

Daglegar rennslisspár með notkun hliðstæðrar greiningar Harmonie veðurgagna

Morgane Priet-Mahéo, Andréa-Giorgio R. Massad, Sif Pétursdóttir,
Tinna Þórarinsdóttir, Davíð Egilson

Lykilsíða

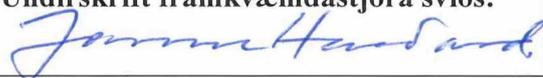
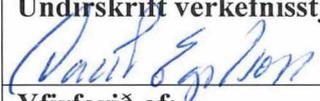
Greinargerð nr. MPM/ofl/2019-02	Dags. Mars 2019	Dreifing: Opin <input checked="" type="checkbox"/> Lokuð <input type="checkbox"/>
		Skilmálar:
Heiti greinargerðar: Daglegar rennslisspár með notkun hliðstæðrar greiningar Harmonie veðurgagna		Upplag: Rafræn útgáfa Fjöldi síðna: 102
		Framkvæmdastjóri sviðs: Jórunn Harðardóttir
Höfundar: Morgane Priet-Mahéo, Andréa-Giorgio R. Massad, Sif Pétursdóttir, Tinna Þórarinsdóttir, Davíð Egilson		Verkefnisstjóri: Davíð Egilson
		Verknúmer: VÍ 4611-0-0002
Gerð greinargerðar/verkstig:		Málsnúmer: 2018-0237
Unnið fyrir: Vegagerðina		
Samvinnuaðilar:		
Útdráttur: Veðurstofa Íslands hefur á undanförnum árum unnið að verkefnum er snúa að flóðagreiningum vatnsfalla á Íslandi í nánú samstarfi við Vegagerðina. Upplýsingar um tíðni og stærð flóða eru nauðsynlegar hönnunarforsendur fyrir vegafamkvæmdir og úrbætur sem og fyrir mat á áhættuviðmiðum og við svæðisskipulag. Í þessu verkefni er sýnt fram á að unnt er að setja fram áreiðanlega spá um rennsli og flóð vatnsfalla með því að nota aðferðafræði sem byggir á hliðstæðri greiningu (e. analogue sorting) gagna. Aðferðin var prófuð á 13 vatnasviðum með góðum árangri þar sem spáð var fyrir um rennsli 1–5 daga fram í tímann. Sérstaklega var forspárgildi fyrsta dags mjög hátt. Niðurstöður hafa verið settar fram á vefsíðu sem er uppfærð daglega með nýrri rennslisspá fyrir hvert vatnasvið.		
Lykilorð: Rennslisspár, hliðstæð greining, Harmonie, flóðaviðvörðun		Undirskrift framkvæmdastjóra sviðs: 
		Undirskrift verkefnisstjóra: 
		Yfirfarið af: SG

Table of content

1	Introduction.....	7
2	Study area.....	8
3	Methodology.....	12
3.1	Principle of analogue sorting.....	12
3.2	Assessments of the method.....	13
3.3	Cluster analysis.....	14
3.4	Statistical tools.....	17
4	Results	19
4.1	Test on reanalysis data.....	19
4.2	Test on operational data.....	20
5	Operational webpage	23
5.1	Operational setup.....	23
5.2	Webpage presenting the results	23
6	Discussion.....	27
7	Conclusions.....	29
8	References.....	30
	Appendix A. Precipitation, air temperature and discharge data used for the test on operational data	32
	Appendix B. Best sets.....	39

Höfundar skýrslunnar bera ábyrgð á innihaldi hennar. Niðurstöður hennar ber ekki að túlka sem yfirlýsta stefnu Vegagerðarinnar eða álit þeirra stofnana eða fyrirtækja sem höfundar starfa hjá.

1 Introduction

Floods in Iceland have led to significant material destruction and disruption in the past. They pose a threat to populations and travelers as well as infrastructures such as roads and bridges, disrupting traffic. With the rapid development of tourism in the country, we can expect floods to cause further threats, so informing people of imminent dangers remains a top priority.

To forecast potential floods, the Icelandic Met Office (IMO) has historically relied on the daily monitoring of water levels along with weather forecasts, which requires expert knowledge and understanding of the catchments' behaviour.

Analogue sorting methods are traditionally used for meteorological forecasts, but similar methods, often named “k-nearest-neighbor” methods, have been successfully tested for streamflow forecasts (Karlsson & Yakowitz, 1987; Akbari, Van Overloop, & Afshar, 2011; Oyeboode, Otieno, & Adeyemo, 2014). In 2013, Crochet (2013) researched successfully the use of these methods at IMO. Crochet tested four methods (combination of predictors) and three analogy criteria on 6 catchments. This study showed that this method could be used for streamflow forecasting in Iceland, however the predictors were at that point not produced on a daily basis and therefore the setting up of an operational system was a step away.

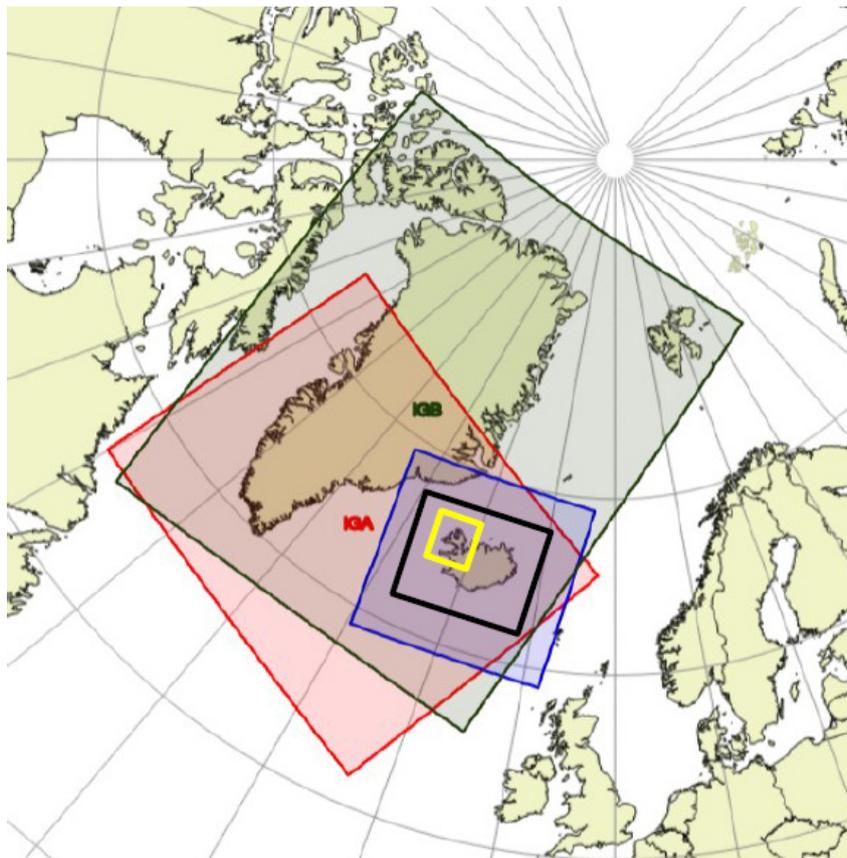


Figure 1. Domains of HARMONIE-AROME runs for the Icelandic Meteorological Office: in red, IGA (interrupted), in green, IGB (operational), in blue HARMONIE, in black HMV/ICRA and in yellow, Westfjords.

The numerical weather prediction model HARMONIE has been in use at IMO since 2011, providing 72 hour forecasts (Figure 1). Additionally, ICRA (2016) reanalysis provided a 30-year record of gridded, 2.5 by 2.5 km meteorological data for the whole of Iceland. In 2018, HVM reanalysis were made available, providing 3-day forecasts. These readily available datasets simplify the development of hydrological forecasts and presented an opportunity to expand on the previous work of Crochet.

This research project aimed at developing a simple operational streamflow forecast system based on an analogue sorting method. Thirteen catchments were used to assess the method. In this report, the different catchments used to test the operational setup are described in a chapter 2; in chapter 3 the methodology used to set-up the analogue forecast and assess the results is detailed; the results are then presented and reviewed in chapter 4 and finally an operational webpage showing the forecast results is presented thoroughly in chapter 5.

2 Study area

Thirteen catchments have been selected all around the country (Figure 2). Their characteristics are diverse as seen on Table 1: The range from simple direct runoff catchments (such as vhm 128) to catchments fed mostly by springs (i.e. vhm 10), lakes (i.e. vhm 204) or glacial rivers (i.e. vhm 150). Their size varies from 38 km² (vhm 19) to 1102 km² (vhm 200) as well as their aspect ratio (between 1.31 and 3.71) and longest flowpaths (from 15 to 131 km).

Table 1. Area, aspect ratio, longest flowpath, mean elevation, glacier percentage and geology (old, young and total bedrock percentages) for the selected catchments.

Vhm	Area	Aspect ratio	Longest flowpath	Mean elevation	Glacier	Old bedrock	Young bedrock	Total bedrock
	km²		m	m.a.sl	%	%	%	%
10	396.1	2.96	55167	527	0	99.2	0.5	99.7
12	164.7	1.74	31960	408	0	96.8	0	96.8
26	266.3	3.36	64555	387	0	61.3	38.7	100
45	458.3	2.37	58072	547	0	67.1	32.9	100
128	513	2.01	58289	338	0	93.7	1.7	95.4
411	387.1	3.71	73405	559	0	97.7	2.3	100
68	201.1	1.33	35345	245	0	6.2	93	99.2
150	225.9	3.03	45563	767	40.23	47	12.8	59.8
19	38.4	1.7	15040	510	0	100	0	100
149	189.4	3.27	37033	609	4.83	91	0	91
198	192.9	1.31	31543	399	0	100	0	100
200	1102.2	3.51	131238	723	0	97.1	0.4	97.6
204	102.3	2.47	28106	466	0	100	0	100

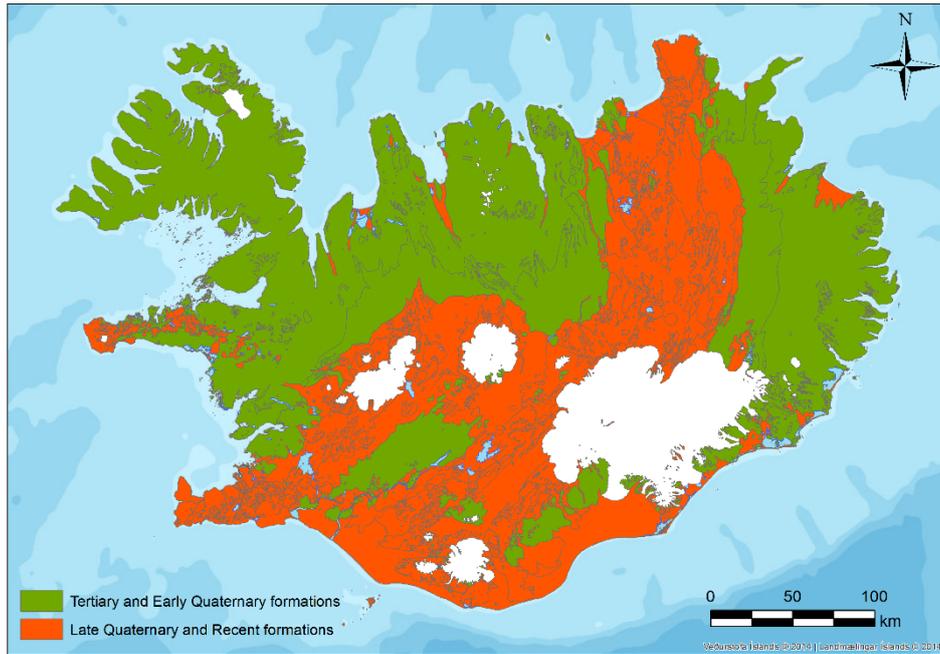


Figure 1. The bedrock in Iceland split into two periods.

