radiation, wind or

ground temperature into the forecast. This more sophisticated forecasting system focuses not only on improving the QPN with a very short lead time, but also on improving forecasts of the NWP with longer lead times. This improvement can be achieved by integrating the multivariable analyses and forecasts of the system into the NWP. These analyses are forced by observation data and remote sensing data. These data contribute substantially to a better prediction as they are able to capture changes in the atmosphere almost instantaneous. The INCA system incorporates a nowcast extrapolation as well as a blending scheme to gradually merge the extrapolation with the NWP. The precipitation extrapolation nowcast is based on Lagrangian persistence. Lagrangian persistence basically assumes that the precipitation fields move persistently along the coordinates over time, and is further described in Zawadzki et al. (1994). The blending scheme in INCA is simple. A weighting function blends the nowcast extrapolation after the first two hours of raw extrapolation nowcast linearly into the NWP. After six hours the full weight is given to the NWP. Obviously, a linear blending scheme is not the most sophisticated method to merge a precipitation extrapolation nowcast with the NWP. However, in the INCA system, other blending schemes like blending the precipitation extrapolation nowcast with the NWP taking into account their uncertainties and weighting them accordingly, was not to be proved beneficial, at least for the variable of precipitation. With respect to other variables, the INCA system could show that with more sophisticated blending schemes, better results can be achieved. If the variable temperature is taken, it can be shown that a more reliable prediction of the temperature can be made by including the stability of the atmosphere in the blending scheme. Thus, the extrapolation nowcast is blended faster with the NWP in a turbulent atmosphere than with stable conditions such as an inversion situation. This is done because in a turbulent atmosphere the nowcast extrapolation loses skill more quickly (Haiden et al., 2011).

Two important findings in relation to FFs predictions can be derived from the INCA system. Firstly, an extrapolation nowcast provides a reliable QPN, which outperforms a raw NWP for very short lead times. Secondly, a linear blending of the extrapolation nowcast with the NWP can be applied to obtain a more reliable QPN. However, it can be assumed that the blending scheme used still has potential for improvement (Haiden et al., 2011).

The recently developed and constantly improved short-term ensemble prediction system (STEPS) incorporates one of these more sophisticated blending schemes and offers great potential for reliable QPN and QPF. STEPS is one of the first prediction system which uses a probabilistic ensemble approach for its nowcast extrapolation and blends it into the NWP. Regarding STEPS, it is possible, therefore, to speak not only of a QPN or a QPF, but also of a

probabilistic quantitative precipitation nowcast (PQPN) or forecast (PQPF) (Bowler et al., 2006). This probabilistic approach is crucial in FFs nowcasting because it allows to assign a probability to a possible precipitation field. The representation of various precipitation fields would not be possible with a deterministic approach because the deterministic approach offers only one possible scenario at the time. This is especially dangerous in small catchments, which are prone to FFs. The deterministic approach, giving no information about other precipitation scenarios, could easily miss a hazardous FF event (Liechti et al., 2013a).

The probabilistic STEPS uses a very high spatial and temporal resolution for the predictions. This is also of great relevance in the field of FFs prediction, since small catchments or subcatchments can only be captured with high resolution prediction systems, and, furthermore, the prediction of FFs must be done with high temporal resolution, since FFs can develop suddenly, and their capture could otherwise be missed. Since the relatively coarse NWP grid, into which the extrapolation nowcast is blended, has not a high enough spatial resolution to consider small precipitation fields, a simple downscale approach had to be incorporated in STEPS. The downscaling approach basically divides a grid of the NWP into four subgrids and reassigns precipitation values for each subgrid. As a result, the precipitation NWP can also be used for small-scale catchments. Furthermore, in the study on STEPS, it was possible to quantify two different uncertainties considering the extrapolation nowcast. Firstly, the uncertainty in the motion of the precipitation field, and secondly, the uncertainty in the evolution of the precipitation field. STEPS was the first system to take into account the uncertainties in the evolution of the precipitation fields, and is, therefore, able to consider important growth and decay processes in the atmosphere. Even more it could be shown that the uncertainties in the evolution of precipitation fields are more important than the uncertainties in the motion of precipitation fields. The blending scheme of STEPS merges the nowcast with the NWP based on their uncertainties. Hence, better predictions could be made with the newly developed blending scheme of STEPS, which takes into account not only the uncertainties in the motion of precipitation fields but also by including the more important uncertainties in the evolution of the precipitation fields (Bowler et al., 2006).

1.2.4. Further Prediction Components

The effort to apply nowcast extrapolation and the use of a suitable blending scheme leading into the NWP is crucial for reliable FFs predictions. Nowcasting FFs is however embedded in a large scientific field, which incorporates various subdisciplines like the use of an appropriate warning system, the correct estimation of the initial precipitation used for the prediction, the hydrological model generating runoff, and the task of communicating comprehensibly the

results between scientists and decision-makers. It is, therefore, essential to give a brief overall view about the interactions of nowcasting FFs with other subdisciplines.

1.2.4.1. Alert System

The hazard potential of a FF arises primarily from the fact that FFs can occur suddenly during and after HPEs. It is evident that FFs are detected with event-based forecasting systems. However, especially in catchments with short response time, there is hardly any time between the HPE and the triggering of the FF to take action. It is, therefore, crucial to have a sophisticated alert system with well-considered thresholds (Antonetti et al., 2019).

Catchments that are prone to FFs are in many cases ungauged. As a result, there are none or only a few runoff measurements of the past. Thresholds are set in most cases regarding the meteorological situation. Or differently spoken the threshold is set for the rainfall input for the hydrological model. The fact that the catchments are ungauged makes it difficult to set meaningful thresholds for the rainfall, because only descriptive reports are available instead of accurate measurements (Alfieri et al., 2015).

A very sophisticated way to find a rainfall threshold for a catchment is to use the European precipitation index based on simulated climatology (EPIC). EPIC is incorporated in the European flood awareness system (EFAS) and uses the NWP from the consortium for small-scale modelling (COSMO) called COSMO limited-area ensemble prediction system (COSMO-LEPS). COSMO-LEPS calculates for each of the 1 km² grids 16 members, which have a value from zero, corresponding to no rain, to one, corresponding to the average of the annual maximum. It is also possible that the value is greater than 1. Criteria can be set for each grid, since, for example, 4 of the 16 members must have at least the value 1. If these criteria are fulfilled for a larger grid area, a probability density function (PDF) can be used to determine the probability that a certain return period will be exceeded. EPIC is only one possibility to determine meaningful thresholds. However, the ensemble approach is very convincing, since it makes possible to consider uncertainties of where and when exactly the FF will appear (Alfieri et al., 2015).

Another approach, which is mainly used in the United States, is the flash flood guidance (FFG). The threshold is not determined by the meteorological input, but by a hydrological model which is run backwards. Thus, the intensity of rainfall can be calculated, which would be needed to trigger a FF. In this variant of determining the threshold, however, a sophisticated hydrological model is the basic prerequisite (Georgakakos, 2006).

1.2.4.2. Radar Technique

The improvement in the application of radar techniques is of great importance in FFs prediction systems. Radar images and the associated radar technology provide the QPE, which is the initial data for the nowcast extrapolation or the NWP. In the last decades, the possibilities in the application of radar images have improved substantially. Nowadays, radar measurements are not only dependent of one single radar, but also of an entire radar measuring network (Thorndahl et al., 2017). Furthermore, error sources like ground echos or ground clutters could be mitigated, which increases the quality of radar measurements (Sideris et al., 2014a). On the basis of the enhanced radar precipitation estimations, the lead times of precipitation nowcasting could be enlarged from less than one hour to a few hours (Germann and Zawadzki, 2002). Radar measurements of precipitation have the great advantage that they can cover a large area without being maintained intensively. It is, therefore, possible to have an estimation of precipitation in regions which are not covered by rain gauges, either because they are not accessible or because the maintenance is too intensive for rain gauges (Sideris et al., 2014a). Rain gauging stations, which have the advantage of providing highly accurate data for the QPE, are widely spread all over Switzerland. However, the network density of the rain gauges is rather low. In addition, it has to be assumed, that the rain gauging network will not densify in the future, due to the considerable effort required for installation and maintenance (Liechti, 2013). However, the great advantage of the rain gauge measurements high accuracy, improve the quality of the QPE remarkably, if they are combined with radar precipitation estimations (Sideris et al., 2014a).

1.2.4.3. Hydrological Model

For the use of FF prediction, it is necessary to have a hydrological model which is forced by the input precipitation data. Various hydrological models exist with different levels of complexity and different application purposes. The decision of which model to use, however, does not depend on its degree of complexity, but on the purpose for which it is needed. As a result, it would be misleading to speak of less advanced or more advanced models (Beven, 2012). The process-based runoff generation model (RGM-PRO), which is also used in this thesis, will be further explained in section 2.2. To put the RGM-PRO in a scientific context it is necessary to separate the hydrological model from others. There are many ways to differentiate between different hydrological model approaches. A first very general differentiation in hydrological models can be made by differentiating between event-based models or continuous models (Beven, 2012). Event-based models serve the purpose to simulate single events. Since

FFs occur during rare heavy, and therefore extreme, rainfall-runoff events RGM-PRO is set up as an event-based model. This means, that the hydrological model simulates the runoff of a stream only for a predefined short period of time. These models are in contrast to the continuous models, which continuously model the runoff and are, consequently, not well suited for capturing single extreme events (Addor et al., 2011). The RGM-PRO is, furthermore, a process-based model. Process-based models rely on the assumptions made by the modeler and their implementation into a hydrological model. The implementation of the hydrologist's assumptions is made by using continuity equations, which contain different parameters. These parameters can either be measured in the field, conscientiously adjusted by calibration, or, as in the hydrological model RGM-PRO, empirically predefined. Additionally, RGM-PRO is able to generate multiple runoff simulation by using not only one parameter value but rather by using parameter values, which are slightly perturbed for each simulation in a predefined range (Antonetti et al., 2017). Each simulation run represents one member of the hydrological model, and can be seen as one possible scenario. This ensemble or probabilistic approach stands in opposition to a deterministic approach, which provides only one single simulation.

Countless studies in recent years have used the probabilistic approach for their simulations. Liechti et al., 2013a, 2013b; Zappa et al., 2008, 2013 are only a few relevant studies using the ensemble approach. The reason why the ensemble approach is so popular is quite simple. The spread of the different members represents the uncertainty in the forecast and thus gives an estimation about the probability, whereas this is not the case by using a deterministic approach. A disadvantage of using the ensemble approach in a process-based runoff generation model are the high computation costs, which are needed to run the model for each member. In addition, the highly complex processes which have to be calculated by the model, likewise increase the computation time of the model (Addor et al., 2011). In the past, it could be shown that the level of complexity in the hydrological model can be decreased without losing significant performance. This more conceptual approach, which is also used by RGM-PRO, has the advantage that it requires less computation time but yet remains reliable (Viviroli et al., 2009). Furthermore, it is worth mentioning that computation efficiency and hydrological understanding has continuously improved, allowing the use of highly sophisticated models within useful computation time (Beven, 2012).

1.2.4.4. Communication

Due to these various progresses in the field of rainfall-runoff predictions, nowcasting has proven to be not only an interesting scientific field for research, but also an early warning system with great potential. (Berenguer et al., 2005). However, the best FFs warning system

is useless if decision-makers and the affected people are not informed correctly. Early FFs warning systems which are capable of giving information not only about a threshold exceedance, but also about more specific information such as the potential risk for infrastructure or possible mitigation methods are necessary for a FFs warning system that operates holistically (Ahmad and Simonovic, 2006). It is, thus, not surprising that plenty of findings have been collected until today on how to communicate results from experts to end-users in the most efficient and useful way possible (Romang et al., 2011). Nowcasting is, therefore, just one tool incorporated into the interdisciplinary field of research of FFs warning systems (Zappa et al., 2011).

1.3. Objectives

The review of literature shows that many components are necessary to consider for a useful FFs nowcast. To deal with all the components would be an impossible task, especially for a master thesis, since each individual component has its own scientific field. It is, however, important to know which components belong to the extensive discipline of FFs nowcasting. Only by understanding the properties and functions of the individual components and their interrelations, it is possible to draw conclusions about the NPWs and to understand and analyse the complexity of the discipline of FFs nowcasting. The schematic illustration in figure 2 summarises the individual components that are part of FFs nowcasting and forecasting. The rectangles represent the products included in FFs nowcasting and forecasting, while the ellipses represent the processes. The schematic illustration clarifies the complexity of FFs nowcasting and forecasting. However, the focus of this thesis will mainly be on the so-called flash flood nowcasting chains (subsequently referred to as nowcasting chains). These nowcasting chains combine the meteorological model with the hydrological model (Zappa et al., 2011).

In this thesis, the nowcasting chains incorporate (i) the up-to-date meteorological initial conditions for the nowcast in the form of the rainfall measured by radar and/or rain gauges, (ii) a probabilistic or deterministic nowcast product, which extrapolates the rainfall into the very short-term prediction providing a PQPN or QPN respectively, (iii) a NWP, producing a PQPN and a QPN as well as a PQPF and a QPF in respect to longer lead times, (iv) a blending scheme, which merges the extrapolation nowcast into the NWP by a predefined weighting function and (v) a probabilistic hydrological model, which provides runoff estimations. The individual components, which belong to the nowcasting chain, are framed in red in figure 2. The runoff as the actual result of the nowcasting chain is marked as a red and black dashed rectangle.

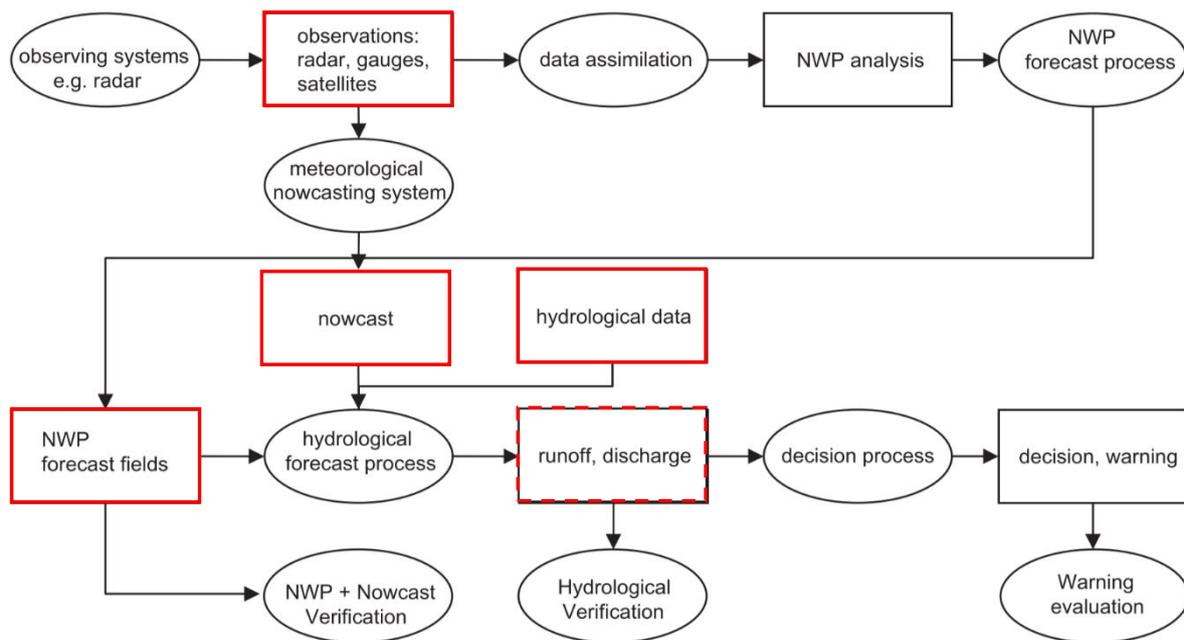


Figure 2: Schematic illustration of products (rectangles) and processes (ellipses) included in FF nowcasting and forecasting. Products of the nowcasting chain are marked in red. The runoff as a result of the nowcasting chain is marked as a red and black dashed rectangle According to Rossa et al. (2011).

In the study by Zappa et al. (2008), the mesoscale alpine programme demonstration of probabilistic hydrological and atmospheric simulation of flood events (MAP D-PHASE) is introduced. In the programme, different flood nowcasting chains were created by combining different meteorological and hydrological models in several catchments. With MAP D-PHASE it could be shown that some nowcasting chains were superior to others and added considerable value to the prediction (Zappa et al., 2008).

Similarly to the MAP D-PHASE study, different nowcasting chains are created and compared in this thesis. The focus lies on the comparison of nowcasting chains, which blend an extrapolation precipitation nowcast into a NWP with nowcasting chains exclusively using a NWP for the prediction of the precipitation. The main research question that runs through the whole thesis is the following:

- < (1) Are nowcasting chains, which blend an extrapolation nowcast into a NWP, superior to nowcasting chains using raw NWP?

The analysis includes both deterministic and probabilistic nowcasting chains. Derived from this circumstance the sub-research question is formulated as:

- < (1.1) Are nowcasting chains, using a blending scheme, superior to nowcasting chains using raw NWP considering only deterministic flood nowcasting chains and likewise considering only probabilistic nowcasting chains?

When analysing whether the nowcasting chains using extrapolation and blending are supe-

rior to nowcasting chains using raw NWP, it has to be checked whether the used nowcasting chains are skilful at all. This results in the sub-research question:

◁ *(1.2) Are the nowcasting chains skilled enough to make useful predictions?*

A last research question concerns the update cycle of initial conditions of nowcasting chains. The initial conditions for a specific event are not only taken from the beginning of the event simulation, but updated frequently. This results in additional runoff simulations, which each have more current initial conditions than the antecedent simulation. This leads to the research question:

◁ *(2) Is it possible to increase the skill of nowcasting chains by updating the initial conditions frequently?*

This introduction is followed in section 2 by the data and methods used to answer the research questions. The following section shows the results, while they will be discussed in section 4. Finally, the most important findings are summarised in section 5.

2. Data and Methods

The data and methods section describes the study catchment as well as the components which are incorporated in the nowcasting chains. Furthermore, the experimental set-up and the verification methods to answer the research questions are explained in detail.

2.1. Study Catchments

Two study catchments were selected to analyze the nowcasting chains. The first and larger study catchment is the Emme catchment, which lies in the prealpine region, for the most part in the canton Bern. The catchment is divided into the subcatchment Trueb (55 km²), which is nested into the subcatchment Ilfis (184 km²), which is in turn nested in the main catchment Emmenmatt (445 km²). The last subcatchment Eggwil (125 km²) is nested in the main catchment Emmenmatt (figure 3). Each subcatchment as well as the main catchment is equipped with a runoff gauge at the outflow. In addition, the Trueb subcatchment has a rain gauge. The elevation of the whole catchment ranges from 638 to 2213 m a.s.l.. Regarding the land use types, the catchment consists of 52 % meadows and 44 % forests. The last small part is settlements with 4 %. The geology of the catchment consists mainly of Flysch and Cretaceous at higher elevations and of Freshwater and Marine Molasse at lower elevations. Additionally, Molasse can be found at lower elevations but only to a small extent (Antonetti and Zappa, 2017).

The second catchment is the Verzasca catchment. The catchment lies in the canton of Ticino in the southern alps of Switzerland and consists of the nested subcatchment Pincascia (44 km²) and the main catchment Verzasca (186 km²), both of which are equipped with a runoff gauge (figure 4). The elevation of the catchment ranges from 490 to 2870 m a.s.l. and the land use types are forest (30 %), shrub (25 %), rocks (20 %) and alpine pastures (20 %). The catchment is hardly affected by anthropogenic activities (Liechti et al., 2013b; Wöhling et al., 2006). The catchment is built of a crystalline gneiss bedrock with sporadically occurring calcareous schists (Bündnerschiefer). Dominant is mainly the lithology of the Pennine units of the central Alps (Horat et al., 2018).

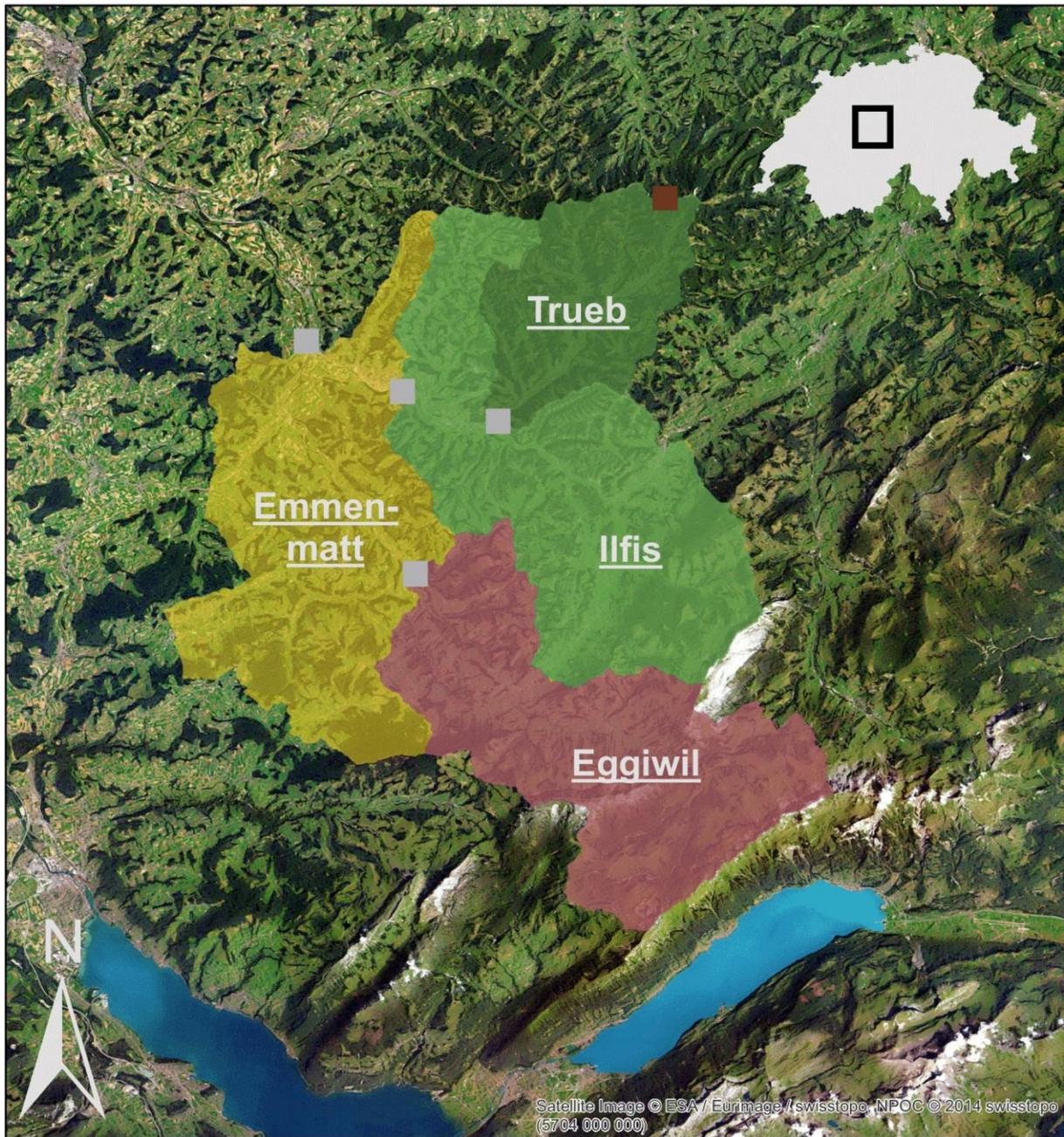


Figure 3: Emme Catchment including the maincatchment Emmenmatt and the subcatchments Eggiwil, Iflis and Trueb (Horat 2017).

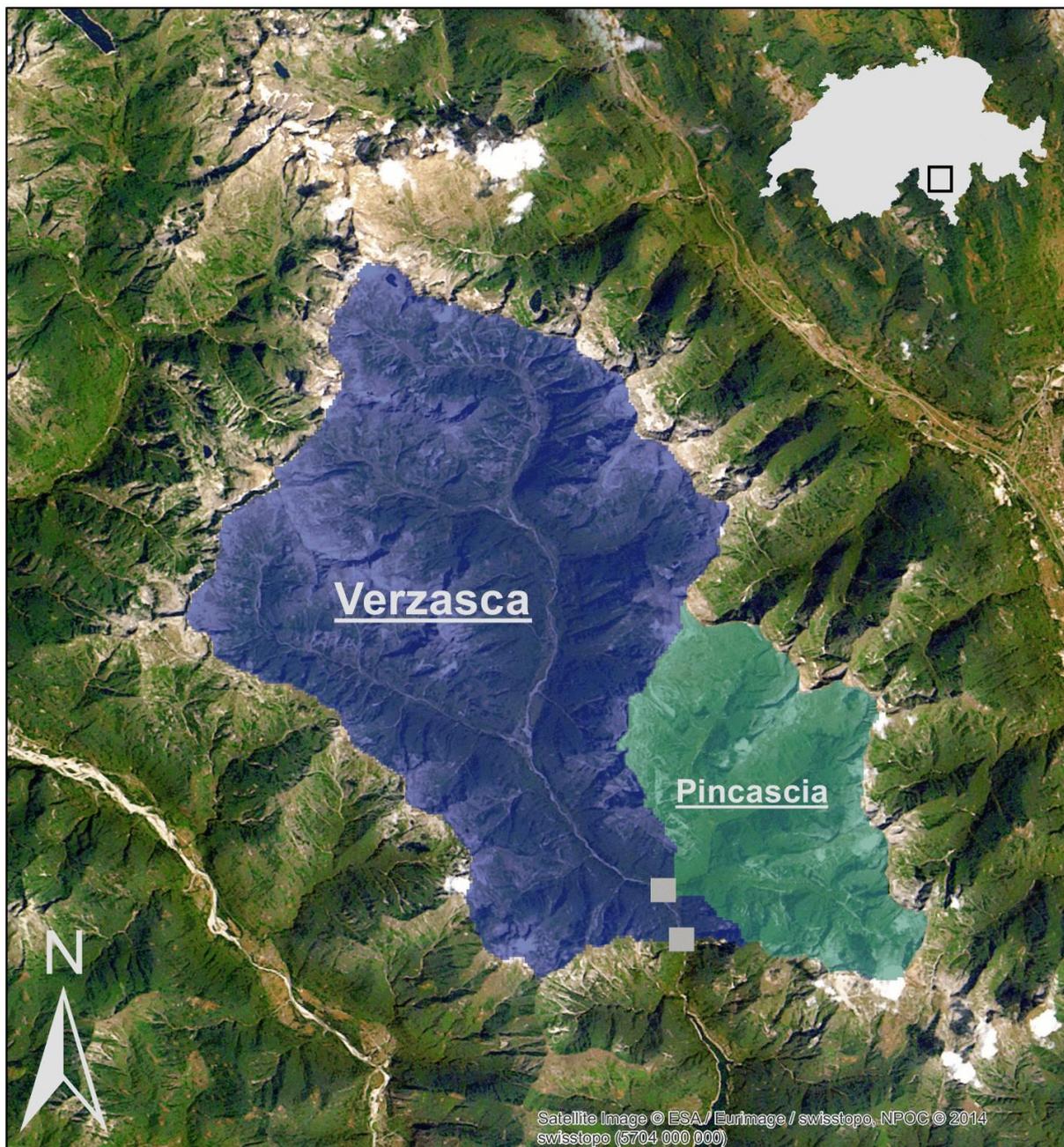


Figure 4: Verzasca Catchment including the maincatchment Verzasca and the subcatchment Pincascia (Horat 2017).

2.2. Hydrological Model

The used hydrological model in this thesis is the process-based runoff generation module (RGM-PRO). This model derives from the hydrological modelling system PREVAH (Precipitation-Runoff-Evapotranspiration HRU Model, where HRU stands for hydrological response unit). Each HRU contains physiographic information and can be clustered into areas with similar hydrological behaviour (Viviroli et al., 2009). RGM-PRO uses similarly so-called maps of runoff types (RTs). The runoff type is determined based on the dominant runoff processes (DRPs). More about the determination of the DRP can be found in Scherrer and Naef (2003). The spatial resolution of the DRP map, and therefore the RT map, is 25x25 m. However, the spatial resolution of RGM-PRO is only 500x500 m. As a result, the percentage of RTs in each RGM-PRO cell is taken into account (Antonetti et al., 2017). Each RT can be assigned to a runoff coefficient. The runoff coefficient describes the relationship between runoff and rainfall and can thus be described in a function curve. The runoff coefficients were determined a priori by sprinkling experiments for each RT (Kienzler and Naef, 2008). Additionally, a parameter range, consisting of the best 1 % Monte Carlo simulations, was assigned to each runoff coefficient curve to take uncertainties into account.

The great advantage that the precalibrated RGM-PRO has over PREVAH is that no classical calibration has to be made using runoff measurements. This allows runoff estimations in ungauged catchments, which are prone to FFs. Soil moisture initial data, which have to be considered in RGM-PRO, are adopted from PREVAH at the same spatial resolution as RGM-PRO (Antonetti et al., 2017).

Finally, RGM-PRO is set up as an ensemble approach. By random perturbation within the uncertainty ranges of the runoff coefficient functions, 21 members are generated. These members, representing different scenarios in runoff, allow to make an assumption about the hydrological model uncertainty.

2.3. Events Selection

The selection of events must be made under consideration of the research question to be answered. To answer the research question whether nowcasting chains that include a blending scheme are superior to nowcasting chains that use a raw NWP only, an optimal set of events would have to be as large as possible in order to capture the diversity of FF events in its complexity. However, when selecting the events, some limitations have to be taken into account. Firstly, FFs only occur during and after heavy precipitation events (Antonetti et al., 2019). Since HPEs are extreme events, they are rare, limiting the choice of events. Secondly, in the

winter season, precipitation in the form of snow does not contribute to runoff as the differentiation between snow and rainfall as well as snow melt leading to runoff is not considered in the nowcasting chains. As a consequence, only events from April to August in the Emme catchment and from April to September in the Verzasca catchment are considered. It is assumed that in these months the precipitation falls in the form of rain. Finally, a third limitation is given by the extensive computer power and the resultant computation time used to run the nowcasting chains. Therefore, only the years 2016 and 2017 are considered.

The event selection is based on the warning system NowPAL (nowcasting of precipitation accumulations) operated by the Swiss national weather service MeteoSwiss. This system accumulates the observed rainfall with the nowcasted rainfall over a given warning region. In case that the accumulated rainfall exceeds a predefined threshold, a possible FF event is given. The accumulation time of the observation, the accumulation time of the nowcast, as well as the threshold are adjustable features (Panziera et al., 2016).

The quantitative precipitation estimation (QPE) is provided by the so-called CombiPrecip (CPC) scheme. CPC combines hourly radar accumulations with rain gauges measurements using geostatistical interpolation techniques to obtain an optimal rainfall estimation (Sideris et al., 2014b). For sub-hourly accumulations, CPC is not available and the radar-only QPE product RZC is used instead as input to NowPAL. The used nowcasting tool for rainfall estimation is the prescribed INCA system (Haiden et al., 2011). The QPE provided by radar observations uses five radar stations, which can be seen as white triangles in figure 5. Additionally, the figure shows the Emme catchment in blue and the Verzasca catchment in red on the Swiss map. The advantage of a radar QPE is that all warning regions are fully covered by the radar network, whereas not all of the warning regions are covered by rain gauges. Further information about the radar precipitation measurement and its uncertainties can be found in Germann et al. (2006a, 2006b).

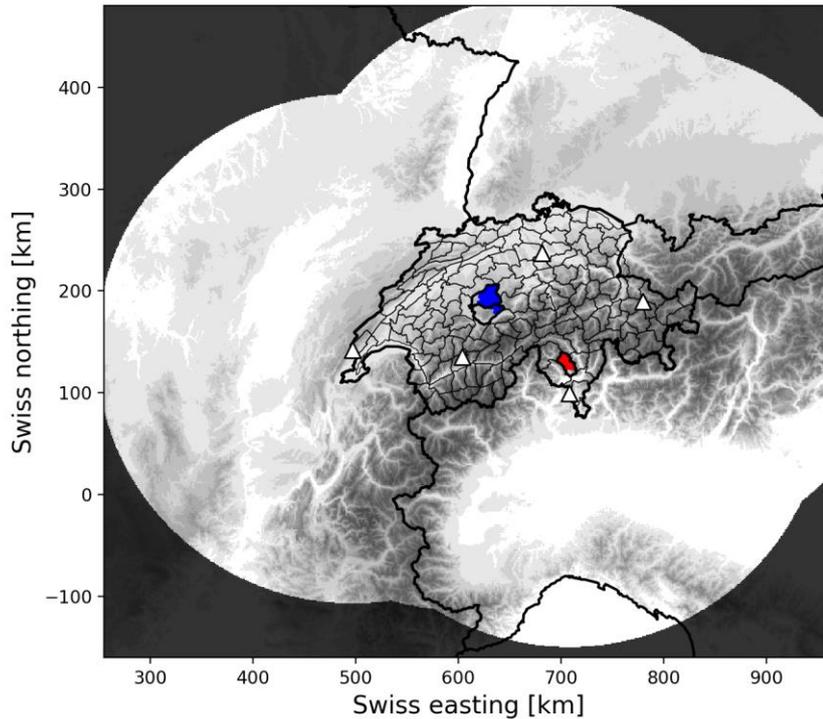


Figure 5: The used radarstations are shown providing radar data for the QPE. Radarstations are shown as white triangles, whereas the Emme catchment is marked in blue and the Verzasca catchment is marked in red.

Two events selection techniques are applied, which are described in the following and summarised in table 1. In a first experimental run CPC120INCA240 was used, which stands for 120 minutes of CPC observation accumulated with 240 minutes of INCA nowcast. The triggering threshold was set to 15 mm accumulated rainfall for the Emme catchment and 20 mm for the Verzasca catchment, which corresponds to the 98 % quantile of the summer rainfall over the last 12 years (Panziera et al., 2016). In both catchments, 10 events were randomly selected to test the calculation time of the nowcasting chain. Taking the computation time as a limitation into consideration, a second event selection run called RZC30INCA60 was launched, which stands for 30 minutes radar observation by RZC, and 60 minutes of INCA nowcast. Thereby, the lead time of NowPAL is reduced, which makes the warning system more applicable to nowcasting. In addition, the triggering threshold in the Emme catchment was increased to 20 mm. This new setting was chosen because the INCA nowcast quickly became inaccurate, and therefore a lead time of 240 minutes was overconfident. Further settings such as the required number of consecutive threshold exceedances is set to four to reduce false alarms. The trigger latency, which is the earliest time to issue a new event was set to 24 hours and the total lead time of the forecast was set to 24 hours after the event. Finally, 41 events for the Emme catchment and 40 events for the Verzasca catchment were chosen both including the 10 CPC120INCA240 events.

Table 1: Characteristics of selection techniques for both catchment regions. The abbreviation Verz. stands for Verzasca the catchment. Further Information about the characteristics can be found in the text.

Characteristics of selection techniques	Obs. Product/ Time [mins]	Nowcast Product/ Time [mins]	Total NowPAL accum. Time [mins]	Triggering Threshold [mm]	Number consecutive exceedances	Tigger Latency [h]	Total forecast Lead time [h]	Selcted Events
CPC120 INCA240	CPC / 120	INCA / 240	360	15 Emme / 20 Verz.	4	24	24	10 Emme 10 Verz.
RZC30 INCA60	RZC / 30	INCA / 60	90	20 Emme / 20 Verz.	4	24	24	31 Emme 30 Verz.

Figure 6 shows the RZC30INCA60 output of NowPAL for the year 2017 in the Emme catchment. Further outputs including the year 2016 and the Verzasca catchment are added in the appendix (figures A.1–A.3). For the performance analysis of the nowcasting chains with frequently updated initial conditions, an event was chosen with an unambiguous peak in runoff to ensure the occurrence of a FF.

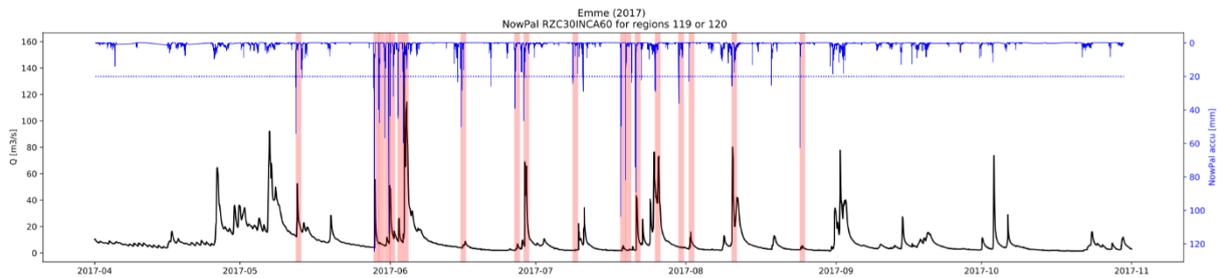


Figure 6: NowPAL RZC30INCA60 output for the year 2017 in the Emme catchment. The blue axis shows the accumulated precipitation by NowPAL. The black axis shows the observed runoff. The blue dotted line represents the threshold triggering an event and the red bars cover the 24-hour forecast.

2.4. Nowcast Extrapolation and Blending

Precipitation input data in form of radar and/or rain gauges observations serve as initial conditions for the probabilistic or deterministic nowcast extrapolation or the raw NWP within the nowcasting chain. The precipitation input is provided by the prescribed CPC.

The nowcast extrapolation is done by Lagrangian persistence, which basically means that the radar images are extrapolated by using the same motion pattern of the precipitation fields as in the antecedent radar images. A detailed description of the Lagrangian persistence can be found in Germann and Zawadzki (2002) The nowcast extrapolation is set up as a probabilistic approach in order to consider the uncertainties. The first member of the ensemble is a unperturbed control member. The remaining 20 members are generated by adding stochastic perturbation to the extrapolation nowcast. This perturbed part of the extrapolation nowcast represents the unpredictable growth and decay processes in the atmosphere, which are not taken into account by the Lagrangian persistence approach. The unpredictable component is

quantified by the autocorrelation coefficient with respect to the predictable part. Thereafter, an auto-regressive model of order 2 is used to generate a nowcast. This leads to a rather small member size of 21 for the probabilistic nowcast extrapolation. More about the stochastic perturbation can be found in the study of Foresti et al., (2016) and Pulkkinen et al. (2019).

The nowcast extrapolation is seamlessly blended with the NWP within 5 hours. The blending is done under consideration of two conditions. Firstly, the blending is proportional to the uncertainty of the nowcast extrapolation and the uncertainty of the NWP, whereas the uncertainty is defined as the spread of the ensemble (i.e. its variance). Differently spoken, the extrapolated nowcast and the NWP are weighted inversely to their uncertainty during the 5 hours nowcast. Secondly in the beginning of the nowcast, the entire weight is assigned to the extrapolation nowcast and, after 5 hours of nowcast, the entire weight is assigned to the NWP. A more detailed description of the blending scheme can be found in Nerini et al. (2019).

2.5. Numerical Weather Prediction

This thesis uses the deterministic numerical model COSMO (Baldauf et al., 2011) as operated by MeteoSwiss since early 2016 for the Alpine region (COSMO-1), as well as its ensemble counterpart, COSMO-E (Klasa et al., 2018). The COSMO-E model incorporates 21 members, whose spread quantifies the uncertainty. Both COSMO models are suitable for medium range forecasts in the alpine region due to their high spatial resolution. The COSMO-1 model has a horizontal spatial resolution of approximately 1.1x1.1 km, while COSMO-E has a spatial resolution of 2.2x2.2 km (MeteoSwiss, 2019). Regarding the temporal resolution, table 2 summarises the forecasts of COSMO-1 and COSMO-E. The time between the forecast initialisation and the forecast availability is the time needed to compute the NWP. The computation times for nowcasting applications is reduced, since there is no need to compute the whole lead time for short-term predictions.

Table 2: Summary of the temporal resolution of COSMO-1 and COSMO-E.

COSMO model	Forecast initialisation [UTC]	Lead time [h]	Forecast available at [UTC]
COSMO-1	00:00	33	01:00
	03:00	45	04:00
	06:00	33	07:00
	09:00	33	10:00
	12:00	33	13:00
	15:00	33	16:00
	18:00	33	19:00
	21:00	33	22:00
COSMO-E	00:00	120	02:00
	12:00	120	14:00

The ensemble members of the COSMO-E model are generated by the so-called stochastically perturbed parametrisation tendencies scheme (SPPT)(Buizza et al., 1999a). The first of the 21 members of the ensemble is an unperturbed control member. Furthermore, the initial and later boundary conditions are perturbed (Klasa et al., 2018).

2.6. Experimental Set-up

Two types of nowcasting chains are set up: nowcasting chains using a blending scheme, which blends the extrapolated nowcast into a NWP, and non-blended nowcasting chains, whose nowcast and forecast is done by a NWP only. The resulting hydrographs are compared to the hydrograph of the pseudo observation nowcasting chain, which consists of a CPC hindcasted QPE, which forces the hydrological model. This CPC QPE is called pseudo observation, because the values are taken from the combined product of radar and rain gauge precipitation measurements CPC instead of real observation values, which in fact would be available. The reason for this approach is that catchments prone to FFs are often ungauged and no observations are available. This makes the results shown in this thesis more meaningful for operational purposes of FFs nowcasting. Furthermore, the contribution of uncertainties related to the precipitation initial conditions and the hydrological model are excluded to focus on the nowcast and forecast uncertainties only.

In each nowcasting chain the runoff in the initialisation phase is provided by CPC and the soil moisture data from PREVAH (section 2.2). The start of initialisation for each nowcasting chain is at the time when minimum runoff is observed within the last five days before the event. At the time of the event the CPC QPE switches into the nowcast of the different nowcasting chains. Three nowcasting chains called nowcast product COSMO-1 (NP1), nowcast product control member (NPC), and nowcast product COSMO-E (NPE) are seamlessly blended into the NWP within a five-hour blending phase. The nowcast products differ in such a way that NP1 is blended into the NWP COSMO-1 and NPE is blended into the NWP COSMO-E. While the 21 members of NP1 all lead into the deterministic prediction of COSMO-1, the 21 members of NPE lead into the corresponding 21 members of COSMO-E. The NPC is the first and unperturbed member of the NP1 nowcasting chain.

After the blending phase the NWP takes over the forecast for a lead time of 19 hours, which results in a total lead time since the start of the event of 24 hours. Additionally, two nowcasting chains are set up as raw NWP without using an extrapolation or a blending scheme for the same 24 hours. These are the mentioned COSMO-1 and COSMO-E NWPs (section 2.5). These nowcasting chains are called CO1 and COE respectively. The resulting

QPF from each nowcasting chain forces the hydrological model RGM-PRO. The hydrographs of each nowcasting chain are compared to the hydrograph from RGM-PRO forced by the pseudo observation QPE from the CPC hindcast. The set-up of all nowcasting chains is shown in figure 7.

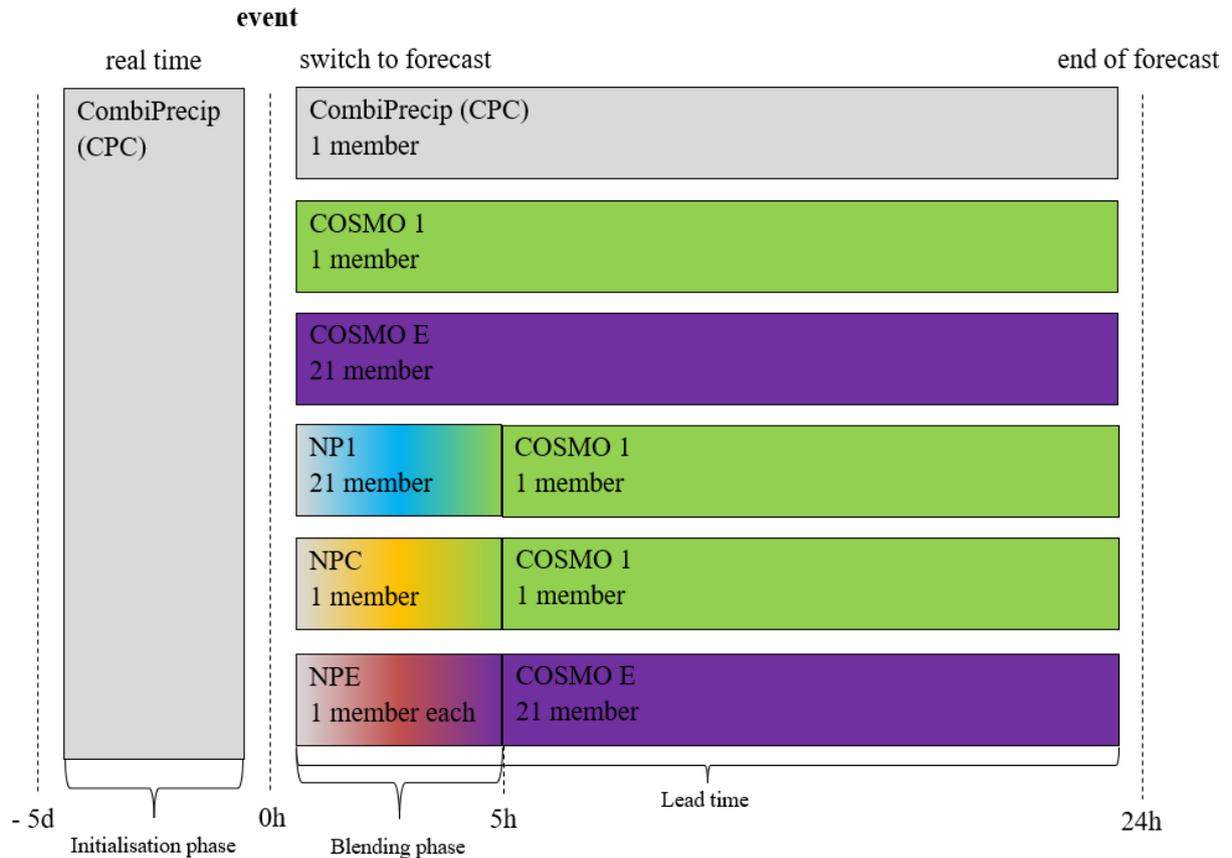


Figure 7: Nowcasting chains providing precipitation forcing the hydrological model RGM-PRO.

As described in section 2.4 the members of the nowcast extrapolation are blended into the NWP according to the uncertainty in the ensembles of the nowcast extrapolation and the NWP. However, NPC and COSMO-1 are deterministic predictions without an ensemble spread, and, therefore, without a quantification of their uncertainty. In order to enable blending for NPC and COSMO-1, the spread of NP1 for the NPC and the spread of COSMO-E for COSMO-1 were adopted respectively.

The generated precipitation prediction from the nowcasting chains force the hydrological model RGM-PRO. The 21 members of each PQQF and the 21 members of the hydrological model would result in a 21^2 ensemble in the resulting hydrograph. Except from the fact that such a hydrograph would be rather confusing, the focus of this thesis lies on the uncertainties coming from the nowcasting chains and not from the hydrological model. Therefore, only the median of the 21 members generated by the hydrological model is considered. This leads to 21 simulations for a PQQF and one simulation for the NPC and COSMO-1 deterministic QPF.

2.7. Verification Methods

The hydrographs of the different nowcasting chains are compared with each other in respect of their goodness to fit to the pseudo observation coming from the CPC QPE hindcast. The GOF in the hydrographs of the nowcasting chains are determined based on statistical methods from Wilks, (2006), Nash and Sutcliffe, (1970), Gupta et al., (2009) and Kling et al., (2012). The focus lies on the Kling-Gupta efficiency score (KGE). Furthermore, the Nash-Sutcliffe efficiency score (NSE) as well as the mean absolute error (MAE), will complement the evaluation of the individual nowcasting chains. The components of which the KGE is composed, (i.e. KGE r , KGE β and KGE γ) are also used for evaluation and described in section 2.7.1. This allows conclusions to be drawn about their relative importance to the KGE.

The selection of these skill scores is supported by the so-called heatmap (table 3). The heatmap displays the correlation between the different skill scores, which were computed for each event and for every member of the nowcasting chain. Negligible correlations are present when comparing the skill scores of the KGE, NSE and MAE with each other. It can therefore be assumed that these skill scores will complement each other in a suitable way and provide additional verification of the results. Skill scores like the root mean square error (RMSE), the correlation (r) and the anomaly correlation coefficient ($Anom\ r$) are not taken into account because the RMSE is correlating too much with the MAE and the correlation r and the $Anom\ r$ are correlating too much with the KGE r . Consequently, these skillscores do not add value to the already used skill scores. Each of the used skill score is further described in section 2.7.1.

Table 3: Heatmap showing the correlation between the different skill scores.

Correlation	<i>MAE</i>	<i>RMSE</i>	<i>r</i>	<i>Anom r</i>	<i>NSE</i>	<i>KGE</i>	<i>KGE r</i>	<i>KGE β</i>	<i>KGE γ</i>
MAE	1.00								
RMSE	0.91	1.00							
<i>r</i>	0.03	0.06	1.00						
<i>anom r</i>	-0.02	-0.01	0.58	1.00					
NSE	0.01	0.00	0.03	0.07	1.00				
KGE	0.00	0.00	0.23	0.12	0.24	1.00			
KGE r	0.03	0.06	1.00	0.58	0.03	0.23	1.00		
KGE β	0.15	0.15	0.02	0.13	-0.13	-0.17	0.02	1.00	
KGE γ	-0.01	-0.01	-0.06	-0.09	-0.77	-0.35	-0.06	0.22	1.00

In order to answer the research question of which nowcasting chain performs the best, the skill scores for each nowcasting chain were computed for each event and for each subcatchment. The best performing nowcasting chain receives one point for the corresponding efficiency score in the corresponding subcatchment. In the so-called stacked barplots, the percentage of total points for each nowcasting chain in each subcatchment is shown to assess the overall performance of the nowcasting chains.

According to the same procedure, the best performing nowcasting chain is determined for only deterministic nowcasting chains (i.e. CO1 and NPC) and for only probabilistic nowcasting chains (i.e. COE, NP1 and NPE).

The stacked barplots will only answer the research questions of which nowcasting chain performs the best. Stacked barplots do not take into consideration how good a model actually performs. Therefore, an overview about the performances of all events will be shown in the results section, where a specific event will be analysed in detail. In addition, the initial conditions of this distinct event are then updated every half an hour and the efficiency scores of the more recent predictions are determined. This approach will show whether the efficiency scores of the nowcasting chains can be improved with more current initial conditions. The Taylor diagram explained in section 2.7.2 and the peak-box approach explained in section 2.7.3 will provide additional information on how the runoff simulations of the different nowcasting chains change during this update cycle.

2.7.1. Efficiency Scores

The MAE, NSE and KGE are deterministic efficiency scores. Since some nowcasting chains are probabilistic, the median member is taken as the deterministic runoff simulation. Furthermore, these efficiency scores determine in different ways how much the simulated runoff differs from the pseudo observed runoff. The time span over which the efficiency scores are computed is from the start of the initialisation until 24 hours after the event.

The MAE can be described as the arithmetic average of the absolute difference between the simulated and the pseudo observed runoff. A perfect value for the MAE would therefore be zero (equation 1):

$$- \tag{1}$$

Whereas y_k and o_k represents the k^{th} forecast-observation pair (Wilks, 2006). The NSE describes the improvement of the simulated runoff forecast over a reference runoff forecast. The reference forecast is represented by the discrepancy of the observed runoff to the mean runoff (Nash and Sutcliffe, 1970). Equation 2 describes the NSE:

(2)

Whereas R_t and O_t is the modelled and observed runoff at time t respectively and \bar{R} is the mean runoff (Gupta et al., 2009).

The KGE is calculated from a linear correlation between the simulated and observed runoff (r), a bias component (β), which is calculated by the ratio between the mean simulated and the mean observed runoff and the ratio between the coefficient of variance (CV) (γ), which is calculated by equation 3 (Kling et al., 2012):

$$KGE = \frac{r \cdot \beta}{\gamma} \quad (3)$$

Whereas CV is the dimensionless coefficient of variance, σ and μ are the standard deviation of the runoff and the mean runoff respectively. The indices s and o stand for simulation and observation respectively.

Finally, equation 4 describes the KGE:

(4)

The perfect value for the NSE and KGE is one. A NSE value of zero states that the simulated forecast does not bring any improvement compared to the reference forecast. A negative value of the NSE indicates that the simulated forecast is even worse than the referenced forecast (Gupta et al., 2009; Kling et al., 2012).

2.7.2. Taylor Diagram

Another tool to illustrate the GOF of a simulated runoff from the different nowcasting chains is the Taylor diagram. The Taylor diagram summarises multiple efficiency scores in one diagram. It illustrates the correlation, standard deviation, and root-mean-square error of the simulated runoff in respect to the observed runoff. Contrary to the NSE, no direct statements about the skill of the prediction compared to a reference forecast are made in the Taylor diagram. However, the Taylor diagram quantifies the deviation of the simulated from the observed efficiency score by their distances in the diagram (Taylor, 2001). This is especially useful in the analysis of the update cycle to track the behaviour of the different skill scores with frequently updated initial conditions for all nowcasting chains.

2.7.3. Peak-box

When dealing with the analysis of FFs, the peak-box evaluation approach of an event has to be mentioned. The Peak-Box enables to analyse the two most important characteristics of a FF event, which is firstly the peak runoff and secondly the peak timing of the simulation for a

FF event. In contrast to other evaluation approaches, the Peak-Box is able to depict the uncertainty of the probabilistic simulation. Like the Taylor diagram, the Peak-Box does not give any information about the skill of the forecast with respect to a reference forecast. However, the Peak-Box is a useful tool to capture to which extent a probabilistic nowcasting chain is capable of simulating a FF event in terms of magnitude and timing (Zappa et al., 2013).

The Peak-Box approach is simple. It allows to estimate the best combination in magnitude and timing of the FF event for probabilistic forecasts. In order to do so, the highest peaks of each member of the ensemble is marked. Each peak contains information about magnitude and timing of the corresponding member. Thereafter, boxes for the quantiles in magnitude and timing are drawn. The intersection of the medians of peak magnitude and peak timing is assumed to be the best estimate. Especially for simulations with large ensemble spread, the Peak-Box approach is extremely useful to summarise the information of the individual members, and thus get the best estimate possible (Liechti and Zappa, 2016).

3. Results

The results section is divided into two parts corresponding to the two research questions. The first part will answer the question if the nowcasting chains using extrapolated nowcasts that blend into a NWP (i.e. NP1, NPC or NPE) are superior to the nowcasting chains using only the raw NWP (i.e. CO1 and COE). The second part is concerned with the question of whether constantly updated initial conditions can increase the efficiency skill and thus the reliability of the predictions.

3.1. Comparison of Nowcasting Chains

The differences in the runoff output of the different nowcasting chains arise from the different precipitation nowcasting schemes. Figure 8 and 9 show the precipitation accumulation from the COSMO-1 related (i.e. CO1, NP1, NPC) and the COSMO-E related (i.e. COE, NPE) nowcasting chains respectively. This precipitation accumulation, forcing the hydrological model RGM-PRO of the different nowcasting chains, is shown for the first event in the Emme catchment. The figures depict the initial phase, which is in this case 24 hours, and the nowcasting phase of 5 hours. The precipitation is accumulated for each hour and shown stepwise.

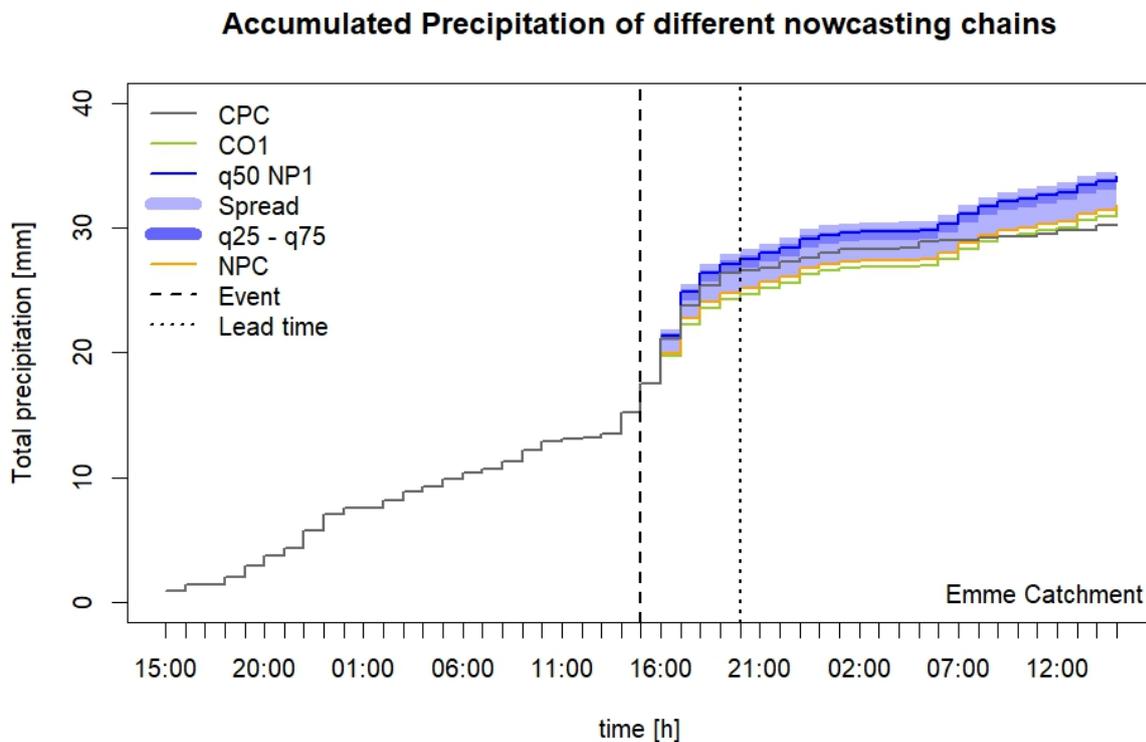


Figure 8: Accumulated precipitation of COSMO-1 related nowcasting chains.

Accumulated Precipitation of different nowcasting chains

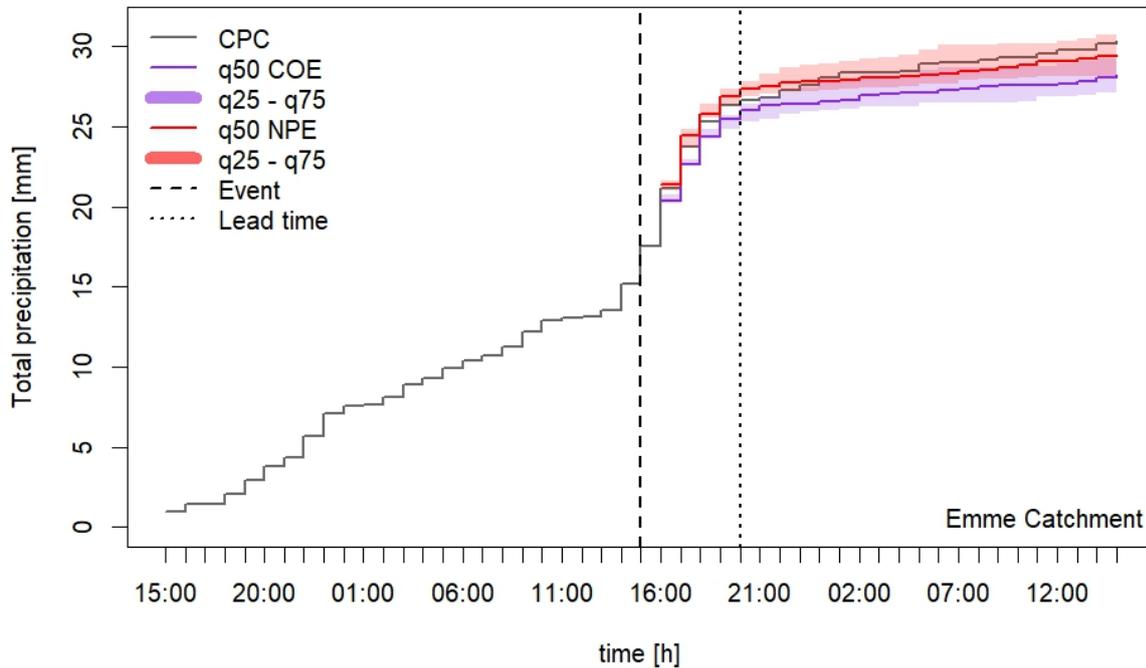


Figure 9: Accumulated precipitation of COSMO-E related nowcasting chains.

The figures show how the different nowcasting chains lead to different amounts in accumulated precipitation, and how probabilistic nowcasting chains are capable of including uncertainty in the form of spread or the interquantile range of the ensemble. After the nowcast, the NWP continues the accumulated precipitation forecast. The deterministic COSMO-1 forecasts of the different nowcasting chains are parallel to each other but differ in their total amount of accumulated precipitation due to their different antecedent nowcasts. The probabilistic COSMO-E forecasts furthermore differ in their total amount of accumulated precipitation due to their different antecedent nowcasts, but do not behave exactly the same, because the medians of the spread in the forecast are slightly different due to the different total amount of spread. Considering the spread and the interquantile range of the probabilistic nowcasting chains, it is visible that both increase with increasing lead time.

Furthermore, the figures depict that in this event the NP1 and the NPE nowcasting chains overestimate and the NPC as well as CO1 nowcasting chains first underestimate then overestimate precipitation. Only the COE nowcasting chain constantly underestimates the precipitation. In this event it can be seen that the precipitation simulations of the nowcasting chains are all very close to the CPC hindcast and the uncertainties are small. This indicates a reliable precipitation prediction of all nowcasting chains, which in terms provide the basis for reliable runoff estimations.

The precipitation input from the different nowcasting chains forces the hydrological model. The output is a hydrograph showing the estimated runoff for each nowcasting chain as well as the runoff based on the precipitation estimation from the CPC hindcast representing the pseudo observation.

In the figures 10, 11, 12 and 13, the hydrological model output forced by the precipitation input, which is shown in figure 8 and 9, are represented in a hydrograph for the Trueeb, Ilfis, Eggiwil and Emmenmatt catchment respectively. The four catchments show with increasing catchment size increasing amounts of runoff. The hydrographs demonstrate that, based on the reliable precipitation predictions of all nowcasting chains, the nowcasting chains in the hydrograph are likewise close to the pseudo observation and therefore provide a reliable prediction. As has already been mentioned before, the CPC hindcast is used as pseudo observation and is, as a result, identical to the observation shown in the hydrographs.

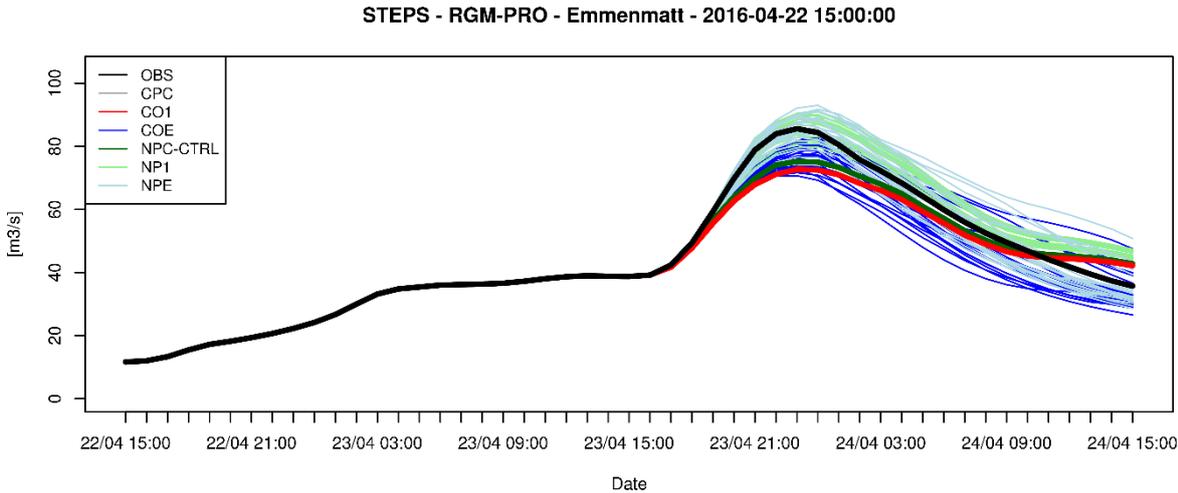


Figure 10: Hydrograph showing all nowcasting chains for the Emmenmatt catchment. OBS and CPC are identical.

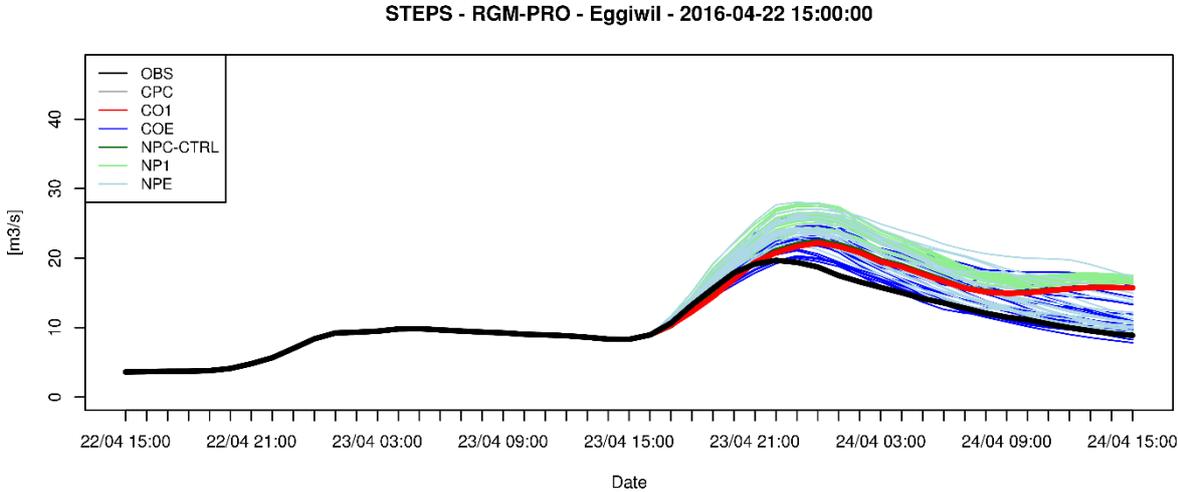


Figure 11: Hydrograph showing all nowcasting chains for the Eggiwil catchment. OBS and CPC are identical.

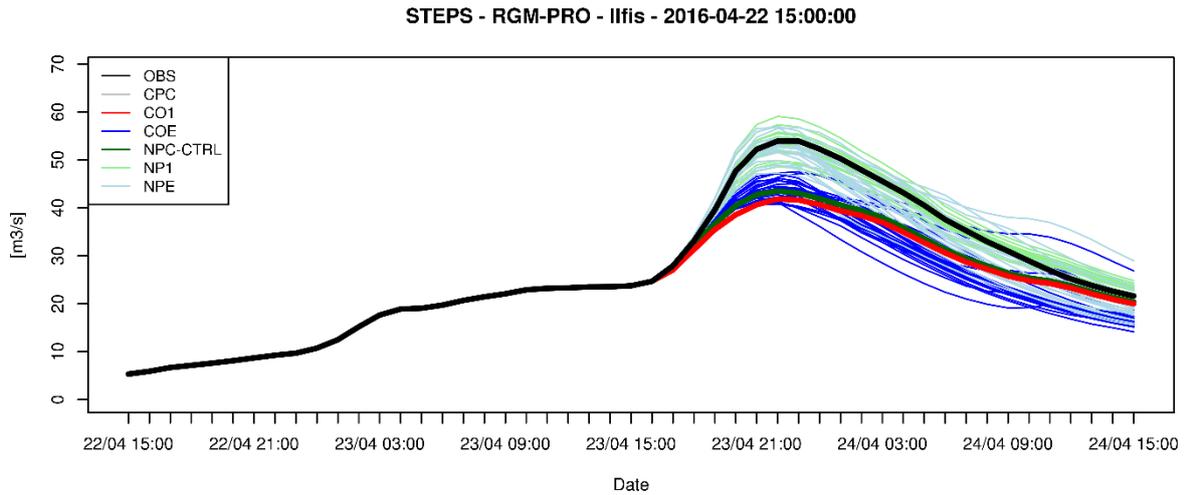


Figure 12: Hydrograph showing all nowcasting chains for the Eggiwil catchment. OBS and CPC are identical.

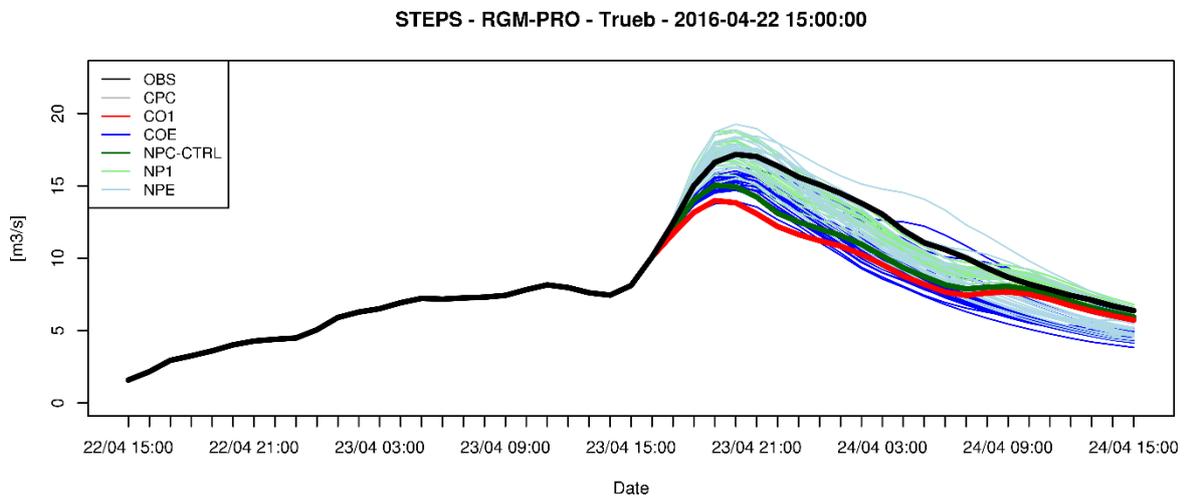


Figure 13: Hydrograph showing all nowcasting chains for the Eggiwil catchment. OBS and CPC are identical.

The skill scores for the different nowcasting chains can be computed with respect to the CPC pseudo observation hydrograph. The KGE values for each catchment are shown as boxplots in figure 14. As expected, the KGE values are high. Additionally, this event shows that the nowcasting chains, which are blended into the NWP, mostly show better values than the raw NWP. Especially NP1 and NPE outperform COE and CO1. Only in the Eggiwil catchment, the raw NWPs are slightly better than the blended nowcasts. However, the reliability of the prediction in the Eggiwil catchment is lower than in the other catchments, shown by the overall lower KGE values. In addition, the range of the KGE values is largest for all nowcasting chains also indicating a lower reliability. This is caused by the range in the KGE values stemming from the spread of the ensemble precipitation prediction. The ensemble spread is, in turn, an estimation of uncertainty in real time applications.

This first impression of an event analysis in the Emme catchment shows already the potential of blended nowcasting chains. It becomes clear that the performance of the different

nowcasting chains in runoff prediction depends on the precipitation prediction serving as input for the hydrological model. Furthermore, differences in accumulated precipitation of the nowcasting chains, which arise in the blending phase, are still visible in the later forecast. It should be mentioned again that uncertainties originating from the hydrological model are not considered.

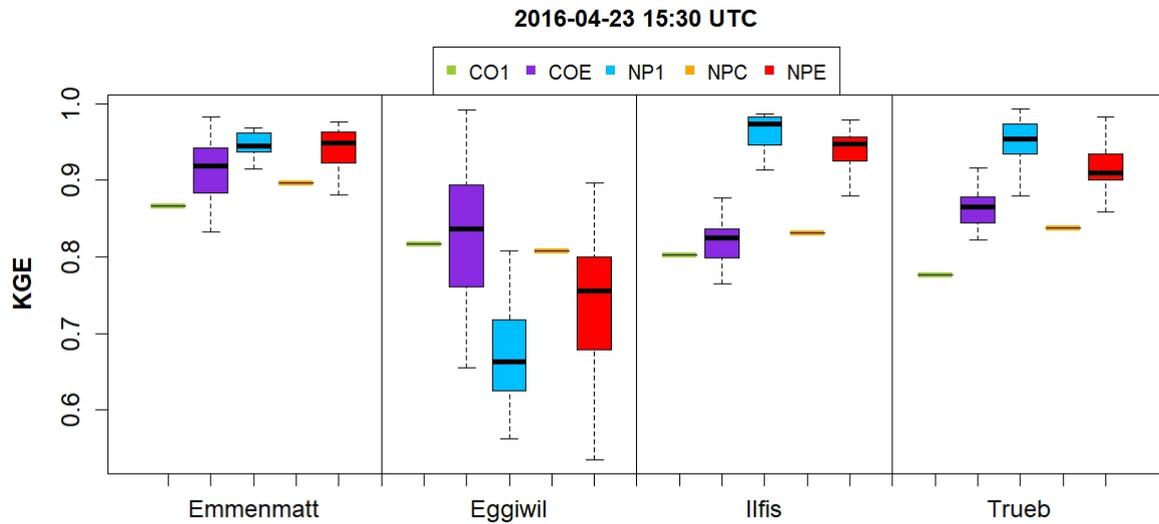


Figure 14: KGE boxplot showing all nowcasting chains for the first event in the Emme catchments. The bold horizontal bar of the boxplot represents the median, while the box represents the interquartile range. The whiskers display the range of extreme values.

3.1.1. Performance Overview

In the same way as shown in section 3.1, the skill scores for the remaining events in the Emme and Verzasca catchment are computed. Figures 15 and 16 summarise the KGE values of all investigated events for the Emme and Verzasca catchment respectively. The KGE shows how the individual nowcasting chains perform in the 41 Emme and 40 Verzasca events. The Y-axis is limited to a value of minus one. The time span used for initialisation is shown as spinup time in hours before the event. The spinup time must be taken into account, because the skill scores are computed regarding the entire hydrograph including the spinup time. Due to the fact that during the spinup time the CPC hincast and the nowcasting chains are identical, the skill score values are always optimal. However, as can be seen in figures 15 and 16, the spinup time does not noticeably influence the overall skill score. This means that events with a long spinup time can have low skill scores, as well as events with a short spinup time can achieve high skill scores. Additionally, in terms of defining the best nowcasting chain per event, the spinup time is the same for each nowcasting chain. Nevertheless, it should be noted that the skill scores are positively biased.

The corresponding performance overviews of the NSE and the MAE for the events are shown in the appendixes (figures A.4–A.7). There are very few simulation results missing in the Trueb catchment. However, due to the high number of events, this circumstance will not considerably influence the analysis.

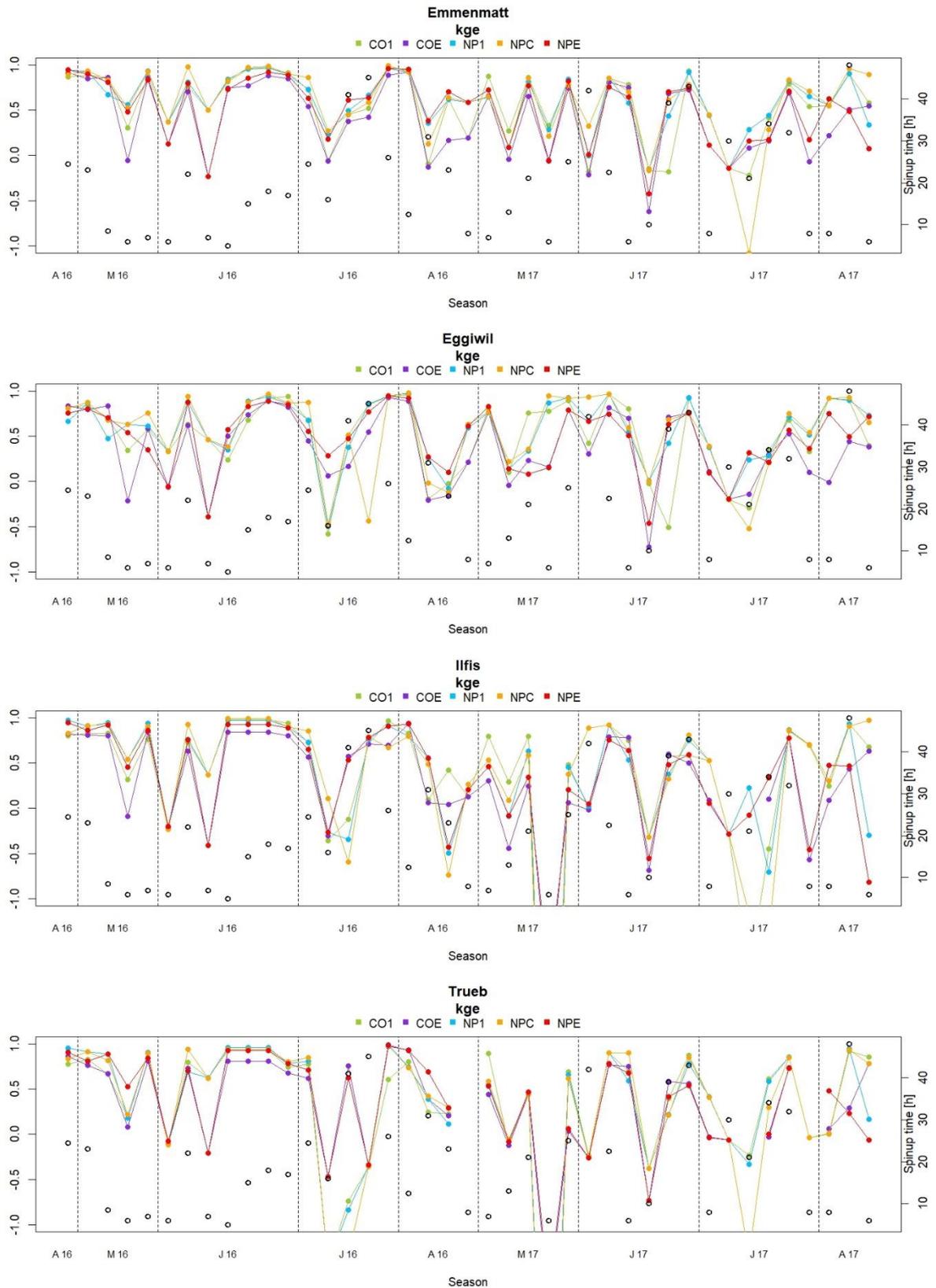


Figure 15: Overview of the KGE medians for all events in the Emme catchments including spinup time. Letters and numbers in the X-axis representing months and years respectively.

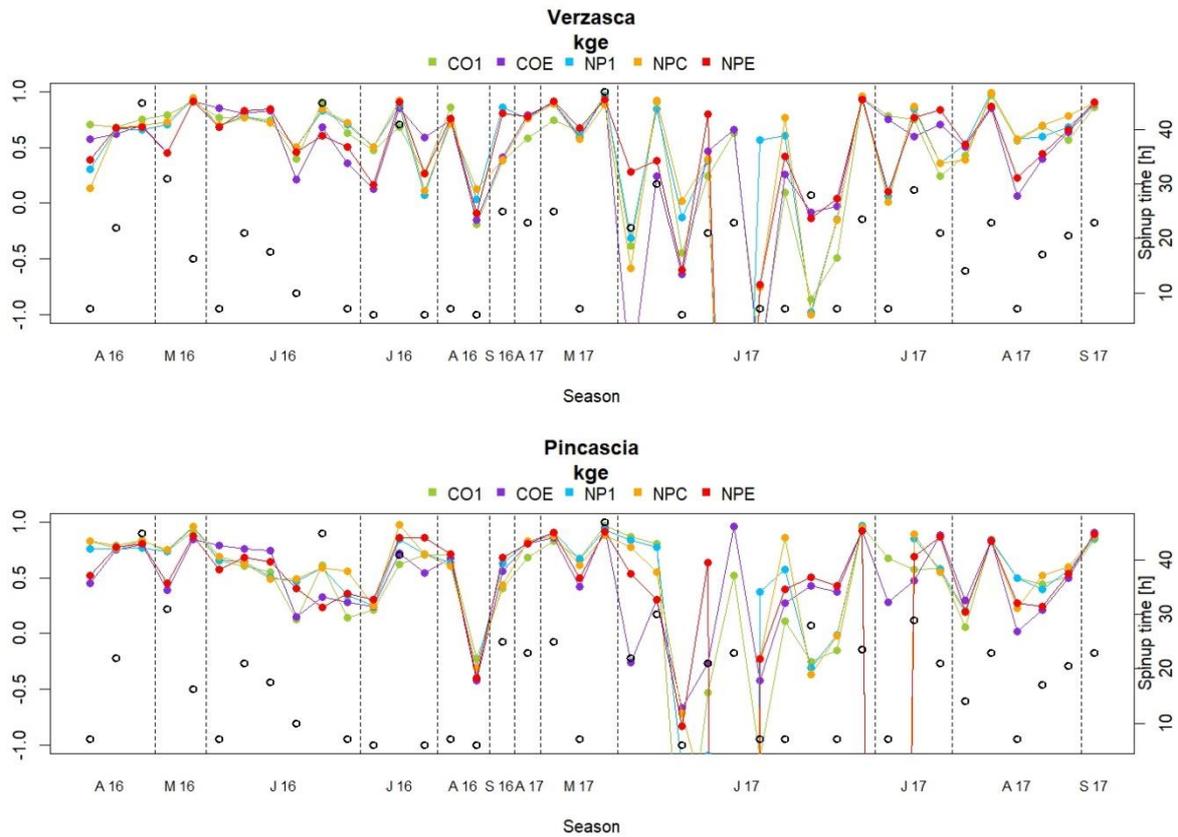


Figure 16: Overview of the KGE medians for all events in the Verzasca catchments including spinup time. Letters and numbers in the X-axis stand for month and years respectively.

Figure 17 shows the KGE values of each nowcasting chain as boxplots for all events in both main catchments Emmenmatt and Verzasca. In the appendix, the boxplots in terms of NSE and MAE are shown including all subcatchments (figures A.8–A.13). The boxplot illustration captures, in addition to the median KGE value, the spread in KGE values stemming from the ensemble members, in probabilistic nowcasting chains, each having a single KGE value. This allows it to analyse the skill score of each nowcasting chain in more detail, because conclusions can be drawn about the performance range of the ensemble members in the runoff prediction. This performance range of the different ensemble members is missing in deterministic nowcasting chains as they have only one member. Having only one member is a crucial disadvantage for deterministic nowcasting chains, since the spread in the precipitation prediction of the probabilistic nowcasting chains is an estimation of the uncertainty already in real time, whereas the prediction uncertainty of deterministic nowcasting chains can only be determined by the KGE value retrospectively.

Figure 17 depicts that the NP1 nowcasting chain in most events shows a smaller performance range in KGE values than the nowcasting chains COE and NPE. Furthermore, it is visible that the KGE values of the different nowcasting chains in the Verzasca catchment are overall lower than the values in the Emmenmatt catchment.

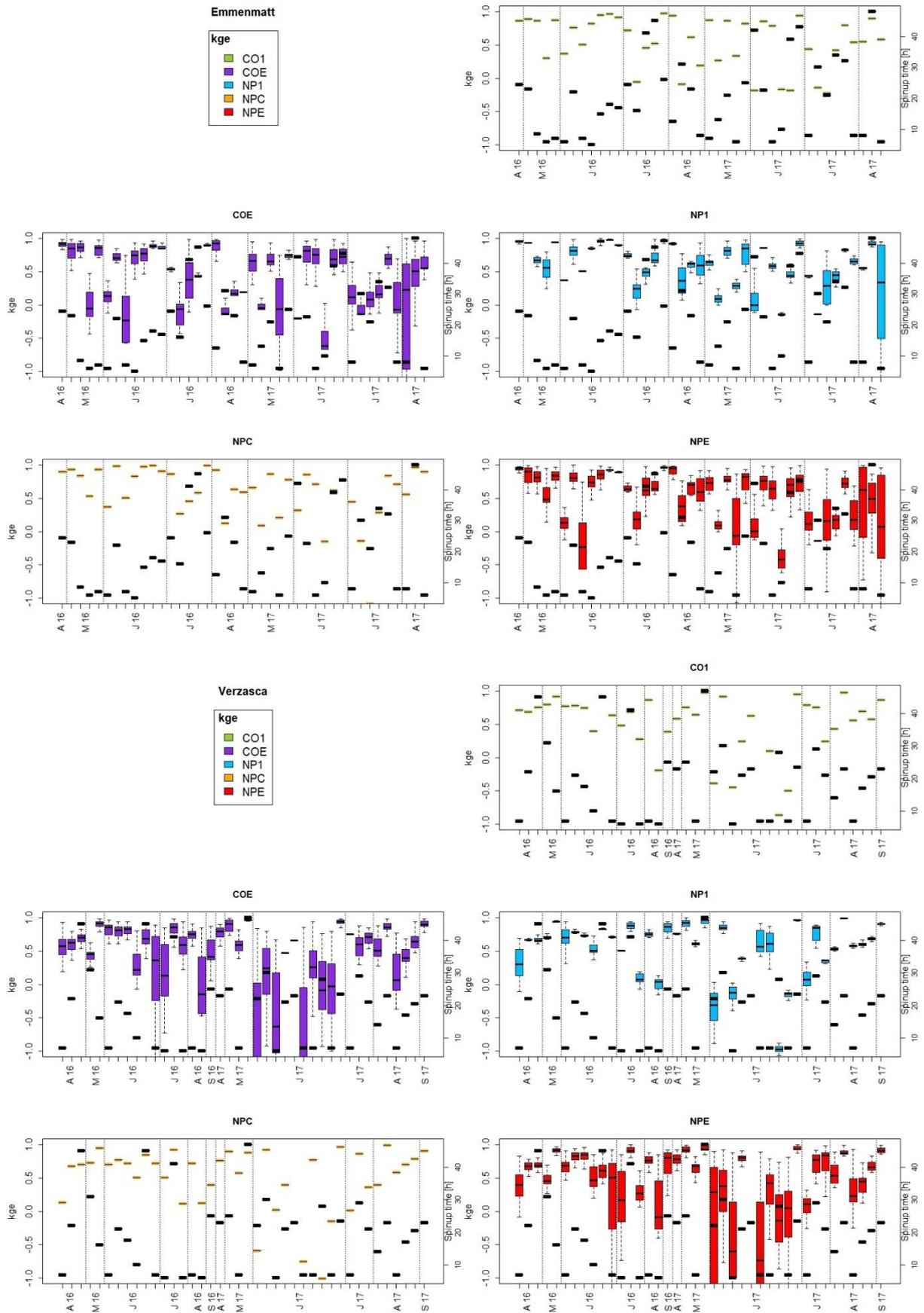


Figure 17: Boxplot overview of the KGE values for all events in the main catchments Emmenmatt and Verzasca including spinup time (bold bar). Letters and numbers in the X-axis representing months and year respectively.

3.1.2. Ranked Events

To find out whether the nowcasting chains even have skill and are, as a consequence, useful for a comparison figure 18 shows all events for both main catchments Emmenmatt and Verzasca ranked by their KGE performance. In the appendix, the ranked events in terms of NSE and MAE are shown including all subcatchments (figures A.14–A.19). The zero KGE value is marked by a horizontal line. All nowcasting chains performing above the zero KGE value line are considered as skilfully predicted events. The barplot shown in figure 19 summarises the number of skilful predicted events for all catchments.

The event-ranked plot in figure 18 underpins the assumption that the overall performance of the nowcasting chains, for the 41 events in the Emme catchment and the 40 events in the Verzasca catchment, is in most cases skilful. Therefore, an analysis to find the best performing nowcasting chain is justified. Moreover, half of the events in the Emmenmatt catchment and more than half of the events in the Verzasca catchment exhibit a KGE value above 0.5. These events are considered as reliably predicted. While the COE nowcasting chain in the Emmenmatt catchment loses a considerable amount of skill after half of the events, the remaining nowcasting chains are able to keep the KGE value high for almost all events. In the Verzasca catchment, the skill decrease is weakly pronounced for all nowcasting chains. However, there are more events which are not skilful in relation to the KGE value.

The barplot in figure 19 depicts that the Emme catchments differ in the number of skilful predicted events. Most events are predicted skilfully by the different nowcasting chains in the Emmenmatt and Eggiwil catchment. Noticeably, fewer events are skilfully predicted by the nowcasting chains in the Ilfis and Trueb catchment. Furthermore, it shows that in the Emme catchment the COE nowcasting chain predicts the least events skilfully. In the Verzasca catchment, on the other hand, no large difference in the amount of skilfully predicted events is visible between the main and the subcatchment. Furthermore, none of the nowcasting chains shows significantly higher or lower amounts of skilfully predicted events.

The large amount of skilfully predicted events in both catchments and of all nowcasting chains indicate that all nowcasting chains have the potential to detect FFs. However, it has to be noticed that the skill scores are positively biased and that errors of the hydrological model are neglected. Moreover, the nowcasting chains are compared to a CPC hindcast pseudo observation instead of real observations. Nevertheless, because of the high percentage in skilfully predicted events, it is assumed that most events deliver useful FFs predictions. Moreover, the comparison between different nowcasting chains or different catchments in terms of skilfully predicted events is unaffected by the aforementioned disadvantages.

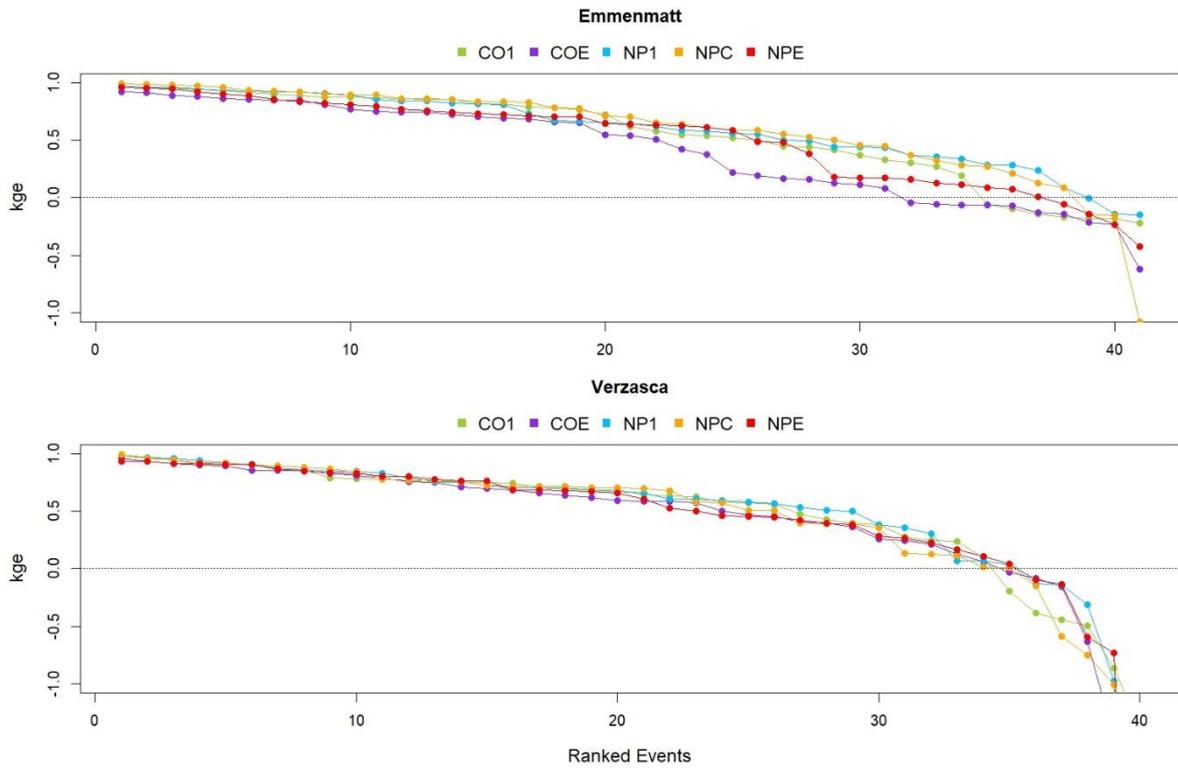


Figure 18: Ranked KGE values shown for all nowcasting chains in the two main catchments Emmenmatt and Verzasca. Values above the horizontal line at a value of zero indicate skilfully predicted events.