The bedrock of Iceland can be roughly divided into two periods based on time of formation, (Figure 1), The Tertiary and Early Quaternary formations contain bedrock with relatively low permeability, and the Late Quaternary and Recent formations contain bedrock with a relatively high permeability. In the Late Quaternary Zone the infiltration is relatively high and surface runoff is limited. On the borders of this formation, or in connection with fissure swarms cutting through it, numerous springs are found with a very strong discharge. In the Tertiary and Early Quaternary Zone the infiltration is minimal, where most of the precipitation flows off as surface runoff, especially during spring and early summer when the snow is melting (Sigurdsson & Einarsson, 1988).

Nine out of thirteen catchments contain >90% old bedrock from Tertiary and Early Quaternary formations. It is only in watershed vhm 68 where young bedrock from Late Quaternary and Recent cover its major part (Figure 2).

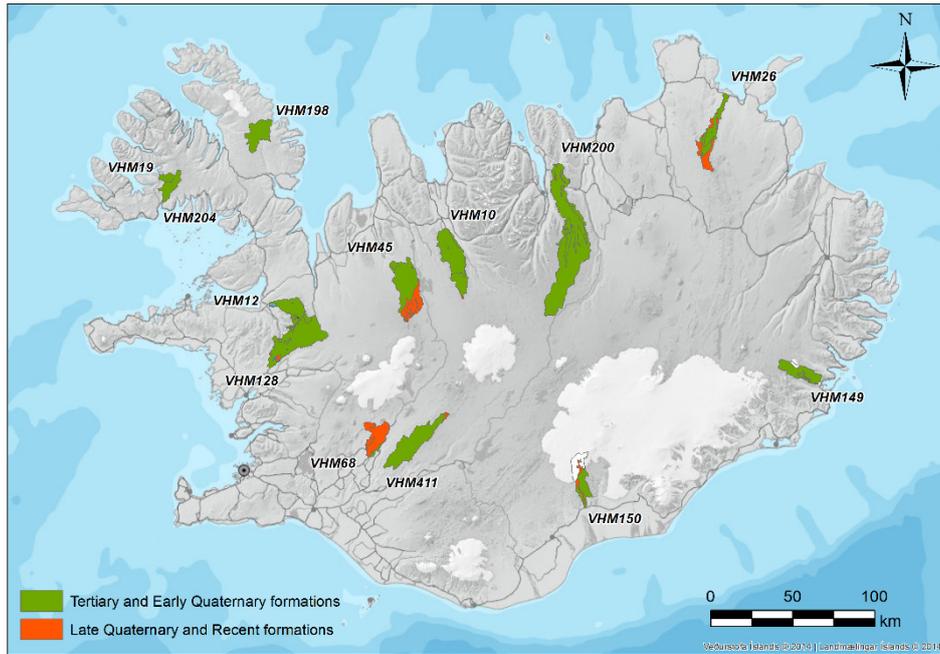


Figure 2. Thirteen catchments have been selected all around the country.

Southwestern region:

Stóra-Laxá (vhm 411) is mainly direct runoff river that has its origin from the north of Geldingafell (Veðurstofa Íslands, 2017). Stóra-Laxá is situated in the south-central highlands of Iceland, where vegetation is sparse.

Tungufljót (vhm 68) is situated in the south of the glacier Langjökull. Most of the bedrock is considered young, or younger than 0.8 million years old. Tungufljót is mainly a spring fed river with a glacier water fed part (Rist, 1990).

West of Iceland:

Norðurá (vhm 128) has direct runoff and runs into the fjord Borgarfjörður. The watershed lies on old bedrock with swampy ground and loess soil fields (Veðurstofa Íslands, 2019).

Haukadalsá (vhm 12) runs through Haukadalur into the fjord Hvammsfjörður in Breiðarfjörður. Haukadalsá is a direct runoff river passing through lake Haukadalsvatn which controls the runoff characteristics to a great extent, (Veðurstofa Íslands, 2017).

Westfjords:

Vatnsdalsá (vhm 204) is a direct runoff river originating in lake Vatnsdalsvatn which controls the main features of the watershed (Veðurstofa Íslands, 2017).

The origin of Dynjandisá (vhm 19) river is at the lakes of Dynjandisheiði Highlands where it runs from a mountain ridge into the spectacular waterfall, Dynjandi, before it reaches sea. Dynjandisá is a direct runoff river with a small spring-fed part and is regulated by the small heathland lakes (Veðurstofa Íslands, 2016).

The source of Hvalá (vhm 198) river is at the numerous lakes of Ófeigsfjarðarheiði Highlands. Ófeigsfjarðarheiði is sparsely vegetated with depleted loess soil fields. Hvalá is a direct runoff river and is slightly regulated by a lake and is one of the largest rivers by volume in the West-fjords (Veðurstofa Íslands, 2018).

North of Iceland:

Vatnsdalsá (vhm 45) is a direct runoff river highly heath regulated where water seeps from wetlands and small lakes into the river (Veðurstofa Íslands, 2017).

Svartá (vhm 10) is a direct runoff river and spring-fed river. Svartá and Vatnsdalsá (vhm 45) run parallel to each other and are similar in size and have similar characteristic discharges (Zóphóniásson & Kristinsson, 2008).

Fnjóská (vhm 200) is a voluminous direct runoff river that responds quickly to high precipitation but has also a considerable baseflow (Veðurstofa Íslands, 2019).

Sandá (vhm 26) river in Þistilfjörður is located in the northeast of Iceland and has an area of 268 km². Sandá is a heath regulated river with a considerable spring-fed contribution (Þórarinsdóttir, 2012).

East of Iceland:

The source of Geithellnaá (vhm 149) river is on the south side of a lava field above the valley Getihellnadalur and from the south- and west-side of Þrándarjökull glacier. Geithellnaá is a direct runoff river and has an area of 189 km² (Veðurstofa Íslands, 2019).

Djúpá (vhm 150) is a glacier fed river with a constant spring-fed part on an average 4 – 15 m³/s. The average daily value can go up to 300 m³/s during the warmest summer days or wet winter days. Djúpá runs down the valley Djúpárdalur where it runs into a gully forming a high waterfall (Veðurstofa Íslands, 2019).

3 Methodology

Analogue sorting method for the forecast of natural phenomena is based on the assumption that any natural event in a system results of a combination of forces acting on this system, and that comparing present conditions of the system with past conditions will give us good insights on future states of the system. This simplistic approach is also computationally-efficient due to its simplicity. The statistics involved in the method are described in 3.1. These assumptions suggest that it is unnecessary to know in detail the physical characteristics of the catchment, however it can be useful to know the main characteristics of the basin to define the most relevant variables (predictors) of the system (3.2). Classification (3.3) can help identifying the most relevant predictors and was used to categorize the catchments. Error measures used to evaluate the performance of the predictions are described in 3.4.

3.1 Principle of analogue sorting

Considering the temporal evolution of a dynamic system X composed of k variables (observations, measured or computed), it is estimated that the future evolution of the system X at time t ($X(t)$) can be estimated from the subsequent states of the most similar past state of the system $X(u)$ at past time u . As it is unlikely in complex system to find past states exactly identical, it is possible to evaluate the future evolution of the system from a combination of different past states whose similarities are determined by the use of analogy criteria, either displaying ensemble forecasts or computing a deterministic forecasts, weighing each past state by its distance to the present state.

In this project, the Mahalanobis distance, an Euclidian distance which considers the correlation between the different predictors, was used in order to discriminate the best past states of the system. It can be computed as follow:

$$MD(u) = \sqrt{(X(t) - X(u))^T S^{-1} (X(t) - X(u))}$$

Where X is the dynamical system, t is the time of last observations, u the time of previous state of the system and S the covariance matrix.

This distance was selected based on past work by Crochet (2013) at the Meteorological Office that showed that the Mahalanobis distance was fit for this purpose and the results from this distance were slightly better than the results from the simple Euclidian distance. Once past states have been identified as analogue, it is possible to infer some ensemble prediction for the discharge of following days such as

$$Q(i, t + T) = Q(u_i + T)$$

Where T between 1 and 5 and u_i the date of the i^{th} analogue. However, in case of exceptional discharge levels, the past states of the system might not be able to provide appropriate discharge values, it can be therefore valuable to rescale the past event. The new discharge value will then be:

$$Q(i, t + T) = \frac{Q(t)}{Q(u_i)} Q(u_i + T)$$

A deterministic forecast can be computed from the ensemble prediction by weighing each ensemble prediction by the analogy criterion. A deterministic discharge prediction can be computed for each day such as:

$$Q(t + T) = \frac{\sum_{i=1}^N w_i Q(i, t + T)}{\sum_{i=1}^N w_i}$$

Where w_i is 1/MD. The number of past events used in the computation of the deterministic prediction needs to be evaluated.

3.2 Assessments of the method

The method is not calibrated per se as there are no adjustment of parameter, however it is important to select a combination of predictor that is relevant for the catchments and allows us to capture discharge, both timely and quantitatively, which is crucial for accurate flood prediction. The procedure for the selection of the best predictor set was twofold:

1. in order to evaluate the performance of each set all year round it was decided to run the model retroactively over year 2016, the meteorological variables are however in this case all ICRA reanalysis (with correction of air temperatures) and retain less uncertainties than actual operational meteorological data. In addition, all discharge values have been corrected.
2. The sets are then tested on an operational mode, with actual 3-day forecast data (HVM forecasts) and uncorrected discharge measurements. 3-day operational forecast data was stored from September 23rd, 2018 and the sets were tested on an approximate period of 130 days (23/09/2018 – 05/02/2019). Appendix A presents some of the climatic data along with the discharge data for each catchment studied, revealing the periods of data gaps associated with data losses.