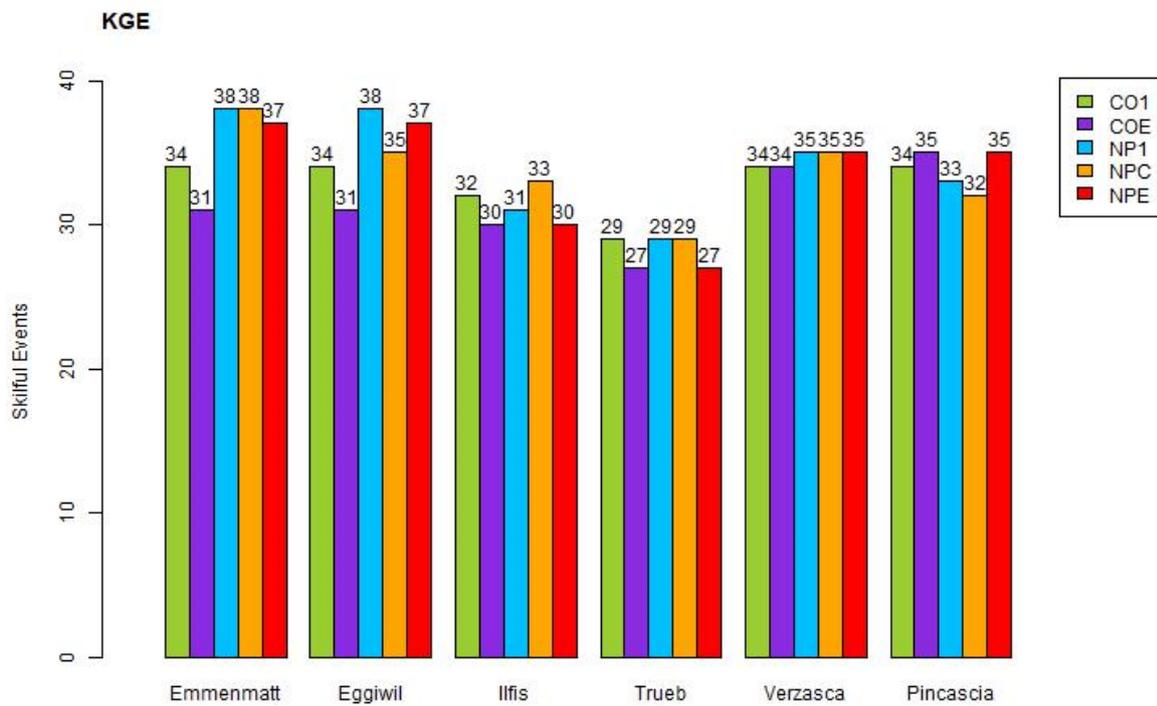


Figure 19: KGE barplot showing the amount of skilfully predicted events for all nowcasting chains in all investigated catchments.

3.1.3. Best Performing Nowcasting Chains

The stacked barplots in the figures 20 and 21 depict the proportion of best performing nowcasting chains in the total number of events for the KGE and MAE respectively. The proportion in best performing nowcasting chains in terms of the NSE is shown in the appendix (figure A.20), since the result is very similar to the result of the best performing nowcasting chain in terms of the KGE.

Figure 20 and 21 show that the ratio of the COSMO-1 related nowcasting chain NPC is the largest in most catchments in terms of KGE and MAE. Only in the small subcatchments Trueb and Pincascia the NPC nowcasting chain is outperformed by the COSMO-E related NPE nowcasting chain. The COSMO-1 related nowcasting chains NP1 and CO1 are not able to take a 40 % proportion in best nowcasting chains in any catchment, neither for the KGE nor the MAE. The raw NWP COE nowcasting chain outperforms its deterministic counterpart CO1 in every catchment except for the small Pincascia catchment.

Overall, nowcasting chains which are blended into the NWP outperform raw NWP nowcasting chains. In the two main catchments Emmenmatt and Verzasca, the blended nowcasting chains are in over 70 % of all events better than the raw NWP in terms of KGE and MAE. In the subcatchments, the superiority of the blended nowcasting chains continues. Even if the superiority in the subcatchments is not so pronounced as in the main catchments, the blended nowcasting chains lead to better results than the raw NWP in more than 60 % of all events in all subcatchments in terms of the KGE and the MAE. If a single nowcasting chain had to emerge as the best performing nowcasting chain, then this would be the NPC nowcasting chain in the Emmenmatt, Eggiwil, Ilfis and Verzasca catchment. In the small catchments Trueb and Pincascia, the NPE nowcasting chain outperforms all the others in terms of both the KGE and the MAE.

In figure 22, the ratios of the best performing nowcasting chain in terms of the three KGE components are shown. It can be seen that, especially in terms of the correlation component, the blended nowcasting chain outperforms the raw NWPs. In the bias component, the blended nowcasting chains mostly outperform the raw NWP. Only in the small catchments Trueb and Pincascia, the blended nowcasting chains have about the same ratio of best performing nowcasting chains as the raw NWPs. Considering the ration between the coefficient of variance, the blended nowcasting chain outperforms the raw NWPs in the Emmenmatt, Eggiwil and Ilfis catchment. Due to the COSMO-1 nowcasting chain, it is possible for the raw NWPs to keep up with the blended nowcasting chains in terms of the ratio between the CV in the Trueb, Verzasca and Pincascia catchment. It can be seen that firstly, the superiority of the

blended nowcasting chains in terms of the KGE is mainly coming from the correlation component of the KGE. Secondly, the COSMO-1 nowcasting chain achieves a high proportion as best performing nowcasting chain in terms of bias and in terms of the ratio between the CVs, especially in small catchments.

For further analyses a differentiation in deterministic and probabilistic nowcasting chains can be done. A more detailed account of deterministic and probabilistic nowcasting chains only is given in the following section 3.1.3.1.

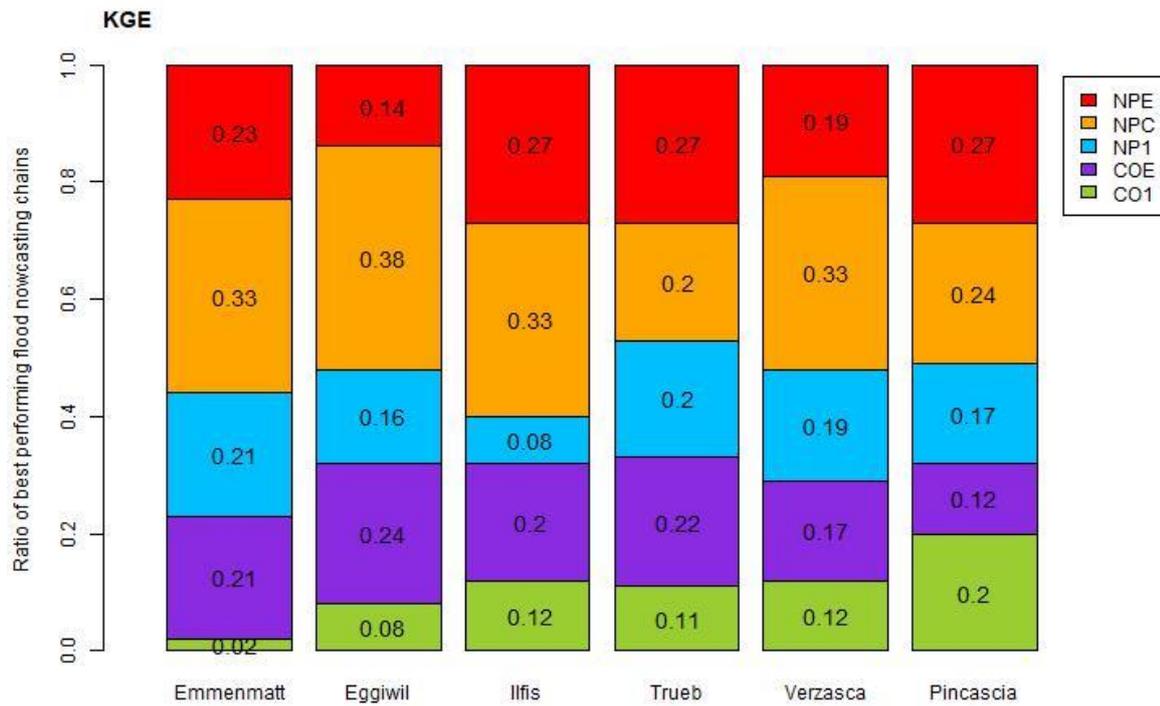


Figure 20: Ratios of the best performing nowcasting chains in terms of the KGE.



Figure 21: Ratios of the best performing nowcasting chains in terms of the MAE.

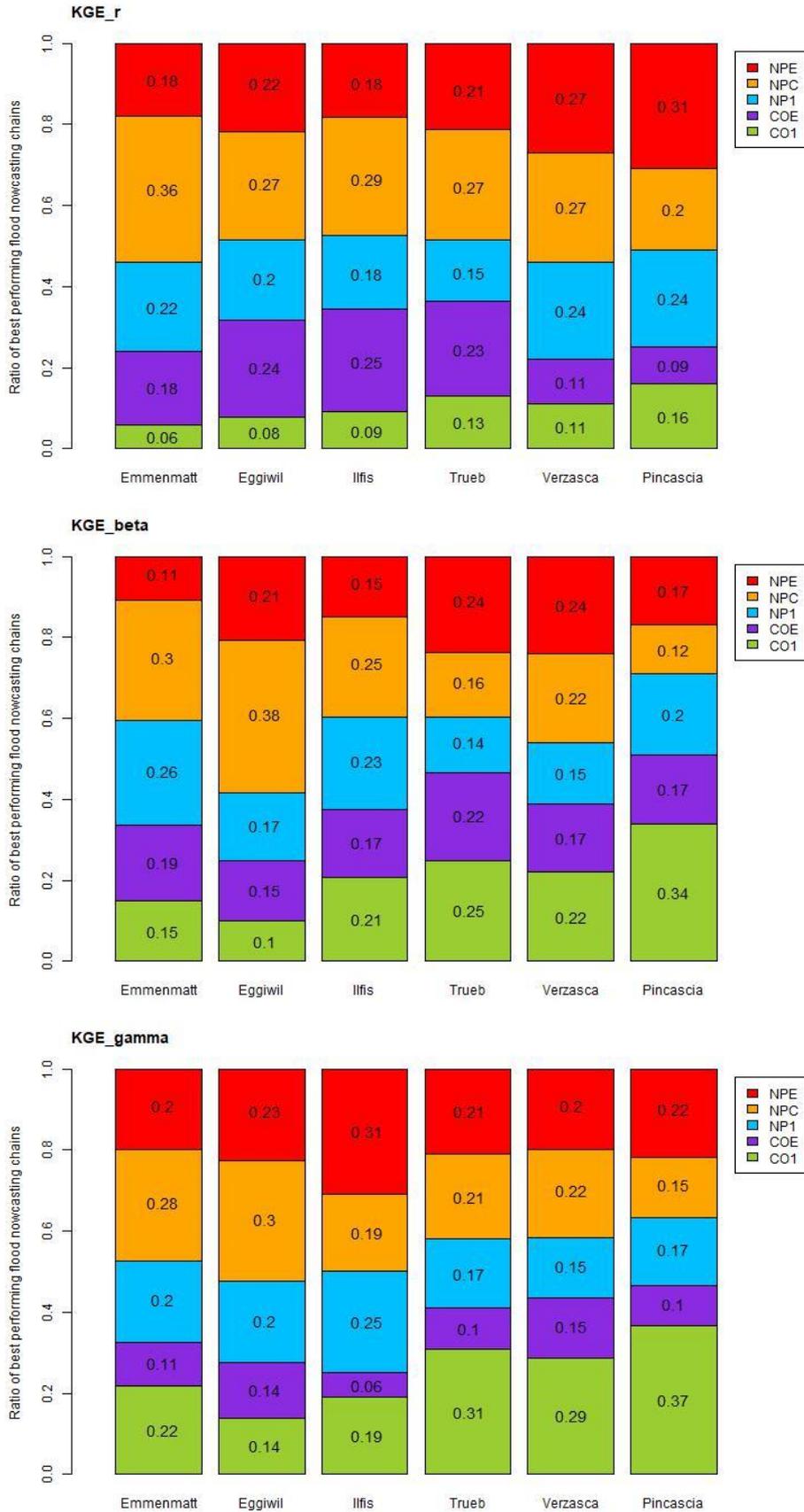


Figure 22: Ratios of the best performing nowcasting chains in terms of the KGE components.

3.1.3.1. Deterministic and Probabilistic Nowcasting Chains

The stacked barplots in the previous section show the ratios of all best performing nowcasting chains. However, figures 23 and 24 show the ratios in terms of the KGE of the best performing nowcasting chains for only deterministic and probabilistic nowcasting chains respectively. Very similar results for the best performing deterministic and probabilistic nowcasting chains are also obtained in terms of NSE, and MAE are shown in the appendix (figures A.21–A.24).

Regarding the best performing deterministic nowcasting chains in figure 23, the NPC nowcasting chain outperforms the CO1 nowcasting chain in all catchments. It should be mentioned, nevertheless, that the CO1 nowcasting chain in the catchments Ilfis, Trueb, Verzasca and Pincascia can almost keep up with the NPC nowcasting chain. This indicates that the CO1 nowcasting chain is often outperformed by probabilistic nowcasting chains, because compared with the figures figure 20 and 21 the proportion of the CO1 nowcasting chain as best performing nowcasting chain increases in figure 23 for each catchment.

The best performing probabilistic nowcasting chain is NP1, as can be seen in figure 24. Neither the NPE nor the COE nowcasting chain is capable of outperforming the NP1 nowcasting chain in any catchment. The reason why the NP1 nowcasting chain, in comparison with all nowcasting chains, does not dominate to the same extent is that it is often outperformed by the deterministic nowcasting chains. Regarding both probabilistic COSMO-E related nowcasting chains NPE and COE, the blended nowcasting chain outperforms the raw NWP COE nowcasting chain in all catchments, except from Ilfis.

Comparing the ratio of the NPE nowcasting chain considering only probabilistic nowcasting chains with the ratio in respect to all nowcasting chains, the NPE nowcasting chain has a higher ratio in the catchments Emmenmatt, Eggiwil, Verzasca and Pincascia. In the Ilfis and Trueb catchment, however, the ratio of the NPE nowcasting chain decreases in comparison to the ratio considering all nowcasting chains. The ratio of the COE nowcasting chain as best performing nowcasting chain increases for the Emme catchment and decreases for the Verzasca catchment considering only probabilistic nowcasting chains.

In the figures figure 23 and 24, as in the previous figures 20 and 21, it can be seen that the nowcasting chains which are blended into the NWP outperform the raw NWP. In consideration of only probabilistic nowcasting chains, the superiority of the blended nowcasting chains is very pronounced. In consideration of only COSMO-E related nowcasting chains in figure 24, the blended nowcasting chain NPE is still superior to the raw NWP COE.

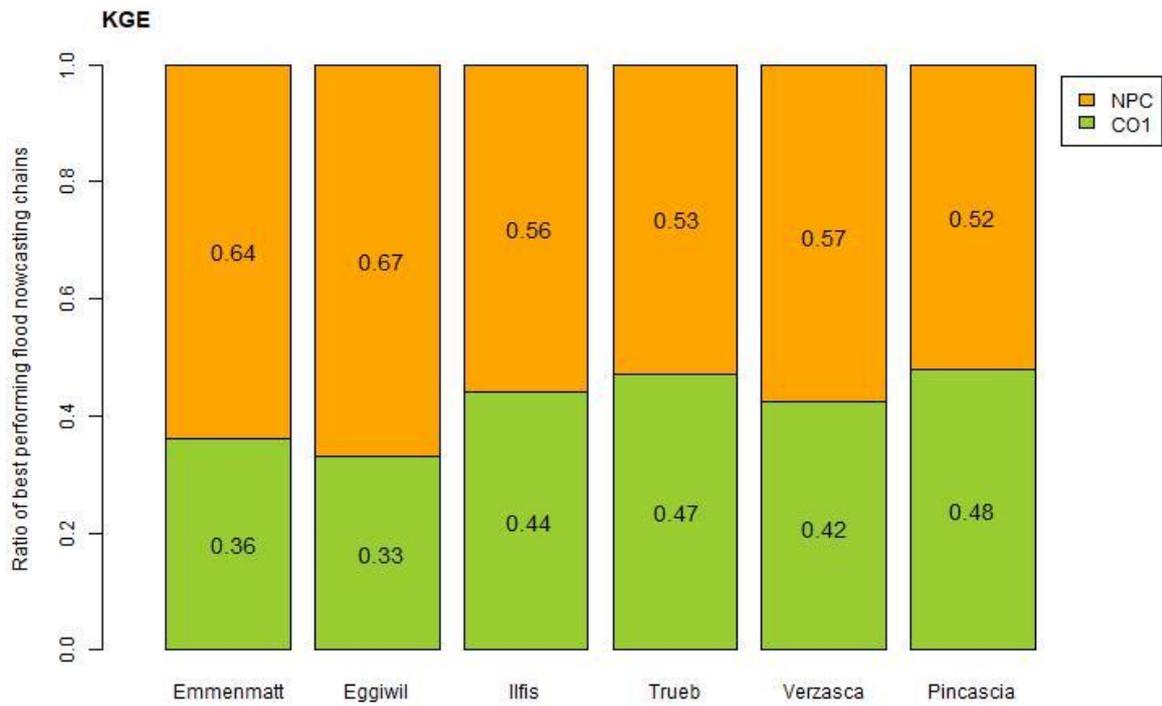


Figure 23: Ratios of the best performing deterministic nowcasting chains in terms of the KGE.

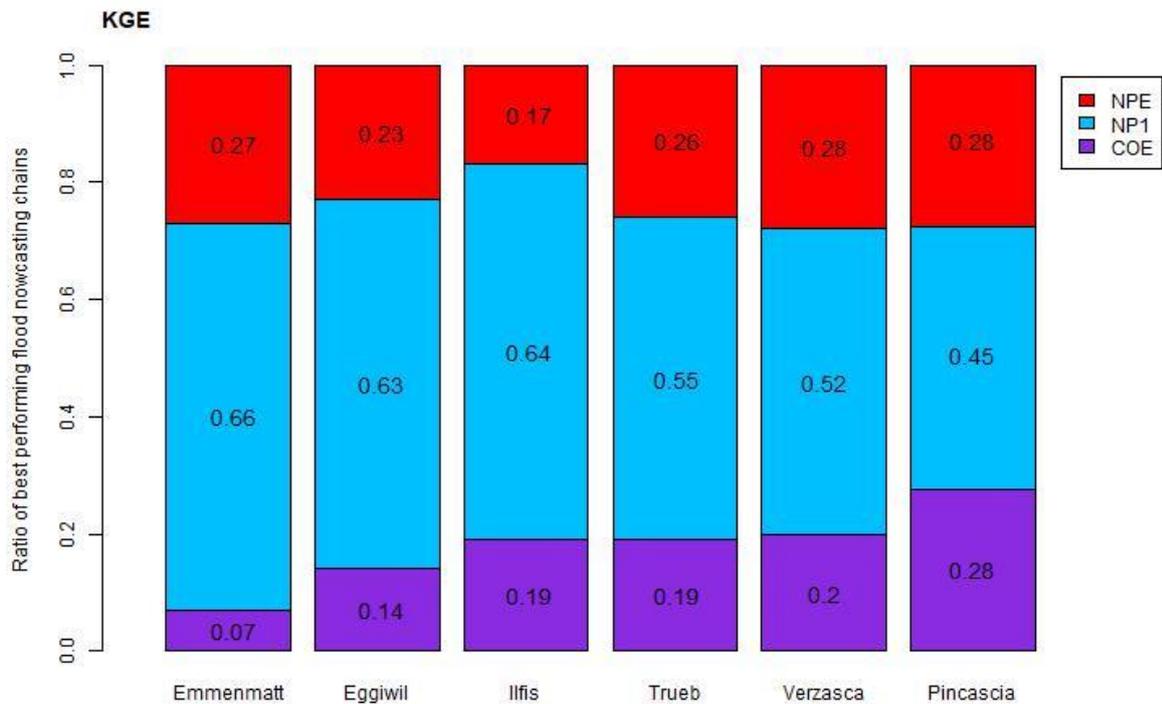


Figure 24: Ratios of the best performing probabilistic nowcasting chains in terms of the KGE.

3.2. Impact of an Update Cycle

The obtained results shown in section 3.1 are computed only once over the time span from initialisation to 24 hours after the event. To analyse the effects of an update cycle, the forecasts are now recomputed every 30 minutes during the whole event. The chosen event takes place in May 2017 during a HPE in the Emme catchment. Figure 25 shows the predicted accumulated precipitation from NowPal and the observed runoff. After the threshold of 20 mm precipitation is exceeded, the first forecast starts at 13:00 UTC.

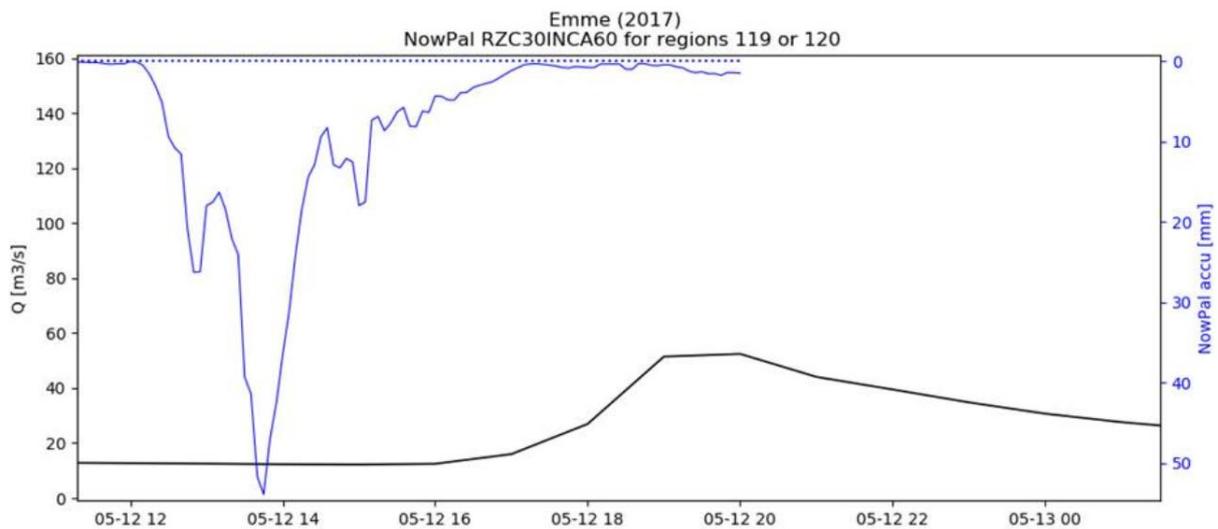


Figure 25: NowPal output showing the predicted accumulated precipitation according to RZC30INCA60 and observed runoff for the investigated heavy precipitation event.

Due to the near real time radar QPEs and nowcast extrapolations, continuously updated runoff predictions are generated every 30 minutes from 13:00 up to and including 19:00. Figure 26 shows the first (13:00), middle (16:00) and last (19:00) precipitation nowcast and continuing forecast for all nowcasting chains, as well as the CPC observation. In addition, the corresponding hydrographs are shown in figure 27 for the main catchment Emmenmatt. The entire update cycle is shown in the appendix (figure A.25 and A.26).

The first precipitation prediction at 13:00 is strongly underestimated by all nowcasting chains. Not even the spread of the COSMO-1 related or the interquantile range of the COSMO-E related nowcasting chains capture the CPC hindcast. This underestimation can be seen in the corresponding hydrograph for the Emmenmatt catchment. Only three members of the COE nowcasting chain overestimate the peak in runoff. These three members approach to the peak in magnitude, but miss it in terms of timing. One COE member captures the peak in terms of timing, but underestimates it in terms of magnitude. Only one member of the NPE nowcasting chain can approximately predict the peak in magnitude, but predicts the peak in runoff too late. Overall the prediction skill of all nowcasting chains is low.

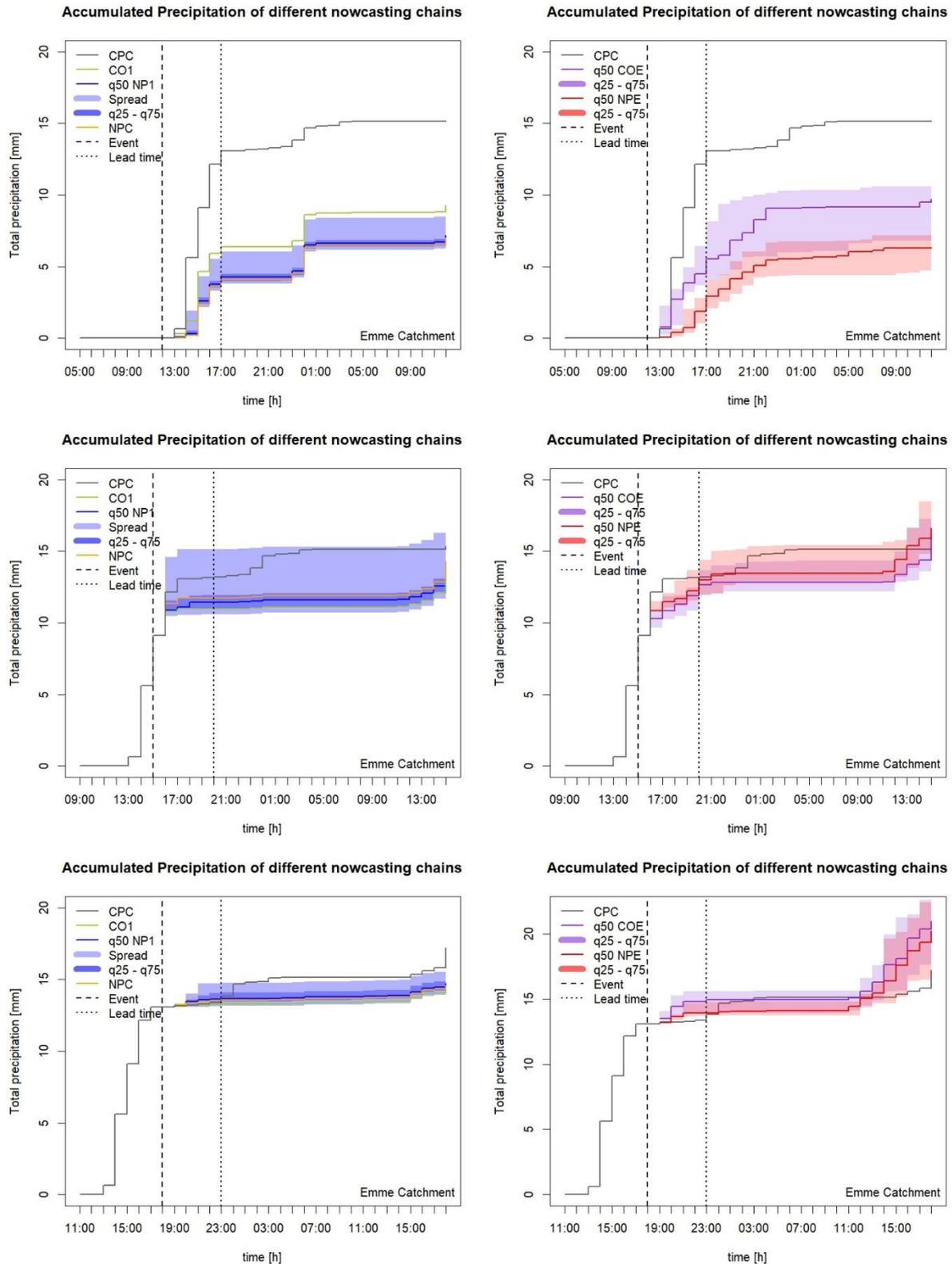


Figure 26: Accumulated precipitation for COSMO-1 related nowcasting chains (left) and COSMO-E related nowcasting chains (right). From top to bottom: The first, middle and last prediction of the update cycle is shown for the Emme catchment.

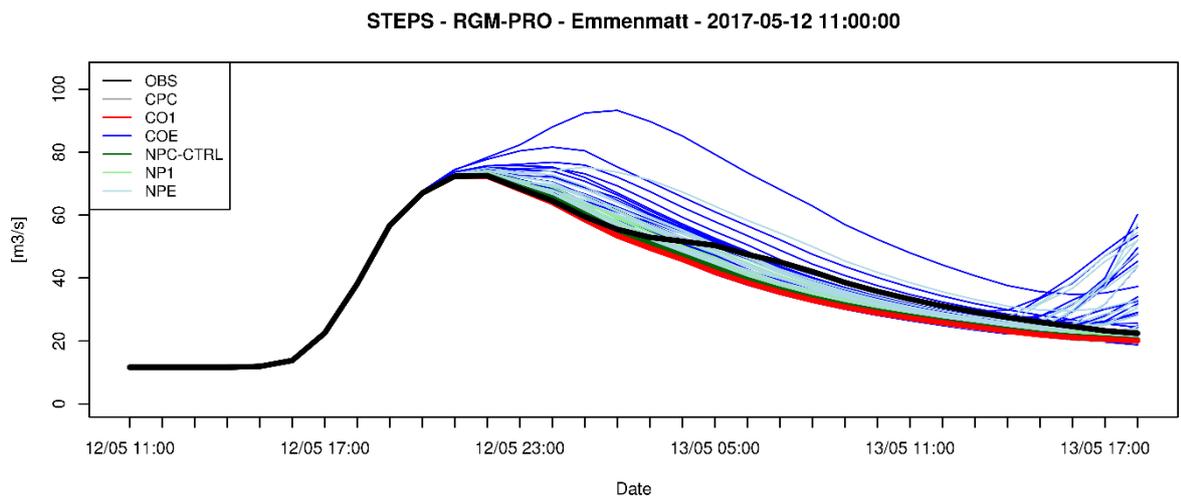
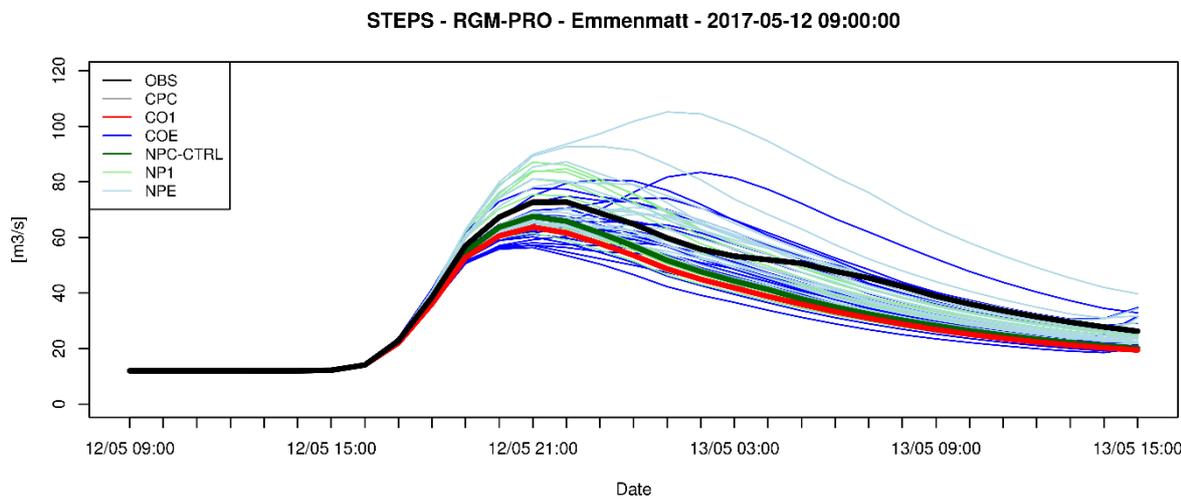
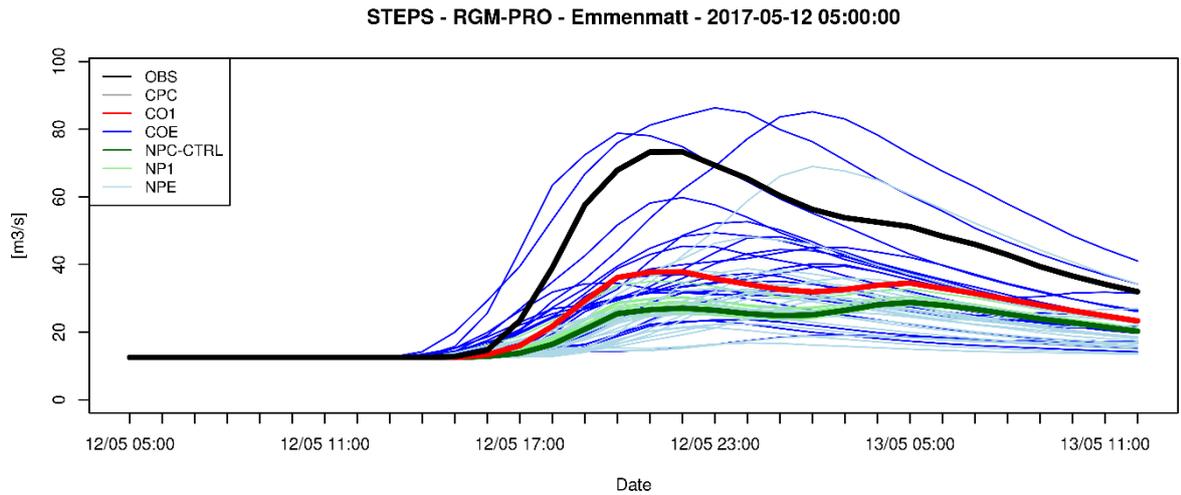


Figure 27: Hydrographs showing runoff predictions for all nowcasting chains in the main catchment Emmenmatt. From top to bottom: The first, middle and last prediction of the update cycle is shown.

The updated precipitation forecast three hours later at 16:00 shows that the accumulated precipitation approaches the CPC pseudo observation. The medians of the COSMO-1 related nowcasting chains still underestimate the amount of precipitation for the entire forecast, while the medians of the COSMO-E related nowcasting chains underestimate the accumulated precipitation in the beginning of the forecast and overestimate it in the end. In the corresponding Emmenmatt hydrograph, most curves of the nowcasting chains are able to predict the timing of the peak in runoff reliable. Considering the magnitude of the peak, the ensemble spread of the probabilistic nowcasting chains is large. Overall, the runoff predictions show a considerable improvement from the 13:00 forecast to the 16:00 forecast.

The last accumulated precipitation at 19:00 of the CPC pseudo observation is simulated similarly by the nowcasting chains. The corresponding hydrograph captures the peak in timing and magnitude of the runoff for all nowcasting chains, but this is because the peak runoff is only about 2 hours after the start of the nowcasting chains' predictions (figure 27), and additionally because the precipitation nowcast is started after the peak rainfall (figure 26). At the very end of the forecast, the COSMO-E related nowcasting chains simulate a new increase in runoff, which does not occur. This is because although the accumulated rainfall increases again, this increase is overestimated by the COSMO-E related nowcasting chains. Taking everything into account, a forecast improvement with updated initial conditions is clearly discernible.

3.2.1. Progression Analysis

Shown in figure 28 and 29 is the update cycle progression of the median values of the different nowcasting chains in terms of the KGE and MAE respectively. The update cycle progression of the median values in terms of the NSE is shown in the appendix (figure A.27). The deviation of the simulated and the observed accumulated precipitation by CPC, shown in figure 26, is now reflected in the value of the KGE. The KGE values for the blended nowcasting chains in the first forecast at 13:00 are low. The raw NWP nowcasting chains achieve higher values in terms of the KGE. In the Ilfis and Trueb catchment, this difference is less pronounced. The same applies in terms of the MAE. In addition, figure 29 shows that the value of the MAE is associated with the catchment size. With increasing catchment size, the MAE increases. The performance of the CO1 nowcasting chain improves in the first hour of the update cycle, while the performance of the COE nowcasting chain stagnates or even worsens in the first hour of the update cycle. In all catchments, the COE nowcasting chain catches up with the others within the second hour of the update cycle. Apart from the Trueb catchment,

the blended nowcasting chains improve considerably after the first hour of the update cycle. The improvement in the Trueb catchment sets in after two hours of update cycle.

Overall, the prediction skill of all nowcasting chains improves remarkably during the first two hours of the update cycle. The KGE values of all nowcasting chains and in all catchments are above 0.5 after this time, except the KGE value of the NP1 nowcasting chain in the Trueb catchment. Throughout the whole update cycle, an improvement in all nowcasting chains is to be found.

Like figure 28, the boxplots in figure 30 show the considerable improvement in the KGE medians for the Emmenmatt catchment within the update cycle, but now shown for each nowcasting chain separately. Additionally, the boxplots depict the range of skill in terms of the KGE within the ensemble of the probabilistic nowcasting chains. The boxplots in terms of the KGE values of the further subcatchments are shown in the appendix (figure A.28). It shows that the range of skill has decreased enormously after the first two hours of the update cycle for all probabilistic nowcasting chains. Especially for the COE nowcasting chain, the range of skill in the first two hours of the update cycle is large. The interquartile range of the KGE values covers a range of more than 0.2 value points of the KGE. The NPE nowcasting chain behaves in a similar way, but the range of skill is rather less pronounced at the beginning compared to the COE nowcasting chain. The same applies for the NP1 nowcasting chain. This range of skill of the KGE values, which evaluate the runoff simulations of the individual members, shows how difficult it is to decide which simulations can be trusted. While some of them achieve high KGE values, and thus provide reliable predictions, other members with low KGE values are useless for predictions. As has already been mentioned before, KGE values can only be determined retrospectively, and are, therefore, no help of deciding which simulations should be trusted in real time application. A sophisticated method of how to deal with an ensemble approach in runoff predictions, and as a result to decide on the occurrence of a FF, is the peak-box approach (Zappa et al., 2013), which is analysed in more detail in the following section 3.2.2.

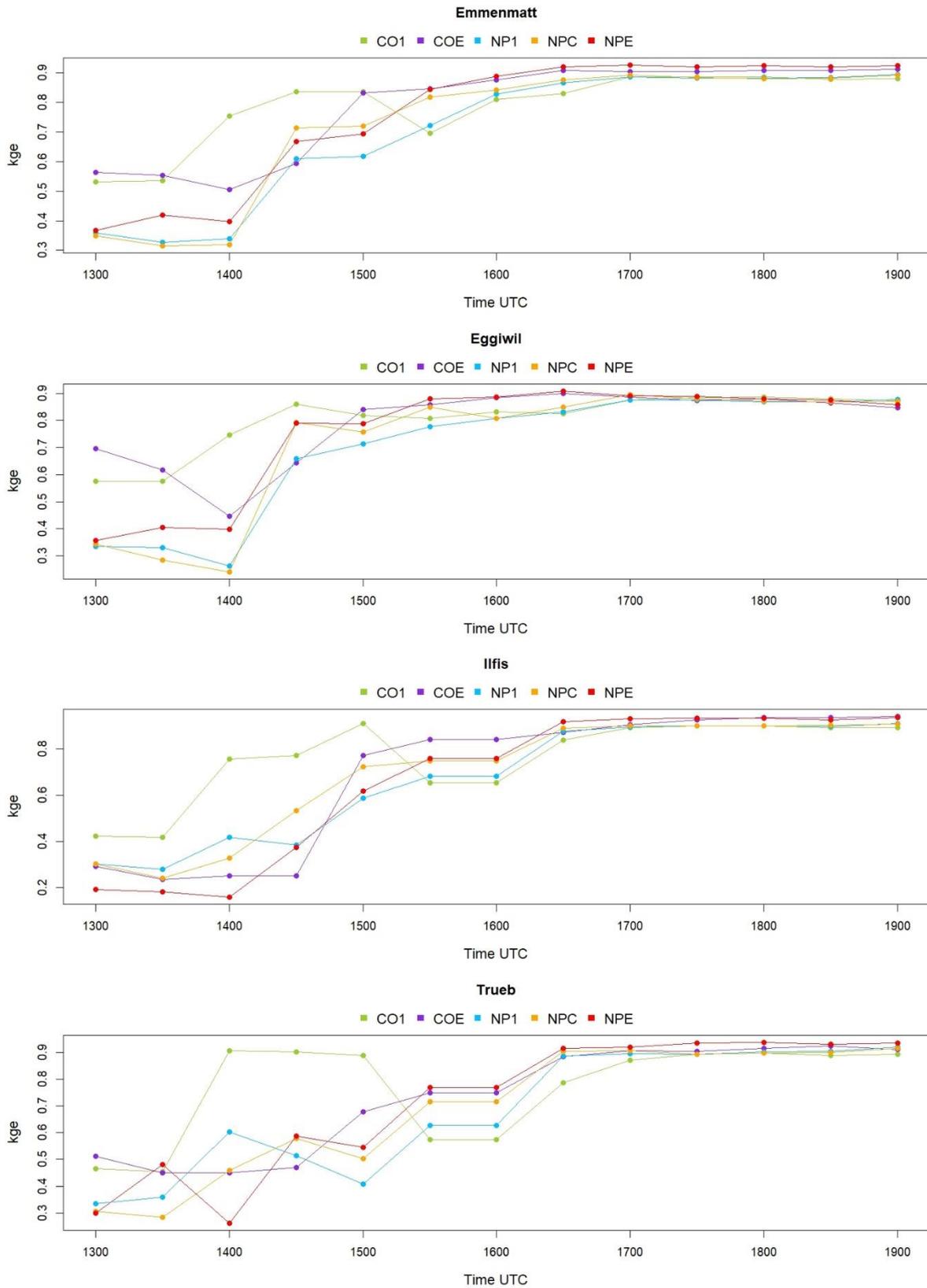


Figure 28: The median progression of the update cycle is shown for all nowcasting chains in the Emme catchments. The slope of the lines between the points indicates whether the skill in terms of the KGE increases or decreases within the update cycle.

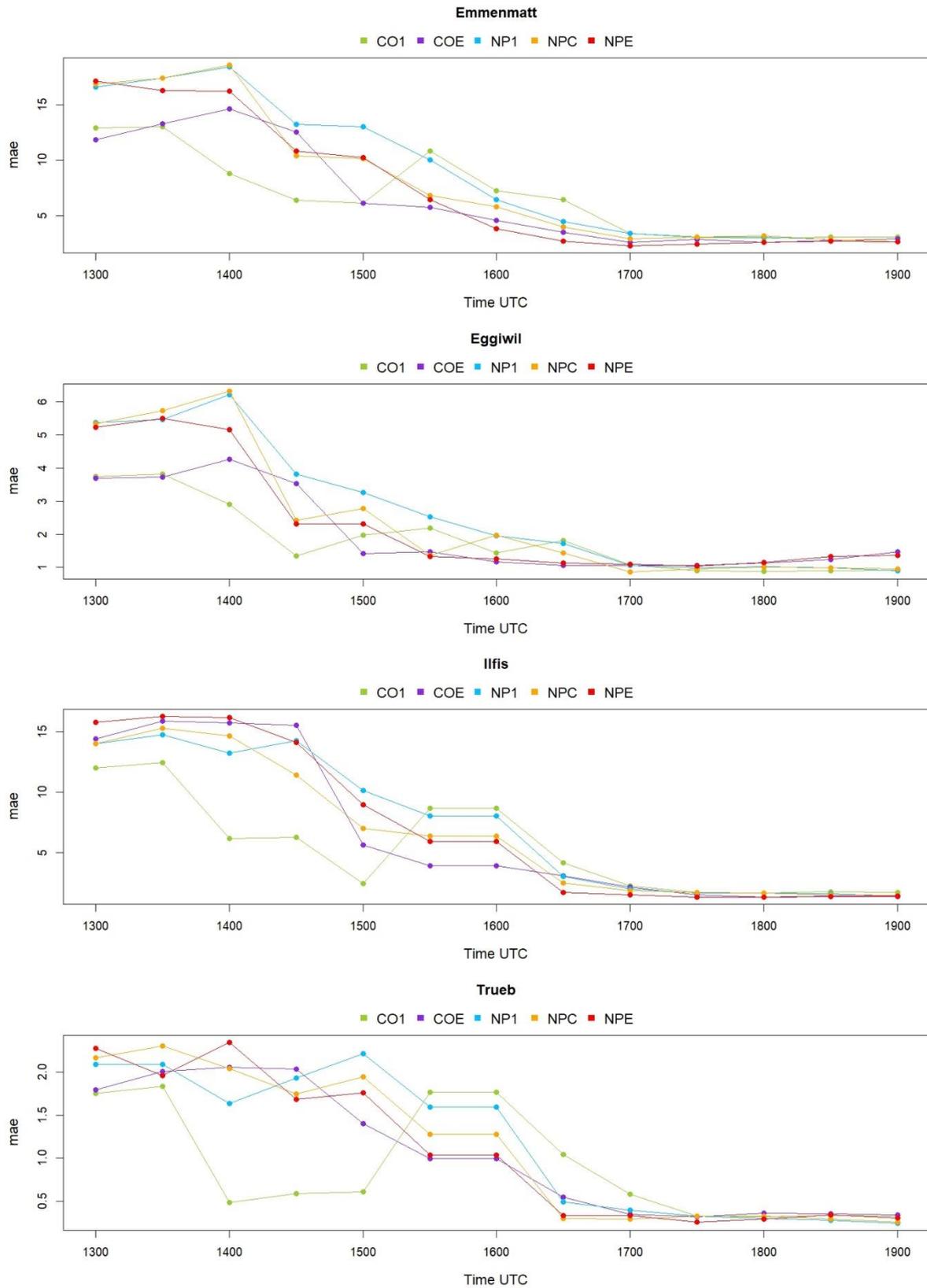


Figure 29: The median progression of the update cycle is shown for all nowcasting chains in the Emme catchments. The slope of the lines between the points indicates whether the skill in terms of the MAE increases or decreases within the update cycle.