Discharge and meteorological history of the catchments are combined into predictor-sets. Three meteorological variables, computed by Harmonie, are selected to evaluate this method: precipitation, air temperature and snow-water equivalent (SWE). In addition, new variables based on these four variables are computed and summarized in Table 2.

Table 2. Description of variables used as predictors.

Name	Description
Q	Daily discharge
T	Daily temperature (mean over the catchment)
P	Daily precipitation (mean over the catchment)
SWE	Daily snow-water equivalent (mean over the catchment)
T2/3	Average daily temperature of last 2/ 3 days
P2/3	Cumulated mean daily precipitation of last 2/ 3 days
TS1/2/3	Daily temperature for day+1/2/3 (mean over catchment), forecast values from Harmonie for last day of observations
PS1/2/3	Daily precipitation for day+1/2/3 (mean over catchment), forecast values from Harmonie for last day of observation
SS1/2/3	Daily snow-water equivalent for day+1/2/3 (mean over catchment), forecast values from Harmonie for last day of observation
Q1/2	Daily discharge of day -1/2
S(eason)	Season of year (0 for summer (june to september)/1 for winter (october to may)
M(onth)	Month of year (01- 12)

Originally an additional product from Harmonie was tested (melt) but in order to simplify the process, it was suppressed in the second round of testing and it won't be discussed here. In total, 22 predictor sets were evaluated (Table 3), set 15, 16, 19, 20 and 21 were added in the second stage after consideration of the first results.

Table 3. Predictor sets. Variables used in each set are indicated by a cross.

	Q	T	P	SWE	P2	P3	T2	T3	TS1	PS1	SS1	TS2	PS2	SS2	TS3	PS3	SS3	Q1	Q2	A	M
set 1	x	x																			
set 2	x		x																		
set 3	x			x																	
set 4	x	x	x																		
set 5	x	x	x	x																	
set 6	x																		x		
set 7	x																		x	x	
set 8	x		x			x															
set 9	x		x																x		
set 10	x		x		x																
set 11	x	x	x																		x
set 12	x	x	x																		x
set 13	x	x	x						x	x											
set 14	x	x	x						x	x		x	x								
set 15	x	x	x	x					x	x											
set 16	x	x	x	x					x	x		x	x								
set 17	x	x	x	x					x	x		x	x								x
set 18	x	x	x	x					x	x		x	x					x	x		x
set 19	x	x	x				x		x		x	x		x	x		x	x			
set 20	x	x	x	x			x		x	x	x	x	x	x	x	x	x	x			
set 21	x	x	x	x		x		x	x	x		x	x					x	x		x
set 22	x	x	x	x	x	x	x	x	x	x		x	x					x	x	x	x

3.3 Cluster analysis

Several classifications have been developed in order to group rivers according to their nature. Stefánsdóttir, Björnsson, Magnússon, & Egilson (2014) classified the rivers based on the geology of the catchment and the presence of lakes and mars while Hróðmarsson and Þórarinsdóttir (2018) based their classification on observations made over more than 50 years of field observations. By using measurements of discharge recorded in all 13 stations and complementary information regarding the watersheds, a new classification is proposed here using hierarchical cluster analysis in a similar way to Crochet (2012). Hierarchical cluster analysis is an analytic method that allows us to categorize rivers in groups that share more similarities than with any other rivers from other groups.

According to Demirel and Kahya (2007), Ward's method based on euclidean distances is best adapted for clustering of hydrological data and has been applied to a dataset comprised of both discharge measurements in all the stations and several independent catchment characteristics. By combining them, the stations are clustered according to both the river types and the watershed's environment. The selection of the data is presented in the following paragraphs.

Discharge data

The first gauging stations in Iceland started recording the rivers heights in 1951 (Rist, 1990). In order to perform the cluster analysis over an homogeneous set of data, a period of 10 hydrological years has been selected, spreading between October 2007 and October 2017. Over this timespan, the only missing values are to be found for vhm 26 with only 7 days of unavailable data which can be disregarded compared to the length of the selected timeperiod. Discharge data have been combined in in three different ways, each method giving a different indication of the general behaviour of the rivers:

- *Seasonality*: for each station, discharges have been averaged over the whole timeseries by Julian day. By doing so the seasonal trend of the rivers can then be observed and will give informations on the river type as described in Einarsson (1994). For each station, only monthly-averages discharges are kept so that only the general trend is reflected in the dataset and not weekly variations that are unrelated to the type of river and more reflective of punctual weather events.
- *Duration curves*: for each station, duration curves have been created using a 10% exceedence step. Those curves expresses how often a discharge level is exceeded, providing a good indication of the river's power.
- *Mass curves*: for each station, the yearly averaged discharge have been summed cumulatively over Julian day and the differences between the monthly averaged values and the values that would be obtained if the discharge remained constant all year long are computed.

An example of seasonality, duration curve and mass curve is given for station vhm 411 on Figure 2.

In addition to those measurements, the catchment characteristics (already listed in Table 1) have been added before performing the cluster analysis: total area of the watershed (km²), aspect ratio to reflect the shape of the catchment, longest flowpath (m), mean elevation (m), percentage of glacial coverage as well as the proportions of old bedrock, young bedrock and total bedrock coverages within the watersheds.

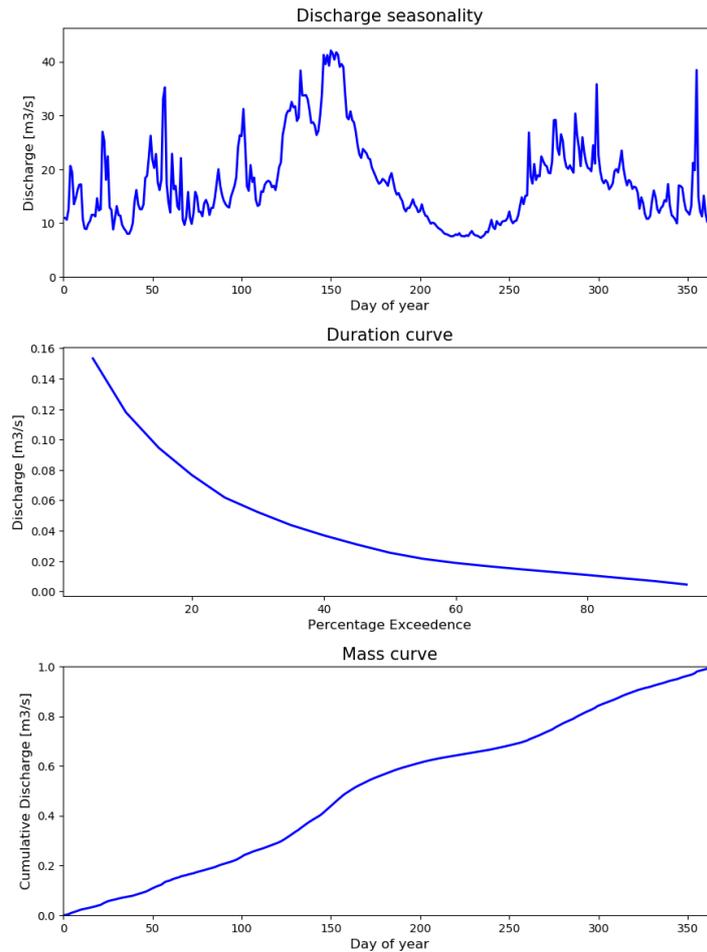


Figure 2. Seasonality (top), duration curve (middle) and mass curve (bottom) obtained for discharge values at station vhm 411. Median has been used to find the Julian days values and for the mass curve the discharge values have been normalised.

Results from the analysis

Figure 3 shows results from the hierarchical cluster analysis in the form of a dendrogram. For the seasonality and the mass curves, only one monthly value has been kept: both the mean and the median have been tested. The mean would give a more accurate estimate of average values over years while the median gets rid of extreme values that could potentially distort the results. Both methods lead to similar results but while using the median (as on Figure 3), the cophenetic correlation coefficient, defined as a measure of how faithfully the analysis preserves the distances between the data, suggests a better clustering with a value of 0.85 against 0.83 with the mean (the closest to 1, the best).

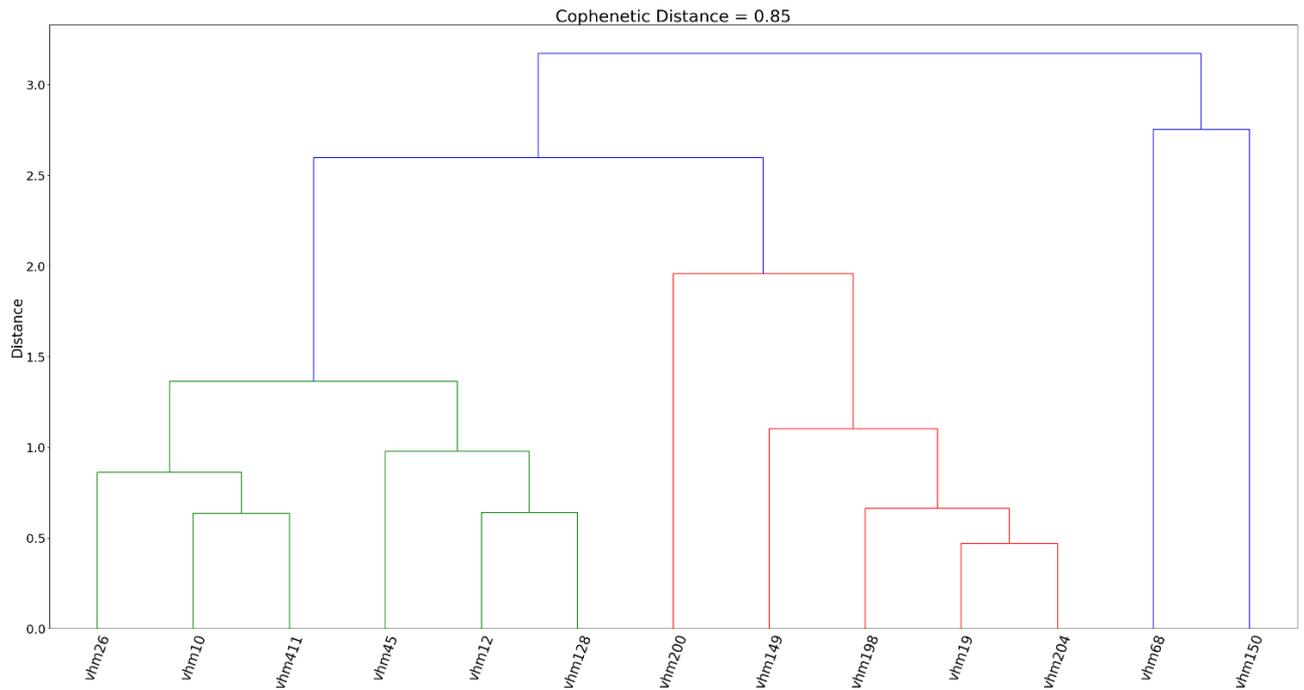


Figure 3. Results from the cluster analysis performed using seasonality, duration curves, mass curves and independent variables from Table 1.