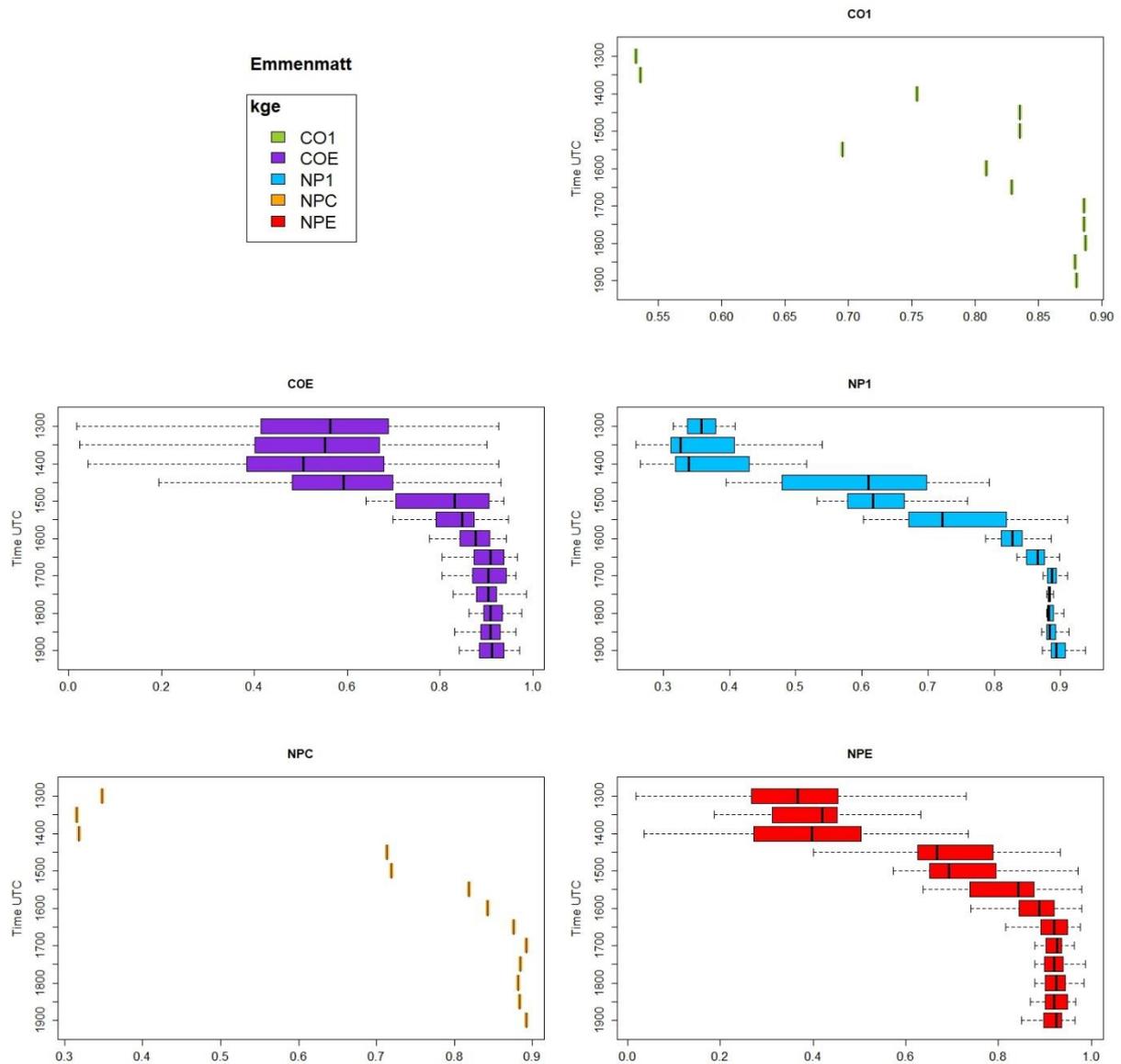


Figure 30: Update cycle boxplots showing all nowcasting chains in the main catchment Emmenmatt in terms of the KGE. The bold vertical bar of the boxplot represents the median, while the box represents the interquartile range. The whiskers display the range of extreme values.

3.2.2. Flood Event Assessment

The Taylor diagrams (Taylor, 2001) and the peak-boxes (Zappa et al., 2013) in figure 31 show from top to bottom the first, middle and last forecast (i.e. 13:00, 16:00, 19:00) of the update cycle in the Emmenmatt catchment. The entire update cycle is shown in the appendix (figure A.29). Both the Taylor diagram and the peak-box are used to assess a FF event. While the Taylor diagram provides important information on the correlation, standard deviation and root mean square error, the peak-box provides a user-friendly estimate of the magnitude and timing of the FF peak.

Regarding the Taylor diagram, the correlation of the simulation observation pair is depicted in the azimuthal angle. The standard deviation is defined by the X- and Y-axes and shown as blue dotted quadrants in the diagram. The RMSE is determined by the grey semicircles. The grey point at a correlation of one, a standard deviation of one, and a RMSE of zero represents the CPC hindcast runoff, that is the pseudo observation. The distances of the remaining colored points show the deviations of the members of the nowcasting chains from the observation in relation to the mentioned skill scores. Simulated members of the nowcasting chains, which are closer to the reference point, perform better in terms of correlation and RMSE. Members which are closer to the black solid quadrant have a more similar standard deviation to the reference.

The peak-box shows on the X-axis the timing of the peak runoff in hours after the initialization and on the Y-axis the magnitude of the peak runoff. The lower and upper horizontal lines of each probabilistic nowcasting chain show the minimum and maximum runoff magnitude of their members respectively. The bold horizontal line inbetween shows the median value. The same applies to the runoff peak timing with respect to the X-axis and the vertical lines. The intersection of the medians in peak runoff magnitude and timing represents the best estimate of the FF peak. In contrast to the previous figures, the peak-box shows the actual runoff observation as well as the CPC hindcast. The spread in the CPC hindcast arises from the ensemble generated by the hydrological model.

With respect to the Taylor diagram, the low standard deviation in the 13:00 forecast of all nowcasting chains, but especially of the NPE nowcasting chain, indicates that the nowcasting chains underestimate the pseudo observation. In the 16:00 forecast, the COE nowcasting chain underestimates the pseudo observation, whereas the NPE nowcasting chain overestimates it. In the 19:00 forecast, all members tend to overestimate the pseudo observation. Regarding the RMSE, the behavior of the COE and NPE members is similar. While in the 13:00 forecast, the COE members have a lower RMSE than the NPE members, in the 16:00 forecast, this is the opposite. In the 19:00 forecast all members, for except a few outliers, exhibit a low RMSE. In terms of correlation, no clear trend is visible. Except for the 19:00 forecast, it can be seen that the CO1, NPC and NP1 chains have higher correlation values than the COE and NPE nowcasting chains. The members of the NP1 nowcasting chain have similar values in correlation, standard deviation, and RMSE. This can be seen in the small spread in the Taylor diagram, which, as can be seen in the 13:00 forecast, does not necessarily imply a reliable prediction. Overall the Taylor diagram makes clearly visible that all members with the update cycle progressing approach the observation.

The peak-box gives information about the magnitude and timing of the FF peak. Magnitude and timing are the two most important variables when predicting FFs. In the Taylor diagrams, it is only possible to estimate how the members of the nowcasting chain perform in terms of statistic values. However, even for hydrologist, it would be discussable, which nowcasting chain performs best. The peak-box offers with the determination of the intersection between the medians in magnitude and timing of each probabilistic nowcasting chain a method to provide a deterministic statement about the expected magnitude and timing of the FF. This makes the probabilistic nowcasting chains comparable.

The peak-boxes underpin the statements of the Taylor diagrams. The ensemble spread is large in the first two forecasts for the COE and NPE nowcasting chain, while the spread of the members in the NP1 nowcasting chain is constantly smaller compared to the COE and NPE nowcasting chain. Furthermore, it can be seen that the NPE nowcasting chain can reduce its ensemble spread in the last 19:00 forecast, whereas this is not the case for the COE nowcasting chain. The peak-box provides a statement about the quality of the CPC hindcast. Since this runoff estimation is always very close to the actual runoff observation, it can be assumed that the CPC hindcast is a reliable reference.

In the 13:00 forecast, the COE nowcasting chain performs best in terms of magnitude. It must be mentioned, however, that the median value is still strongly underestimating the pseudo observation. In terms of runoff peak timing, both nowcasting chains COE and NPE predict the peak at the same time and outperform the NP1 nowcasting chain. However, the timing of the peak is predicted clearly too late and the prediction skill is as a result low. In the 16:00 forecast, the NPE nowcasting chain outperforms the COE and NP1 nowcasting chains in terms of magnitude. The COE outperforms the NPE and NP1 nowcasting chain in terms of timing. Above all, however, it becomes apparent that all nowcasting chains have approached the pseudo observation and increased therefore their performance skill. The last forecast at 19:00 shows nearly no difference between the different nowcasting chains. All nowcasting chains hit the timing perfectly and overestimate the pseudo observation only slightly. This is rather unexpected, since the nowcasting chains show strong differences in their spreads.

Overall, the peak-box underpins that as the update cycle progresses, the performance in the probabilistic nowcasting chains, in estimating peak magnitude and peak timing of the runoff, improves considerably.

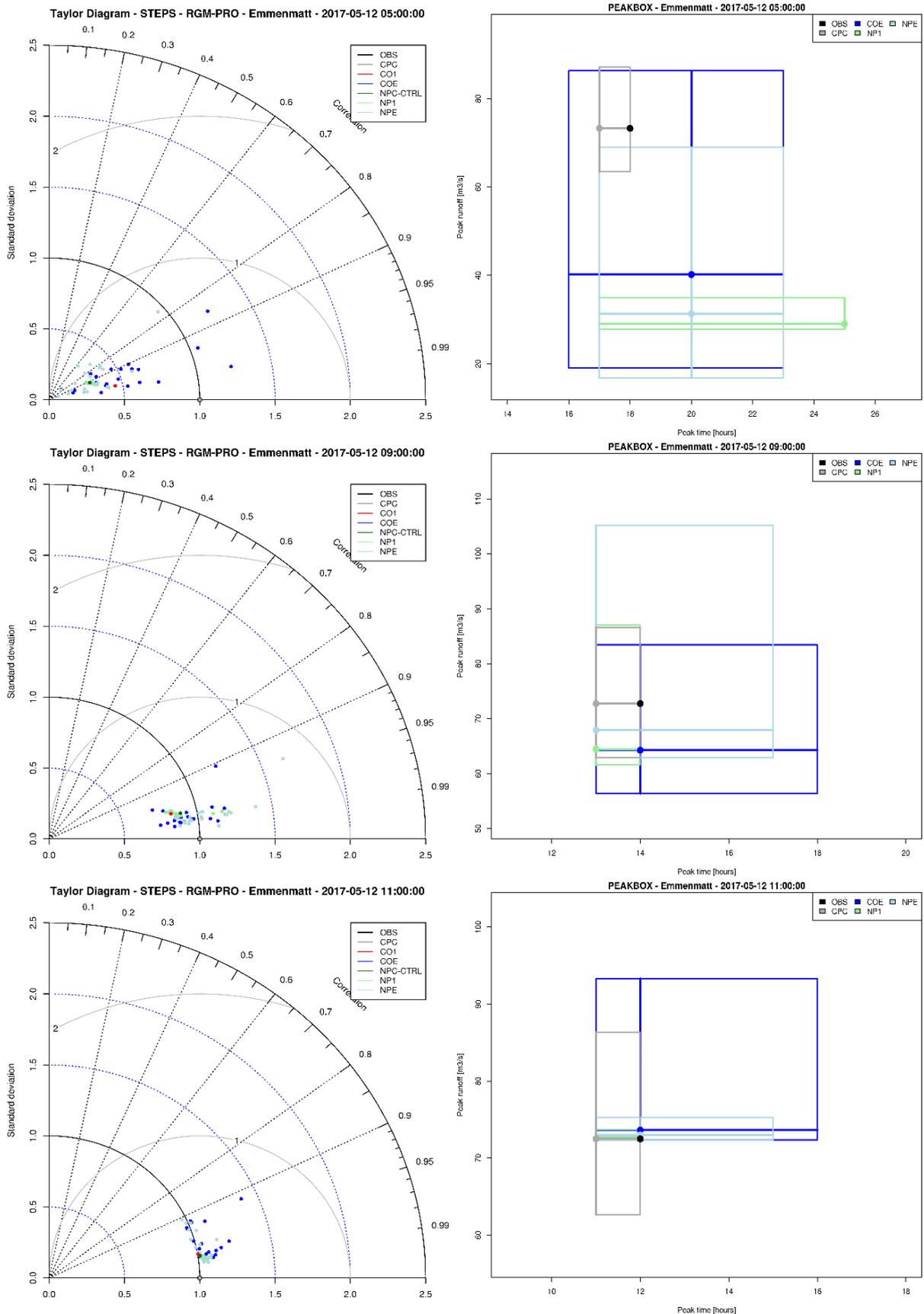


Figure 31: Left: Taylor diagrams showing correlation (azimuthal angle), standard deviation (blue dotted quadrant), RMSE (grey semicircles) and CPC hindcast (grey point). Right: Peak-boxes showing the best flash flood estimation in terms of peak magnitude and timing of the runoff. From top to bottom the first, middle and last prediction of the update cycle is shown for all nowcasting chains in the main catchment Emmenmatt.

4. Discussion

In the first section of the thesis the main research questions are stated as (1) are extrapolated nowcasting chains, which are blended into a NWP superior to nowcasting chains using raw NWP? Moreover, (2) is it possible to increase the skill of a nowcasting chain by frequently updated initial conditions? The results in the previous section clearly show that both main research questions can be answered affirmatively. This discussion section provides a complete analysis of the obtained results and the underlying methods, and places the new findings into the scientific context of relevant scientific works.

4.1. Blending an Extrapolation Nowcast into a NWP

Regarding the research question whether an extrapolated nowcast blended into a NWP outperforms a raw NWP nowcasting chain, Golding (1998) states that the very short-term nowcast shows its best performance at the beginning, while the NWP performs better for longer lead times. Therefore, using a blending scheme, which combines the nowcast extrapolation with the NWP, is capable of mitigating the weaknesses of both forecasting systems by focusing on their strengths and blend them in a suitable manner. The findings depicted in the stacked barplots in figures 20 and 21, showing that in the majority of events the blended extrapolation nowcasting chains outperform the raw NWP nowcasting chains, confirm the statement of Golding (1998). It can even be assumed that the nowcasting chains used in this thesis are more reliable than in the study by Golding (1998), since the study does not use the ensemble approach. The reason for this is most likely the lack of computer power at that time. The extrapolated nowcast and the NWP are deterministic, and the weighting function are determined empirically, whereas in this thesis the weighting is based on the ensemble spread of the extrapolated nowcast and the NWP. This assumption is supported by Bowler et al (2006), describing that single forecast scenarios have to be seen as the main reason for failure of any nowcasting system.

4.1.1. Blending Scheme and Weighting Function

The blending scheme as an important component of the nowcasting chains has a major influence on their performance. It must, therefore, be discussed whether the blending scheme used in this thesis is the most appropriate one. While this thesis uses a weighting function in proportion to the uncertainties of the extrapolated nowcast and the NWP, the operationally used INCA system from Haiden et al. (2011) uses a fixed linear weighting function. This simple approach is particularly suitable for operational use, since it is easy to implement and effi-

cient. According to Haiden et al (2011), the approach used in this thesis, which takes into account the uncertainties in the extrapolated nowcast and in the NWP, has not yet demonstrated its superiority. However, a blending scheme, which considers the spread of the ensemble as an estimation of uncertainty, is adaptive and able of taking into account the forecast skills of the extrapolation and the NWP. The linear weighting function, on the other hand, is only a function of lead time, which neither deals with the extent of the precipitation field nor with its position (Nerini et al., 2019). This can also be seen in the study by Poletti et al. (2019), which analyses different and non-adaptive weighting functions. The study shows that in different events, different weighting functions provide the best hydrological prediction. This underpins the assumption by Nerini et al. (2019) that an adaptive weighting function provides the most reliable hydrological predictions, and is consequently the most recommendable weighting function for the blending scheme.

4.1.2. Extrapolation Nowcast and Ensemble Size

Regarding the extrapolation of the nowcast, the used approach in this thesis is the Lagrangian persistence (Germann and Zawadzki, 2002). Lagrangian persistence is a simple approach to let the precipitation fields of the radar images move based on their antecedent behaviour. The weakness of Lagrangian persistence, however, is that changes such as grow and decay processes are not taken into account over the period of precipitation. Therefore, highly convective cells, which are likely in HPEs, are only considered by the Lagrangian persistence approach if they already exist at the time of the initialisation. In cases where the initialisation is made too early, the extrapolation becomes inaccurate due to the lack of convective precipitation (Nerini et al., 2019). To represent this weakness, stochastic noise is added to 20 of the 21 members of the extrapolated nowcast, which gives an estimation about the uncertainty in the nowcast (section 2.4). Furthermore, a currently introduced extrapolation nowcast system by Pulkkinen et al. (2019) called pySTEPS is able of generating some new random rain in order to improve the consideration of convective precipitation. However, pySTEPS is still not able to reliably capture intense convection. To increase extrapolation reliability during convective processes, frequently updated initial conditions can be used. In real time applications, this option is only limited by the computation time of the update cycle (Nerini et al., 2019).

A final question regarding the extrapolation nowcast which needs to be discussed is whether the size of members in the ensemble of the used nowcasting chains is sufficient. With regard to the number of members, care must be taken to ensure that on one hand the ensemble is large enough to cover the uncertainties in the estimation of precipitation as completely as possible, and that on the other hand the ensemble is no larger as needed, because with increas-

ing member size the computational time increases as well. The study by Pulkkinen et al. (2019) can be used to address the question about the member size. In the study, the relative operating characteristic (ROC) curve, established by Jolliffe and Stephenson (2003), is used to analyse the rate of correctly detected precipitation. Therefore, the probability of detection (POD) is plotted against probability of false detection (POFD). A perfect scenario for a predetermined precipitation threshold reaches a POD of 100 % and a POFD of 0 %. The area under this ROC curve given as a value from 0 to 1 is, as a result, maximum. Pulkkinen et al., (2019) examined the ROC areas for the ensemble sizes of 6, 12, 24, 48 and 96 members for the precipitation thresholds of 0.1 mm/h and 5 mm/h. The study analyses how the ROC area values change with increasing lead time. The findings of the study are that the size of the ensemble is crucial particularly when doing nowcasts, and especially for larger amounts of precipitation. Of the ensemble sizes examined, 24 members are considered optimal. The reason for this is that with ensembles larger than 24 members, the increase in the ROC area is only marginal.

The probabilistic extrapolation nowcasts in this thesis use 21 members to make the extrapolation compatible with the 21 members of the COSMO-E NWP. The ensemble size is quite close to the optimal value of 24 members. It must be considered that for each member of the nowcasting chains, an additional 21 members of the hydrological model are added, which noticeably increases the computation time. It can, as a consequence, be assumed that with 21 members, the uncertainties are sufficiently covered without the need for excessive computer power. This assumption is additionally supported by the figures 18 and 19, which assign skill to the majority of events for each nowcasting chain in terms of the KGE.

4.2. Performance Improvement through Frequently Updated QPEs

In the figures relating to the update cycle in section 3.2, it can consistently be seen that overall the reliability of the forecast increases with more recent QPEs, which are used as initial conditions. Especially figures 28 and 29 show that the performance in all Emme catchments improved conspicuously after the first two hours of the update cycle, and stay at a high performance for the rest of the update cycle.

The question arises why the improvement after the first two hours of the update cycle is so pronounced. A possible explanation for this behaviour can be found in the study by Nerini et al. (2019), which includes the type of precipitation formation, in the form of a case study, in the nowcast analysis. The study states that the NWP's struggle to represent the precipitation of a convective event correctly, especially for short lead times. The event used for the update cycle in this thesis is to be considered highly convective, since the accumulated precipitation, shown in figure 25 and 26, strongly increases until around 17:00. As a result, the KGE values

for the NWP are rather low in the update cycle up to and including 14:30. This is due to the intensive convective activity, which can not be fully captured by the NWP, and the fact that the actual runoff peak is still far away. In addition, the initial conditions of the NWP are outdated until 14:00.

Likewise, it can be seen in figure 30 that the nowcasting chains using an extrapolation nowcast show a low KGE value for the early update cycle. This is because important growth and decay processes during the convective phase are insufficiently simulated for longer lead times. At first glance, one could argue that the nowcasting chains that blend an extrapolated nowcast into an NWP should be able to capture the convective event with frequently updated initial condition already in an early stage of the update cycle. However, in agreement with the study from Poletti et al. (2019), it is to be assumed that the extrapolated nowcast is only superior to the NWP in the first two hours of the nowcast. However, since the convective activities in the updated cycles from 13:00 to 14:30 are not yet completed, even with the additional two hours of reliable extrapolation nowcast, also the nowcasting chains which blend an extrapolated nowcasts with a NWP cannot adequately capture the event. For the update cycle at 15:00, the most active convection phase is completed within the two hours of reliable nowcast extrapolation. Therefore, the overall performance of the blended nowcasting chains start to increase. Similar behaviour can be seen in the subcatchments of the Emme. The corresponding boxplot can be found in the appendix (figure A.28).

Consequently, it can be concluded that in the investigated event, the update cycle is initialized too early. The resulting KGE values of the first update cycles indicate that no reliable prediction can be made about the occurrence of a flash flood. Nevertheless, an early warning system should be able to issue a warning early and reliable enough to take and implement possible measures. Very late update cycles show in fact a very reliable prediction, but are useless for decision-makers, because measures against a FF could no longer be implemented (Romang et al., 2011).

4.3. Nowcasting Chains Specific Characteristics

As has already been demonstrated, the thesis shows that first nowcasting chains which blend an extrapolated nowcast into a NWP are superior to raw NWP, and second frequently updated initial conditions improve the forecast skill of all nowcasting chains within the update cycle. The section that follows discusses the specific characteristics of the used nowcasting chains.

In the figures 8 and 9 showing the accumulated precipitation forecast and in the resulting hydrographs of the Emme catchment shown in the figures 10 to 13, it can be seen that the nowcasting chains NP1 and NPE tend to overestimate the runoff while the nowcasting chains CO1, COE and NPC tend to underestimate the runoff in respect to CPC. The exception is the Eggiwil catchment, where all nowcasting chains overestimate the runoff. Since the Eggiwil catchment, exhibits substantially lower KGE values than the other catchments (figure 14), this catchment is not decisive for the analysis. A similar behaviour of the nowcasting chains can be seen in the Emme runoff of the middle update cycle in figure 27. The runoff estimations of the also shown first and last update cycles are not considered, because the first update cycle is considered too uncertain to make any assumptions, and the last one is useless for any real time application because it is too close to the actual FF. In general, the nowcasting chains underestimate the runoff rather than overestimate it.

A possible explanation for the underestimation in the nowcasting chains using raw NWP (i.e. CO1 and COE) can be found in the study from Buizza et al. (1999b). According to the study, large precipitation events, such as those investigated in this thesis, tend to be underestimated by the NWP. Furthermore, the study from Jenker et al. (2008) argues that precipitation values from catchments located in a valley, such as the Emme and Verzasca catchments studied in this thesis, are underestimated by the COSMO models. The underestimation of the raw NWP nowcasting chains in the investigated events is further supported by the findings of Klasa et al. (2018), who assign a severe underestimation of precipitation to the COSMO-E model during heavy precipitation events. The underestimation in the NPC nowcasting chain is probably due to the fact that in the investigated mountainous catchments, precipitation due to orographic uplift contributes to the heavy precipitation event. These precipitation amounts are not captured by the unperturbed Lagrangian persistence extrapolated member of the NPC nowcasting chain.

In the shown events, the nowcasting chains NP1 and NPE are overestimated. These nowcasting chains differ only in their extrapolation nowcast from the CO1 and COE nowcasting chain, which is therefore the reason for the overestimation. The study from Pulkkinen et al. (2019) identifies the autoregressive model as the reason for the overestimation. Due to the short memory of the second order auto-recessive model, it is not possible to capture correctly the long lifetime of large precipitation features. The model overestimates the lifetime of precipitation in the extrapolation. As a result, the amount of precipitation is overestimated. However, the overestimation of the nowcast extrapolation is mitigated by blending the

extrapolated nowcast with the NWP, which contributes essentially to the superiority of blended nowcasting chains over raw NWPs (Poletti et al., 2019).

The effect of the blending scheme can also be seen in figure 23, which shows that the blended NPC nowcasting chain outperforms the CO1 nowcasting chain in all catchments. In addition, figure 24 shows that the probabilistic nowcasting chains with a blending scheme perform clearly better than the probabilistic COE NWP nowcasting chain, independent of the catchment. Furthermore, the COSMO-1 related NP1 nowcasting chain benefits from the higher resolution.

4.4. Catchment Specific Characteristics

Regarding the catchment specific characteristics, figure 15 and 16 show that especially small catchments like the Trueb or the Pincascia catchment exhibit overall lower KGE values than the other catchments. Due to the low Trueb KGE values, the Ilfis catchment is also partly affected by low KGE values, as Trueb catchment is nested in the Ilfis catchment. The study from Liechti et al. (2013a) shows that especially in small catchments prone to FFs, the amount of precipitation as well as the location of the precipitation features are crucial in estimating the runoff. However, the detection of such small-scale features in small catchments is difficult, as they can only be detected by high-resolution measurement systems.

Figures 15 and 16 also exhibit that the overall performance in terms of the KGE value is rather higher for the Emme catchment than for the Verzasca catchment. This can also be seen in figure 17, which not only shows the median values of the KGE for the Emme and Verzasca catchment for all nowcasting chains separately, but also takes their ranges of KGE values into account. The forecasts in the Verzasca catchment show a larger range in KGE values and, most of all, lower KGE values than in the Emme catchment. One explanation is the aforementioned smaller catchment, which makes a reliable prediction more difficult. A further explanation is that the complex topography in the Verzasca catchment also reduces the reliability of the prediction (Zappa et al., 2011).

In the figures 20 and 21, the proportion of the best performing nowcasting chains per catchment in terms of KGE and MAE is shown respectively. It is visible that in the small catchments Trueb and Pincascia, the NPE nowcasting chain performs best. In the other catchments, the NPC nowcasting chain shows best performance in the most events. Since predictions in small catchments are challenging, the NPE nowcasting chain benefits from having 21 members in the extrapolation nowcast and 21 members in the NWP. The large number of members, in comparison to the other nowcasting chains, is able to capture the events rea-

sonably well, even if, as has already been mentioned, the prediction skills for the Trueb and the Pincascia catchments are rather low, and the approach of the NPE nowcasting chain reminds more of a wild guess. In the other catchments, the NPC nowcasting chain benefits from the higher resolution of the COSMO-1 NWP, in which the NPC is blended. The reason why the NP1 nowcasting chain, which is also blended into the COSMO-1 NWP, cannot keep up with the NPC nowcasting chain is that the control member of the NPC nowcasting chain is unperturbed and can be regarded as the best estimation. The remaining 20 perturbed members of the NP1 nowcasting chain are only 20 equally probable but slightly blurred members. Nevertheless, a probabilistic forecast should be preferred to a deterministic one. Why this is the case is explained in the following section.

Regarding the components of the KGE figure 22 shows that in the Emme catchment the NPC nowcasting chain performs best in terms of correlation. The reason for this is that the deterministic COSMO-1 related nowcasting chain is highly resolved and unperturbed. In the smaller and more challenging Verzasca catchment, the different scenarios of the NPE nowcasting chains ensemble help to better correlate with the CPC hindcast. The better performance of the COSMO-1 related nowcasting chains in terms of bias and the ratio between the CVs is assigned to the better spatial resolution compared to the COSMO-E related nowcasting chains. Furthermore, COSMO-1 uses more frequently updated initial conditions than COSMO-E. Exceptions must be made in the Verzasca catchment in terms of bias, and in the Ilfis catchment in terms of the ratio between the CVs. The reasons for this cannot be conclusively clarified without the individual events being examined in detail in a case study, which is beyond the scope of this thesis.

4.5. The Benefit of Probabilistic Nowcasting Chains

Every prediction is subject to uncertainty. In hydrological modelling, these uncertainties arise from the radar QPE, the extrapolation algorithm, and/or the hydrological model, only to mention a few. In order to achieve reliable forecasts, it is, therefore, part of the hydrological modelling to take these uncertainties into account. Even if the greatest care is taken in dealing with uncertainties, the simulation never completely corresponds to the observation. This residual uncertainty is quite large and must be taken into account for the forecasts by using the ensemble approach (Germann et al., 2006a). This opinion is shared by Addor et al. (2011), who points out that deterministic models sometimes miss an event due to the fact that these models only refer to a single scenario. Deterministic models are, therefore, unsuitable for early warning FFs systems. Even more the study shows that the median hydrograph of the prob-

abilistic COSMO-LEPS model was superior to the single member of the deterministic counterpart COSMO-7, and justifies the use of the median member in probabilistic nowcasting chains.

The Taylor diagrams shown in figure 31 depict skill variables of all nowcasting chains in a single plot. Information on correlation or over- and underestimation of nowcasting chains members can be read from it and discussed, as in section 4.3. However, when it comes to making a decision regarding the occurrence of a possible FF, decision-makers are not sufficiently referring to probabilistic forecasts, even though it is a superior technique than the deterministic one (Antonetti et al., 2019; Bruen et al., 2010). A possible explanation for this is that it requires a more detailed analysis of all members of the probabilistic nowcasting chains to interpret the forecast. Therefore, the peak-box approach introduced by Zappa et al. (2013), which considerably facilitates the analysis of probabilistic nowcasting chains, is used in the update cycle and shown in figure 31. While the peak-box of the first update cycle is too uncertain to estimate the occurrence of a FF, the peak-box of the last update cycle is unusable because it is too close in time to the occurrence of the FF. The peak-box of the middle update cycle is considered to be just early enough and sufficiently certain to make an estimate about the occurrence of a FF. It is difficult to determine the exact time when a FF warning should be issued. Especially in small catchments, which are prone to FFs, the response time oftentimes too short to take action (Liechti et al., 2013b). However, determining this point in time for the event would go beyond the scope of this thesis. Furthermore, it will have to be shown in the future how the operationalisation of such update cycles will develop. Since they require a lot of computer power, they are used only sparsely, and if they are used at all, it is solely with deterministic predictions (Nerini et al., 2019).

4.6. Limitations

Starting with the selection of the events, this thesis never refers to whether an FF occurred in the investigated events or not. To do so, the probability of detection and the false alarm ratio (FAR) of the investigated events would have to be analysed. However, the exclusion of such an analysis is justified, since in this thesis the best performing nowcasting chain is evaluated in relation to different skill scores, and the occurrence of a FF is not decisive in this respect. With regard to the chosen threshold for the events it must be mentioned that more conservative thresholds (e.g. lower thresholds or shorter trigger latency) can be applied to the alert system NowPAL. In operational use, this ensures that no FF event is missed. On the other hand, this causes more events to predict, which is easy for the automated alert system Now-

PAL. This is however not feasible for this thesis since some of the time consuming preprocessing steps were still done manually.

A further limitation is made regarding the radar QPE and the rain gauge measurements used for the initial conditions and the CPC hindcast, which are considered here as observation data. Obviously, this is not the case in reality. However, it can be justified by the fact that in small catchments, which are prone to FFs, in-situ rain gauge measurements are often not possible (Liechti et al., 2013a). Furthermore, as shown in the peak-boxes in figure 31, the CPC hindcast is a good estimation of the runoff observation, which in turn is also never completely error-free. Furthermore, the thesis does not address any analysis of the errors in the radar measurements, or the uncertainties in the CPC model. In this respect, the studies of Germann et al. (2006a, 2006b) give more information about the uncertainties in the radar measurements, and the studies of Sideris et al. (2014a, 2014b) about the uncertainties in the CPC model. In regards to the hydrological model RGM-PRO, no statements are made about the uncertainties in the runoff type maps and its resulting runoff generation. The ensemble approach considered by RGM-PRO, however, ensures that the residual uncertainties of the model can be dealt with. An analysis of the uncertainties in the runoff types of the hydrological model RGM-PRO can be found in Antonetti et al. (2019) and Horat et al. (2018).

The evaluation of the different nowcasting chains refers mainly to the KGE value. Like the NSE, the KGE is well familiar to most hydrologists. Nevertheless, most decision makers are unaware of what a certain value of these efficiency scores actually mean. For instance, the NSE value reaches constantly high values in seasonally dominated catchments. These high values are, however, deceptive, since they are only achieved on the basis of the poor reference forecast. With regard to the efficiency scores, it is, therefore, recommended to set catchment specific benchmarks, which allow to distinguish a useful from a useless prediction (Schaefli and Gupta, 2007). In this thesis, such a benchmark is neither set for the Emme nor the Verzasca catchment, since the nowcasting chains are only compared with each other. Even though, the question remains above which KGE value the performance of the nowcasting chains can be regarded as skilful. It has to be assumed that the KGE values are overestimated, as the spinup time of the events, which have an optimum KGE value of one, artificially increase the KGE value. This does not affect the analysis concerning which nowcasting chain performs best, as each event is examined separately, and all nowcasting chains are equally affected by the KGE overestimation. Furthermore, a finding by Addor et al. (2011) and Liechti et al. (2013) regarding the KGE value must be taken into account. The study states that with increasing lead time, the skill of the prediction decreases. Since this thesis focuses

mainly on the nowcast, and thus on very short lead times, it can be assumed that for the very short lead times, the skill of the prediction is higher than the skill, which is computed over the entire forecast.

It can be concluded that the quantification of the actual forecasting skill values is somewhat blurred by the mentioned issues. Nevertheless, this thesis is able to unambiguously show that blended nowcasting chains are superior to raw NWP nowcasting chains, and that frequently updated initial conditions improve the prediction skill.

5. Summary and Conclusion

Suddenly occurring flash floods induced by heavy precipitation events do not only cause costly damages but are also a threat to life. Very short-term nowcasting systems provide a means of identifying the hazard potential of a flash flood, and thus mitigate hazards. This thesis deals with the analysis of so-called nowcasting chains in the small and prone to flash floods Emme and Verzasca catchments. The used nowcasting chain consists of (i) a QPE provided by the CPC scheme, delivering up-to-date rainfall initial conditions for the nowcast. (ii) a nowcast product, which extrapolates the initial rainfall into the very short-term prediction for the next five hours, (iii) a NWP continuing the prediction for another 19 hours, (iv) a blending scheme, which merges the probabilistic extrapolation nowcast into the probabilistic NWP by adjusting their weights in inverse proportion to their uncertainties, and (v) a hydrological model called RGM-PRO with no need for calibration.

To answer the research question (1) whether nowcasting chains, which blend an extrapolation nowcast into a NWP are superior to raw NWPs, three nowcasting chains are set up using an extrapolated nowcast and a blending scheme, called NP1, NPC and NPE, and two nowcasting chains, using the deterministic COSMO-1 and probabilistic COSMO-E NWP for the entire forecast referred to as CO1 and COE respectively. The NP1 and NPE nowcasting chains are using a probabilistic nowcast extrapolation approach while the NPC nowcasting chain uses the deterministic control member of the NP1 nowcasting chain as nowcast extrapolation. For the blending scheme, deterministic nowcasting chains use the uncertainties of their probabilistic counterparts. For a sub-research question (1.1), deterministic nowcasting chains are split from the probabilistic ones and investigated separately, so as to answer whether the blending of the deterministic and probabilistic extrapolated nowcast into the NWP is superior to the raw deterministic and probabilistic NWP respectively. A further sub-research question (1.2) clarifies whether the investigated nowcasting chains are skilled enough to provide useful statements. A second research question (2) investigates whether frequently updated QPEs (i.e. initial conditions) can increase the nowcasting performances of the individual nowcasting chains.

To determine the performance of each nowcasting chain, their hydrographs are compared to the hydrograph of the pseudo observation, which clarifies if the CPC QPE is forcing the hydrological model. To achieve a representative result, 41 events in the Emme catchment and 40 events in the Verzasca catchment are investigated. Regarding the update cycle of the second research question, one event in the Emme catchment was analysed, in which more recent nowcasts based on updated initial conditions are computed every 30 minutes. To make the

performances of the individual nowcasting chains for each event comparable, the Kling-Cupta efficiency score was used in this thesis. In addition, the Nash-Sutcliffe efficiency score and the mean absolute error support the results. To compare deterministic and probabilistic nowcasting chains, the median of the skill score was used for probabilistic nowcasts.

In the results section, it is shown that (1) the blended nowcasting chains NP1, NPC and NPE outperform the raw NWP nowcasting chains CO1 and COE in the Emme and the Verzasca catchment. The superiority of blended nowcasting chains in the Emme and Verzasca main catchments is quite pronounced with a ratio of best performing blended nowcasting chains to raw NWP nowcasting chains of over 70 %. In the Emme and Verzasca subcatchments, the blended nowcasting chains are still superior but less pronounced, with a proportion of best performing nowcasting chains in more than 60 % of all events. Regarding (1.1) the blended deterministic nowcasting chain, NPC outperforms the raw deterministic NWP CO1 nowcasting chain in all catchments. The blended probabilistic nowcasting chains outperform the raw probabilistic NWP nowcasting chain distinctly. Concerning (1.2) the skill of the nowcasting chains almost all events are predicted skilful by all nowcasting chains, and in all catchments. In terms of (2) the update cycle, the results show a clear improvement in forecasting skill for all nowcasting chains with increasingly up-to-date initial conditions. Furthermore, the predictability in terms of magnitude and timing of the flash flood increases with more recent initial conditions.

The discussion section concludes that the merging of an extrapolated nowcast with a NWP is beneficial, because the extrapolated nowcast outperforms the NWP for very short lead times, whereas the NWP outperforms the extrapolated nowcast for longer lead times. Moreover, the initial time of the extrapolation is crucial because, if the extrapolation is started too early, important convective growth and decay processes in the atmosphere are not considered, whereas a late initialized extrapolation is unsuitable for hazard mitigation. The use of an update cycle is advisable for reliable flash flood prediction. However, its operational use is still sparse due to high computing requirements. In addition, it must be further investigated in which time span of the update cycle a prediction is most beneficial, since early predictions are burdened with large uncertainties, however, late predictions hardly leave time for countermeasures.

Besides, a tendency has been found that the extrapolated nowcast tend to overestimate, and the NWP tend to underestimate the runoff. The blending scheme merging both components together leads, therefore, to a mitigation of both effects. Catchment specific characteristics were addressed, which assign the small catchments major difficulties in their

predictability. To complete the discussion, the limitations of this thesis are discussed. Since this thesis focuses on the comparison of the nowcasting chains, the probability of detection or the false alarm ratio of the events are not considered. A further limitation is that some uncertainties are ignored, which is justified by the fact that they often cannot be measured in reality. The answer of the (1.2) question if the prediction of the nowcasting chains are skilful or not is somewhat unsatisfactory, since the skill score of the individual nowcasting chains is artificially increased by a spinup time in the predictions. Furthermore, the benchmark of a meaningful forecast in this thesis was set at zero. This is somewhat arbitrary, as other benchmarks are also reasonable.

This thesis has demonstrated that in this huge interdisciplinary field of flash flood nowcasting, further research is needed to increase nowcast reliability, lead time, and hazard mitigation. The study of Nerini et al. (2019) shows how future extrapolated nowcasts can be improved using a correction step based on the NWP. Poletti et al. (2019) reveals how different blending schemes affect the nowcast, and thus makes an essential contribution to improving the seamlessly blended nowcasting technique. A new trend in probabilistic precipitation nowcasting is set by Pulkkinen et al. (2019), which uses the advantage of a networked world by establishing the open-source Python library for probabilistic precipitation nowcasting, and thus creates a platform from and for users. It goes without saying that this approach, which involves the entire hydrological community, provides huge potential for the future. Finally, it is to be mentioned that hydrological modelling is not spared by a changing climate. When dealing with nowcasting or forecasting systems, the question of the adaptability of models to changing environmental conditions often remains unanswered. However, changes in landcover or in the atmospheric conditions are by no means to be ignored in the future and must establish their position in future nowcast related research (Thirel et al., 2015).

6. References

- Addor, N., Jaun, S., Fundel, F., and Zappa, M. (2011). An operational hydrological ensemble prediction system for the city of Zurich (Switzerland): skill, case studies and scenarios. *Hydrology and Earth System Sciences* 15, 2327–2347.
- Ahmad, S., and Simonovic, S.P. (2006). An Intelligent Decision Support System for Management of Floods. *Water Resources Management* 20, 391–410.
- Alfieri, L., Berenguer, M., Knechtel, V., Liechti, K., Sempere-Torres, D., and Zappa, M. (2015). Flash Flood Forecasting Based on Rainfall Thresholds. In *Handbook of Hydrometeorological Ensemble Forecasting*, Q. Duan, F. Pappenberger, J. Thielen, A. Wood, H.L. Cloke, and J.C. Schaake, eds. (Berlin, Heidelberg: Springer Berlin Heidelberg), 1–38.
- Andres, N., and Badoux, A. (2018). Unwetterschäden in der Schweiz im Jahre 2017. *Wasser Energie Luft*, 110. Jahrgang, 2018, Heft 1, CH-5401 Baden, 67–74.
- Antonetti, M., and Zappa, M. (2017). How can expert knowledge increase the realism of conceptual hydrological models? A case study in the Swiss Pre-Alps. *Hydrol. Earth Syst. Sci. Discuss.* 1–42.
- Antonetti, M., Scherrer, S., Kienzler, P.M., Margreth, M., and Zappa, M. (2017). Process-based hydrological modelling: The potential of a bottom-up approach for runoff predictions in ungauged catchments. *Hydrological Processes* 31, 2902–2920.
- Antonetti, M., Horat, C., Sideris, I.V., and Zappa, M. (2019). Ensemble flood forecasting considering dominant runoff processes – Part 1: Set-up and application to nested basins (Emme, Switzerland). *Natural Hazards and Earth System Sciences* 19, 19–40.
- Baldauf, M., Seifert, A., Förstner, J., Majewski, D., Raschendorfer, M., and Reinhardt, T. (2011). Operational Convective-Scale Numerical Weather Prediction with the COSMO Model: Description and Sensitivities. *Mon. Wea. Rev.* 139, 3887–3905.
- Bauer, P., Thorpe, A., and Brunet, G. (2015). The quiet revolution of numerical weather prediction. *Nature* 525, 47–55.
- Berenguer, M., Corral, C., Sánchez-Diezma, R., and Sempere-Torres, D. (2005). Hydrological Validation of a Radar-Based Nowcasting Technique. *Journal of Hydrometeorology* 6, 532–549.
- Beven, K.J. (2012). *Rainfall-runoff modelling: the primer*, 2nd ed., Chichester, West Sussex ; Hoboken, NJ. Wiley-Blackwell, 1–455.

- Bowler, N.E., Pierce, C.E., and Seed, A.W. (2006). STEPS: A probabilistic precipitation forecasting scheme which merges an extrapolation nowcast with downscaled NWP. *Quarterly Journal of the Royal Meteorological Society* *132*, 2127–2155.
- Bruen, M., Krahe, P., Zappa, M., Olsson, J., Vehvilainen, B., Kok, K., and Daamen, K. (2010). Visualizing flood forecasting uncertainty: some current European EPS platforms-COST731 working group 3. *Atmosph. Sci. Lett.* *11*, 92–99.
- Buizza, R., Milleer, M., and T. N. Palmer, (1999a): Stochastic representation of model uncertainties in the ECMWF Ensemble Prediction System. *Quart. J. Roy. Meteor. Soc.*, *125*, 2887–2908.
- Buizza, R., Hollingsworth, A., Lalauette, F., and Ghelli, A. (1999b). Probabilistic Predictions of Precipitation Using the ECMWF Ensemble Prediction System. *Weather and forecasting* *14*, 168–189.
- Cristiano, E., ten Veldhuis, M., and van de Giesen, N. (2017). Spatial and temporal variability of rainfall and their effects on hydrological response in urban areas – a review. *Hydrol. Earth Syst. Sci.* *21*, 3859–3878.
- Foresti, L., and Seed, A. (2015). On the spatial distribution of rainfall nowcasting errors due to orographic forcing. *Meteorological Applications* *22*, 60–74.
- Foresti, L., Reyniers, M., Seed, A., and Delobbe, L. (2016). Development and verification of a real-time stochastic precipitation nowcasting system for urban hydrology in Belgium. *Hydrology and Earth System Sciences* *20*, 505–527.
- Georgakakos, K.P. (2006). Analytical results for operational flash flood guidance. *Journal of Hydrology* *317*, 81–103.
- Germann, U., and Zawadzki, I. (2002). Scale-Dependence of the Predictability of Precipitation from Continental Radar Images. Part I: Description of the Methodology. *Monthly Weather Review* *130*, 2859–2873.
- Germann, U., Berenguer, M., Sempere-Torres, D., and Salvade, G. (2006a). Ensemble radar precipitation estimation - a new topic on the radar horizon. *European Conference on Radar in Meteorology and Hydrology*, 559–562.
- Germann, U., Galli, G., Boscacci, M., and Bolliger, M. (2006b). Radar precipitation measurement in a mountainous region. *Q. J. R. Meteorol. Soc.* *132*, 1669–1692.
- Golding, B.W. (1998). Nimrod: A system for generating automated very short range forecasts. *Meteorological Applications* *5*, 1–16.

- Gupta, H.V., Kling, H., Yilmaz, K.K., and Martinez, G.F. (2009). Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. *Journal of Hydrology* 377, 80–91.
- Haiden, T., Kann, A., Wittmann, C., Pistotnik, G., Bica, B., and Gruber, C. (2011). The Integrated Nowcasting through Comprehensive Analysis (INCA) System and Its Validation over the Eastern Alpine Region. *Weather and Forecasting* 26, 166–183.
- Hilker, N., Badoux, A., and Hegg, C. (2009). The Swiss flood and landslide damage database 1972–2007. *Natural Hazards and Earth System Sciences* 9, 913–925.
- Horat, C. (2017). Operational applications of a process-based runoff generation module in the Emme and Ticino areas. Masterthesis. Department of environmental systems science, Swiss Federal Institute of Technology Zurich, Zurich.
- Horat, C., Antonetti, M., Liechti, K., Kaufmann, P., and Zappa, M. (2018). Ensemble flood forecasting considering dominant runoff processes: II. Benchmark against a state-of-the-art model-chain (Verzasca, Switzerland). *Nat. Hazards Earth Syst. Sci. Discuss.* 1–34.
- Houtekamer, P., Lefaiivre, L., Derome, J., Ritchie, H., and Mitchell, L. (1996). A system simulation approach to ensemble prediction. *American meteorological society*, 1225–1242.
- Jenkner, J., Frei, C., and Schwierz, C. (2008). Quantile-based short-range QPF evaluation over Switzerland. *Meteorologische Zeitschrift* 17, 827–848.
- Jolliffe, T., Stephenson, D. (2003). *Forecast verification. A practitioner's guide in atmospheric science.* John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester, West Sussex, England, ISBN: 0-471-49759-2.
- Kienzler, P.M., and Naef, F. (2008). Temporal variability of subsurface stormflow formation. *Hydrol. Earth Syst. Sci.* 257–265.
- Klasa, C., Arpagaus, M., Walser, A., and Wernli, H. (2018). An evaluation of the convection-permitting ensemble COSMO-E for three contrasting precipitation events in Switzerland. *Quarterly Journal of the Royal Meteorological Society* 144, 744–764.
- Kling, H., Fuchs, M., and Paulin, M. (2012). Runoff conditions in the upper Danube basin under an ensemble of climate change scenarios. *Journal of Hydrology* 424/425, 264–277.
- Liechti, K., and Zappa, M. (2016). Verification of Short-Range Hydrological Forecasts. In *Handbook of Hydrometeorological Ensemble Forecasting*, Q. Duan, F. Pappenberger,

- J. Thielen, A. Wood, H.L. Cloke, and J.C. Schaake, eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 1–24.
- Liechti, K., Fundel, F., Germann, U., and Zappa, M. (2012). Flood nowcasting in the southern Swiss Alps using radar ensemble. *Weather Radar and Hydrology* 351, 496–501.
- Liechti, K. (2013). Uncertainty propagation in operational flash flood forecasting systems. Dissertation. Mathematisch-naturwissenschaftlichen Fakultät der Universität Zürich, Zurich.
- Liechti, K., Zappa, M., Fundel, F., and Germann, U. (2013a). Probabilistic evaluation of ensemble discharge nowcasts in two nested Alpine basins prone to flash floods: Evaluation of ensemble discharge nowcasts in two nested alpine basins. *Hydrological Processes* 27, 5–17.
- Liechti, K., Panziera, L., Germann, U., and Zappa, M. (2013b). The potential of radar-based ensemble forecasts for flash-flood early warning in the southern Swiss Alps. *Hydrology and Earth System Sciences* 17, 3853–3869.
- Mass, C. (2011). Nowcasting: The next revolution in weather prediction. Department of atmospheric sciences, University of Washington, 1–23.
- MeteoSwiss, (2019). COSMO forecasting system.
<https://www.meteoswiss.admin.ch/home/measurement-and-forecasting-systems/warning-and-forecasting-systems/cosmo-forecasting-system.html> – [Last access: 19.06.2019].
- Murphy, J.M., Sexton, D.M.H., Barnett, D.N., Jones, G.S., Webb, M.J., Collins, M., and Stainforth, D.A. (2004). Quantification of modelling uncertainties in a large ensemble of climate change simulations. *Nature* 430, 768–772.
- Nash, J.E., and Sutcliffe, J.V. (1970). River flow forecasting through conceptual models part I — A discussion of principles. *Journal of Hydrology* 10, 282–290.
- Nerini, D., Foresti, L., Leuenberger, D., Robert, S., and Germann, U. (2019). A Reduced-Space Ensemble Kalman Filter Approach for Flow-Dependent Integration of Radar Extrapolation Nowcasts and NWP Precipitation Ensembles. *Monthly Weather Review* 147, 987–1006.
- Ochoa-Rodriguez, S., Wang, L.-P., Gires, A., Pina, R.D., Reinoso-Rondinel, R., Bruni, G., Ichiba, A., Gaitan, S., Cristiano, E., van Assel, J., et al. (2015). Impact of spatial and temporal resolution of rainfall inputs on urban hydrodynamic modelling outputs: A multi-catchment investigation. *Journal of Hydrology* 531, 389–407.

- Panziera, L., Gabella, M., Zanini, S., Hering, A., Germann, U., and Berne, A. (2016). A radar-based regional extreme rainfall analysis to derive the thresholds for a novel automatic alert system in Switzerland. *Hydrology and Earth System Sciences* 20, 2317–2332.
- Poletti, M.L., Silvestro, F., Davolio, S., Pignone, F., and Rebora, N. (2019). Using nowcasting technique and data assimilation in a meteorological model to improve very short range hydrological forecasts. *Hydrol. Earth Syst. Sci. Discuss.* 1–28.
- Pulkkinen, S., Nerini, D., Pérez Hortal, A.A., Velasco-Forero, C., Seed, A., Germann, U., and Foresti, L. (2019). Pysteps: an open-source Python library for probabilistic precipitation nowcasting (v1.0). *Geosci. Model Dev. Discuss.* 1–68.
- Romang, H., Zappa, M., Hilker, N., Gerber, M., Dufour, F., Frede, V., Béro, D., Oplatka, M., Hegg, C., and Rhyner, J. (2011). IFKIS-Hydro: an early warning and information system for floods and debris flows. *Natural Hazards* 56, 509–527.
- Rossa, A., Liechti, K., Zappa, M., Bruen, M., Germann, U., Haase, G., Keil, C., and Krahe, P. (2011). The COST 731 Action: A review on uncertainty propagation in advanced hydro-meteorological forecast systems. *Atmospheric Research* 100, 150–167.
- Schaefli, B., and Gupta, H.V. (2007). Do Nash values have value? *Hydrol. Process.* 21, 2075–2080.
- Scherrer, S., and Naef, F. (2003). A decision scheme to indicate dominant hydrological flow processes on temperate grassland. *Hydrol. Process.* 17, 391–401.
- Seed, A. (2003). A dynamic and spatial scaling approach to advection forecasting. Cooperative Research Centre for Catchment Hydrology, Bureau of Meteorology, Melbourne, Victoria, Australia, 381–388.
- Sideris, I.V., Gabella, M., Erdin, R., and Germann, U. (2014a). Real-time radar-rain-gauge merging using spatio-temporal co-kriging with external drift in the alpine terrain of Switzerland: Real-time radar-rain-gauge merging. *Quarterly Journal of the Royal Meteorological Society* 140, 1097–1111.
- Sideris, I.V., Gabella, M., Sassi, M., and Germann, U. (2014b). The combiprecip experience: Development and operation of a real-time radar-rain gauge combination scheme in Switzerland. 1–10.
- Taylor, K.E. (2001). Summarizing multiple aspects of model performance in a single diagram. *J. Geophys. Res.* 106, 7183–7192.
- Thirel, G., Andréassian, V., Perrin, C., Audouy, J.-N., Berthet, L., Edwards, P., Folton, N., Furusho, C., Kuentz, A., Lerat, J., Lindström, G., Martin, E., Mathevet, T., Merz, R., Parajka, J., Ruelland, D., and Vaze, J. (2015). Hydrology under change: an evaluation

- protocol to investigate how hydrological models deal with changing catchments, *Hydrological Sciences Journal*, 60:7-8, 1184–1199.
- Thorndahl, S., Einfalt, T., Willems, P., Nielsen, J.E., ten Veldhuis, M.-C., Arnbjerg-Nielsen, K., Rasmussen, M.R., and Molnar, P. (2017). Weather radar rainfall data in urban hydrology. *Hydrology and Earth System Sciences* 21, 1359–1380.
- Viviroli, D., Zappa, M., Gurtz, J., and Weingartner, R. (2009). An introduction to the hydrological modelling system PREVAH and its pre- and post-processing-tools. *Environmental Modelling & Software* 24, 1209–1222.
- Wilks, D.S. (2006). *Statistical methods in the atmospheric sciences*. volume 91 of *International Geophysics Series*. Academic press, 2nd edition.
- Wilson, C.B., Valdes, J.B., and Rodriguez-Iturbe, I. (1979). On the influence of the spatial distribution of rainfall on storm runoff. *Water Resources Research* 15, 321–328.
- Wöhling, Th., Lennartz, F., and Zappa, M. (2006). Technical Note: Updating procedure for flood forecasting with conceptual HBV-type models. *Hydrol. Earth Syst. Sci.* 10, 783–788.
- World Meteorological Organisation, (2015). *Seamless Prediction of the Earth System: From Minutes to Months*. WMO, 483 pp.
- Zappa, M., Rotach, M.W., Arpagaus, M., Dorninger, M., Hegg, C., Montani, A., Ranzi, R., Ament, F., Germann, U., Grossi, G., et al. (2008). MAP D-PHASE: real-time demonstration of hydrological ensemble prediction systems. *Atmospheric Science Letters* 9, 80–87.
- Zappa, M., Jaun, S., Germann, U., Walser, A., and Fundel, F. (2011). Superposition of three sources of uncertainties in operational flood forecasting chains. *Atmospheric Research* 100, 246–262.
- Zappa, M., Fundel, F., and Jaun, S. (2013). A ‘Peak-Box’ approach for supporting interpretation and verification of operational fensemble peak-flow forecasts: supporting ensemble peak-flow forecasts. *Hydrological Processes* 27, 117–131.
- Zawadzki, I., Morneau, J., and Laprise, R. (1994). Predictability of precipitation patterns: An operational approach. *Journal of applied meteorology*. 1562–17571.

7. Appendix

7.1. Event Selection NowPAL Output

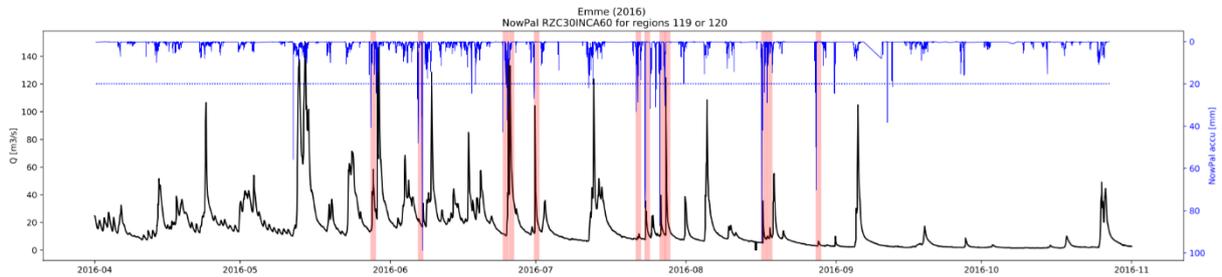


Figure A.1: NowPAL RZC30INCA60 output for the year 2016 in the Emme catchment. The blue axis shows the accumulated precipitation by NowPAL. The black axis shows the observed runoff. The blue dotted line represents the threshold triggering an event and the red bars cover the 24-hour forecast.

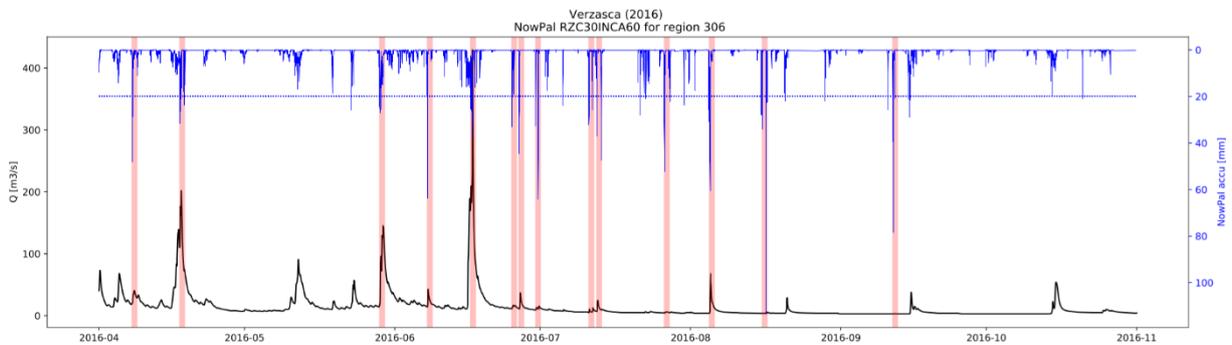


Figure A.2: NowPAL RZC30INCA60 output for the year 2016 in the Verzasca catchment. The blue axis shows the accumulated precipitation by NowPAL. The black axis shows the observed runoff. The blue dotted line represents the threshold triggering an event and the red bars cover the 24-hour forecast.

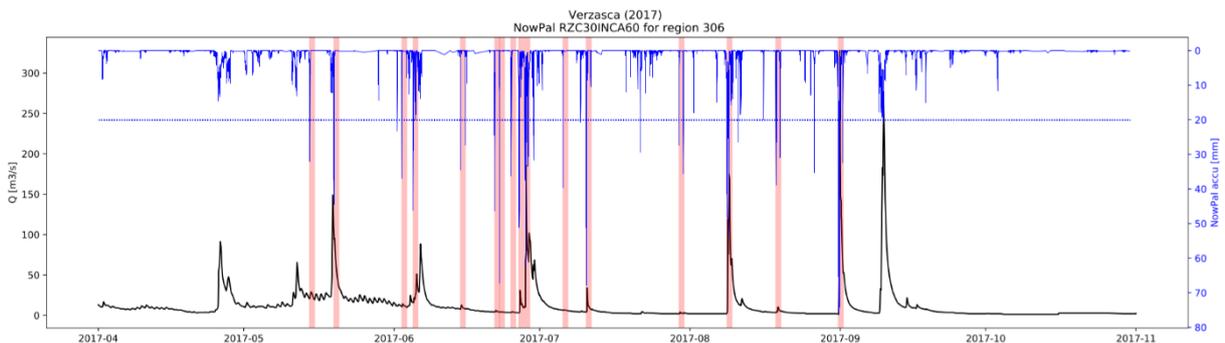


Figure A.3: NowPAL RZC30INCA60 output for the year 2017 in the Verzasca catchment. The blue axis shows the accumulated precipitation by NowPAL. The black axis shows the observed runoff. The blue dotted line represents the threshold triggering an event and the red bars cover the 24-hour forecast.

7.2. Performance Overview NSE and MAE

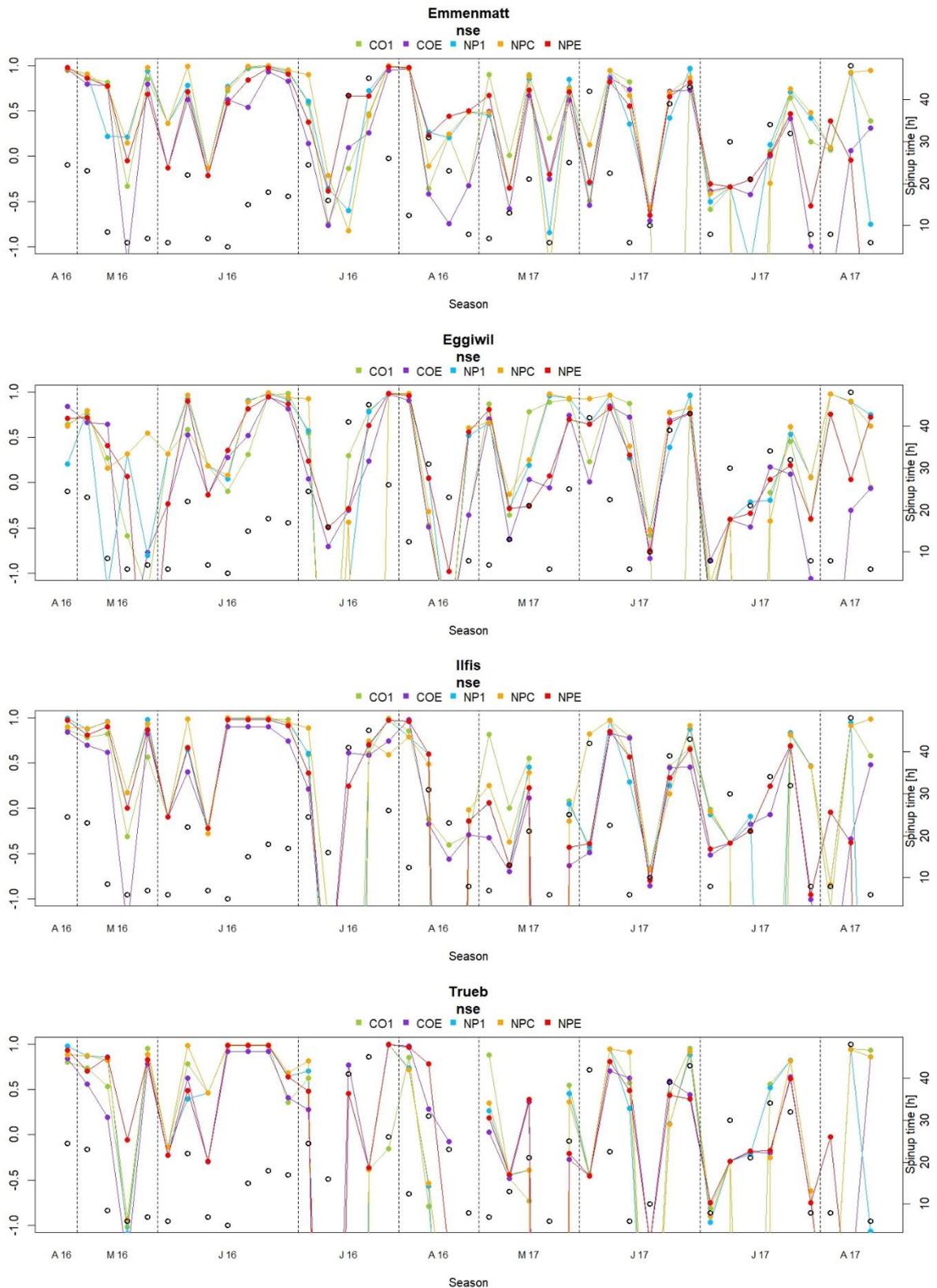


Figure A.4: Overview of the NSE medians for all events in the Emme catchments including spinup time. Letters and numbers in the X-axis representing months and years respectively.

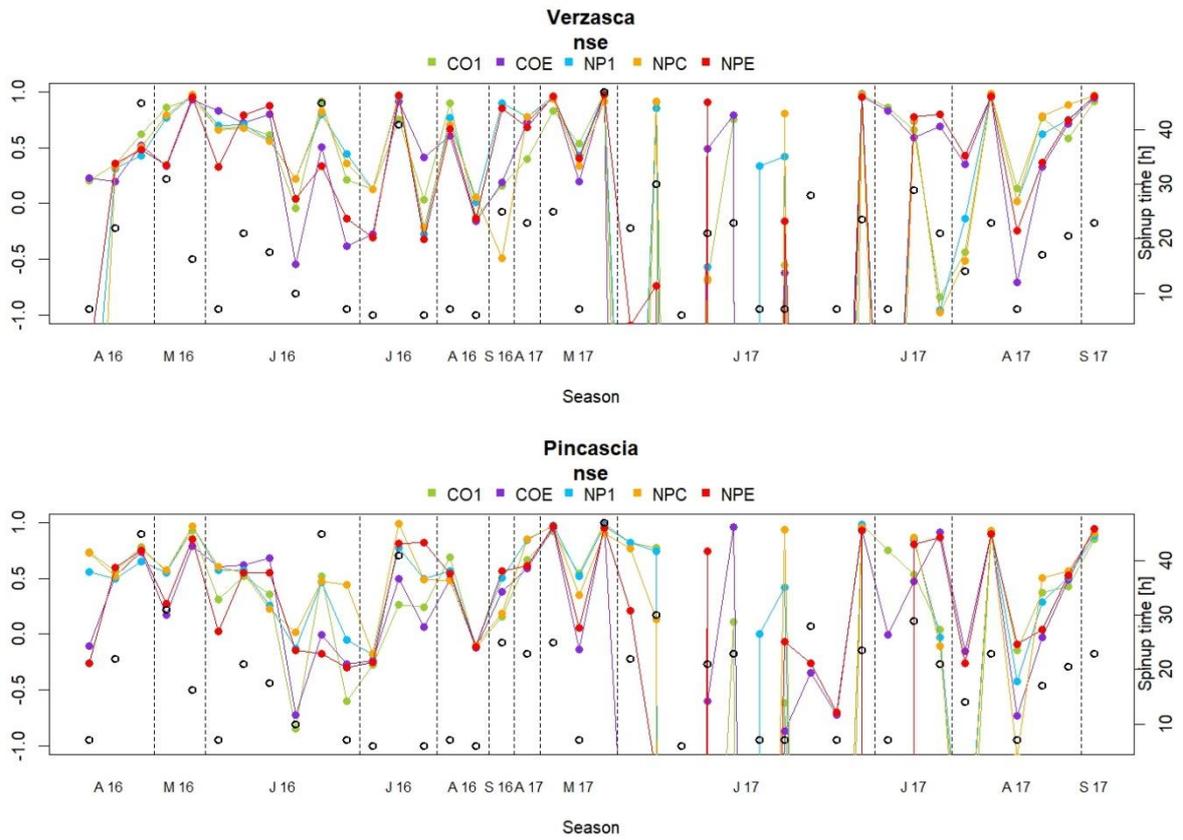


Figure A.5: Overview of the NSE medians for all events in the Verzasca catchments including spinup time. Letters and numbers in the X-axis representing months and year respectively.

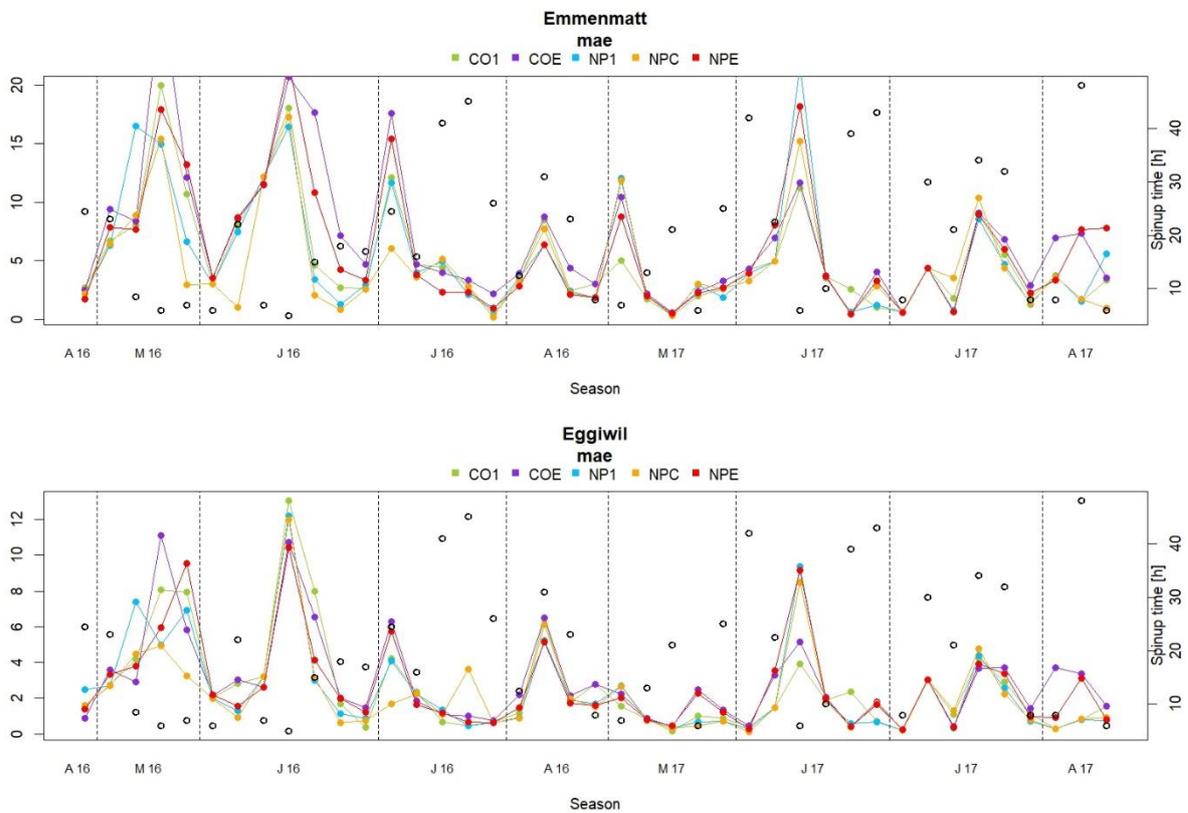


Figure continuous on next page.

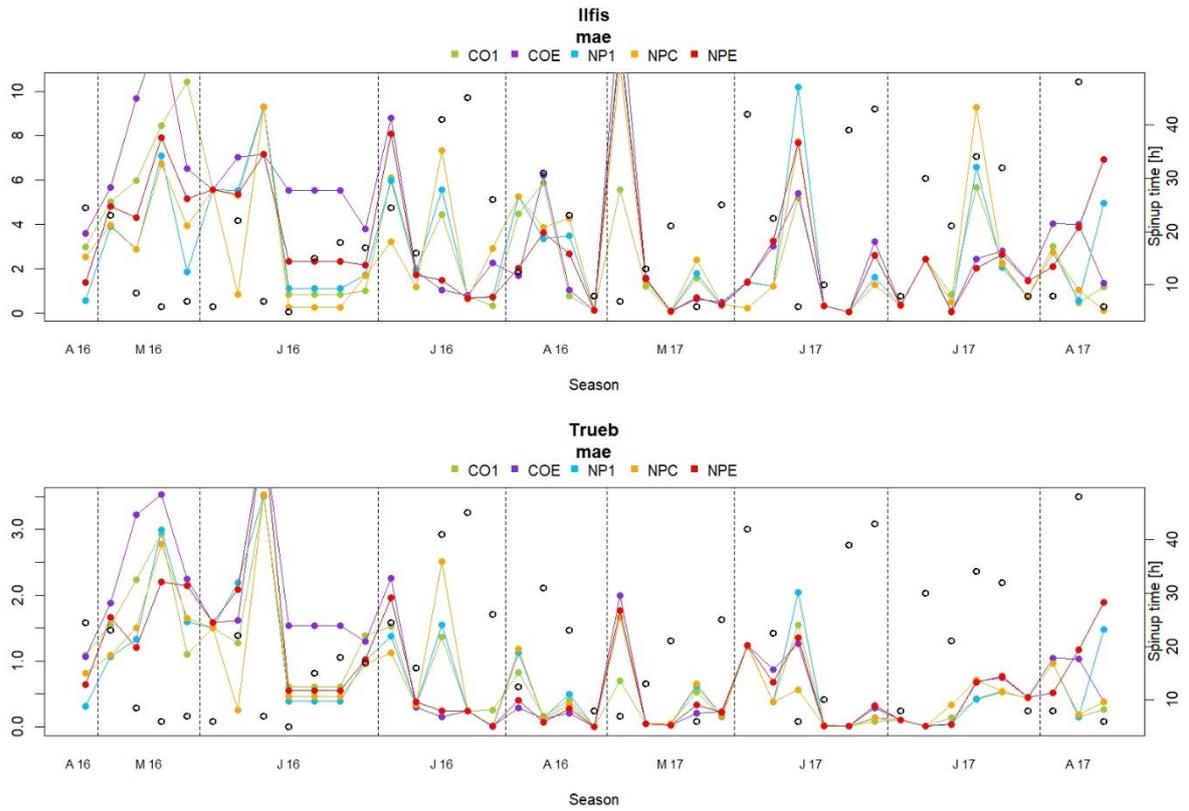


Figure A.6: Overview of the MAE medians for all events in the Emme catchments including spinup time. Letters and numbers in the X-axis representing months and year respectively.

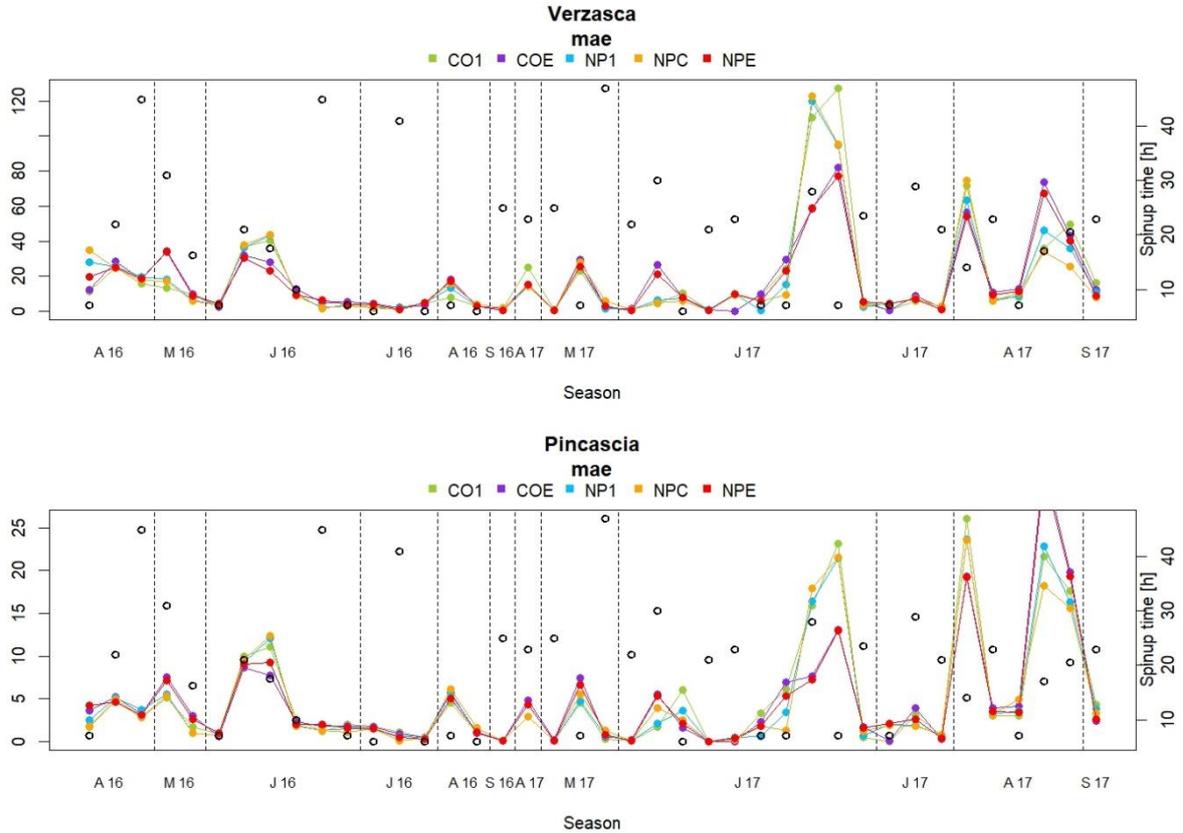


Figure A.7: Overview of the MAE medians for all events in the Verzasca catchments including spinup time. Letters and numbers in the X-axis representing months and year respectively.

7.3. Boxplot Overview KGE, NSE and MAE

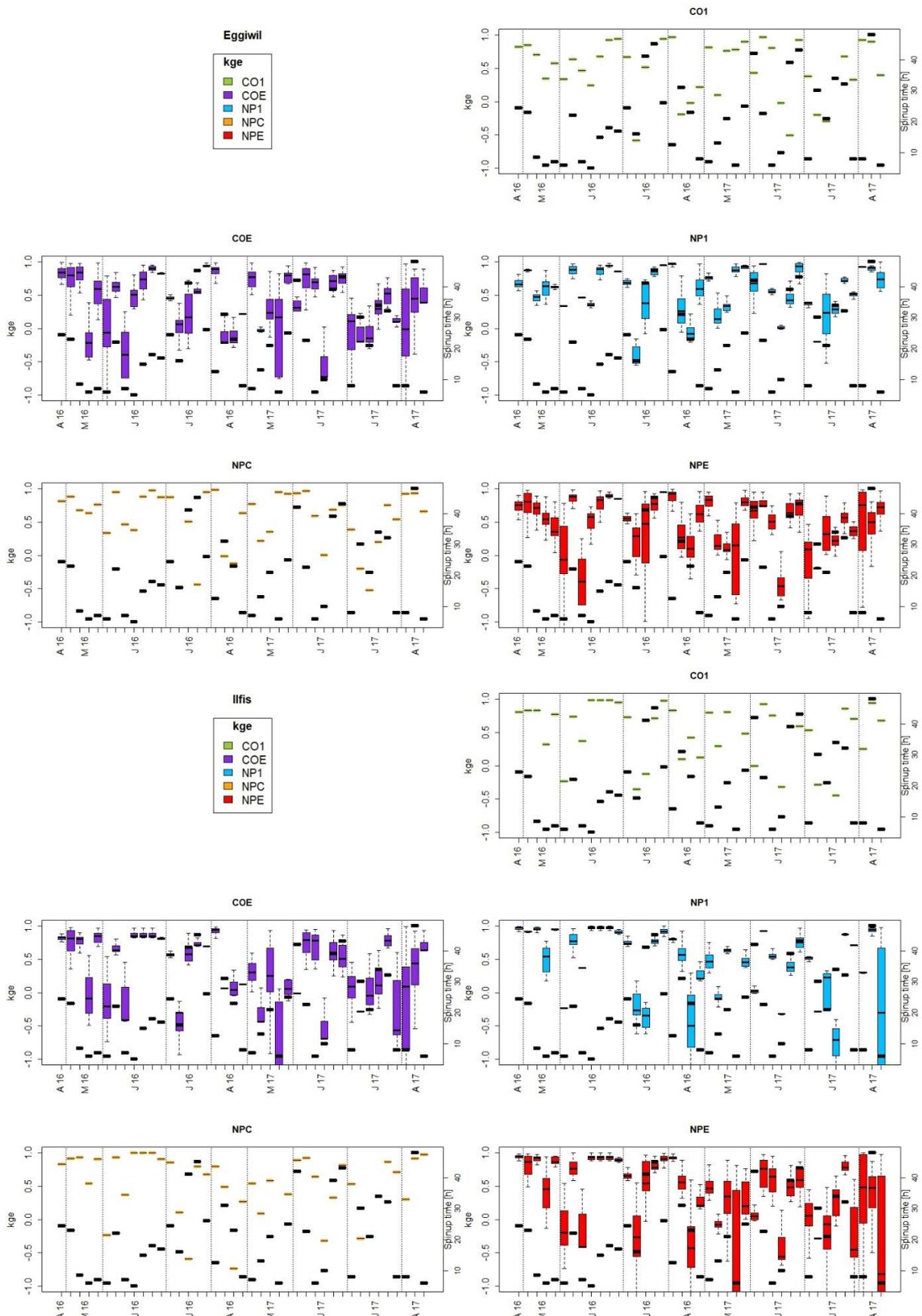


Figure continuous on next page.

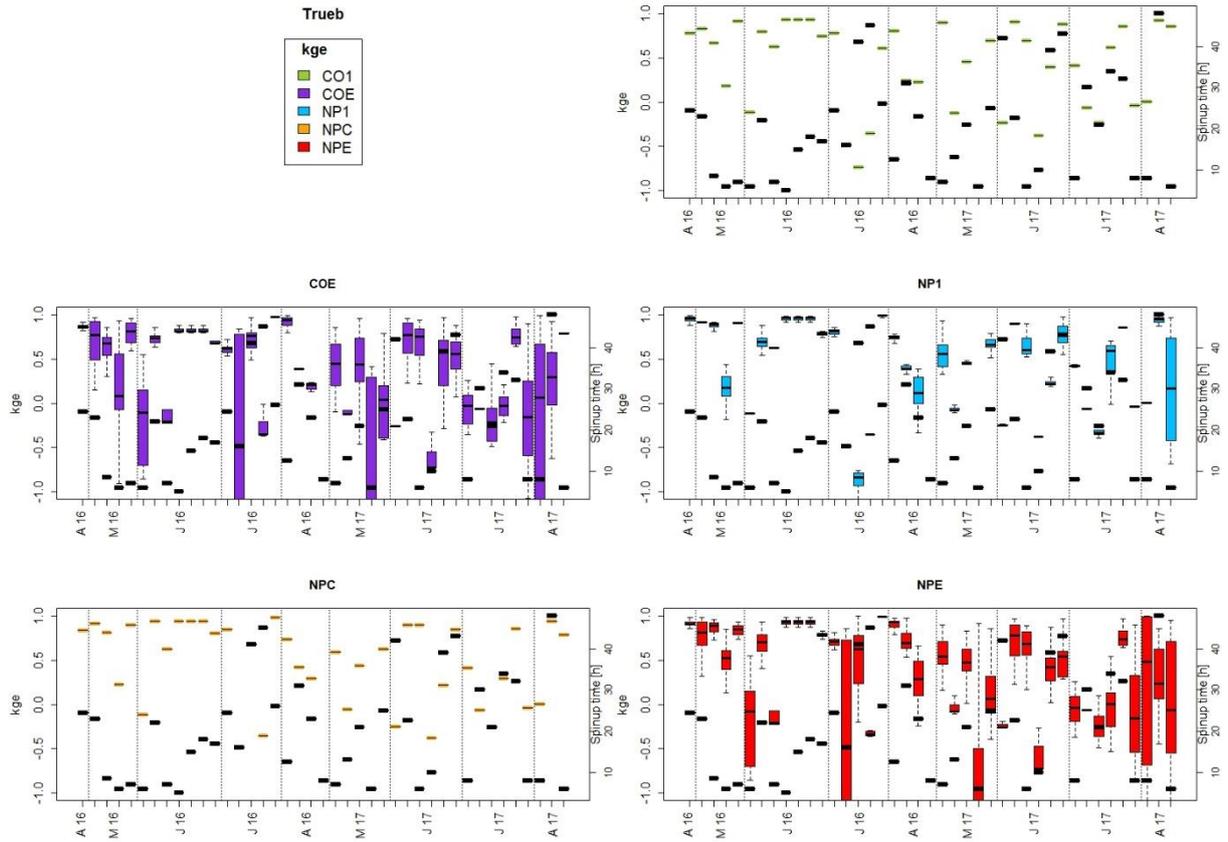


Figure A.8: Boxplot overview of the KGE values for all events in the Emme subcatchments including spinup time (bold bar). Letters and numbers in the X-axis representing months and year respectively.

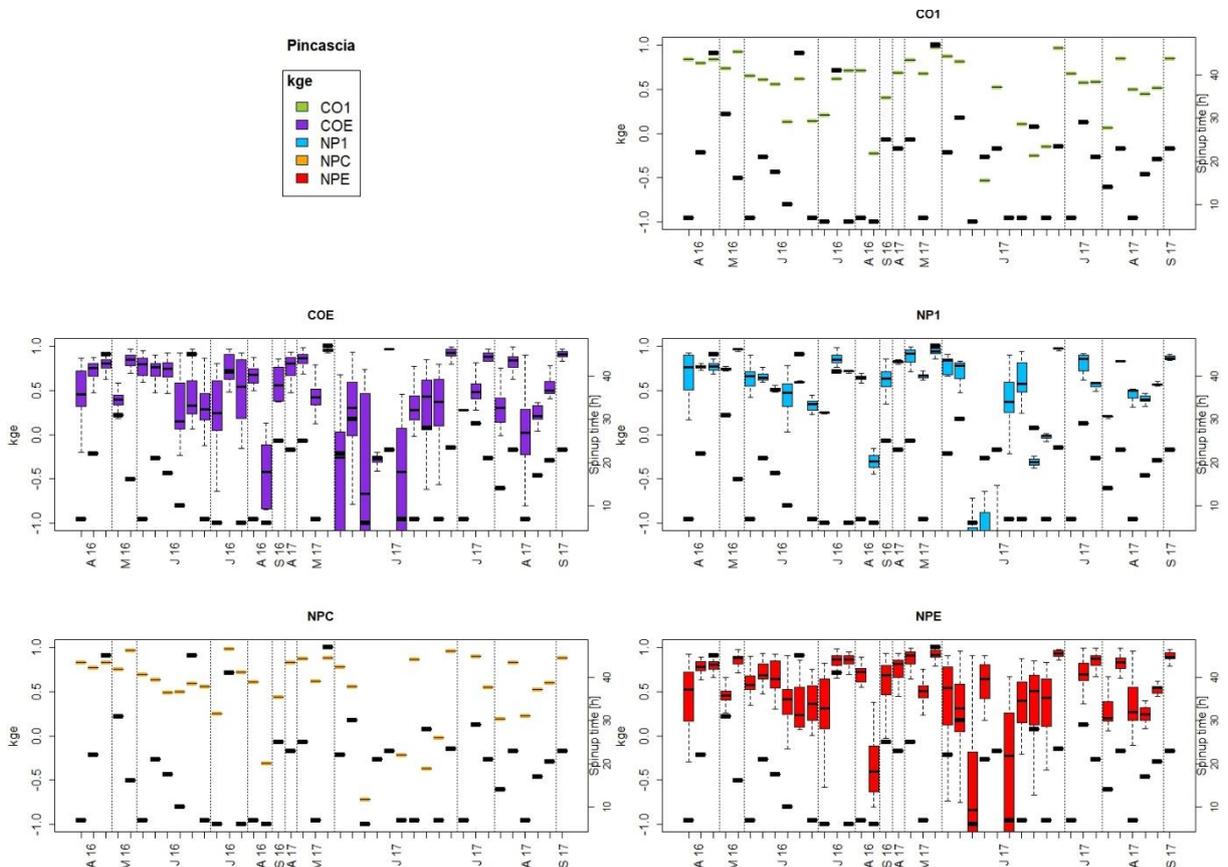


Figure A.9: Boxplot overview of the KGE values for all events in the Verzasca subcatchment including spinup time (bold bar). Letters and numbers in the X-axis representing months and year respectively.

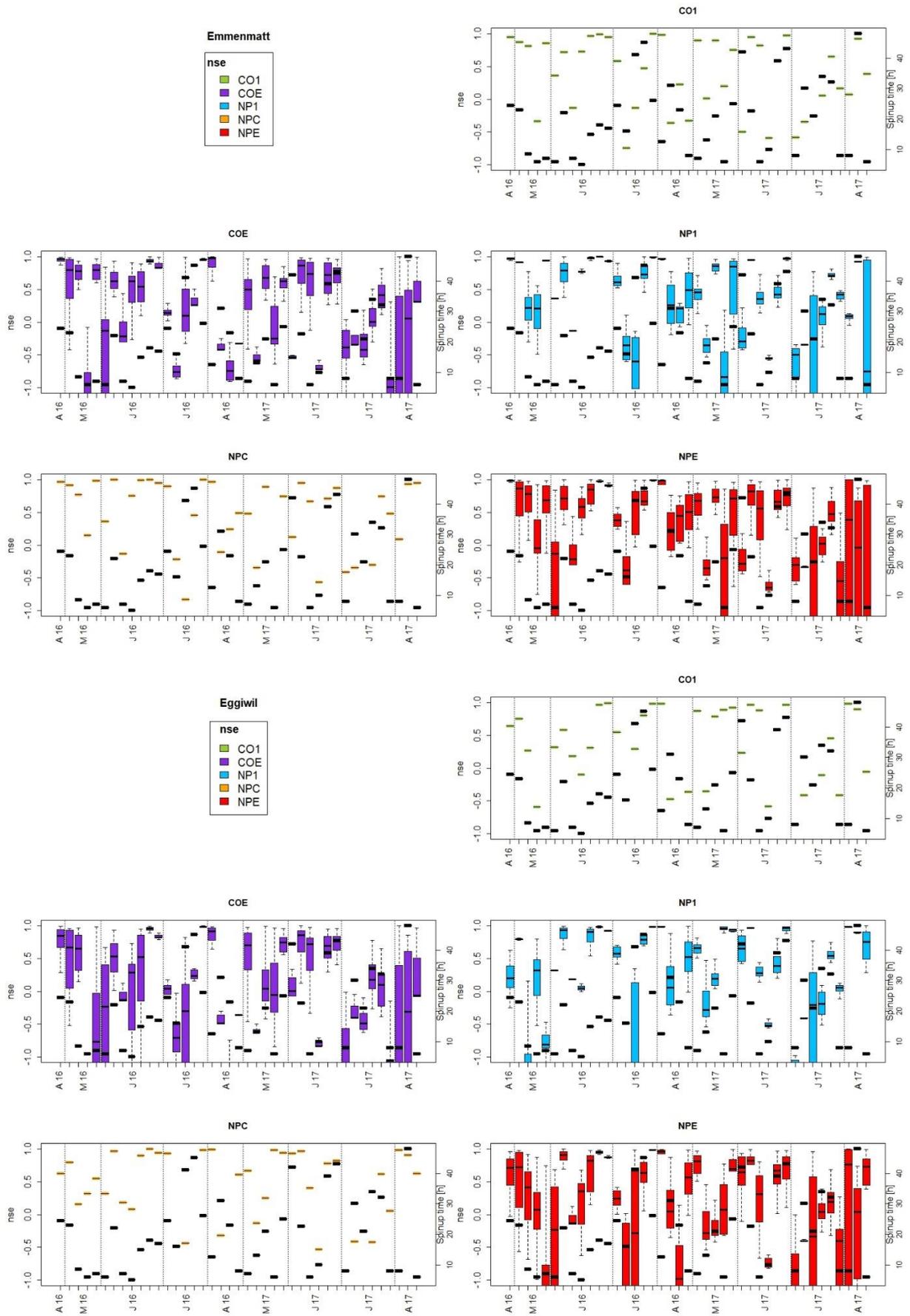


Figure continuous on next page.