On Figure 3, three clusters appear clearly on the diagram:

- Cluster A (green): vhm 45, vhm 12, vhm 128, vhm 26, vhm 10, vhm 411
- Cluster B (red): vhm 200, vhm 149, vhm 198, vhm 19, vhm 204
- Cluster C (blue): vhm 68, vhm 150

From those results, it appears that the cluster analysis grouped stations according to the type of river and weather conditions. Cluster A groups together stations that include storages (groundwater, wetlands, lakes) and except for stations vhm 26 and vhm 411, all stations are in the northwestern part of the country. Cluster B are snow-influenced direct-runoff catchments and vhm 198, vhm 19 and vhm 204, which are all located in the west fjords belong indeed to the same group. Cluster C gathers stations that receive glacial runoff.

3.4 Statistical tools

In order to evaluate the accuracy of the analogue forecasts, different statistical coefficients can be defined to measure the differences between simulated and observed discharges.

Nash-Sutcliffe Efficiency and Modified Nash Sutcliffe Efficiency coefficients

Nash-Sutcliffe (NSE) and Modified Nash-Sutcliffe (MNS) Efficiency coefficients are two of the most commonly used statistical tools in hydrology (Krause, Boyle, & Bäse, 2005; Franz & Hogue, 2011; Gupta, Kling, Yilmaz, & Martinez, 2009). They are defined as follow:

$$NSE = 1 - \frac{\sum_{i=1}^n (O_i - S_i)^2}{\sum_{i=1}^n (O_i - \bar{O})^2} \quad MNS = 1 - \frac{\sum_{i=1}^n |O_i - S_i|}{\sum_{i=1}^n |O_i - \bar{O}|}$$

where O stands for observed values, \bar{O} for mean observed values and S for simulated values.

From this formula, it appears that NSE and MNS values can range between $-\infty$ and 1. Simulations are closer to the actual discharge values when values tends to 1 and are better than using mean discharge values when the coefficients are positive. NSE is very sensitive to peak flows as a result of simulations and observations being squared but is a well-fitted criteria to assess the quality of the predictions when rivers are in low-flow regimes. This problem can be avoided by using the MNS in which the data are not squared hence giving smaller weight to peak values.

Root-Mean-Square Error and Mean Absolute Error

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (S_i - O_i)^2}{n}} \quad MAE = \frac{1}{n} \sum_{i=1}^n |S_i - O_i|$$

where O stands for observed values and S for simulated values.

Root-Mean-Square (RMSE) and Mean Absolute Error (MAE) are used this study as a complement to the NSE. Their values vary between 0 and ∞ . RMSE is the squared roots of the average of squared errors and hence gives as well better results when the variability of the dataset is small, the large errors having a large effect on its value. It is worth noting that comparing those coefficients between rivers whose discharge rates are very different will affect greatly the results as the data are not normalised.

4 Results

4.1 Test on reanalysis data

In order to test the results and determine which combination of predictor-sets are the most effective ones for the selected stations, five-day analogue forecasts have been computed over the year 2017 for each predictor-set and every station. The data used in this test was only based on reanalysis data and not on operational forecasts.

Table 4 presents for each station which predictor-set gave the best result and the corresponding values of NSE coefficients for the first and the fifth day of forecasts. Those tests show rather good results for the first day of forecast with a NSE varying between 0.68 and 0.92. These values drop with each day of forecast and for the fifth day, they range between 0.12 and 0.83. As seen in Table 4, the best predictor-sets associated to the highest NSE vary greatly between stations.

Table 4. Results from the test run on analysis data for the first and the fifth forecast day giving the best predictor-sets and corresponding NSE for each station.

Station number	First day of forecast D+1		Fifth day of forecast D+5	
	Best predictors-set	Best NSE	Best predictors-set	Best NSE
vhm 10	1	0.75	17	0.34
vhm 12	6, 7, 12	0.67	17	0.24
vhm 19	11	0.76	16	0.34
vhm 26	5	0.88	17	0.59
vhm 45	6	0.93	18	0.63
vhm 68	13	0.76	5	0.34
vhm 128	4, 11	0.68	14, 17, 18	0.17
vhm 149	13	0.72	14	0.12
vhm 150	13	0.73	10	0.43
vhm 198	5	0.88	17	0.62
vhm 200	6, 7	0.97	17	0.82
vhm 204	4, 5, 12	0.87	17	0.38
vhm 411	1, 6, 7	0.88	17	0.51

Using the results from the cluster analysis, it is then possible to target for each group of stations which predictor-sets are the most efficient, while taking into consideration that for each day of forecast, the best sets for a given station might change. If we focus on D+1, looking at the NSE results, predictor-sets that contain the discharge of the previous days improve the forecast for the station from Cluster A. Using the temperature and SWE variables gives better results for Cluster B while predictor-sets that include weather forecasts seem

most efficient for Cluster C. Regarding the D+2 - D+5 results, the use of weather forecasts in the predictor-sets gives better forecast for all the clusters.

These preliminary results have then been refined by testing the results from the analogue forecast on operational data.

4.2 Test on operational data

It is found that the best deterministic forecasts using an operational dataset are obtained when the 50 best analogues are used, as previously found in 4.1. In this part, forecast results will be referring to deterministic predictions computed based on the 50 best analogues.

Table 5 shows the performance of the method for the first and the fifth day of forecast, using operational data. When using operational data in the forecast, some drop in the quality of the prediction can be observed, compared to the forecast using data from reanalysis. This is expected since the predictor-sets rely on uncorrected discharge data and weather forecast for the state of reference ($X(t)$). However, the table shows that the forecasts are very reliable for the first day (D+1) and still instructive (>0) for most stations for the last day of forecast (D+5).

Table 5. Best Nash-Sutcliffe efficiency coefficients for raw and rescaled deterministic forecasts D+1 and D+5 for each station.

NSE	A						B				C		
	vhm 010	vhm 012	vhm 026	vhm 045	vhm 128	vhm 411	vhm 019	vhm 149	vhm 198	vhm 200	vhm 204	vhm 068	vhm 150
S1	0.58	0.68	0.52	0.59	0.75	0.56	0.46	0.79	0.66	0.74	0.67	0.49	0.75
Sr1	0.75	0.81	0.52	0.72	0.84	0.69	0.41	0.86	0.77	0.82	0.62	0.54	0.83
S5	0.03	0.30	-0.25	-0.24	0.10	0.07	-0.11	0.29	0.34	0.19	0.17	0.21	0.09
Sr5	-0.02	0.36	-0.46	-0.24	0.05	-0.07	-0.13	0.08	0.34	0.27	0.18	-0.01	-0.01

The rescaled deterministic forecast (Sr) is almost systematically improving the prediction of discharge (Table 5), as it generally results in a more condensed scatterplot along the 1:1 line (Figure 4) and a better representation of the larger flows. However, stations vhm 19 (Figure 5), vhm 204 and vhm 26 do not seem to benefit from this correction.

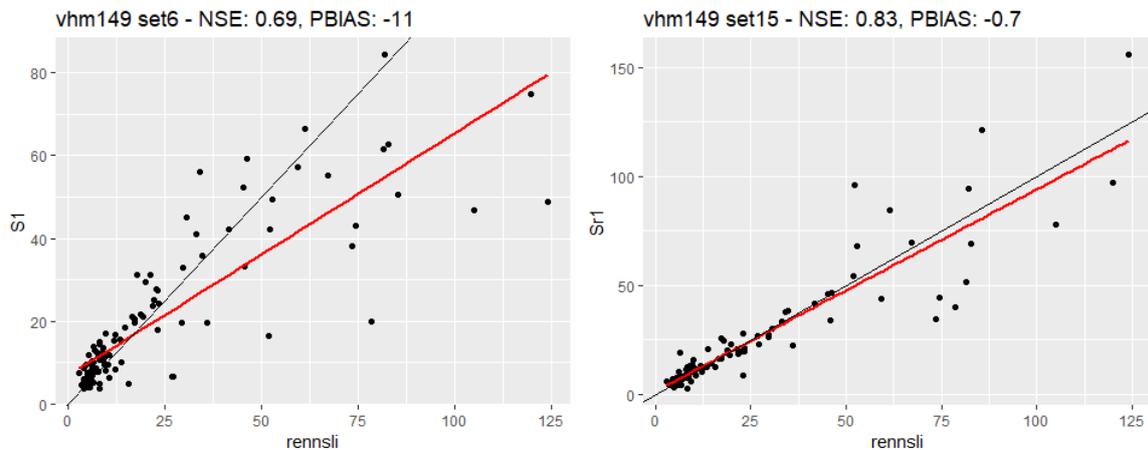


Figure 4. Comparison of one day raw (S1 - left) and rescaled (Sr1 - right) streamflow forecasts with streamflow observations at station vhm. The 1:1 line is represented in black, the linear regression of the data set in red.

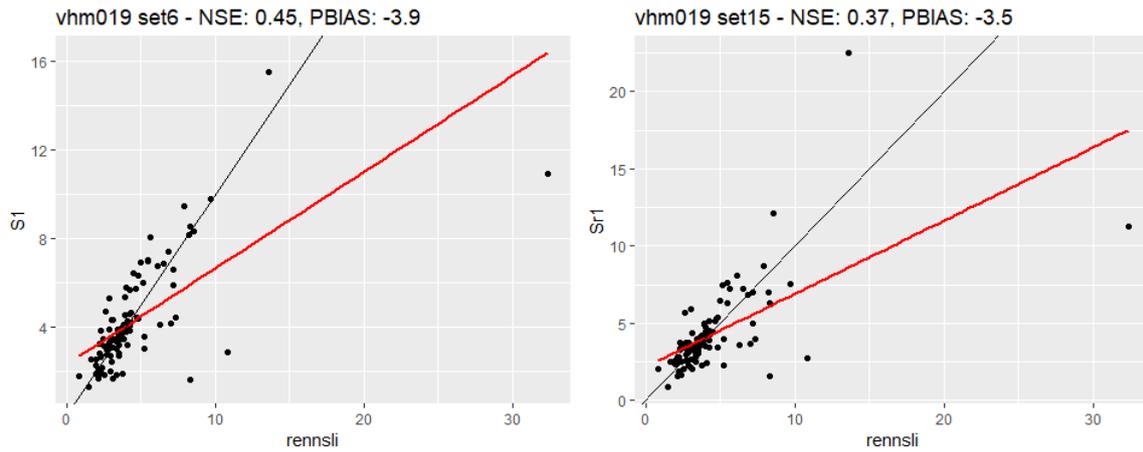


Figure 5. Comparison of one day raw ($S1$ - left) and rescaled ($Sr1$ - right) streamflow forecasts with streamflow observations at station vhm 19. The 1:1 line is represented in black, the linear regression of the data set in red.

Looking at the time series of station vhm 19 (Figure 6) for raw (A) and rescaled (B) discharge, it appears that a delay of about a day is present for some of the largest peaks. This delay is also observed for predictions at other stations (vhm 26, 411, 198 and 200) (see Appendix B).

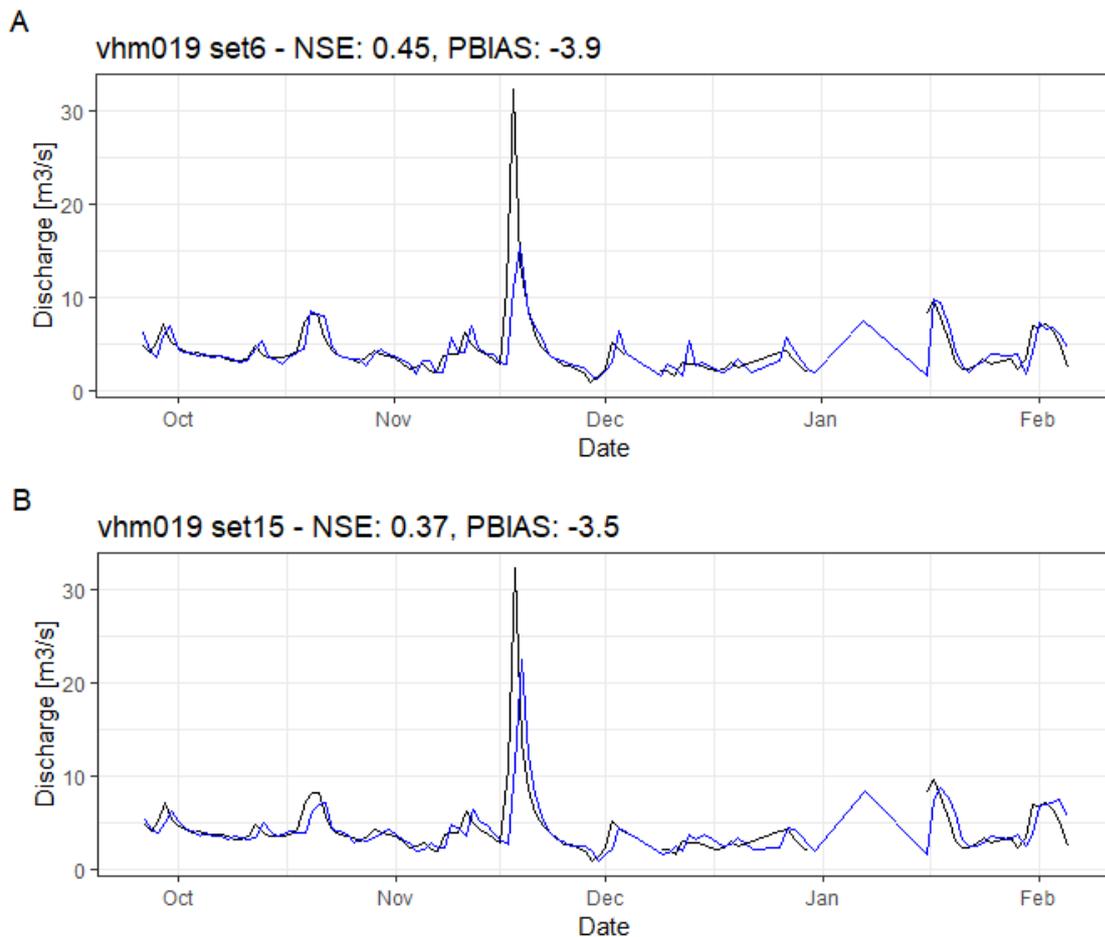


Figure 6. Computed discharge (blue) for vhm 19 with best predictor sets for class C for raw ($S1$) (A) and rescaled ($SR1$) (B) one-day forecast. The observed discharge appears in black.

Table 6 summarizes the best sets for each day of prediction for each class of catchments, both for the raw (S) and rescaled (Sr) discharges. While classes A and C benefit from the presence of forecasted weather variables both for discharge prediction with raw and rescaled discharges, class B is more dependent on past discharges Q1 and Q2 (set 6 and 7) for raw discharge predictions, however weather predictions improve the rescaled deterministic forecast. It is estimated that the best predictions can be achieved by combining the rescaled deterministic streamflow forecasts of the set providing the best NSE for that day (see Table 6).

Table 6. Summary of best predictor sets by class, error measure and day of forecast, e.g. for the first forecast day, class A and raw discharge the best NSE coefficients have been gained using predictor-set 3.

Class	D1				D2				D3				D4				D5				
	NSE	MNSE	RMSE	ME																	
A	S	3	8	10	9	15	15	15	10	20	16	20	10	19	19	19	2	12	16	15	9
	Sr	20	20	15	21	20	20	20	21	20	20	20	20	19	20	19	20	12	5	5	15
B	S	6	6	9	3	7	7	7	2	7	7	1	2	7	7	7	2	7	7	7	2
	Sr	15	15	15	9	19	15	19	15	19	20	19	15	19	20	19	14	19	19	19	19
C	S	15	15	14	11	16	16	20	13	20	16	20	12	19	20	20	2	21	20	20	17
	Sr	15	20	13	22	16	20	14	21	20	20	20	3	20	20	20	3	20	20	20	3

5 Operational webpage

5.1 Operational setup

Two shell scripts have been set up to activate all the programs needed to compute the analogue forecast.

The first shell script, *run_analogue_map*, is triggered by a crontab command every hour and runs the Python script *analogue_sorting_map.py* that creates the flood-warning map displayed on the webpage and described below.

The second shell script, *run_analogue_forecast*, runs once a day at 7:00 UTC and activates several scripts. Firstly, the Python program *spende.py* updates the text-files *spende_VHMXXX.dat* by adding a line to the pre-existing one with the date of the day and the corresponding daily-averaged discharge measured by the corresponding gauging station. Then, the program *harmonie2grid_predictors.py* updates the text-files *predictors_VHMXXX.txt* by rewriting over the last two lines of the file with the newest 24, 48 and 72-hour forecasts from HARMONIE. The script has been written in a way that all the past 24-hour forecasts are conserved. The program *harmonie_grid2predictors48-72h.py* is then run and updates the files *predictors_VHMXXX_48.txt* and *predictors_VHMXXX_72.txt* respectively with the newest 48- and 72-hour forecasts. Afterwards, the R-written program *analogue_forecast_spagaedi.R* generates the 5-day analogue forecast and lastly the Python program *analogue_plots.py* creates the forecasts-figures displayed online.

5.2 Webpage presenting the results

All figures presenting the results from the analogue forecast are hosted on the following website:

http://customer.vedur.is/vegag/analogue_forecast/analogue_sorting.html

The website opens a map of Iceland that shows the location of all the operational catchments selected for this project. The color of each catchment is determined by the latest discharge measurements that have been recorded at the corresponding gauging station. The gauging stations do not record the discharge directly: they measure the water height and convert it into a discharge value using the river's discharge rating curve. Some stations measure the water height hourly (e.g. vhm 012), while vhm 411 records every 30 minutes, vhm 128 every 15 minutes and vhm 150 every 10 minutes. Only the hourly averaged discharge is used here. If the latest discharge measurement is inferior to the value of the 2-year return-period of the river, the watershed is colored in green. If the latest discharge measurement reaches 90% of the value of the 2-year return-period, the catchment will be colored in yellow. It goes on the same way if the observations reach 90% of the 5-year return-period, 10-year return-period and 25-year return-period by taking the colors orange, red and brown respectively. The values of the return-periods have been calculated by using instantaneous discharge observations that were associated with a possible flooding risk. In that sense, the map serves as a warning map that could assist Vegagerðin with travel advisories to the public when the discharge reaches unusually high values. If one of the stations has not been recording for more than a day, a warning sign appears on top of the catchment to warn the reader. An example of map is given on Figure 7.

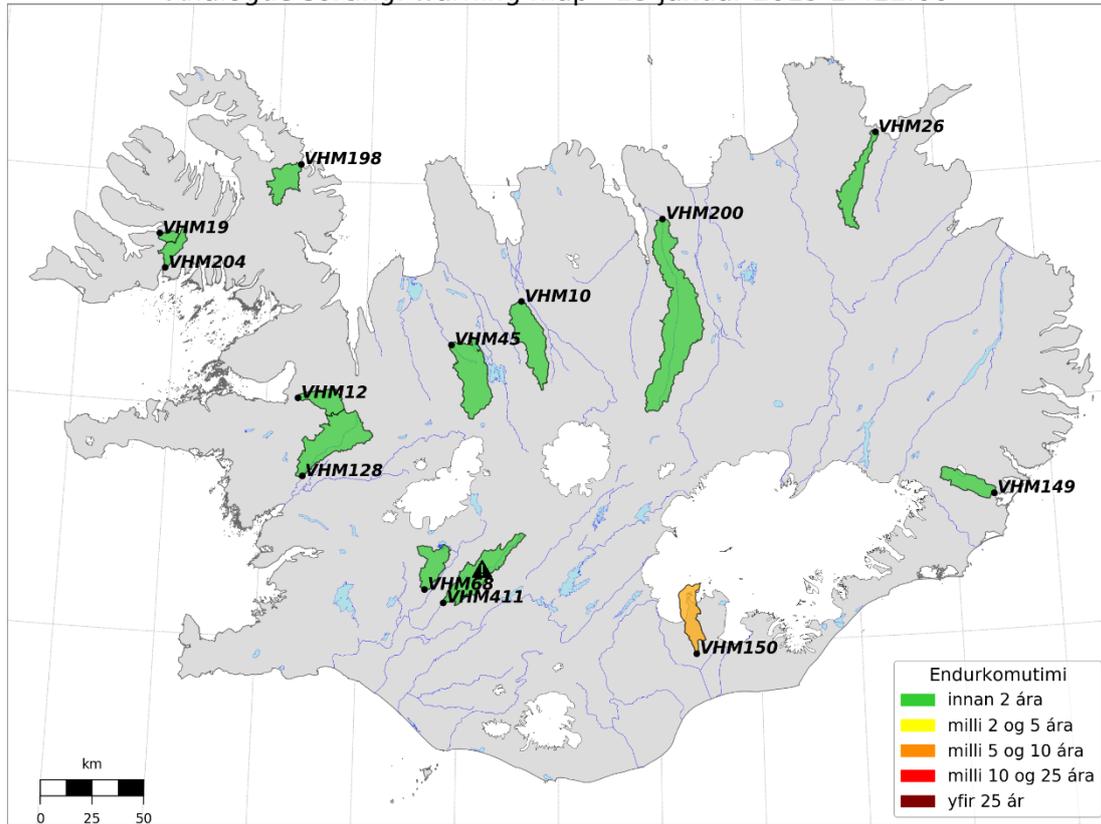


Figure 7. Example of warning map as generated on January 23rd, 2019. Discharge for all stations is under the value of their 2-year return-period, except for vhm 150 where the latest measurements are superior to 90% of the value of the 5-year return-period. A warning sign has appeared on station vhm 411 which indicates that the gauging station has not been recording for more than 24 hours.