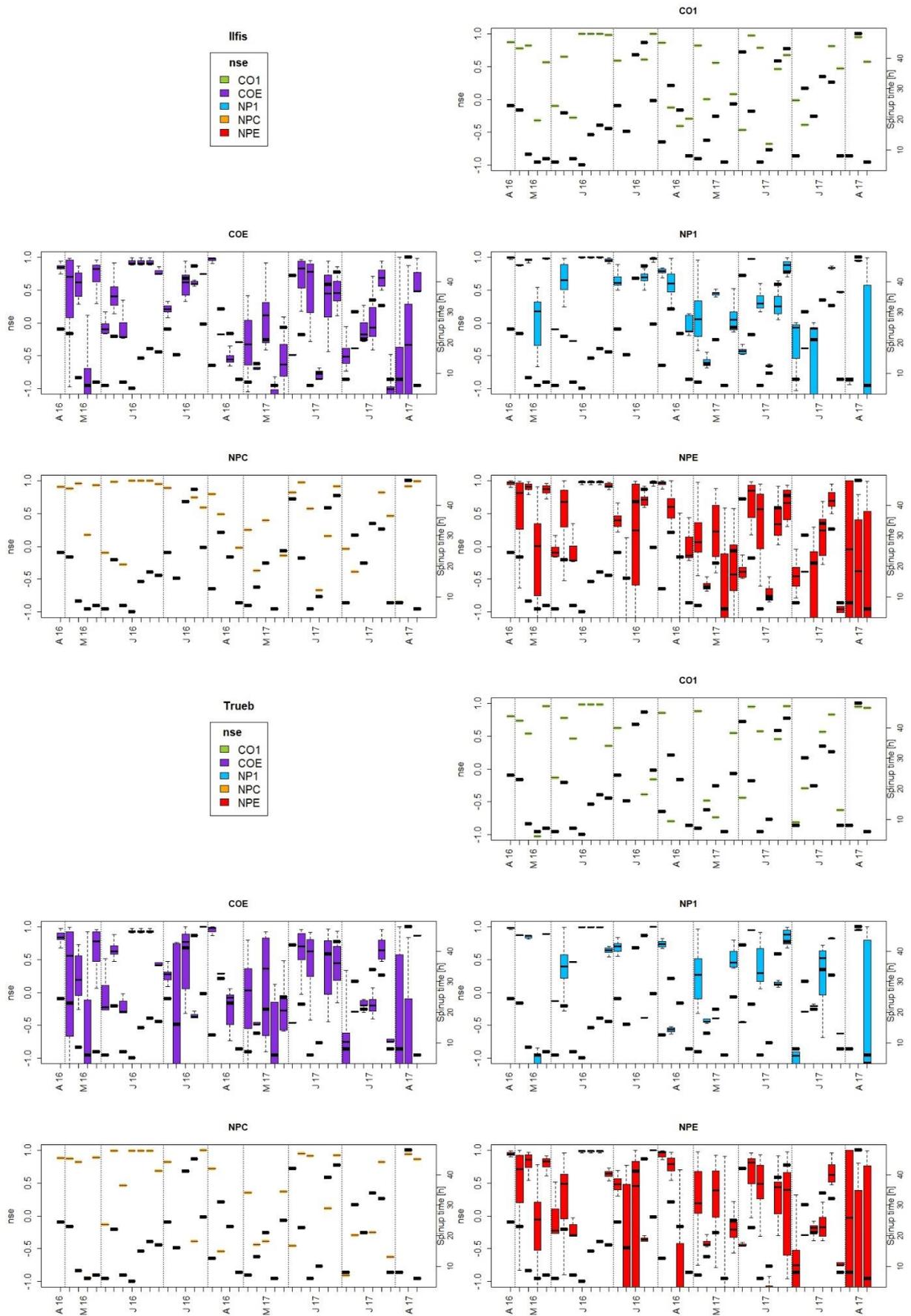


Figure A.10: Boxplot overview of the NSE values for all events in the Emme catchments including spinup time (bold bar). Letters and numbers in the X-axis representing months and year respectively.

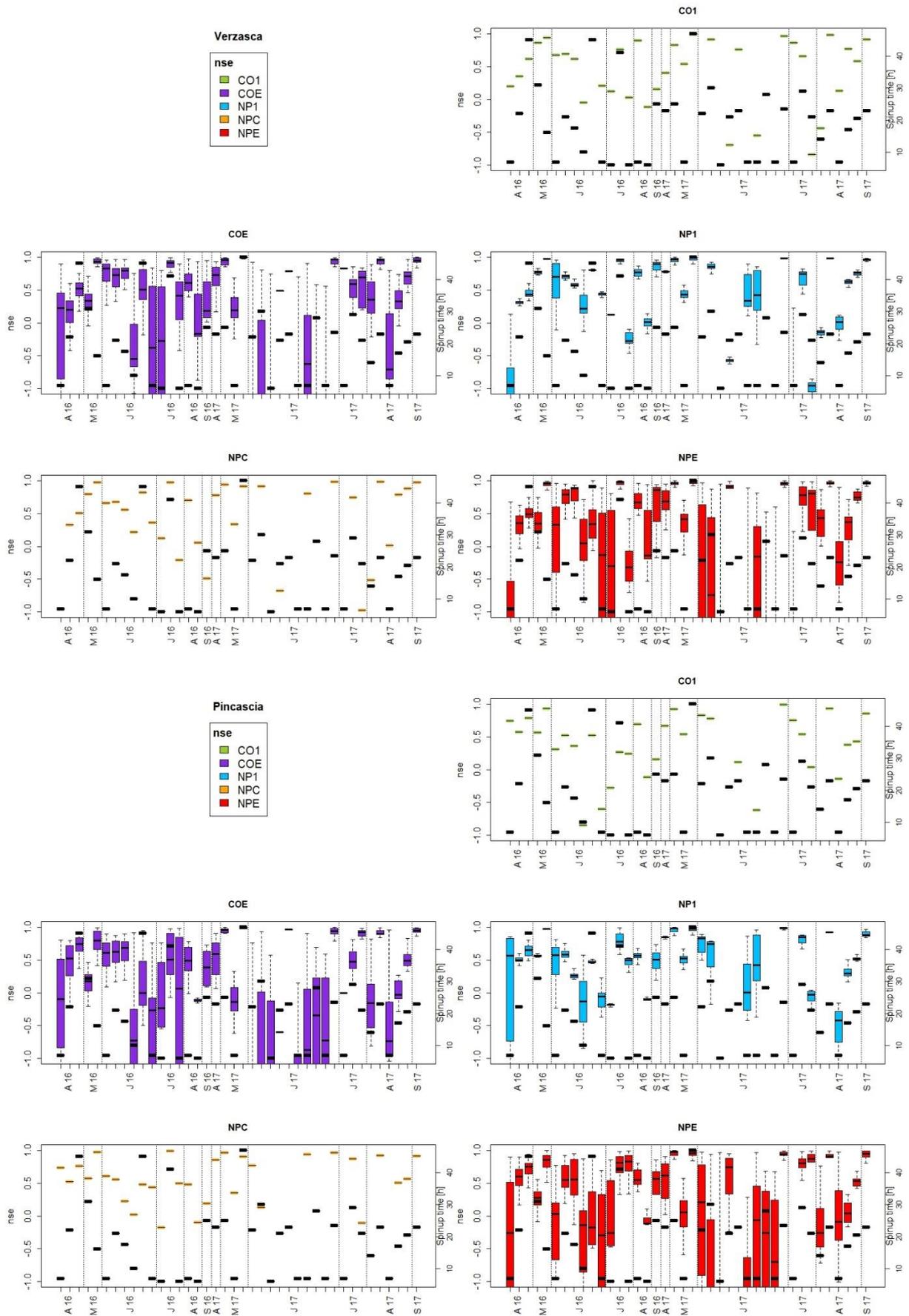


Figure A.11: Boxplot overview of the NSE values for all events in the Verzasca catchments including spinup time (bold bar). Letters and numbers in the X-axis representing months and year respectively.

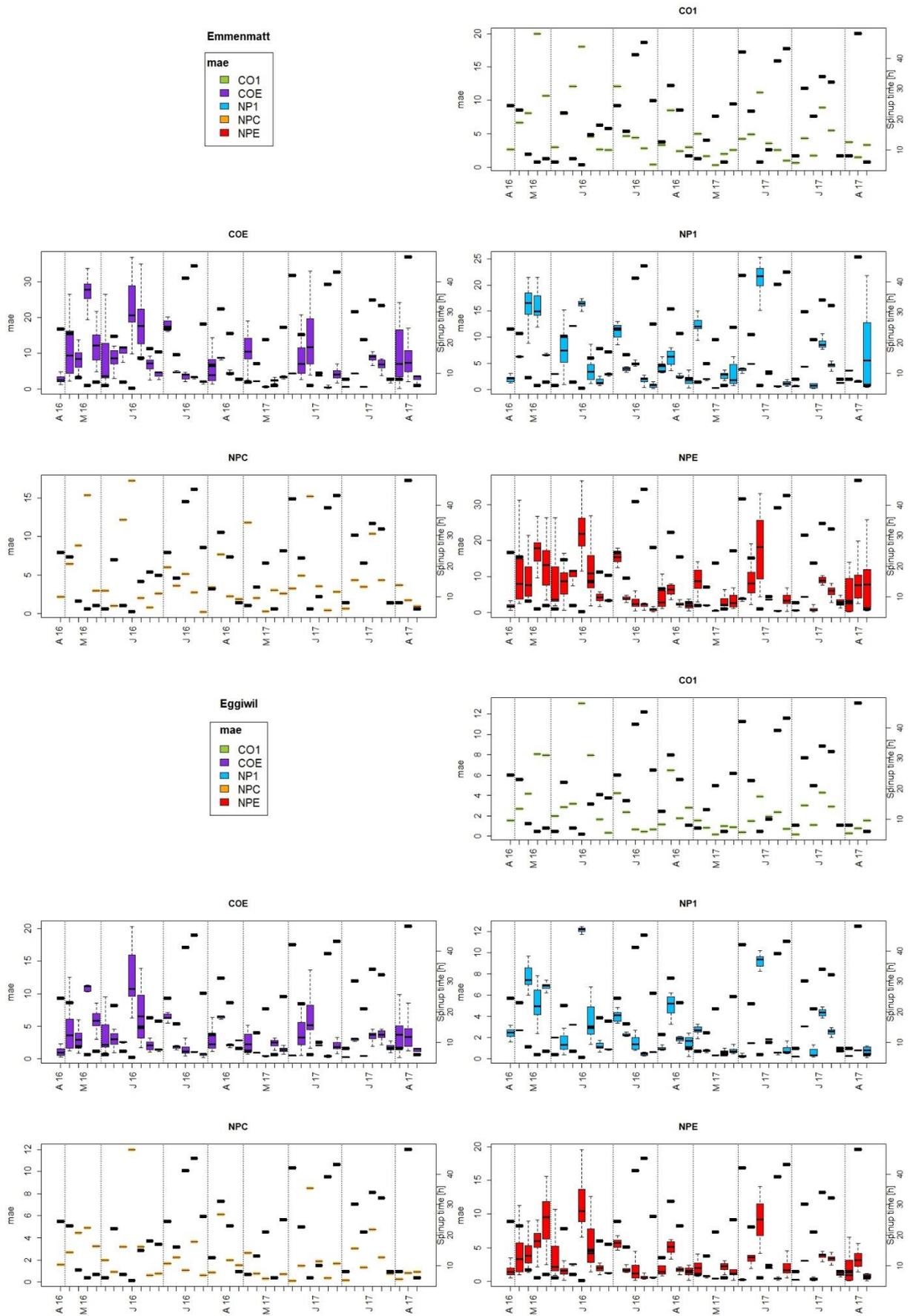


Figure continuous on next page.

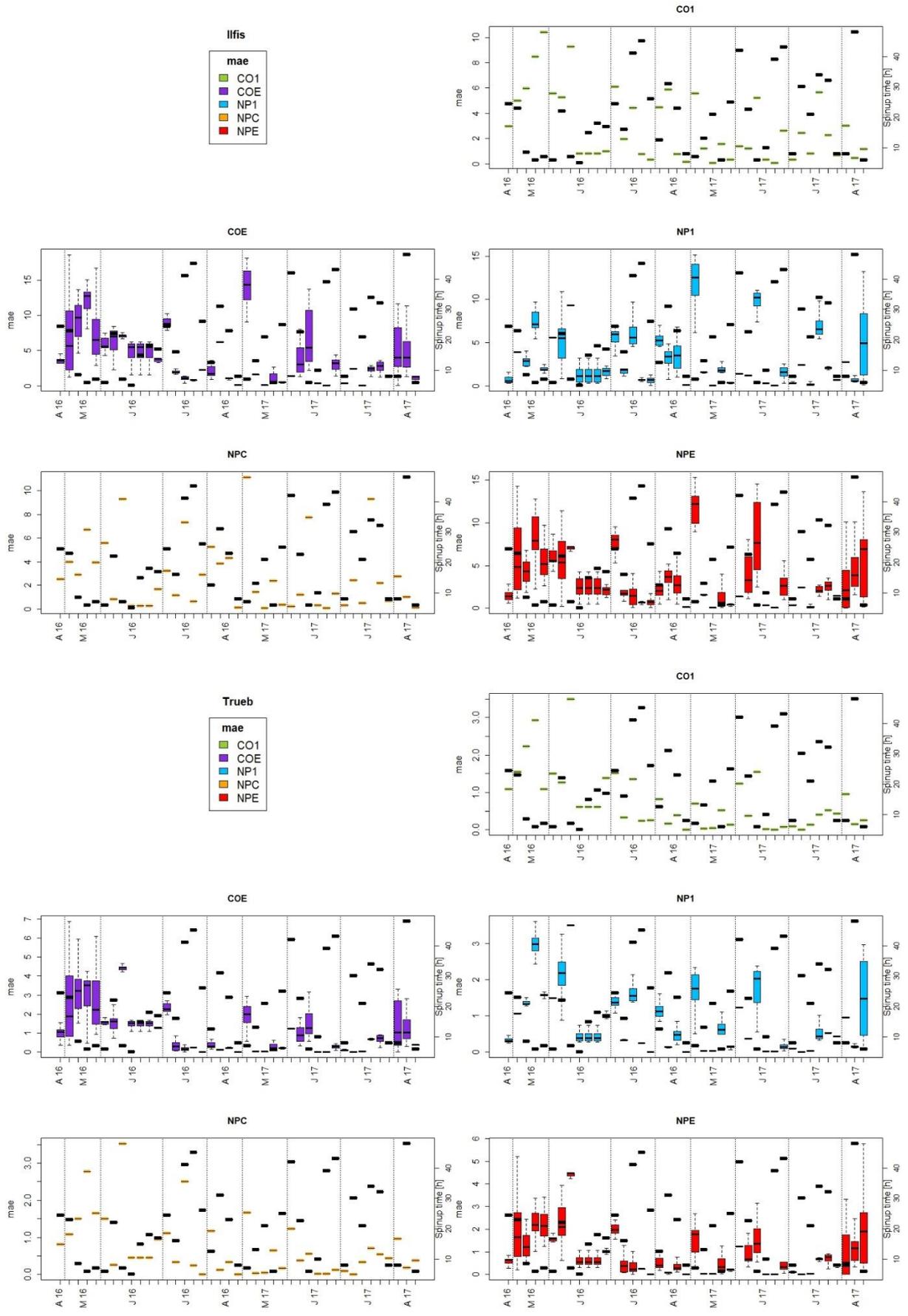


Figure A.12: Boxplot overview of the MAE values for all events in the Emme catchments including spinup time (bold bar). Letters and numbers in the X-axis representing months and year respectively.

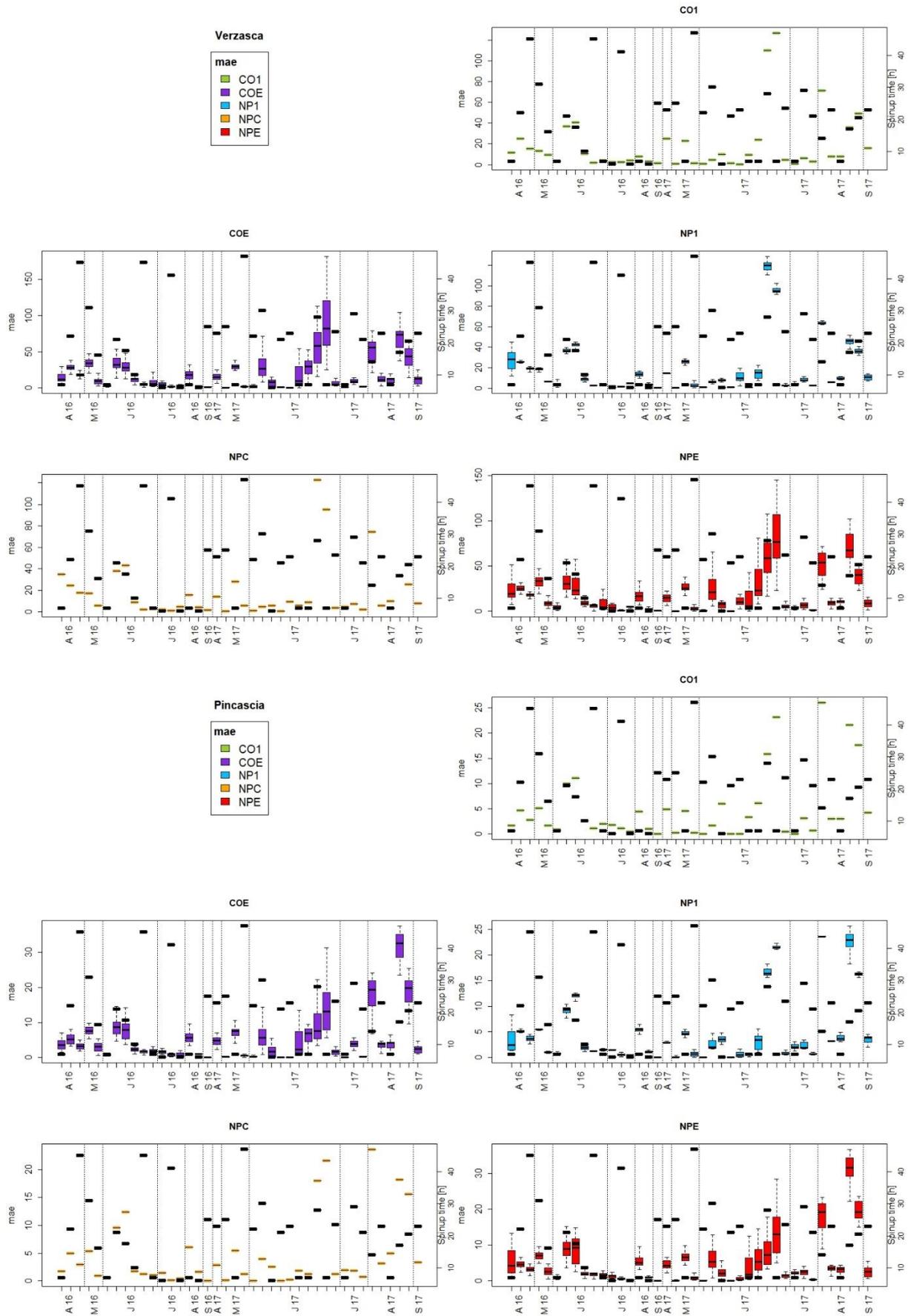


Figure A.13: Boxplot overview of the MAE values for all events in the Verzasca catchments including spinup time (bold bar). Letters and numbers in the X-axis representing months and year respectively.

7.4. Ranked Events KGE, NSE and MAE

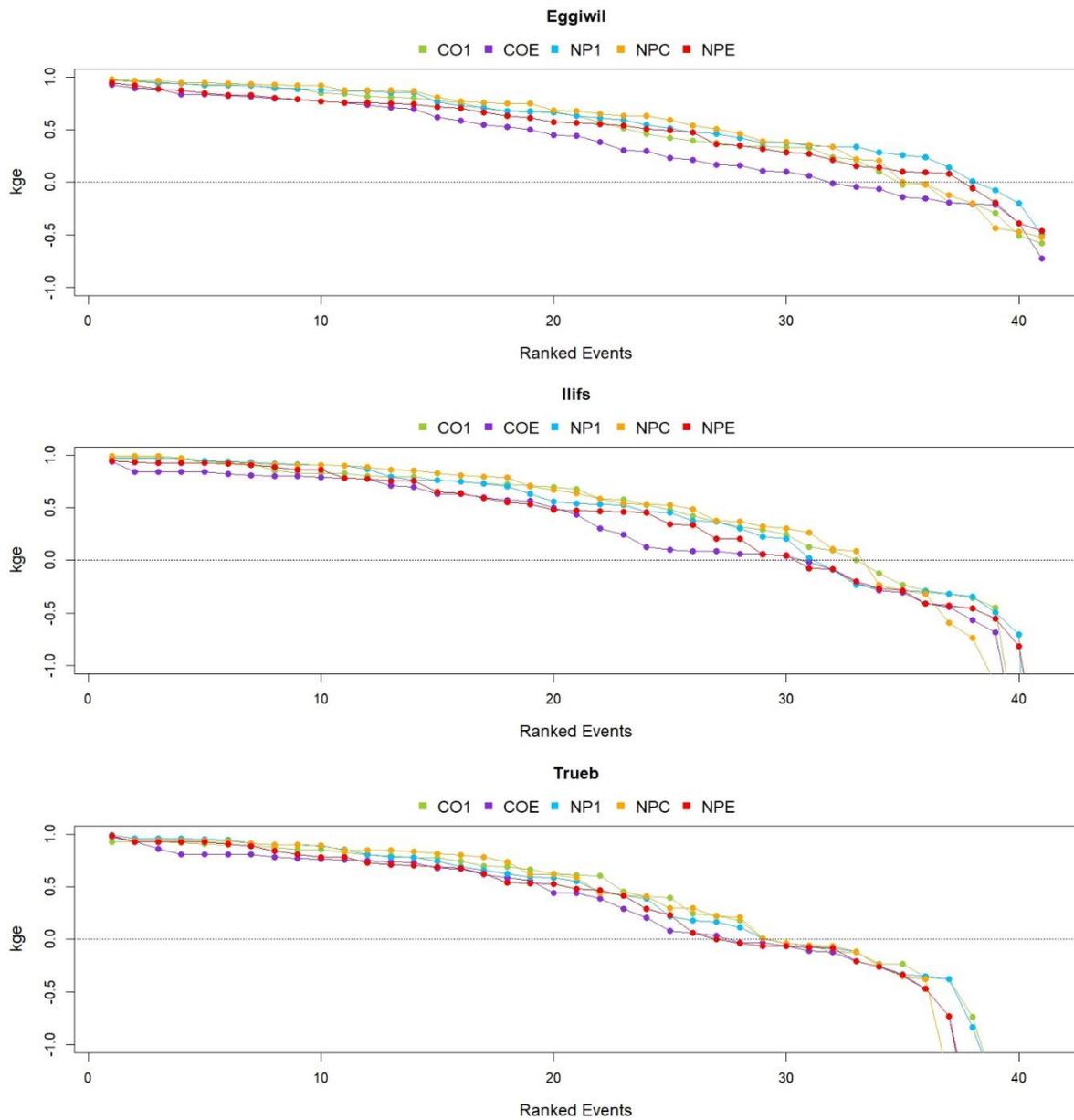


Figure A.14: Ranked KGE values shown for all nowcasting chains in the Emme subcatchments. Values above the horizontal line at a value of zero indicate skilfully predicted events.

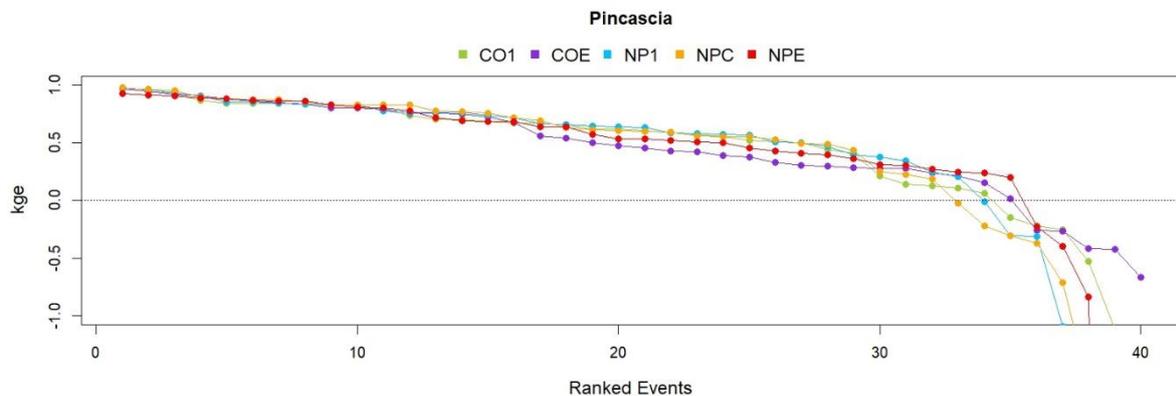


Figure A.15: Ranked KGE values shown for all nowcasting chains in the Verzasca subcatchment. Values above the horizontal line at a value of zero indicate skilfully predicted events.

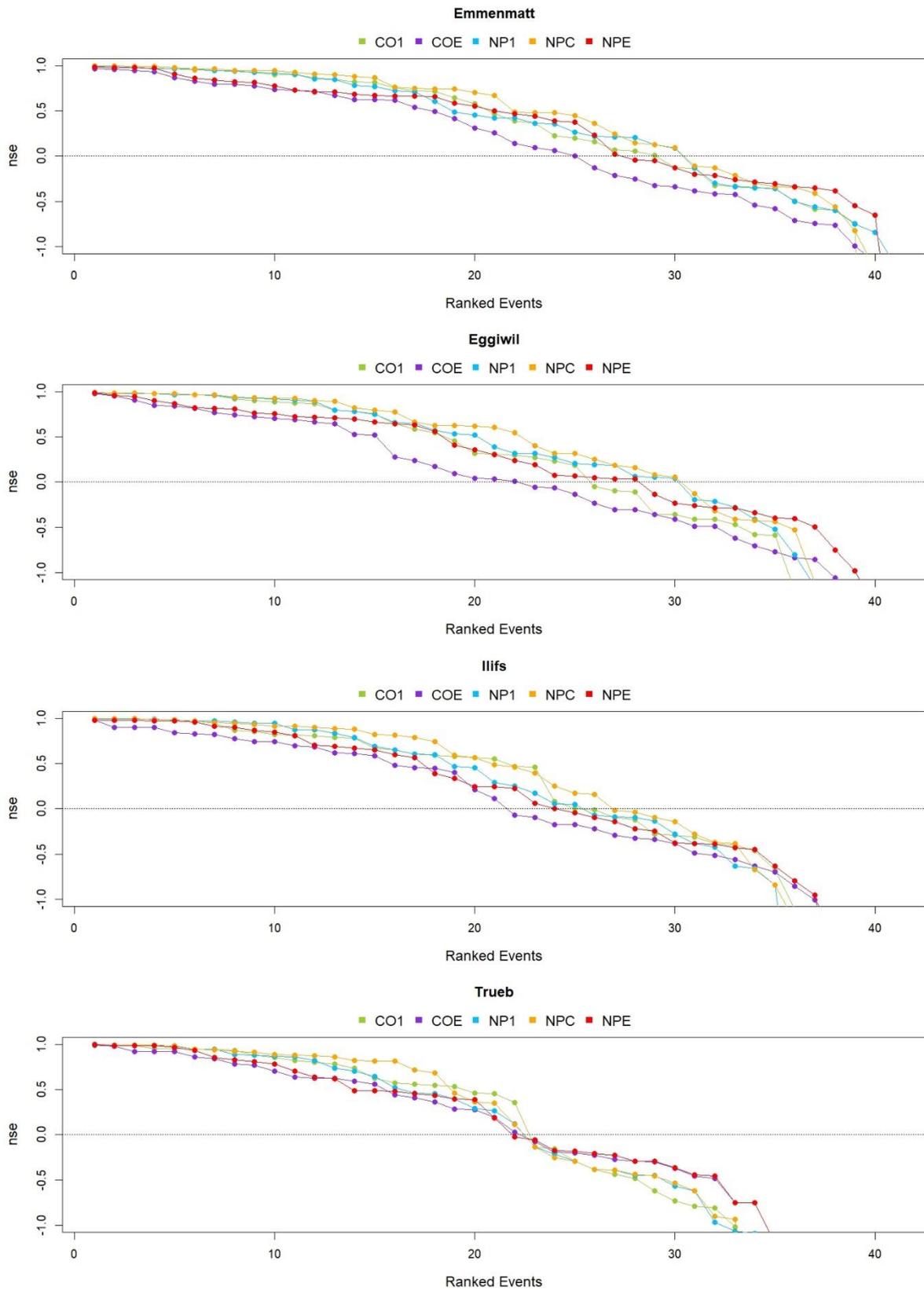


Figure A.16: Ranked NSE values shown for all nowcasting chains in the Emme catchments. Values above the horizontal line at a value of zero indicate skilfully predicted events.

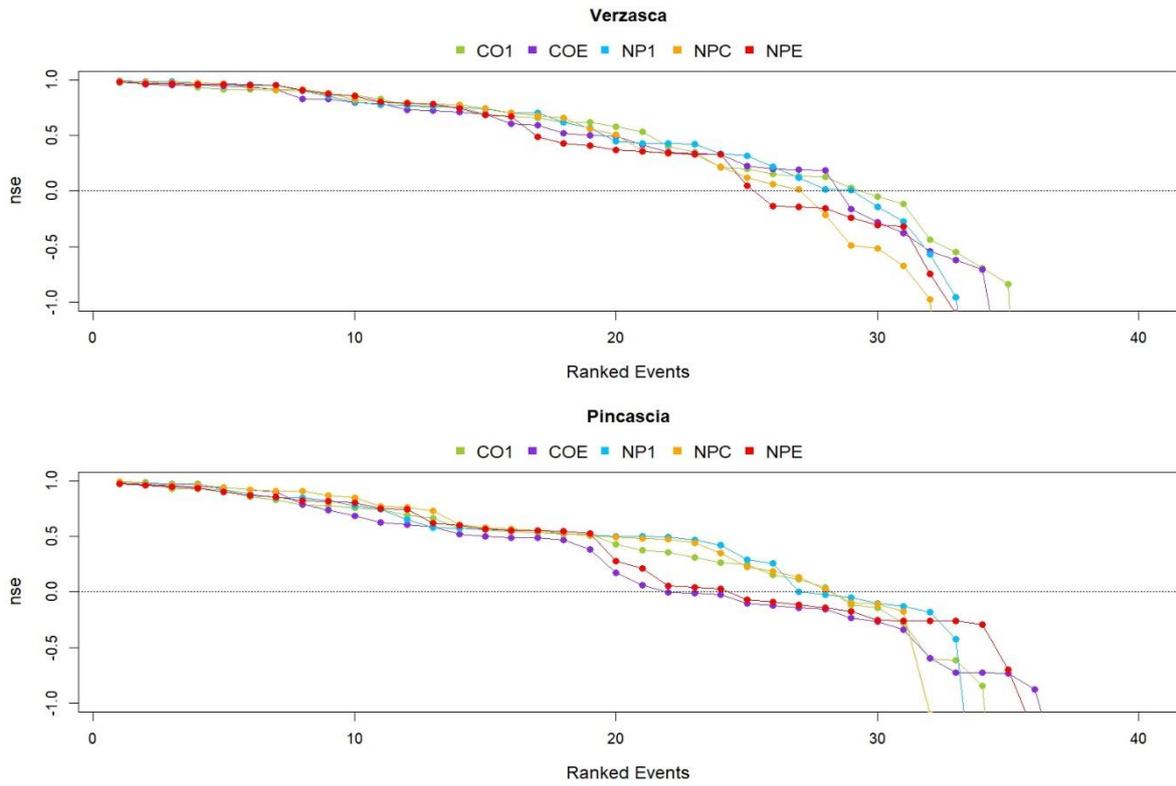


Figure A.17: Ranked NSE values shown for all nowcasting chains in the Verzasca catchments. Values above the horizontal line at a value of zero indicate skilfully predicted events.

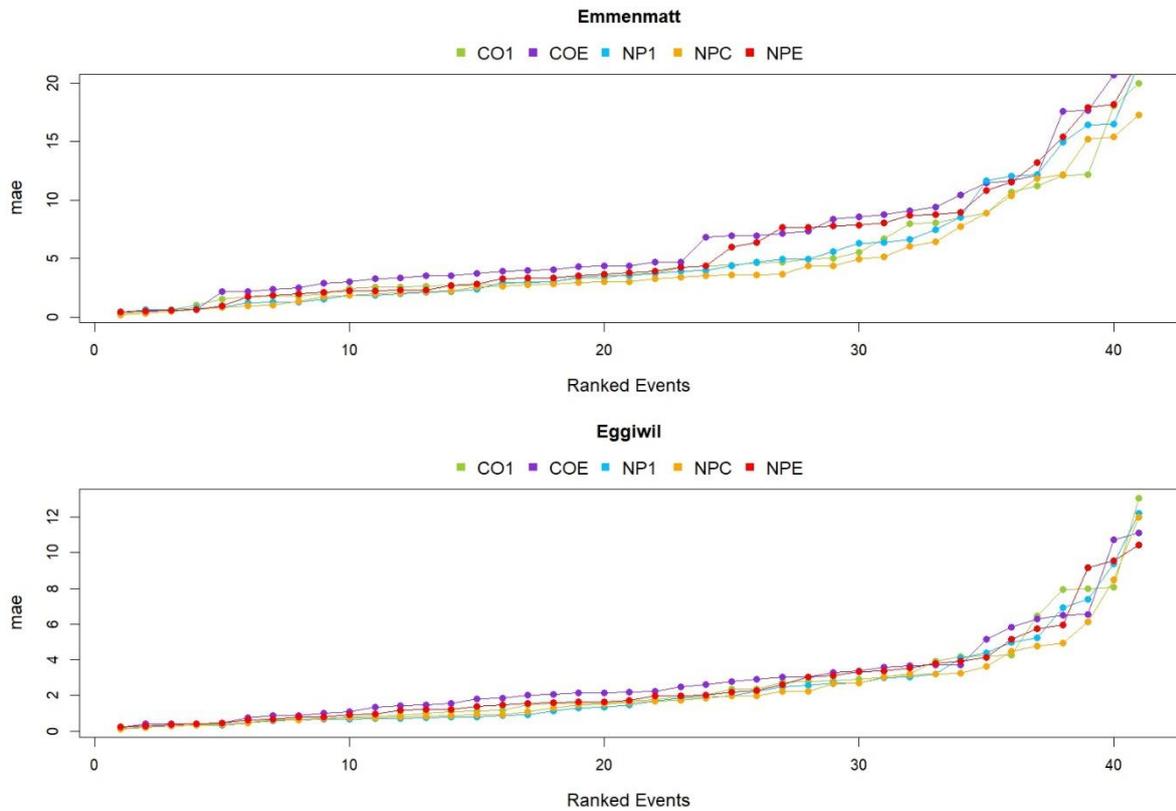


Figure continuous on next page.

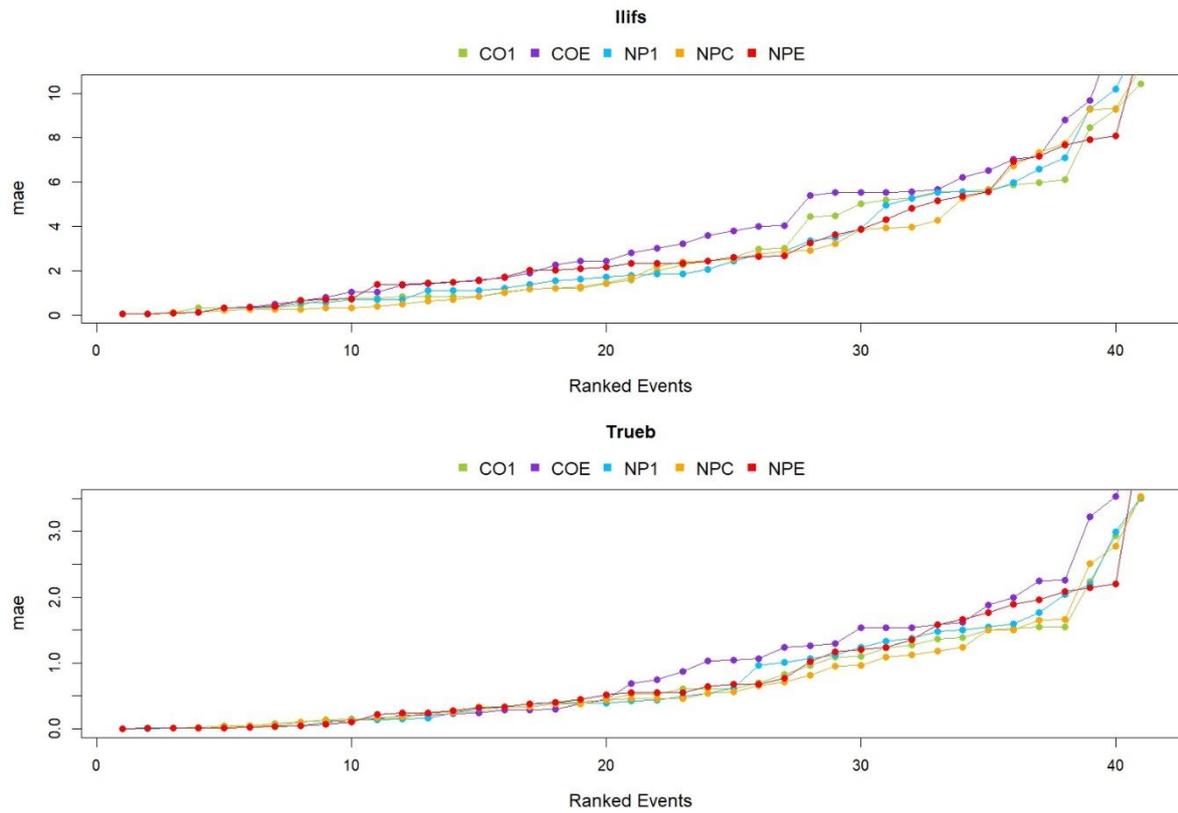


Figure A.18: Ranked MAE values shown for all nowcasting chains in the Emme catchments.

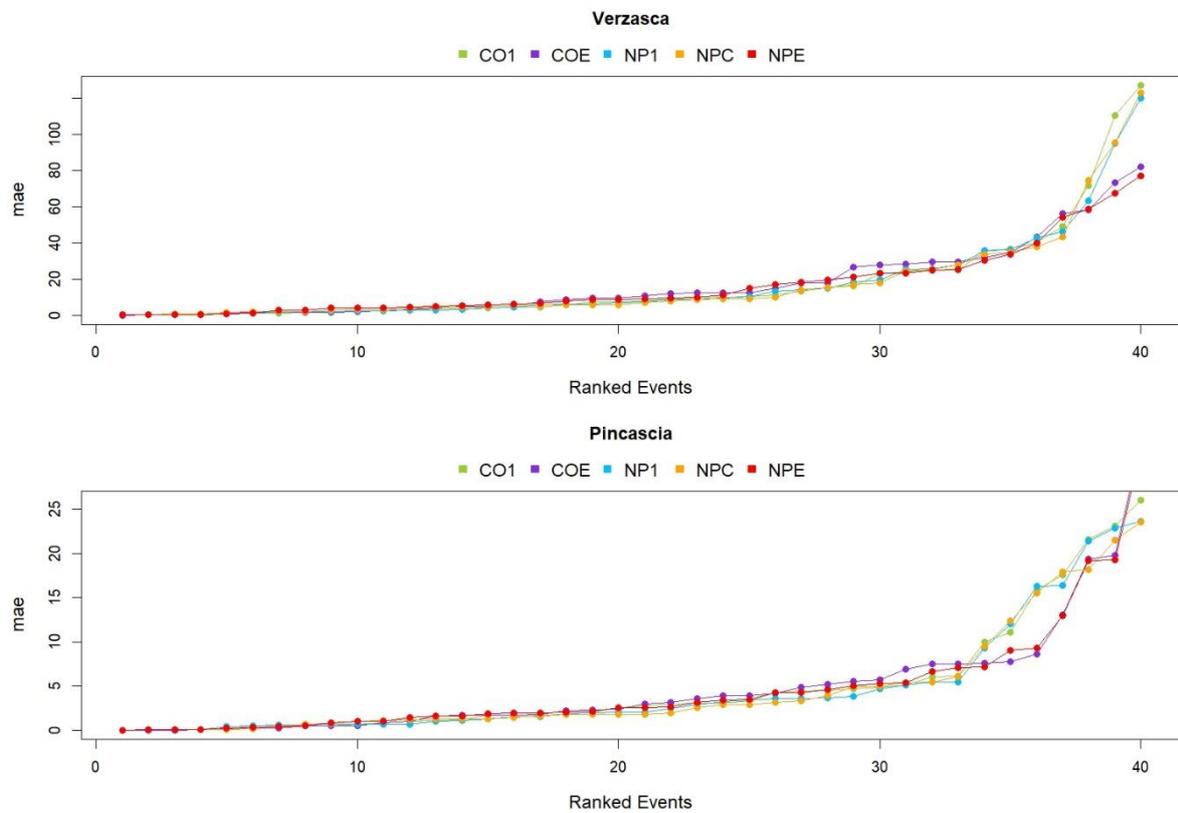


Figure A.19: Ranked MAE values shown for all nowcasting chains in the Emme catchments.

7.5. Best Performing Nowcasting Chains

7.5.1. All Nowcasting Chains NSE

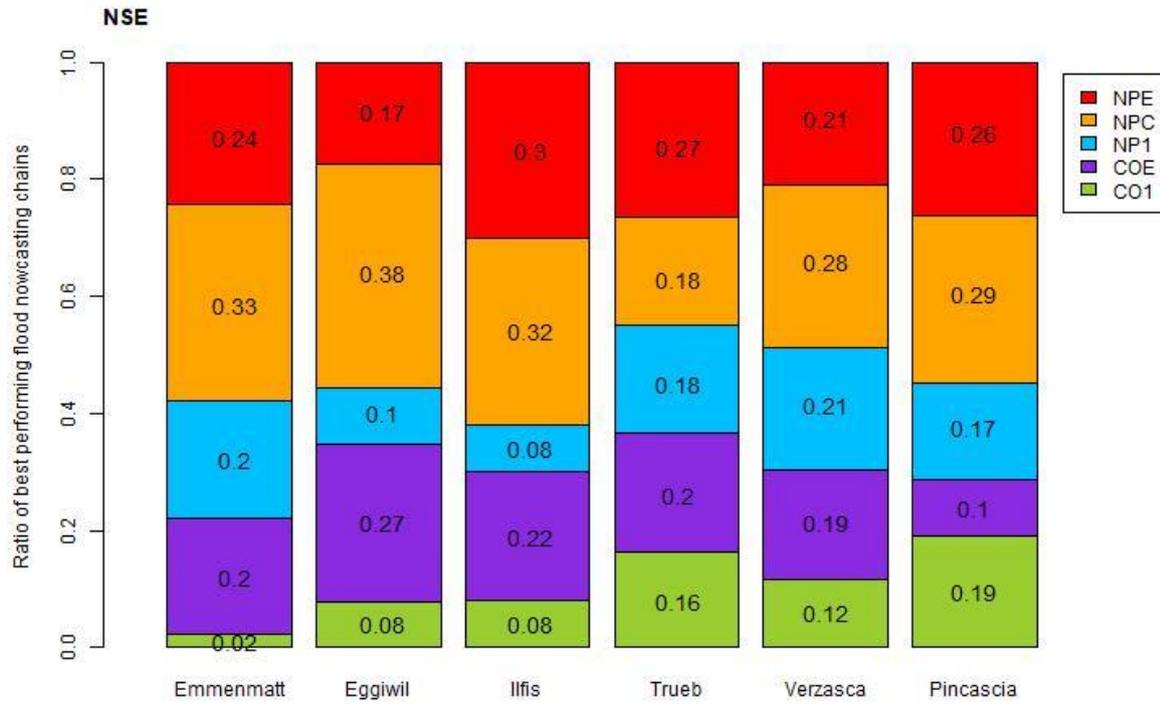


Figure A.20: Ratios of the best performing nowcasting chains in terms of the NSE.

7.5.2. Deterministic Nowcasting Chains NSE and MAE

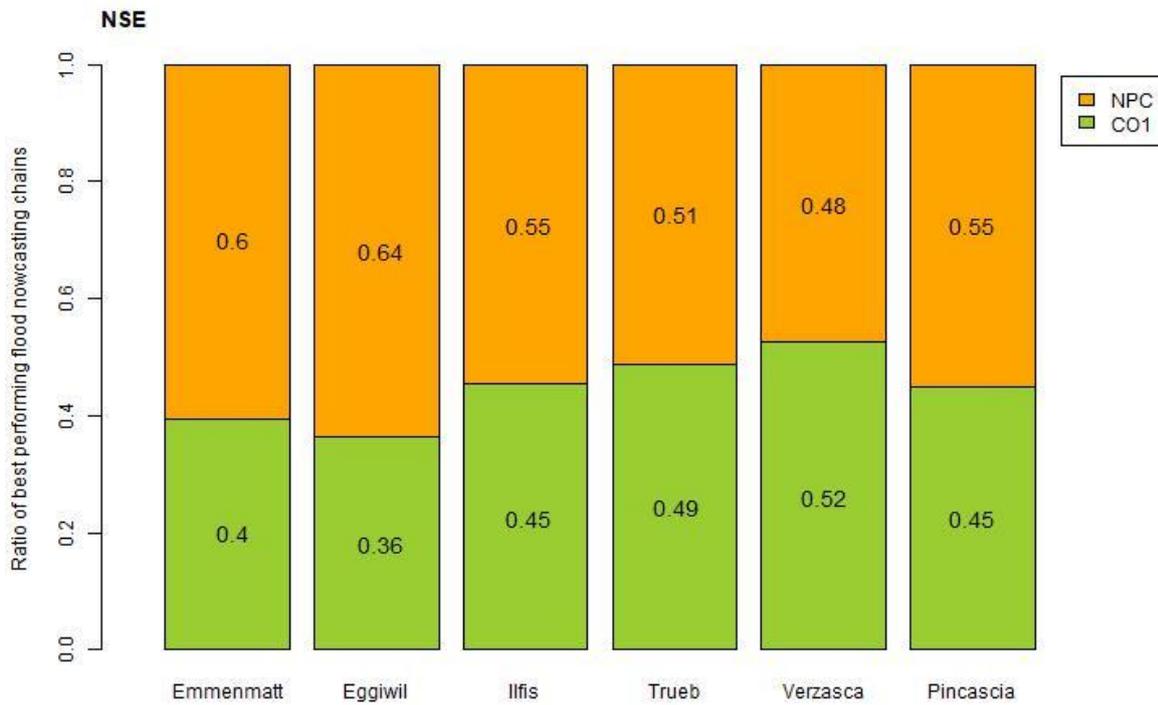


Figure A.21: Ratios of the best performing deterministic nowcasting chains in terms of the NSE.