As the computer mouse hovers over a catchment, results from the analogue forecast appear to the right of the flood-warning map. Those results are presented on two subplots that are created every morning at 07:00 UTC, an example can be seen on Figure 8 for station vhm 26.

The lower subplot shows the measured daily-averaged discharge over the last 30 days (black line) up until the day prior to the current day which is represented by the vertical dashed grey line. Results from the analogue sorting for the next 5 days are represented by a red line that shows for each day the results from the most efficient predictor-set according to its NSE. A green shading area illustrates the minimum and maximum forecast interval for each day of forecast while results from past forecasts are shown by the light grey shaded area for the past 1-day analogue forecast interval. In addition, the horizontal dashed yellow line indicates the 2-year return-period and corresponding discharge threshold calculated from daily and instantaneous discharge values. If the first daily threshold is reached either by the observed discharge or by the forecast, the next one will be displayed and so on until the last threshold (25-year return-period) is reached.

The upper sub-plot shows the simulated daily-averaged temperatures (purple line) and daily-summed rainfalls (grey bars) for the last 30 days and for the next 3 days as predicted by HARMONIE. The 0°C isotherm is represented by the dashed purple line and again the vertical dashed grey line indicates the day prior to the forecast.

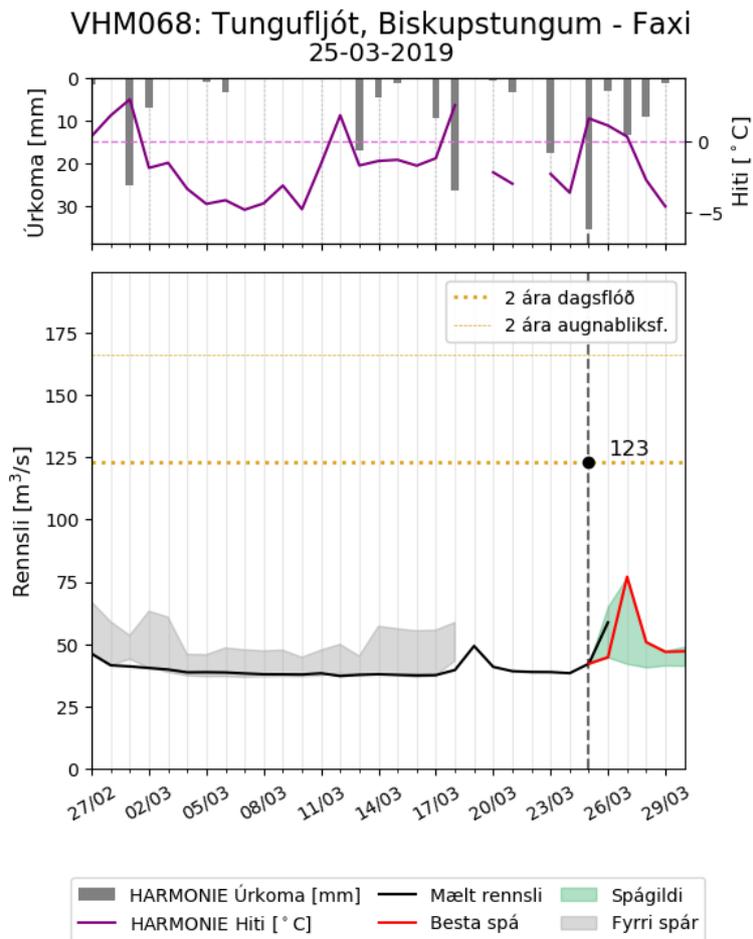


Figure 8. Forecast results for vhm 26 on March 25th, 2019.

It is also possible to click on a catchment and open a new page similar to the screenshot displayed on Figure 9. Results from the analogue forecast are displayed in the same way as on the front page of the website although, in this case past results are shown both for the last 3 months (left) and the last 10 days (right). Three boxplots have also been added and display temperature, precipitation and discharge values of the day (red dot) and compare them to previous values for the same Julian day, giving some indication of where the current values stand compared to statistics from previous years.

VHM 26: Sandá, Þistilfirði - Flögubré

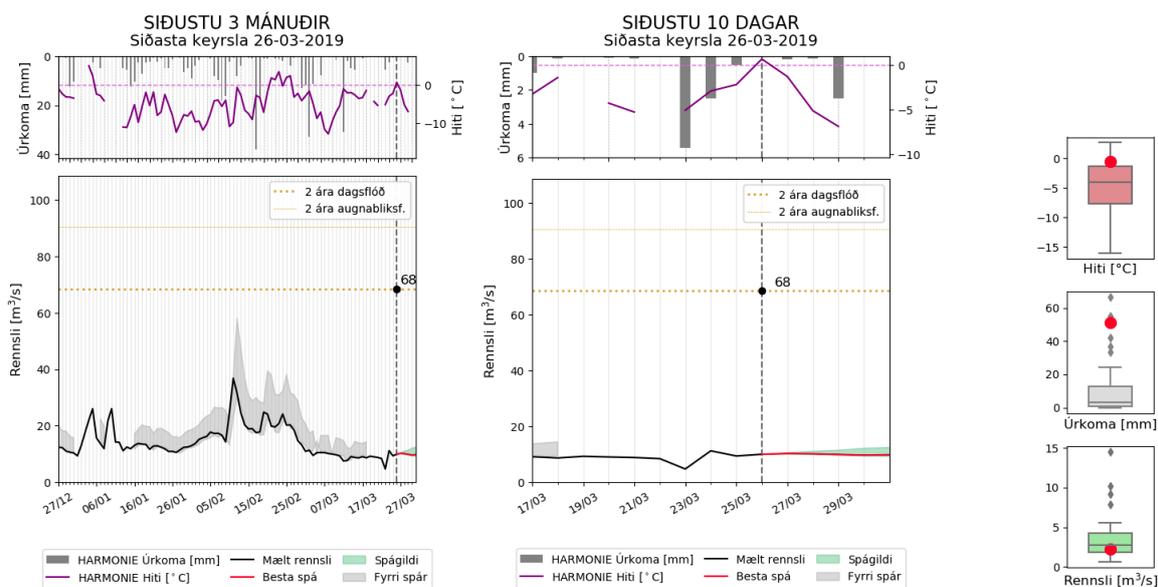


Figure 9. Screenshot presenting the results of the analogue forecast as well as temperature, precipitation and discharge boxplots for vhm 26 on 25 March, 2019.

6 Discussion

The analogue sorting method was able to show its great ability to forecast discharge up to 5 days ahead of time when provided with a homogeneous dataset (from reanalysis). However, in an operational setup, some biases are introduced both by the forecasts and the observed uncorrected discharge that can suggest a sudden increase of streamflow when the meter is covered with ice and results in false forecasts. Raw (S) forecasts of discharge do not always manage to capture the highs as well as the rescaled version (Sr), however rescaled forecasts tend to reproduce the artifacts created by the icing of the meters and can in times result in exaggerated highs. This could be a part of the reason why the rescaled forecast (Sr) does not seem to improve the raw forecast (S) in specific catchments (see Table 5), as well as the effects of snow cover on the HMV model.

Timeliness of peak flow is crucial for forecasting floods and can be captured by the model up to 4 days before the event, however some stations are affected in winter by a snow cover which induces too low air temperatures in the forecasts of the HMV model (Nawri et al., 2017) resulting in delayed streamflow forecasts. The reanalysis weather variables from the HMV and ICRA models were compared over a period of four months in 2017 (Nawri, Pálmason, Petersen, Björnsson, & Þorsteinsson, 2017) in order to establish if there was a continuity between the two models results. Figure 10 shows spatial variability with an underestimation of temperature in September and October 2017 in the north-east, the Westfjords and in the highlands. Nawri et al. (2017) suggests that these differences originate from the earlier accumulation of snow on the ground resulting in lower temperature. In a previous report, Nawri et al. (2017) indicate the tendency of the ICRA model to underestimate the temperature of air over snow cover and propose a correction that minimize this error. The past climatic data in this model is ICRA reanalysis data including post processed temperature, as it is expected to be most accurate. However, forecasted temperature based on the HMV model is expected to underestimate even more the air temperature over snow cover, creating a bias in the computation of the Mahalanobis distance that could explain this one-day delay. In some instances of very large peaks, this delay associated with rescaling can result in an amplified error and the computation of extravagant discharge.