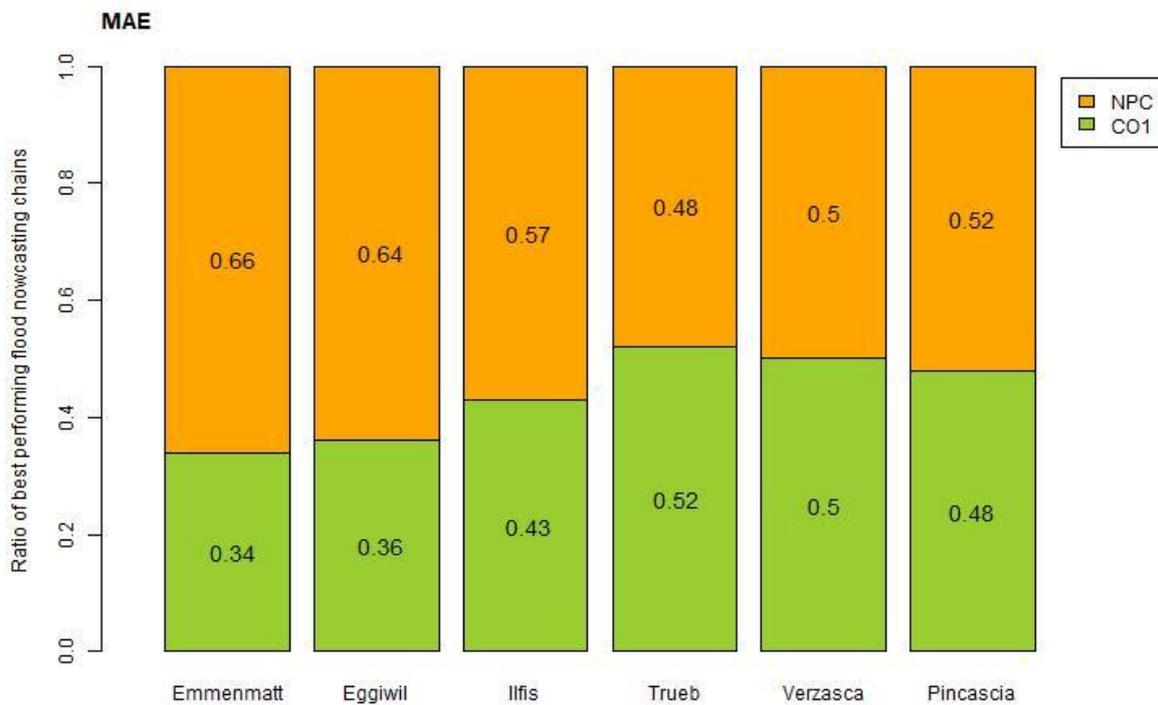


Figure A.22: Ratios of the best performing deterministic nowcasting chains in terms of the MAE.

7.5.3. Probabilistic Nowcasting Chains NSE and MAE

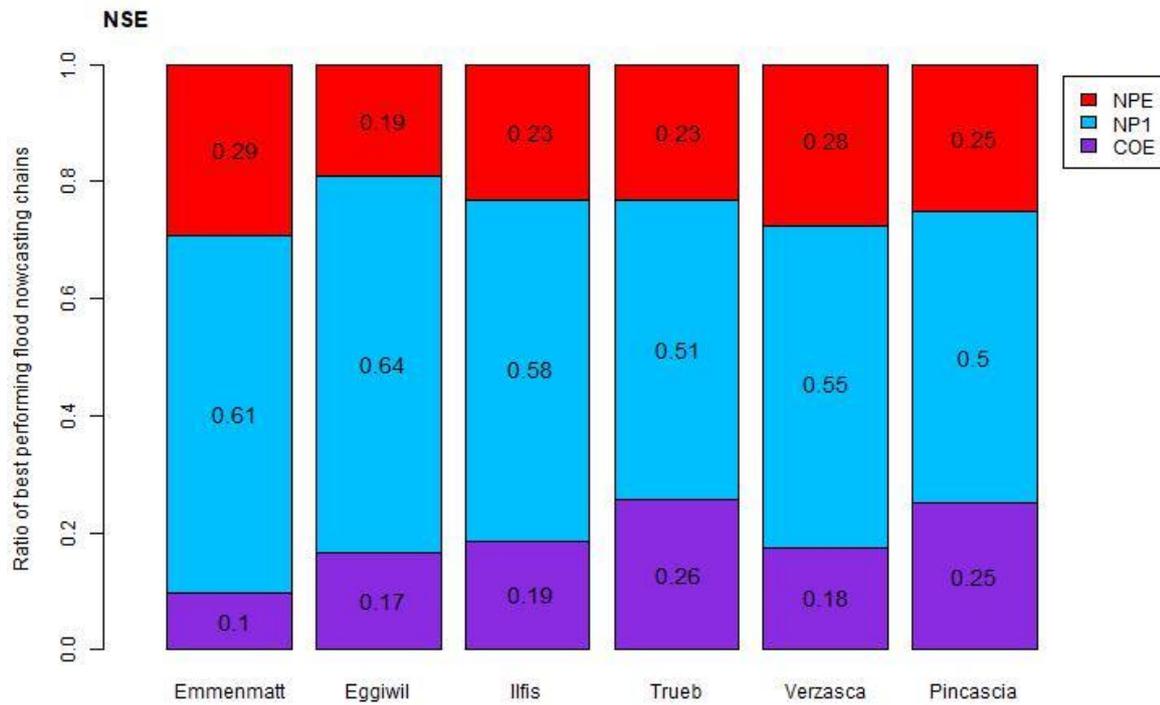


Figure A.23: Ratios of the best performing probabilistic nowcasting chains in terms of the NSE.

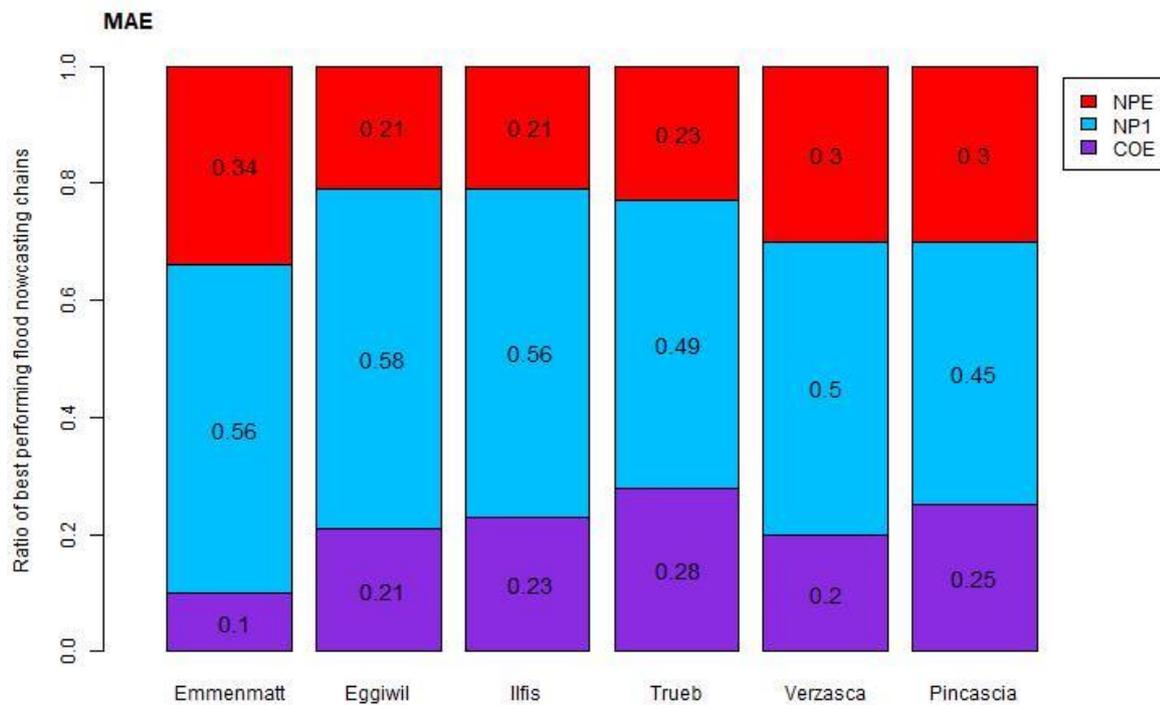


Figure A.24: Ratios of the best performing probabilistic nowcasting chains in terms of the MAE.

7.6. Entire Update Cycle

7.6.1. Accumulated Precipitation Prediction

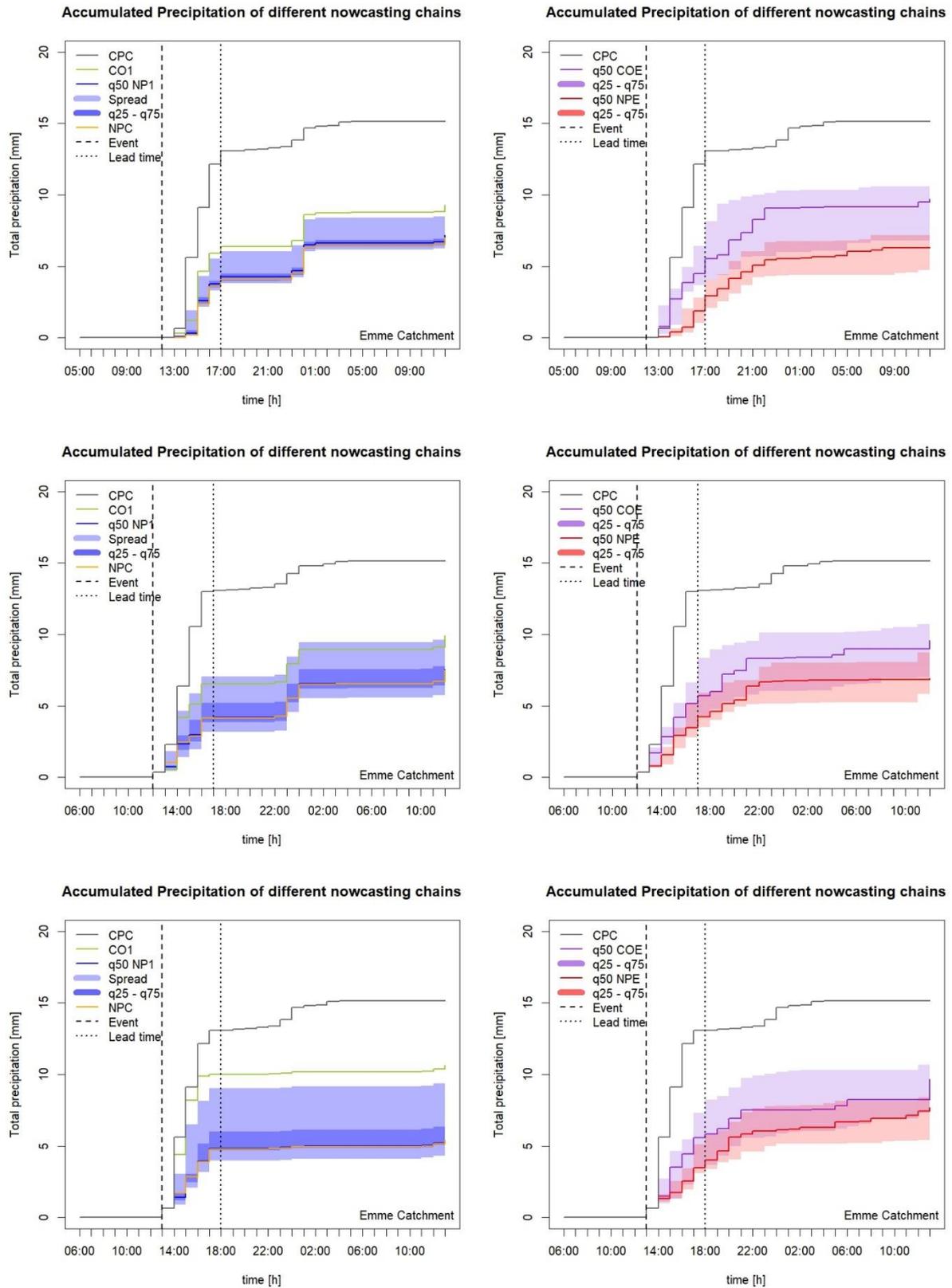


Figure continuous on the next page.

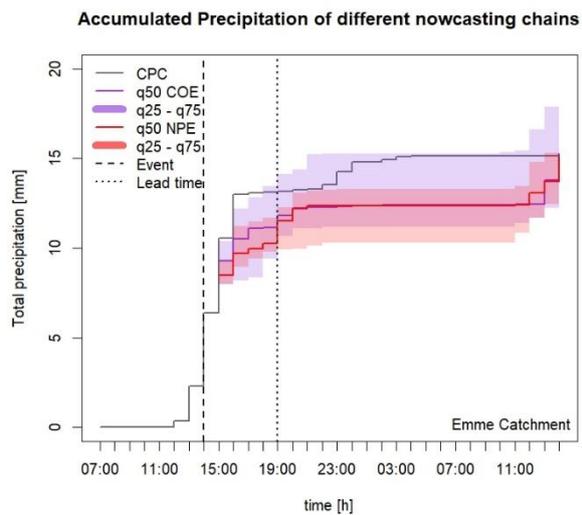
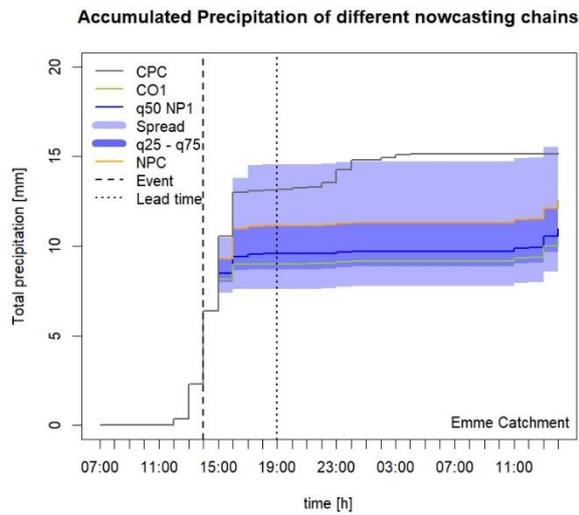
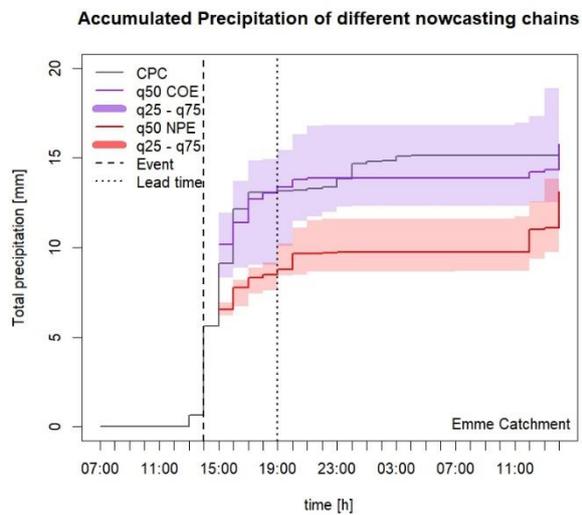
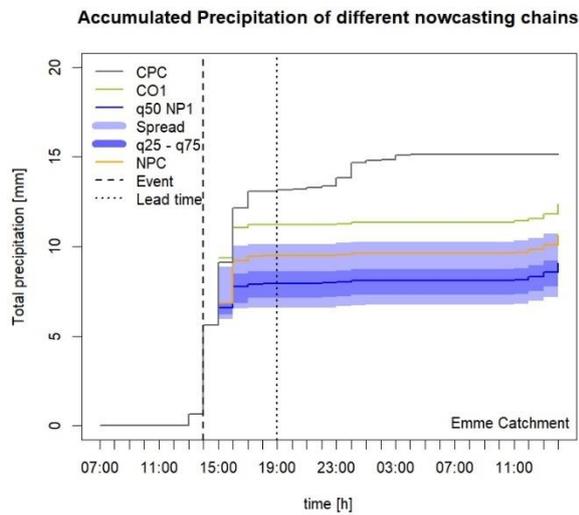
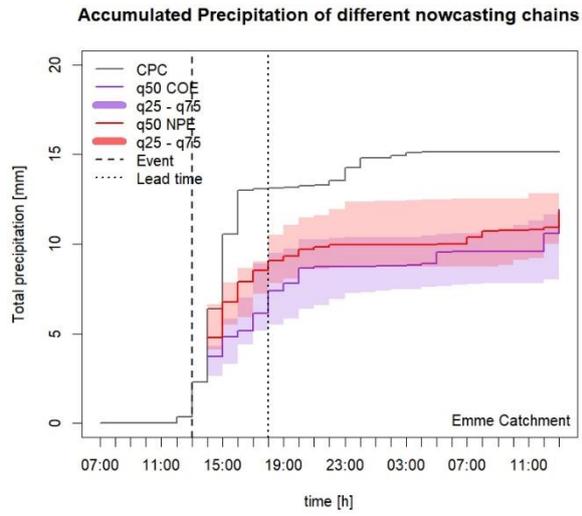
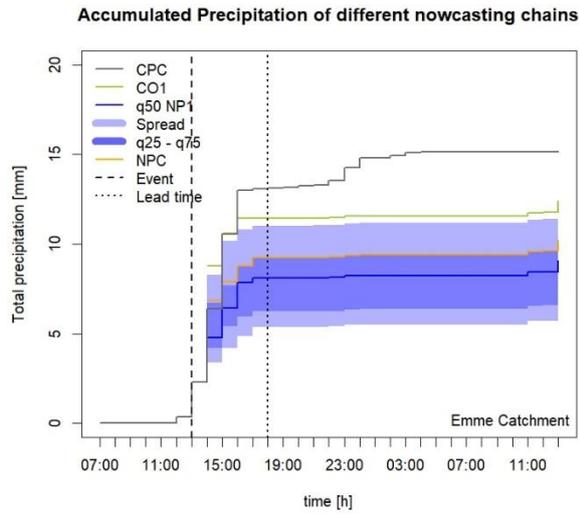


Figure continuous on the next page.

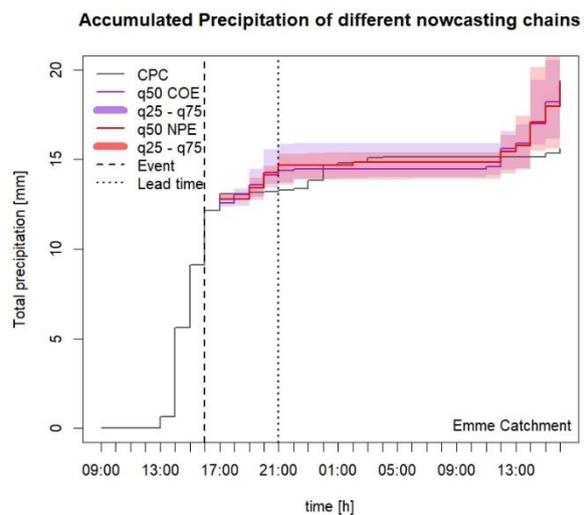
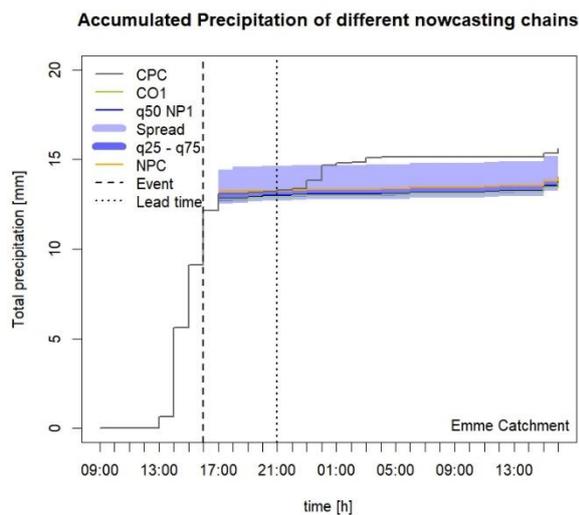
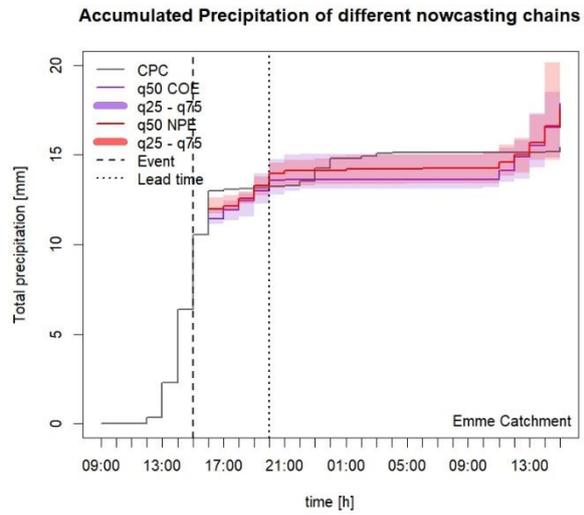
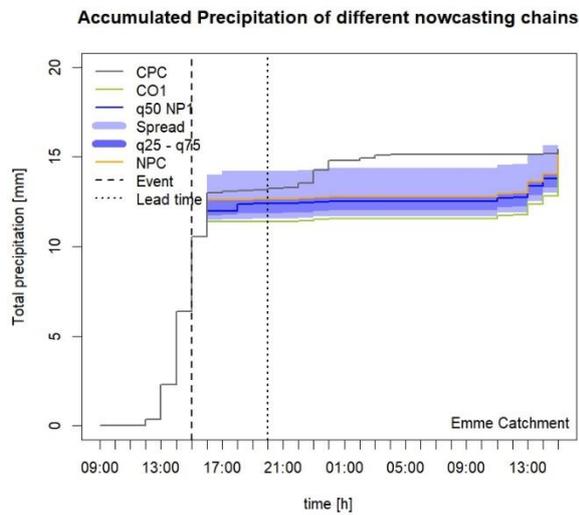
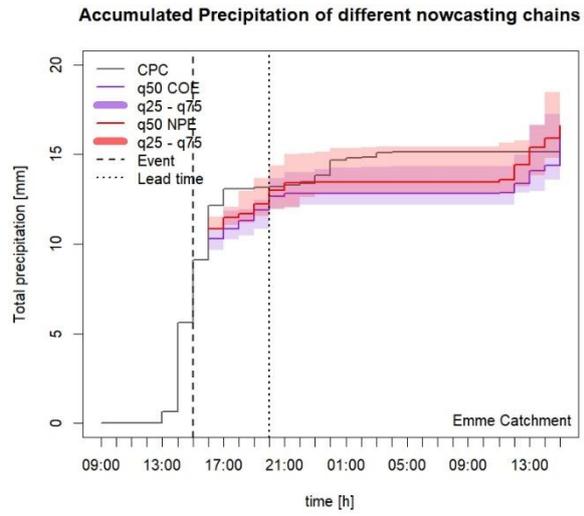
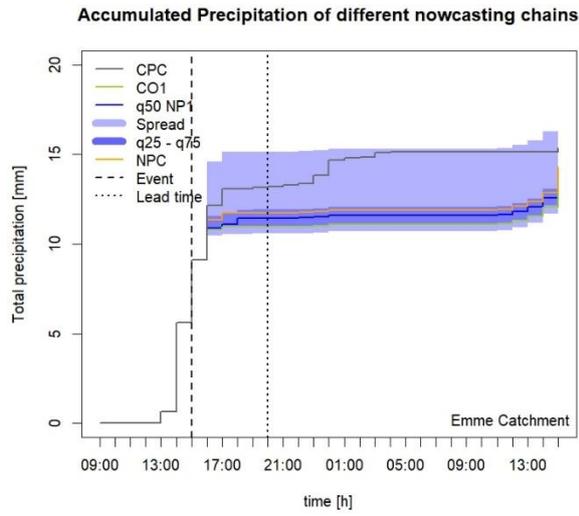


Figure continuous on the next page.

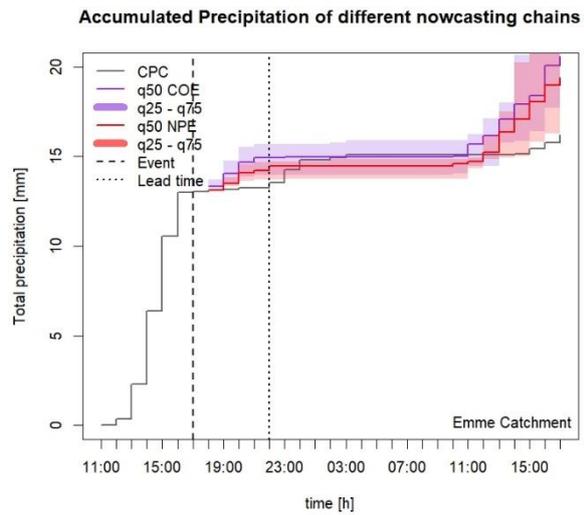
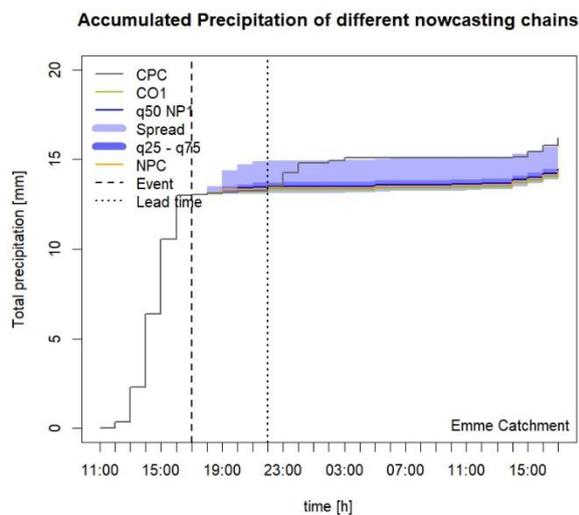
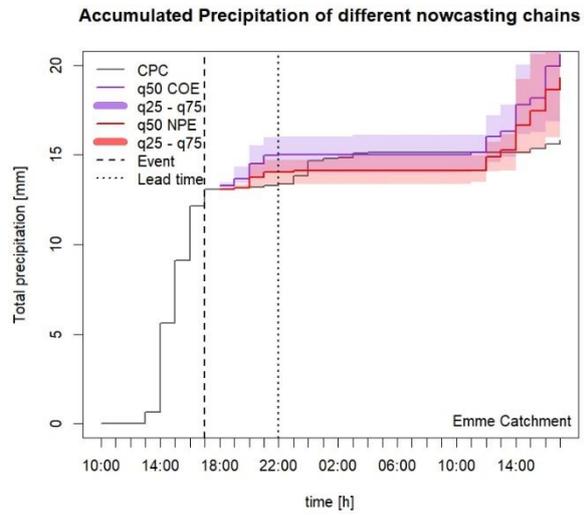
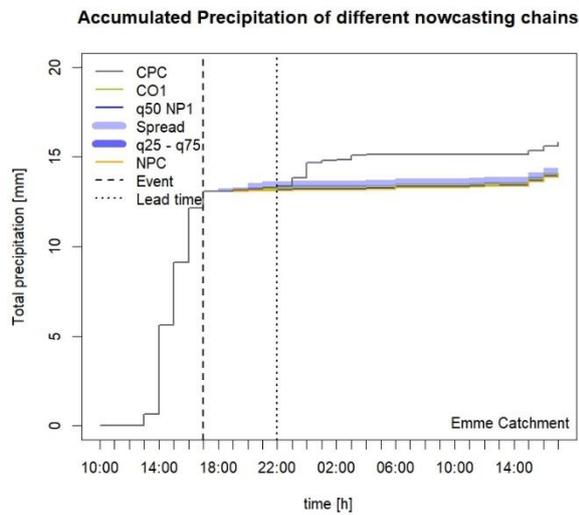
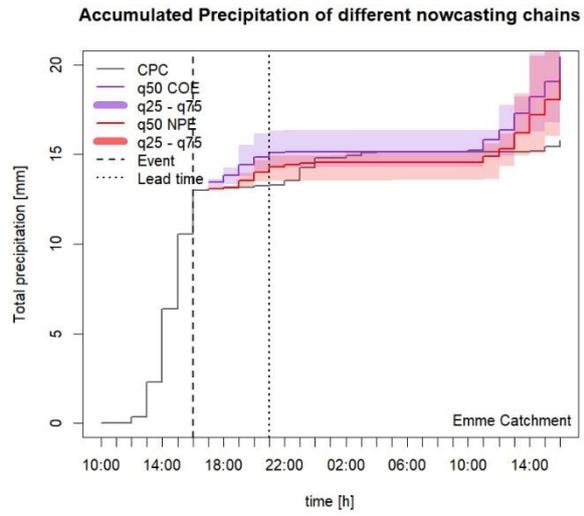
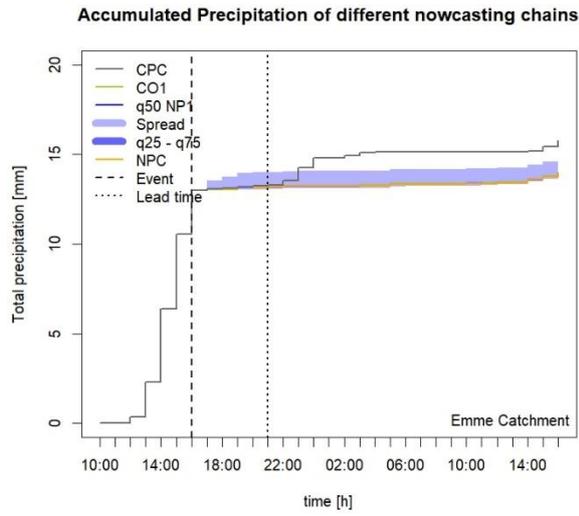


Figure continuous on the next page.

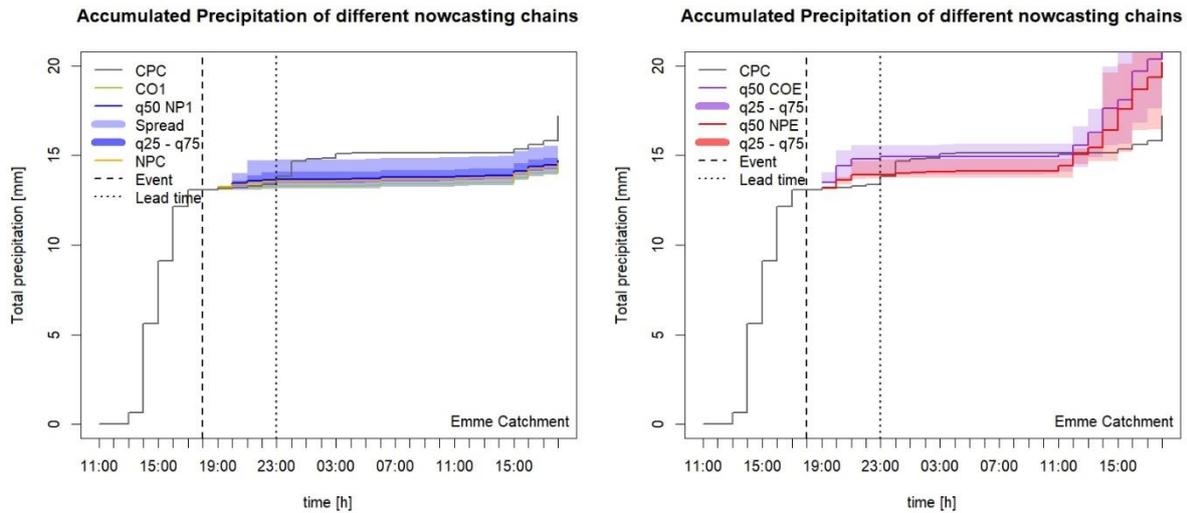


Figure A.25: Accumulated precipitation for COSMO-1 related nowcasting chains (left) and COSMO-E related nowcasting chains (right). From top to bottom: The first prediction at 13:00 to the last prediction at 19:00 of the update cycle is shown for the Emme catchment.

7.6.2. Runoff Simulation

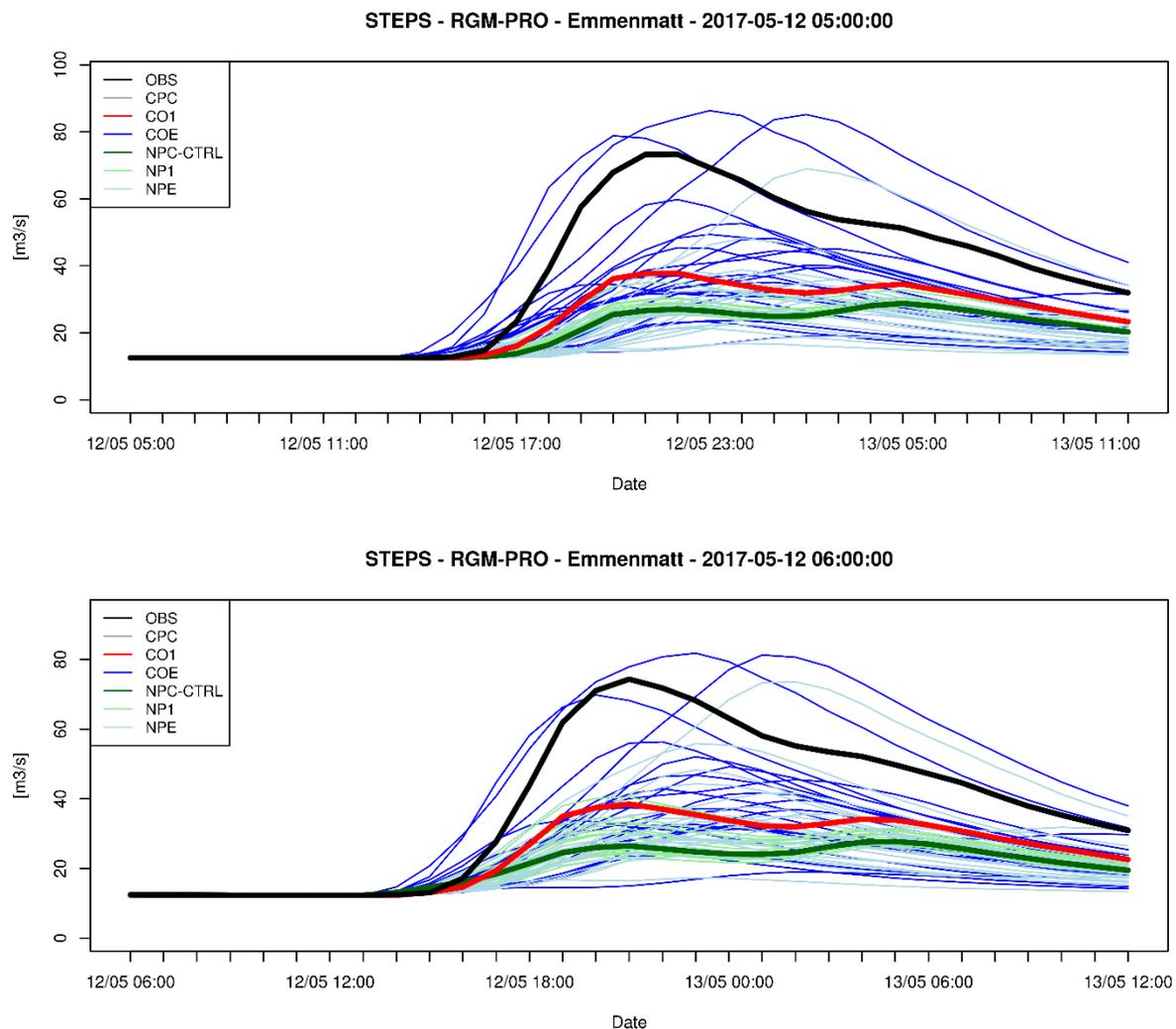
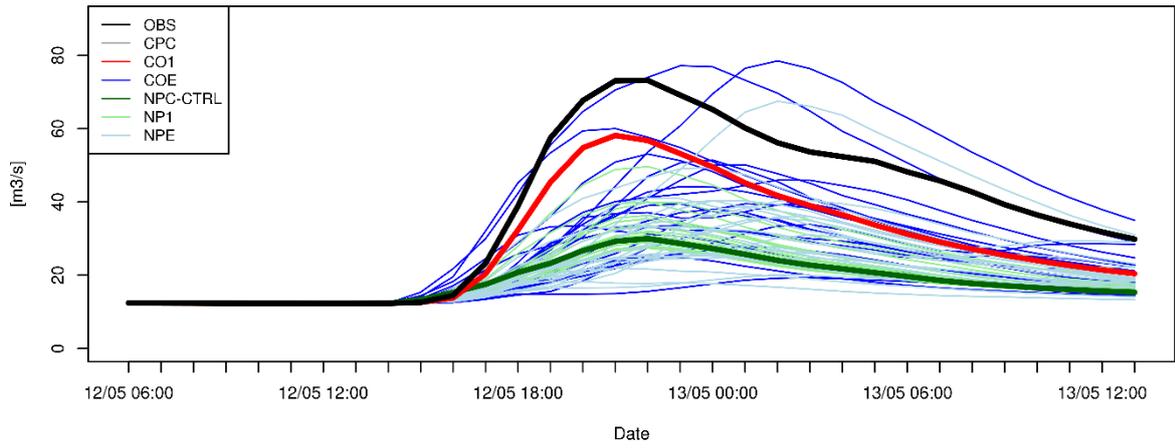
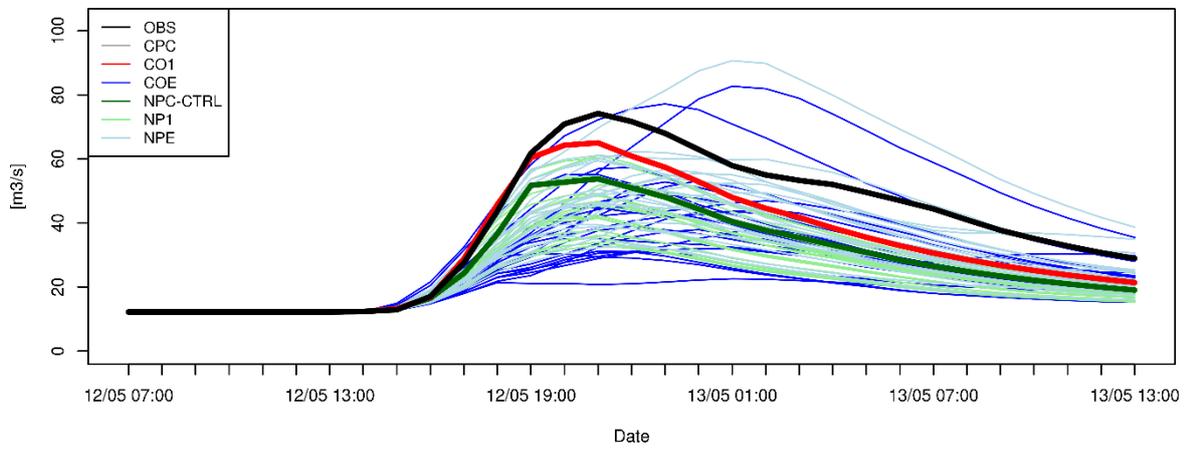


Figure continuous on the next page.

STEPS - RGM-PRO - Emmenmatt - 2017-05-12 06:00:00



STEPS - RGM-PRO - Emmenmatt - 2017-05-12 07:00:00



STEPS - RGM-PRO - Emmenmatt - 2017-05-12 07:00:00

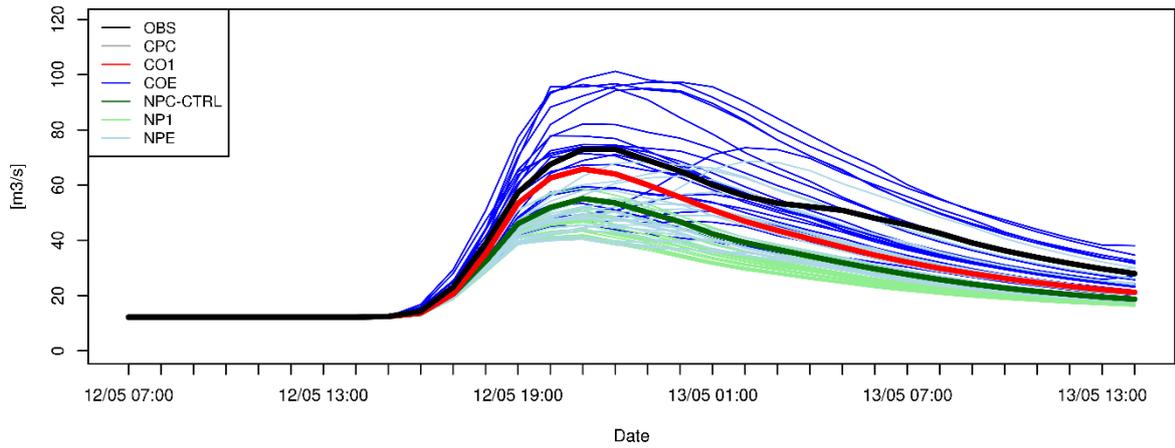
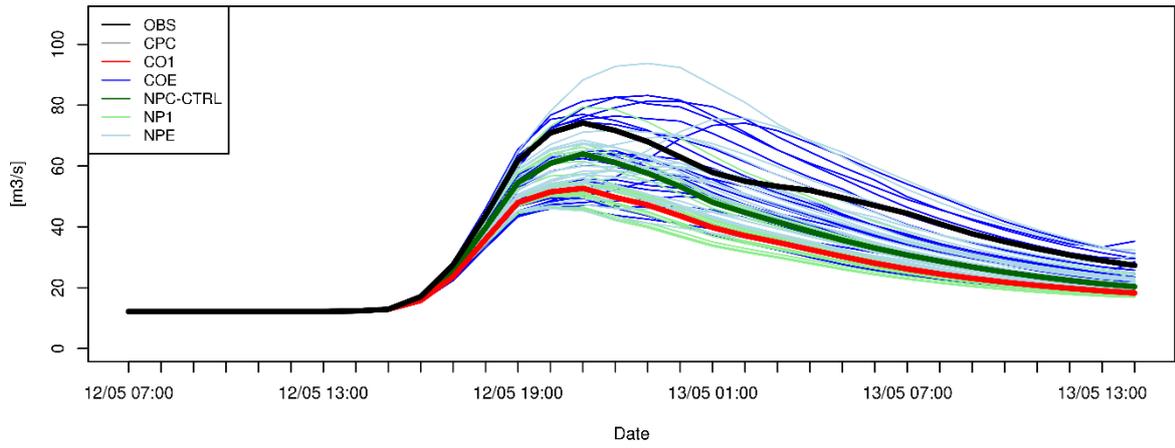
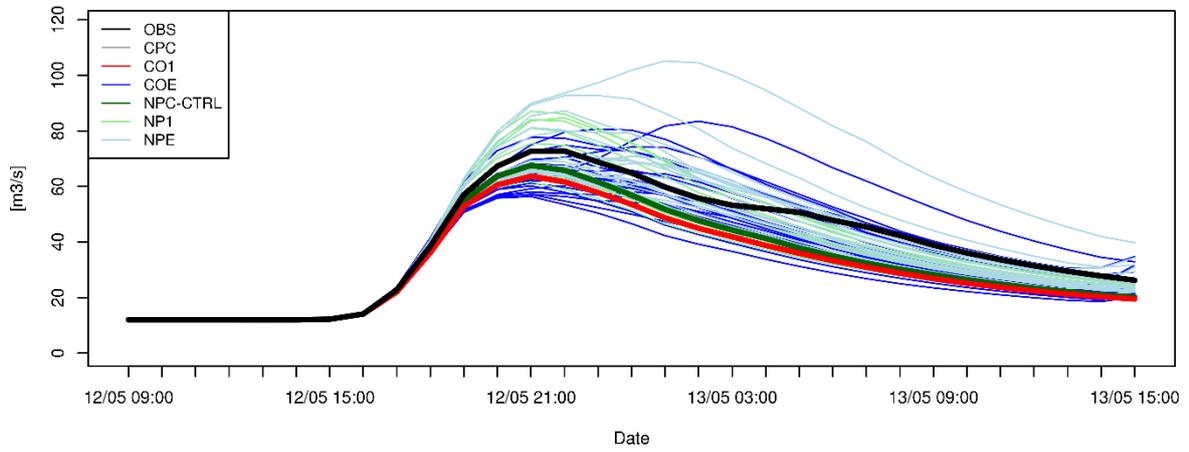


Figure continuous on the next page.