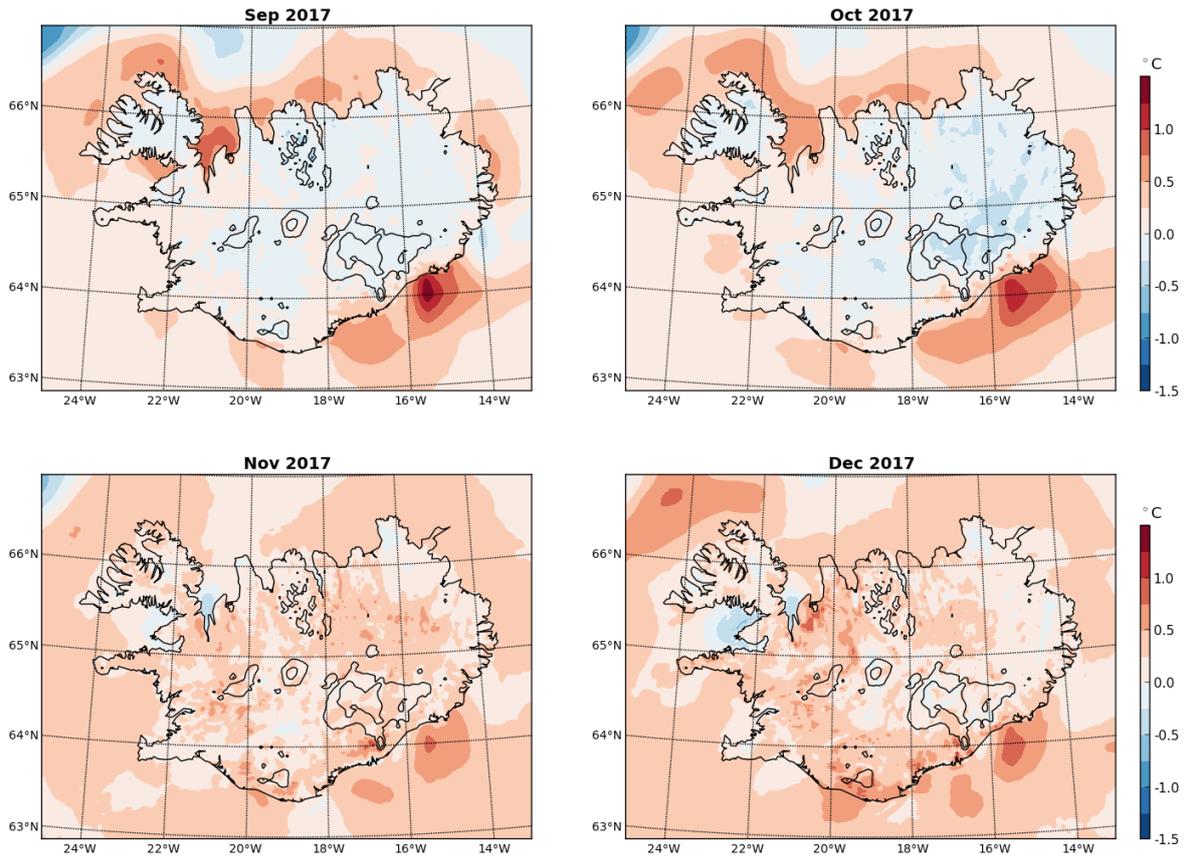


Figure 10 Comparison of air temperatures 2 m computed by HVM and ICRA reanalysis. Red colors represent increased temperatures by hmv compared to icra computations, and inversely for blue colors (Nawri, Pálmason, Petersen, Björnsson, & Þorsteinsson, 2017).

It is therefore important to keep a critical eye and consider the weather conditions when looking at the predictions. For that purpose, temperature and precipitation associated with each catchment were added to the webpage. Absence or missing data for some of the predictors is not yet circumvented and result in absence of streamflow prediction. This will be fixed in future updates of the webpage.

7 Conclusions

This research project aimed at developing a simple operational streamflow forecast system based on an analogue sorting method. Thirteen catchments were used to assess the method. The model has been operational since September 2018 and shows that this simple method can successfully predict streamflow in thirteen diverse catchments around Iceland. The model provides acceptable 1-day forecast. With increased lead time, the quality of the forecast decreases, but remains reasonable for up to 4 days ahead (and up to 5 days in some cases). Cluster analysis of the catchment characteristics in to three categories emphasized the most relevant predictors for each category of catchments. This classification was used to simplify the selection of predictor sets. An operational webpage showing the results of forecasted discharge has been set up and is updated on a daily basis.

Further studies will involve the addition of catchments to the flood forecasting system. Pre- and post-processing routines will also be developed in order to correct biases associated with ice perturbation of the gauging station and air temperature over snow cover. Results of this method is expected to improve with time, since longer timeseries are more likely to include similar scenarios to what is observed/forecasted. Possible optimization regarding the website could include automatic messages sent when a threshold is reached and the addition of a warning sign when very low temperatures are reached to report possible ice perturbations. In this project, discharge has been used as a systematic predictor, but while climatic information is available for the whole country, discharge measurements are sparse. Some investigation will be needed in order to apply this methodology to forecast discharge on catchments with no discharge information. Some biases associated to the datasets are resulting in artifacts in the predictions and will need to be addressed in further studies, but the webpage presenting the forecast can educate the user and help him discriminate the event by combining the streamflow information with climatic dataset for each of the catchment.

A simple operational streamflow forecasting system based on an analogue sorting method has been successfully developed and tested on thirteen catchments in Iceland. The results have been set up on a webpage which shows a map that can serve as a warning map, in order to assist Vegagerðin both with travel advisories to the public as well as for internal information regarding preparedness for possible flood events.

8 References

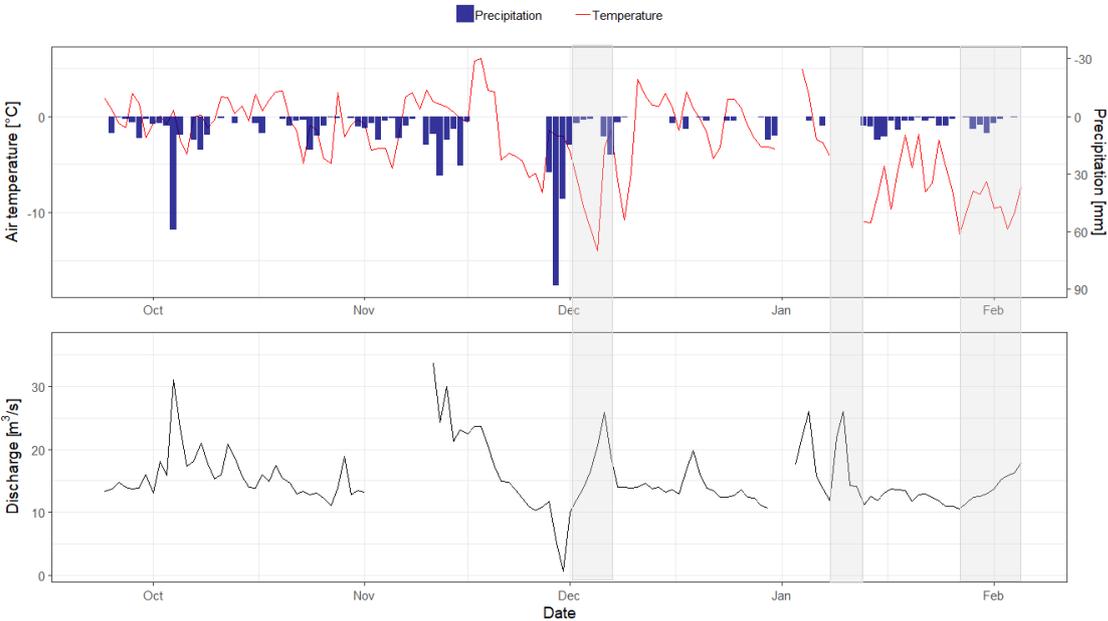
- Akbari, M., Van Overloop, P., & Afshar, A. (2011, March). Clustered K Nearest Neighbor Algorithm for Daily Inflow Forecasting. *Water resources management*, 25(5), 1341-1357.
- Crochet, P. (2012). *Evaluation of two delineation methods for regional flood frequency analysis in northern Iceland*. Veðurstofa Íslands.
- Crochet, P. (2013). *Probabilistic daily streamflow forecast using an analogue sorting method*. Reykjavík: Veðurstofa Íslands.
- Demirel, M., & Kahya, E. (2007). Hydrological determination of hierarchical clustering scheme by using small experimental matrix., (bls. 161-168). doi:10.13140/RG.2.1.1885.6720
- Franz, K. J., & Hogue, T. S. (2011). Evaluating uncertainty estimates in hydrologic models: borrowing measures from the forecast verification community. *Hydrol. Earth Syst. Sci.*, 3367-3382.
- Gupta, H. V., Kling, H., Yilmaz, K. K., & Martinez, G. F. (2009). Decomposition of the mean squared error and NSE performance criteria.: *Journal of Hydrology*, 80-91.
- Hróðmarsson, H. B., & Þórarinsdóttir, T. (2018). *Flóð íslenskra vatnsfalla, flóðagreining rennslisraða*. Reykjavík: Veðurstofa Íslands.
- Karlsson, M., & Yakowitz, S. (July 1987). Nearest-neighbor methods for nonparametric rainfall-runoff forecasting. *Water Resources Research*, 23(7), 1300-1308.
- Krause, P., Boyle, D., & Bäse, F. (2005). Comparison of different efficiency criteria for hydrological model assessment. *Advances in Geosciences*, 89-97.
- Nawri, N., Pálmason, B., Petersen, G. N., Björnsson, H., & Þorsteinsson, S. (2017). *The ICRA atmospheric reanalysis project for Iceland*. Icelandic Meteorological Office, Reykjavík.
- Oyebode, O., Otieno, F., & Adeyemo, J. (July 2014). Review of three data-driven modelling techniques for hydrological modelling and forecasting. *Fresenius Environmental Bulletin*, 23(7), 1443-1454.
- Rist, S. (1990). *Vatns er þörf*. Reykjavík: Bókaútgáfa Menningarsjóðs.
- Sigurðsson, F., & Einarsson, K. (1988). Groundwater, resources of Iceland: availability and demand. *Jökull*, 38, 35-54.
- Stefánsdóttir, G., Björnsson, B. B., Magnússon, S., & Egilson, D. (2014). *Verklokaskýrsla vegna stjórnar vatnamála - Vinna ársins 2013*. Reykjavík: Veðurstofa Íslands.
- Veðurstofa Íslands (2019). *Rennslisskýrsla vatnsárið 2017/2018, V550, Djúpá, neðan Djúparðals*.
- Veðurstofa Íslands (2016). *Rennslisskýrsla vatnsárið 2014/2015, V449, Dynjandisá, Sjóarfoss*.
- Veðurstofa Íslands (2017). *Rennslisskýrsla vatnsárið 2015/2016, V12, Haukadalsá, Hvammsfirði; Haukadalsvatn; útfall*.
- Veðurstofa Íslands. (2017). *Rennslisskýrsla vatnsárið 2015/2016, V204, Vatnsdalsá, Vatnsfirði; Eiði*.
- Veðurstofa Íslands (2017). *Rennslisskýrsla vatnsárið 2015/2016, V411, Stóra-Laxá, Stórhylur*.
- Veðurstofa Íslands (2017). *Rennslisskýrsla vatnsárið 2015/2016, V45, Vatnsdalsá, Forsæludalur*.
- Veðurstofa Íslands (2018). *Rennslisskýrsla vatnsárið 2016/2017, V198, Hvalá, Óp*.
- Veðurstofa Íslands (2019). *Rennslisskýrsla vatnsárið 2017/2018, V128, Norðurá, Borgarfirði; Stekkur*.
- Veðurstofa Íslands (2019). *Rennslisskýrsla vatnsárið 2017/2018, V149, Geithellnaá; Álftafirði; Geithellur*.