STEPS - RGM-PRO - Emmenmatt - 2017-05-12 07:00:00



STEPS - RGM-PRO - Emmenmatt - 2017-05-12 09:00:00



STEPS - RGM-PRO - Emmenmatt - 2017-05-12 09:00:00

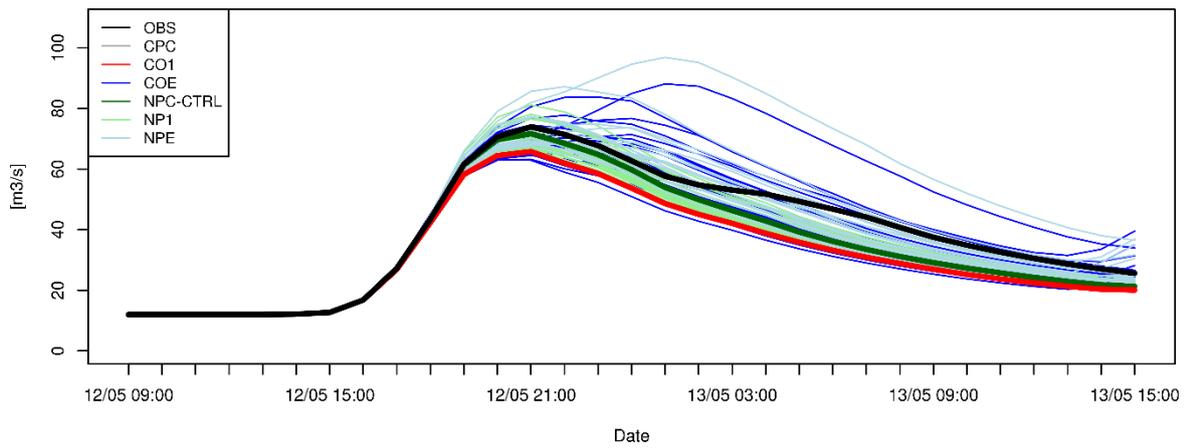
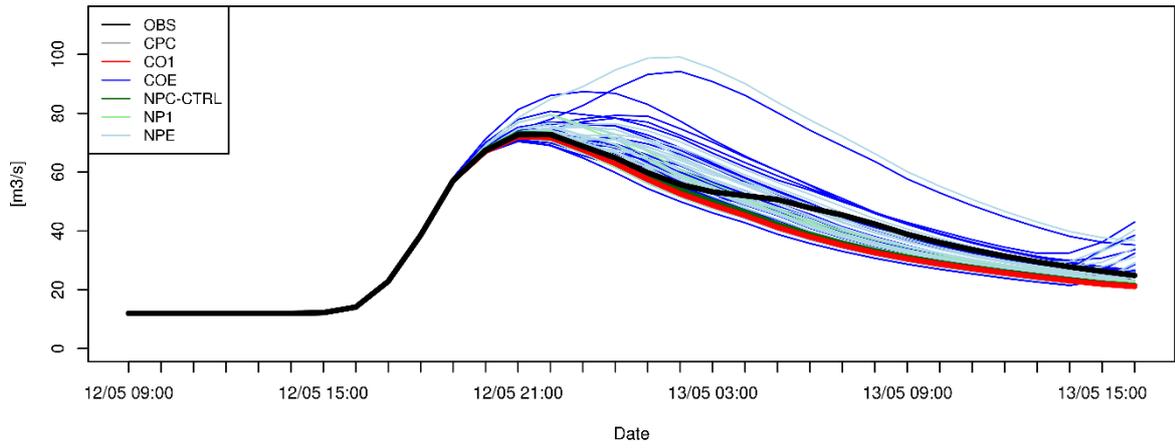
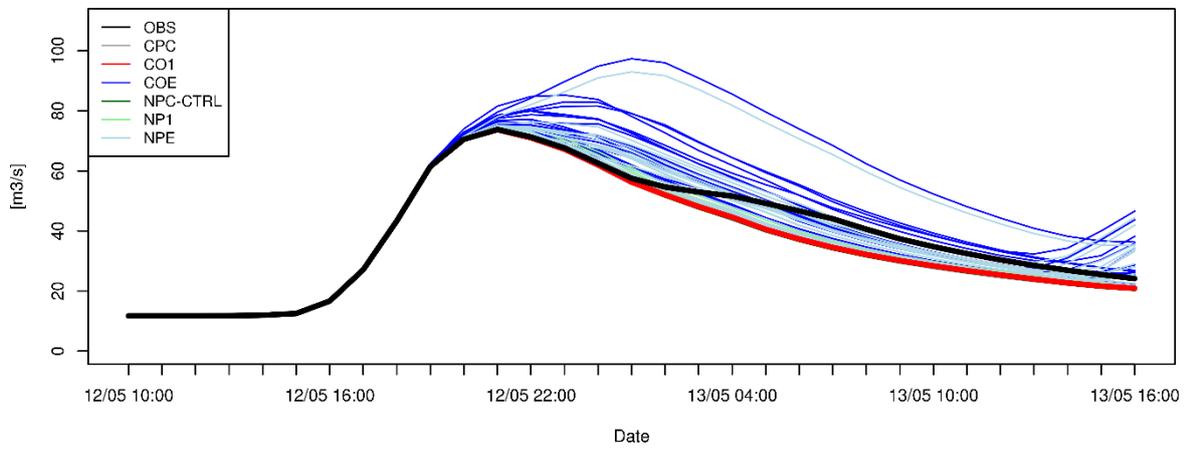


Figure continuous on the next page.

STEPS - RGM-PRO - Emmenmatt - 2017-05-12 09:00:00



STEPS - RGM-PRO - Emmenmatt - 2017-05-12 10:00:00



STEPS - RGM-PRO - Emmenmatt - 2017-05-12 10:00:00

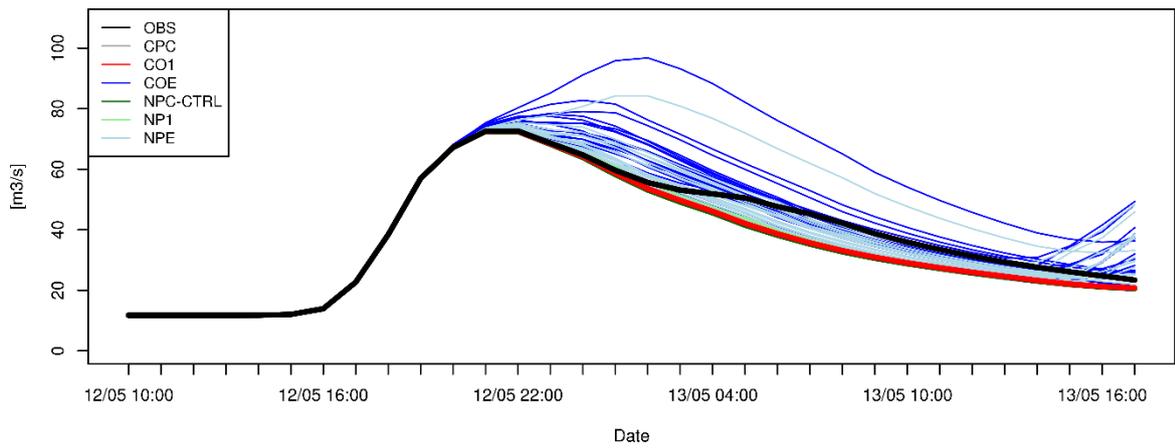


Figure continuous on the next page.

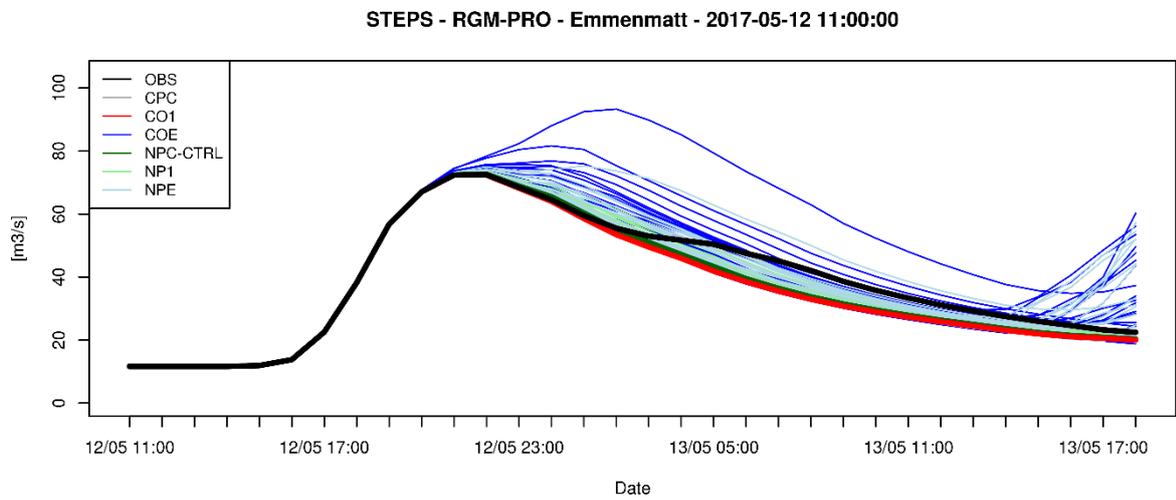
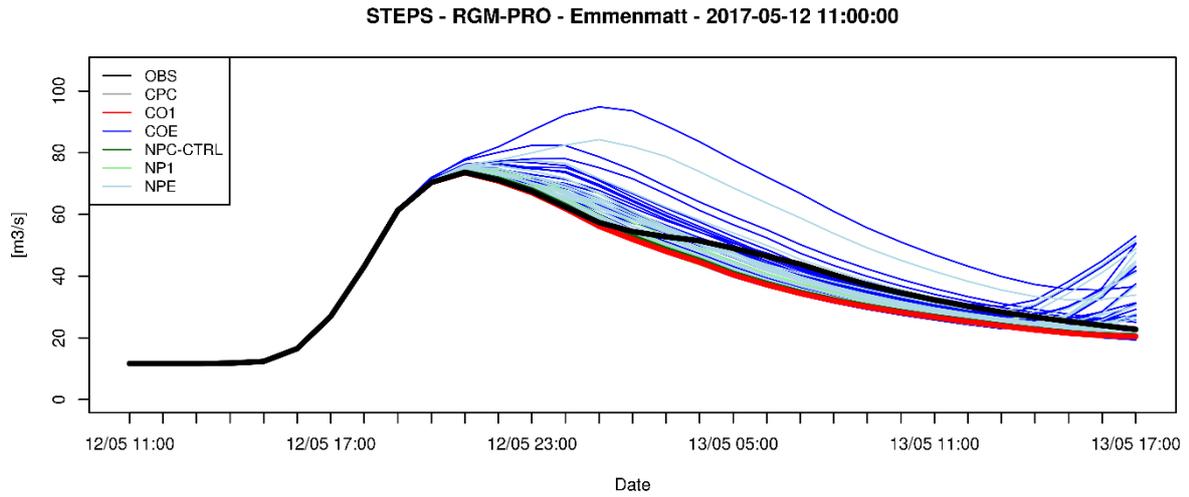


Figure A.26: Hydrographs showing runoff predictions for all nowcasting chains in the main catchment Emmenmatt. From top to bottom: The first prediction at 13:00 to the last prediction at 19:00 of the update cycle is shown.

7.6.3. Median Progrssion NSE

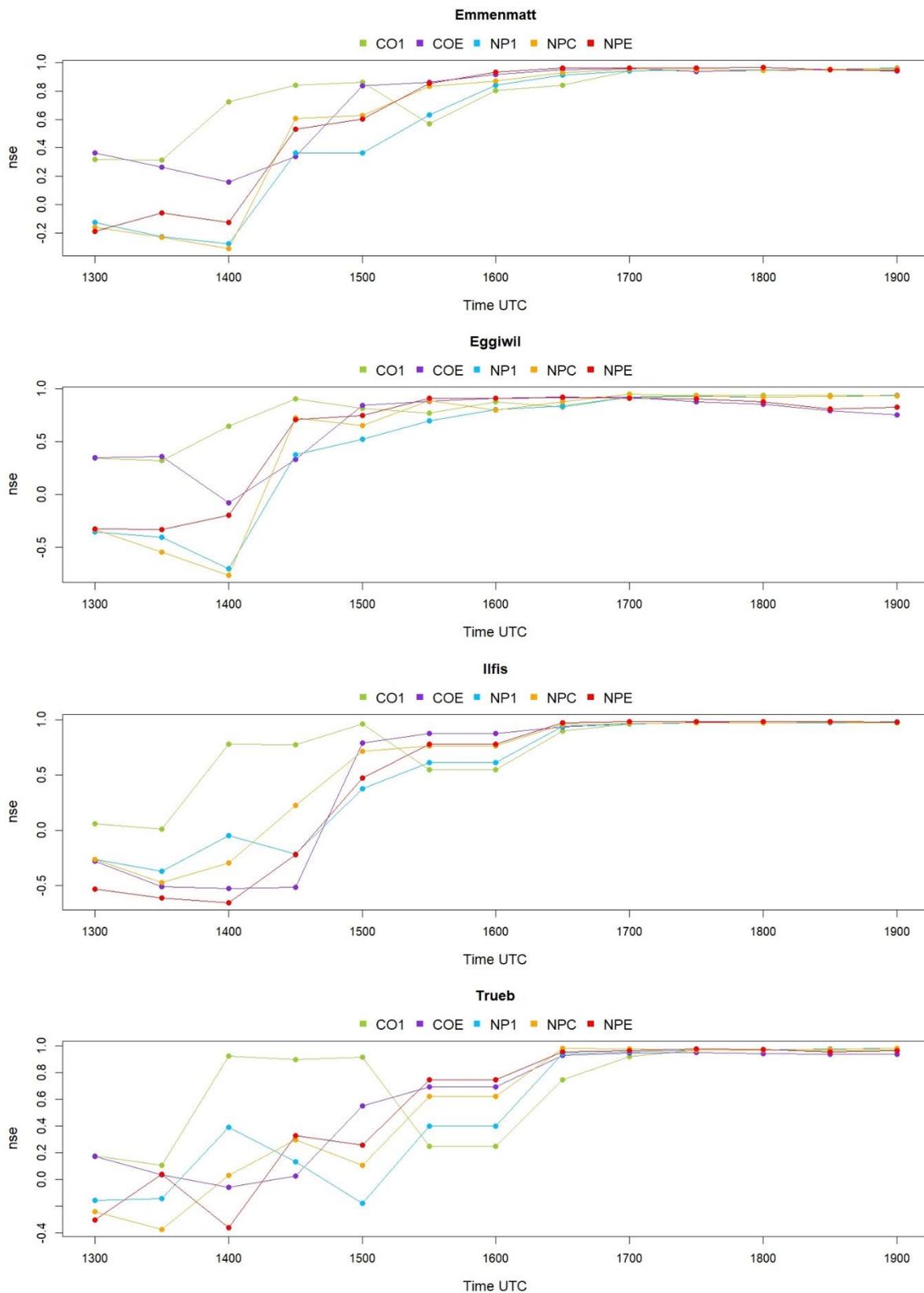


Figure A.27: The median progression of the update cycle is shown for all nowcasting chains in the Emme catchments. The slope of the lines between the points indicates whether the skill in terms of the NSE increases or decreases within the update cycle.

7.6.4. Boxplot Progression Subcatchments

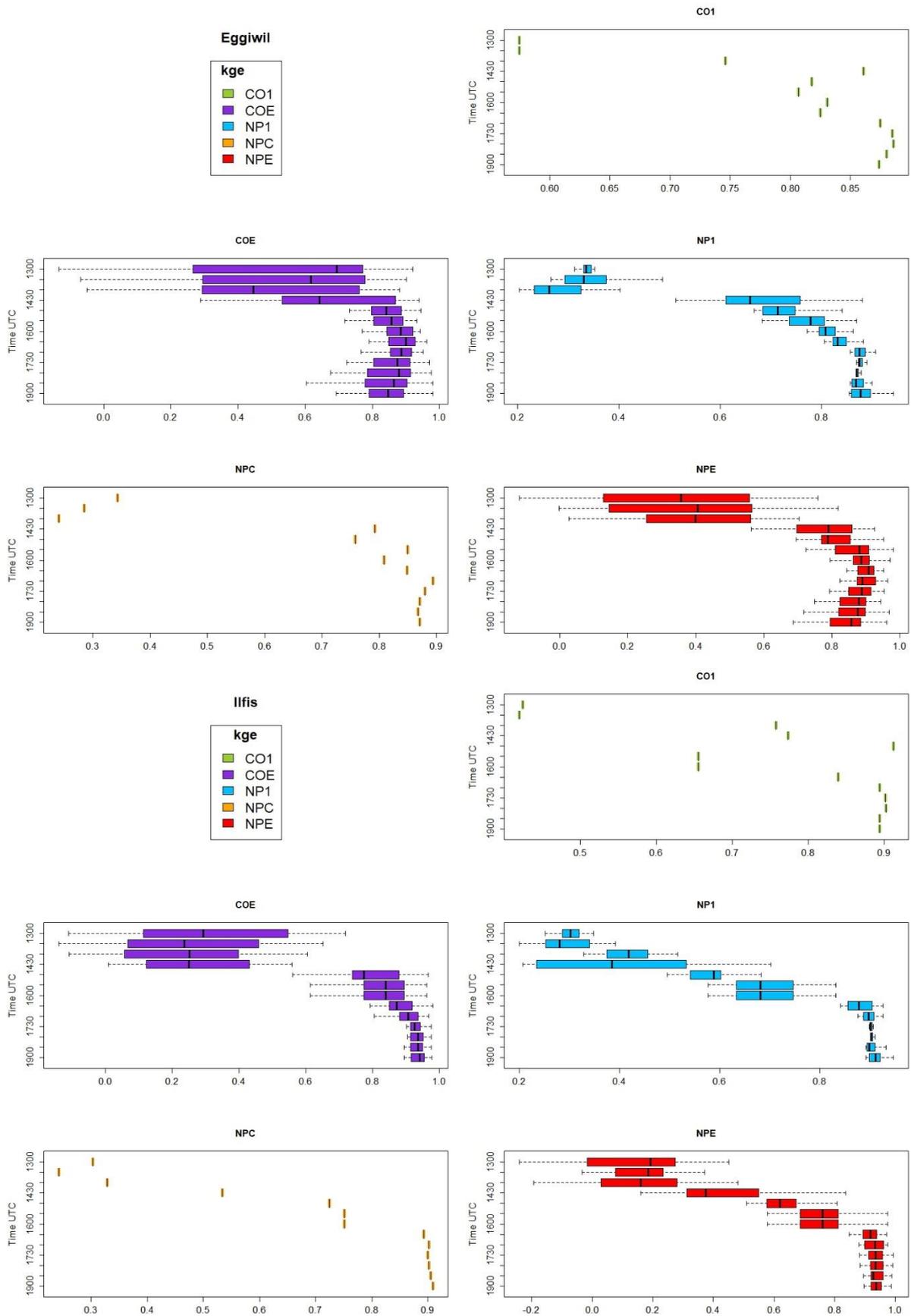


Figure continuous on the next page.

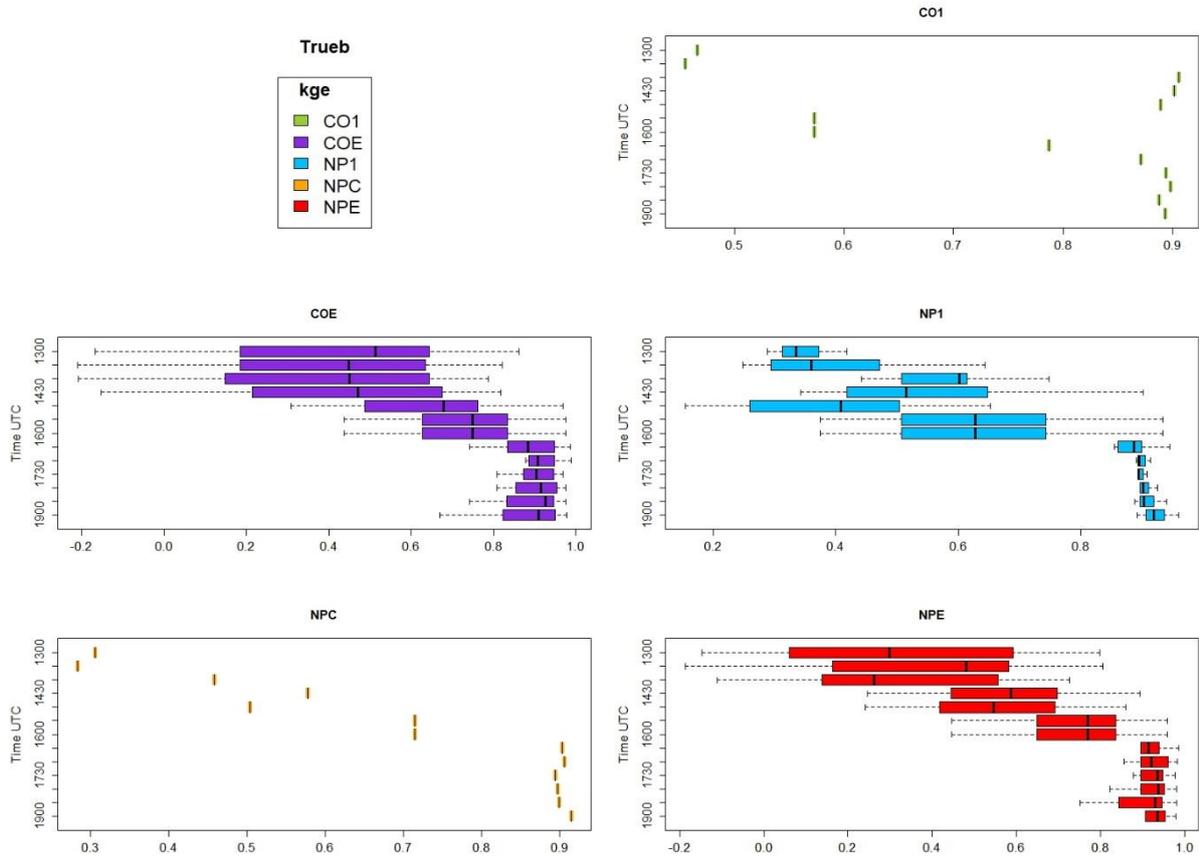


Figure A.28: Update cycle boxplots showing all nowcasting chains in the Emme subcatchments in terms of the KGE. The bold vertical bar of the boxplot represents the median, while the box represents the interquartile range. The whiskers display the range of extreme values.

7.6.5. Taylor Diagram and Peak-box

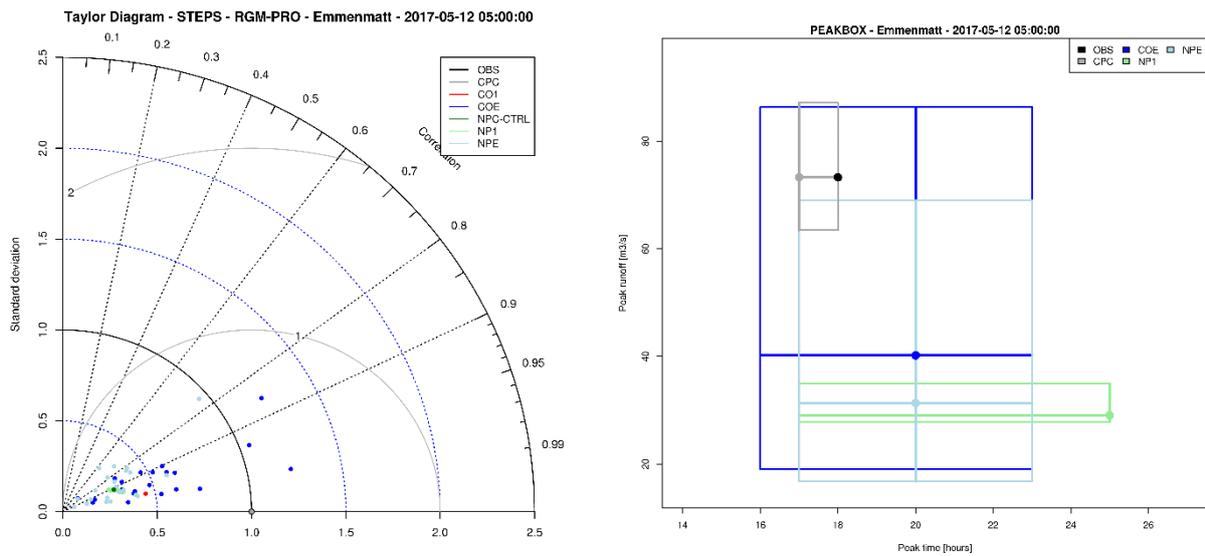


Figure continuous on the next page.

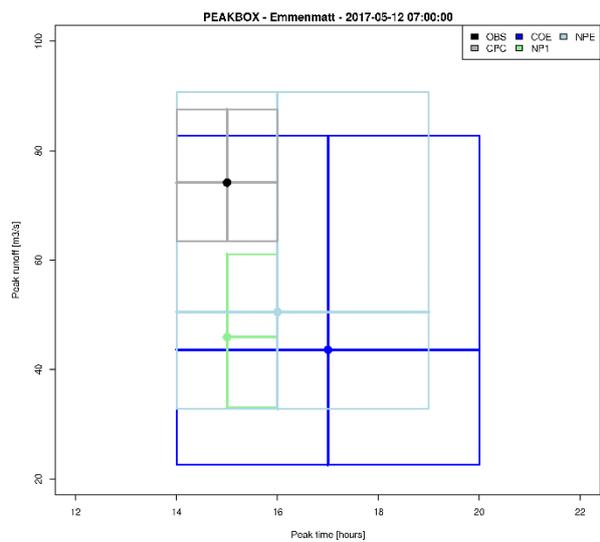
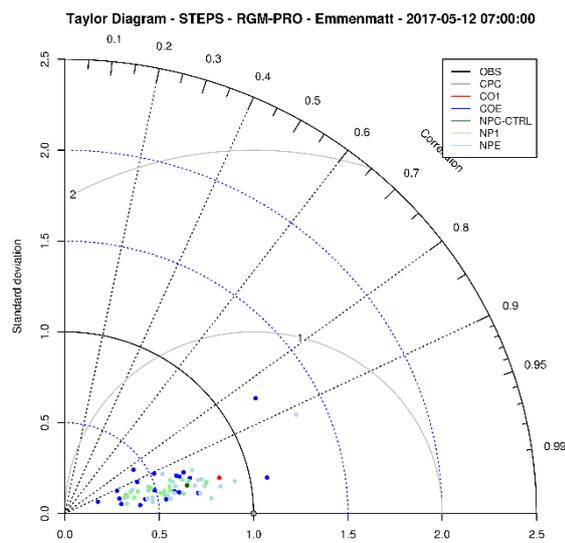
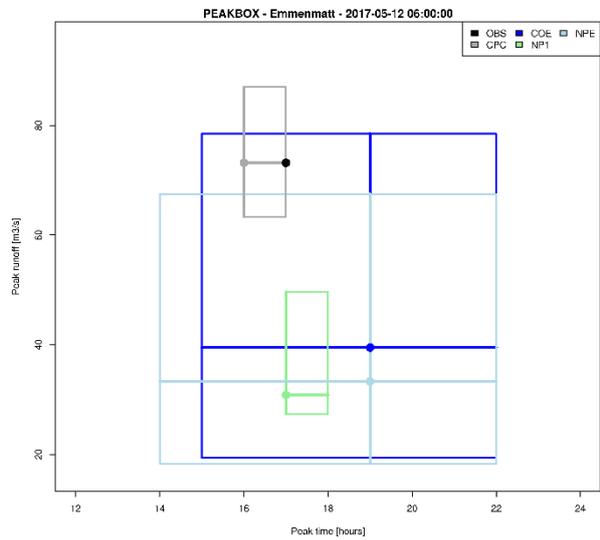
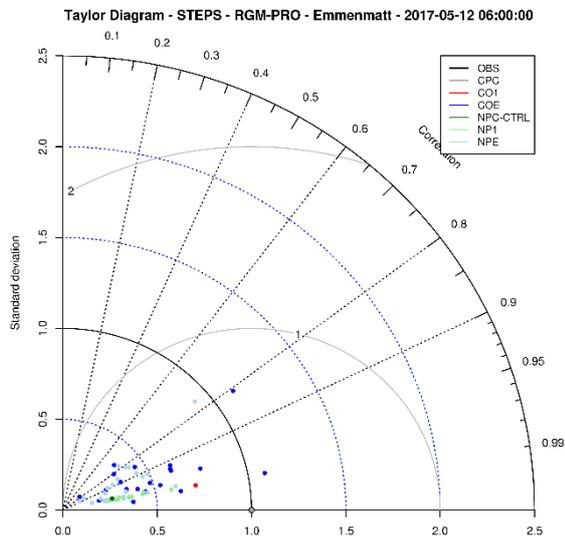
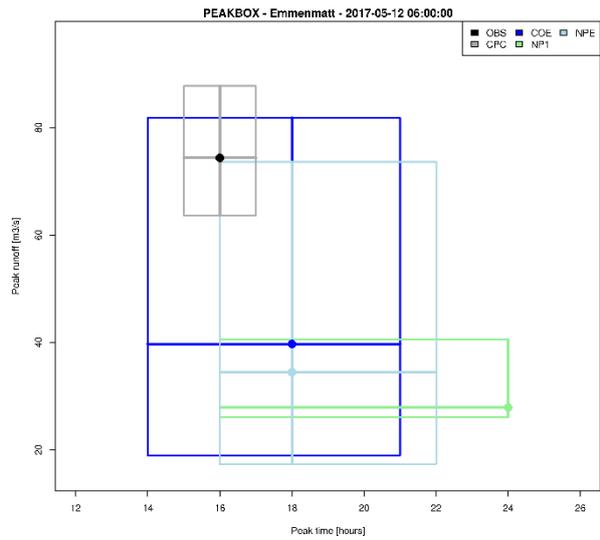
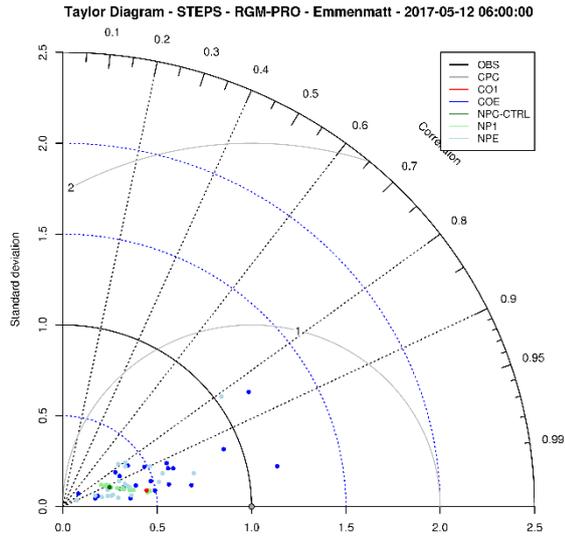


Figure continuous on the next page.

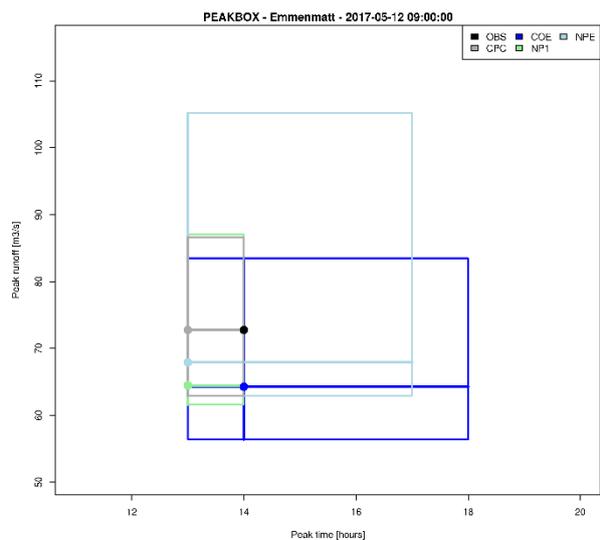
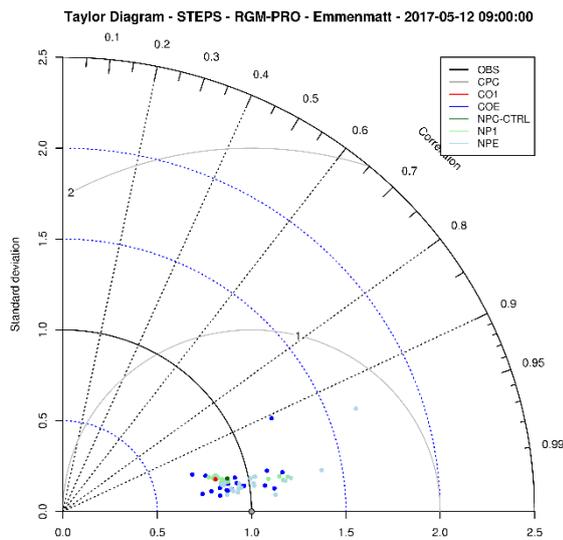
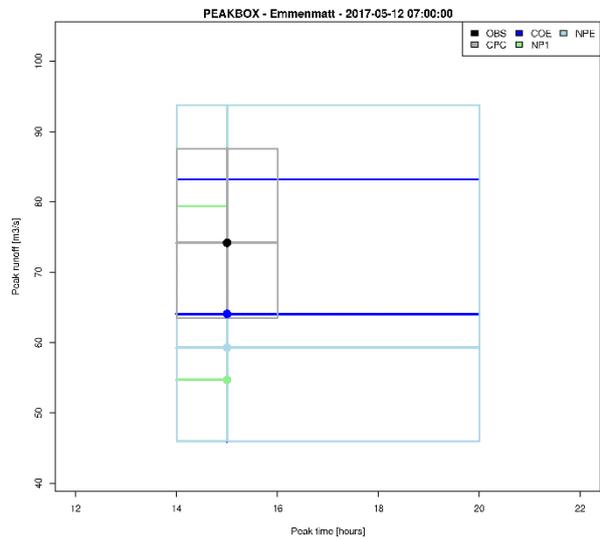
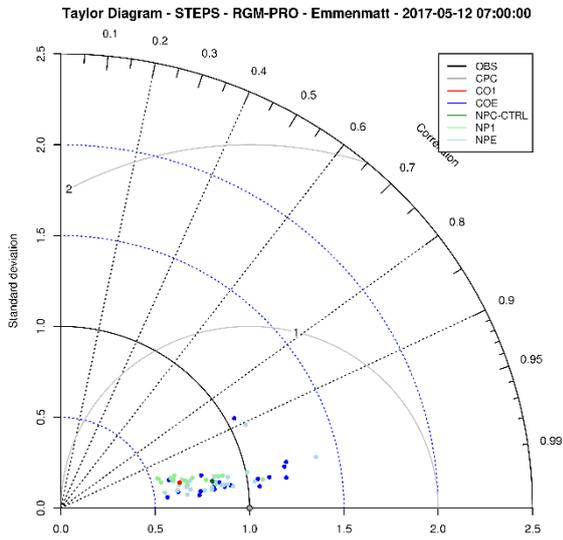
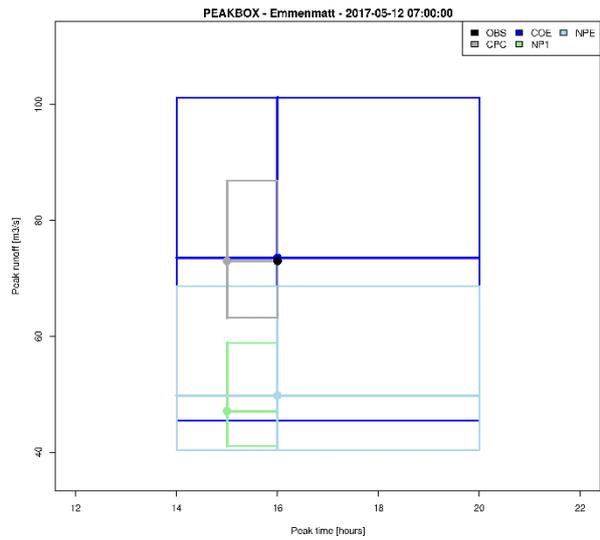
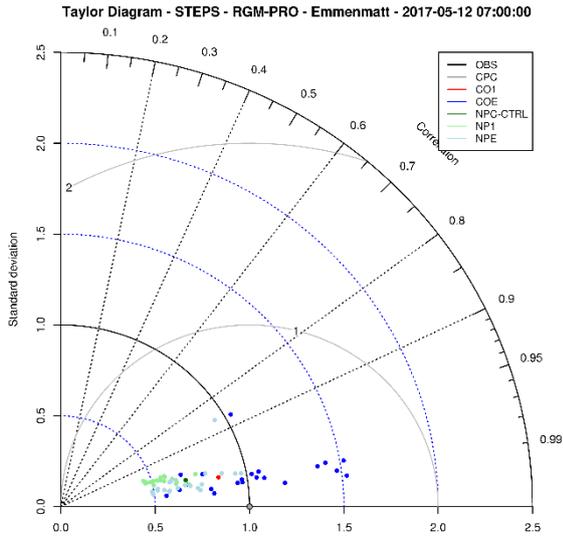


Figure continuous on the next page.

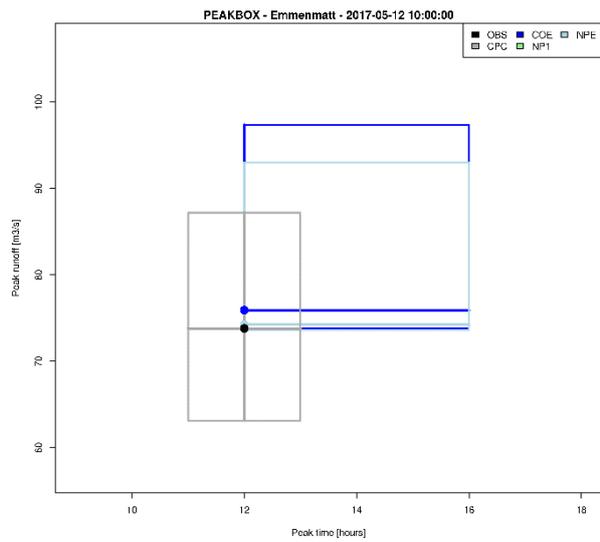
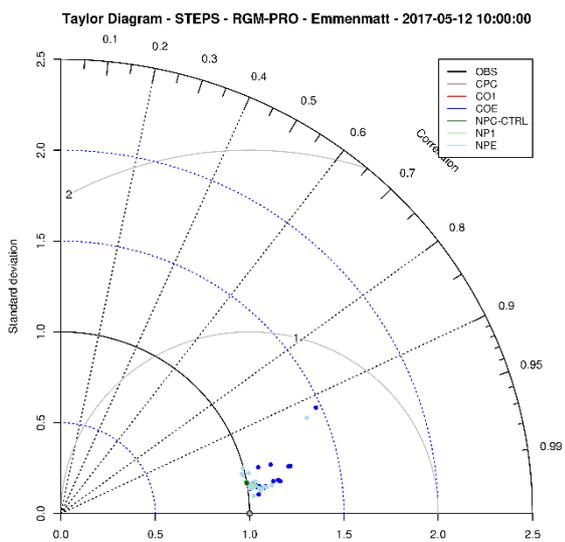
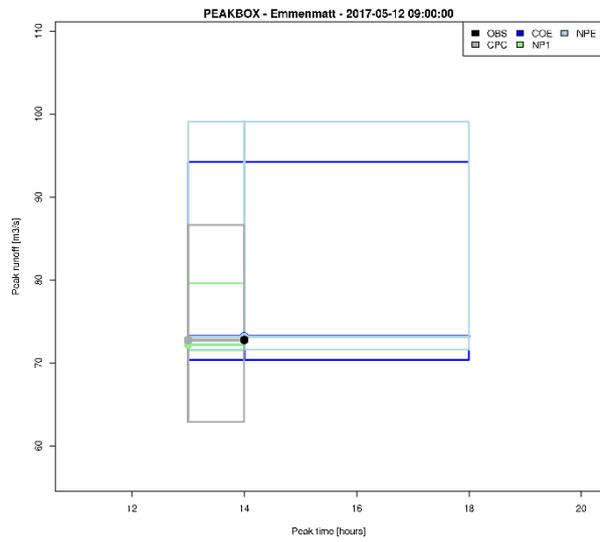
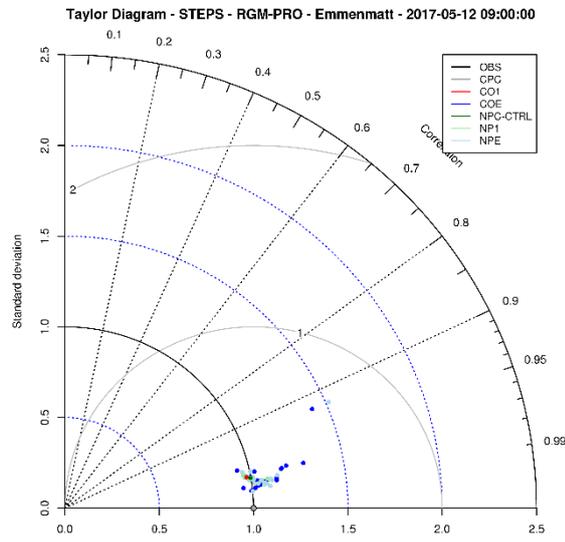
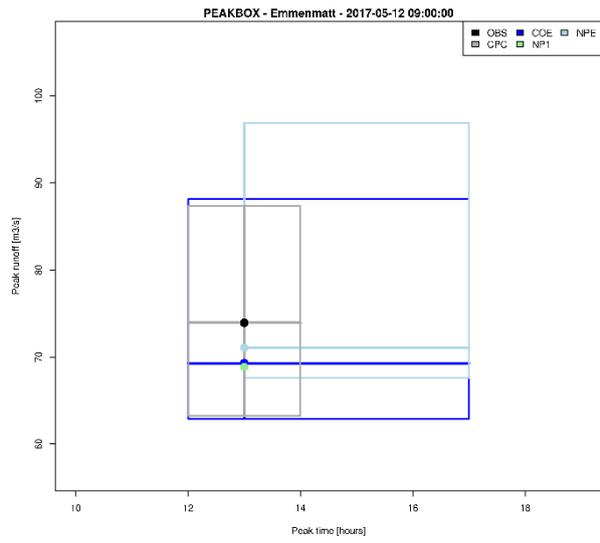
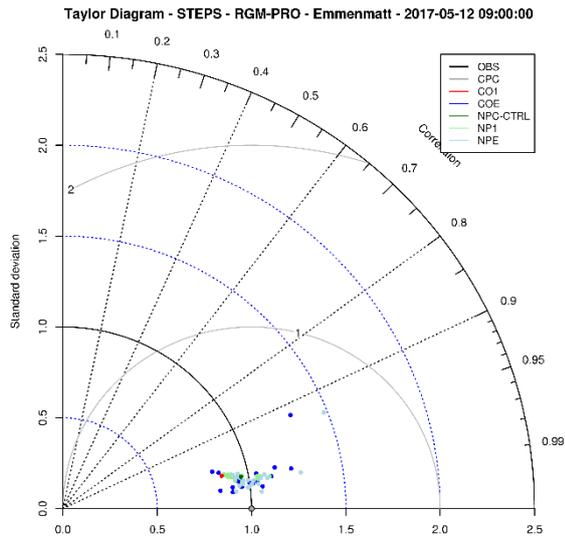


Figure continuous on the next page.

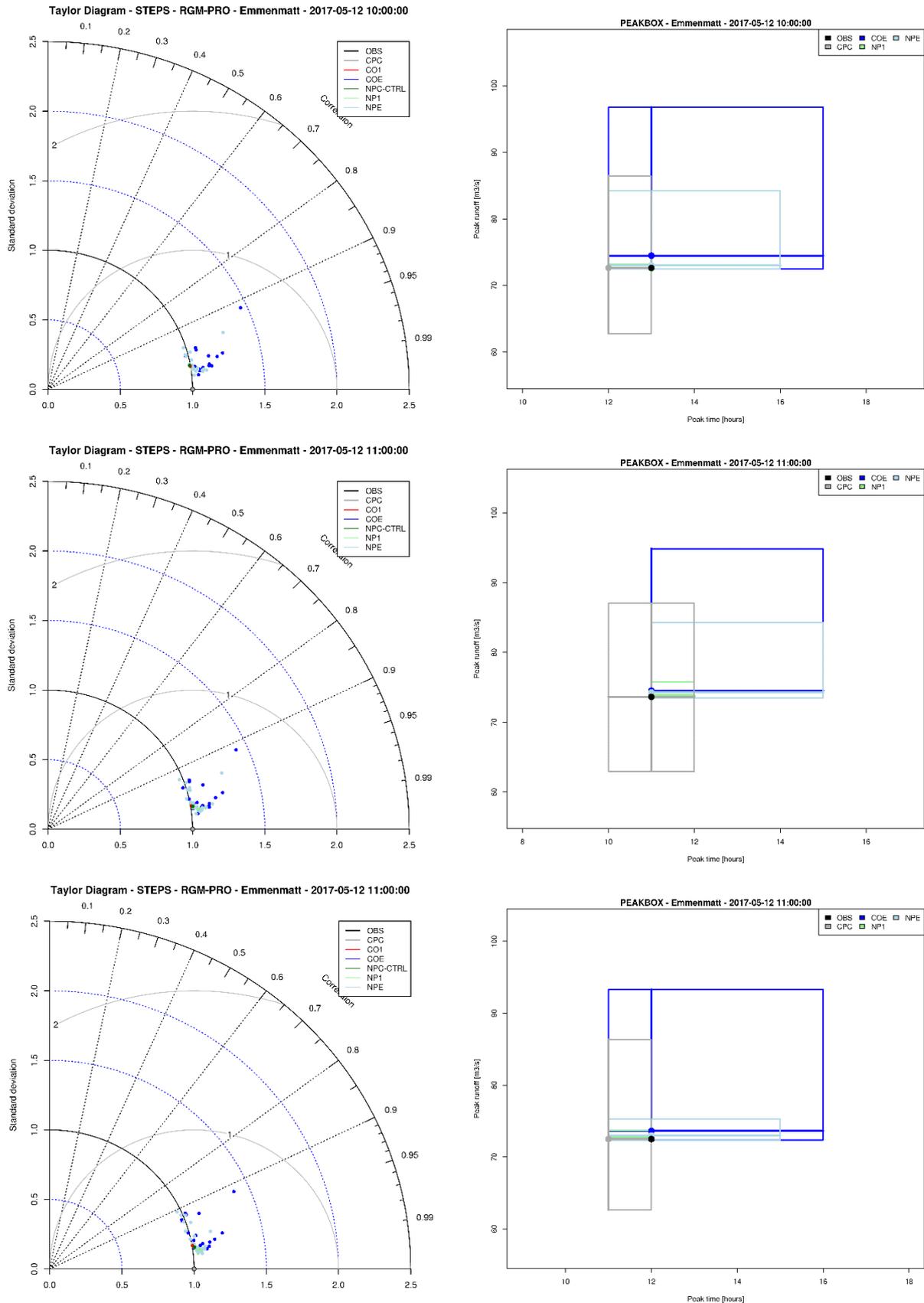


Figure A.29: Left: Taylor diagrams showing correlation (azimuthal angle), standard deviation (blue dotted quadrant), RMSE (grey semicircles) and CPC hindcast (grey point). Right: Peak-boxes showing the best flash flood estimation in terms of peak magnitude and timing of the runoff. From top to bottom: The first prediction at 13:00 to the last prediction at 19:00 of the update cycle is shown for all nowcasting chains in the main catchment Emmenmatt.

Personal Declaration

Personal declaration: I hereby declare that the submitted thesis is the result of my own independent work. All external sources are explicitly acknowledged in the thesis.



Severin Koller