- Veðurstofa Íslands (2019). *Rennslisskýrsla vatnsárið 2017/2018, V200, Fnjóská, ofan Árbugsár.*
- Zalachori, I., Ramos, M.-H., Garçon, R., Mathevet, T., & Gailhard, J. (July 2012). Statistical processing of forecasts for hydrological ensemble prediction: a comparative study of different bias correction strategies. *Advances in Science and Research*, 8, 135-141. doi:10.5194/asr-8-135-2012
- Zóphóniasson, S., & Kristinsson, B. (2008). *Rennslisgögn úr vatnshæðarmæli 10 í Svartá í Skagafirði við Reykjafoss: árin 1962-1997.*
- Pórarinsdóttir, T. (2012). *Development of a methodology for estimation of Technical Hydropower potential in Iceland using high resolution Hydrological Modeling.* Veðurstofa Íslands.

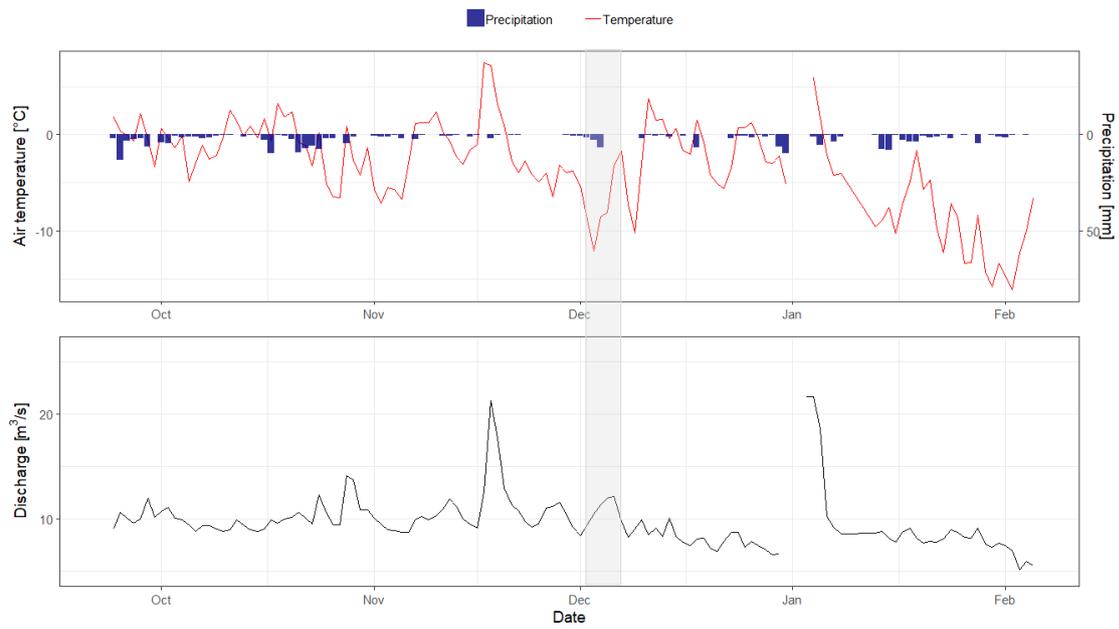
Appendix A. Precipitation, air temperature and discharge data used for the test on operational data

In the following sections, temperatures and precipitation from HMV are presented alongside discharge measurements in order to outline the disruption of the dataset due to data loss (network problem, disruption of computation or transmission issues) and the possible ice perturbations occurring at the gauging stations (emphasized in gray in the following figures) resulting in artifacts in the streamflow observations. These artifacts are often observed when air temperatures fall below -10°C and take the shape of a fast-rising discharge peak which drops as soon as the temperature increases again.

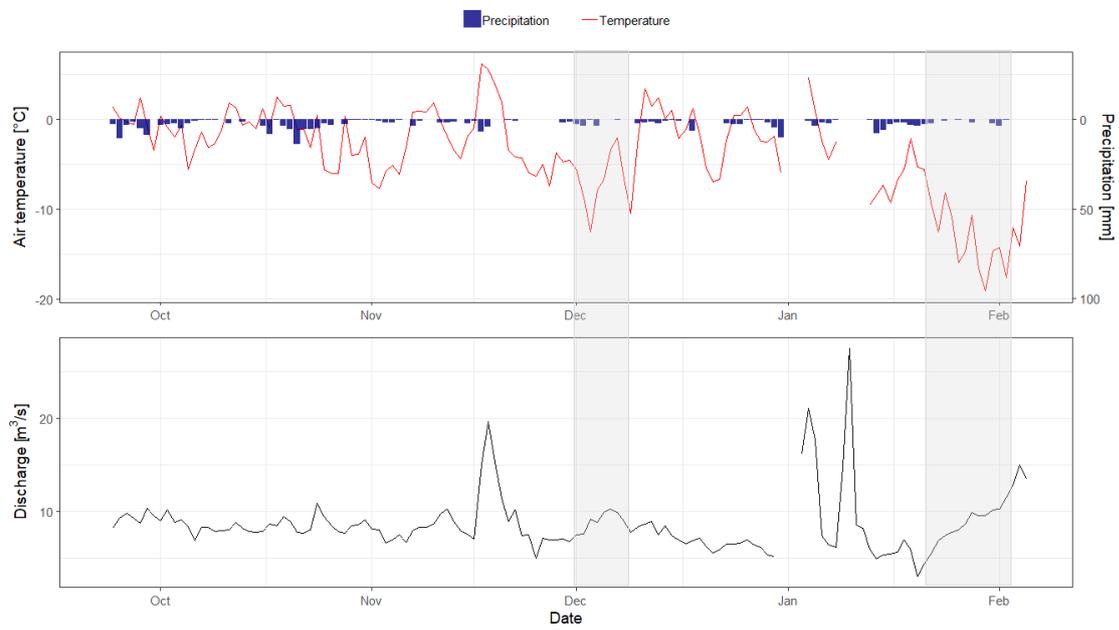
A.1. vhm 26



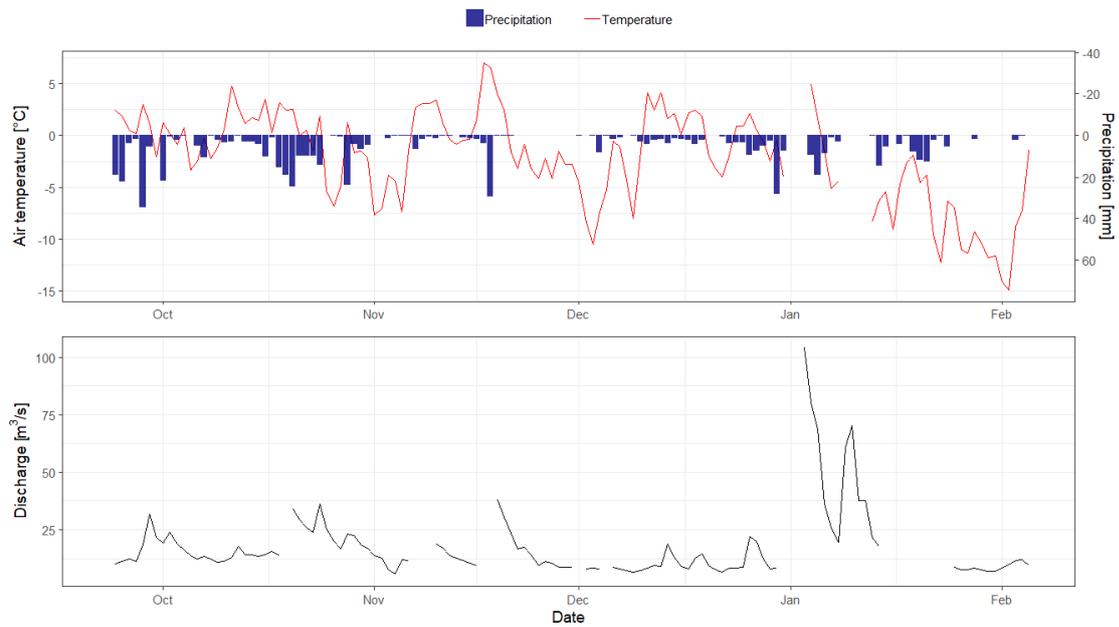
A.2. vhm 10



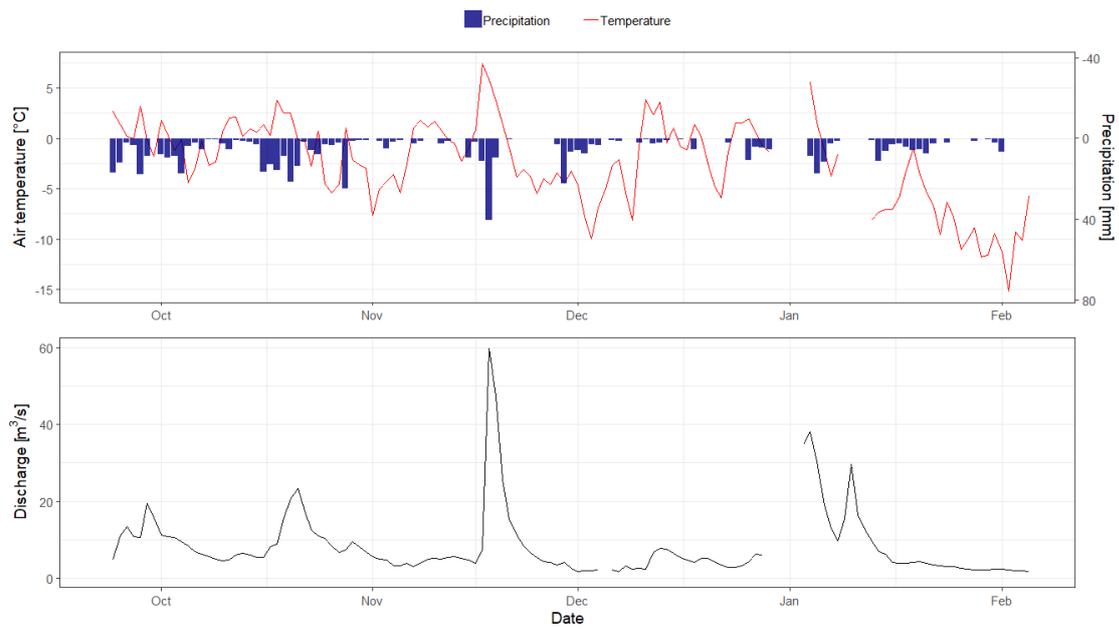
A.3. vhm 45



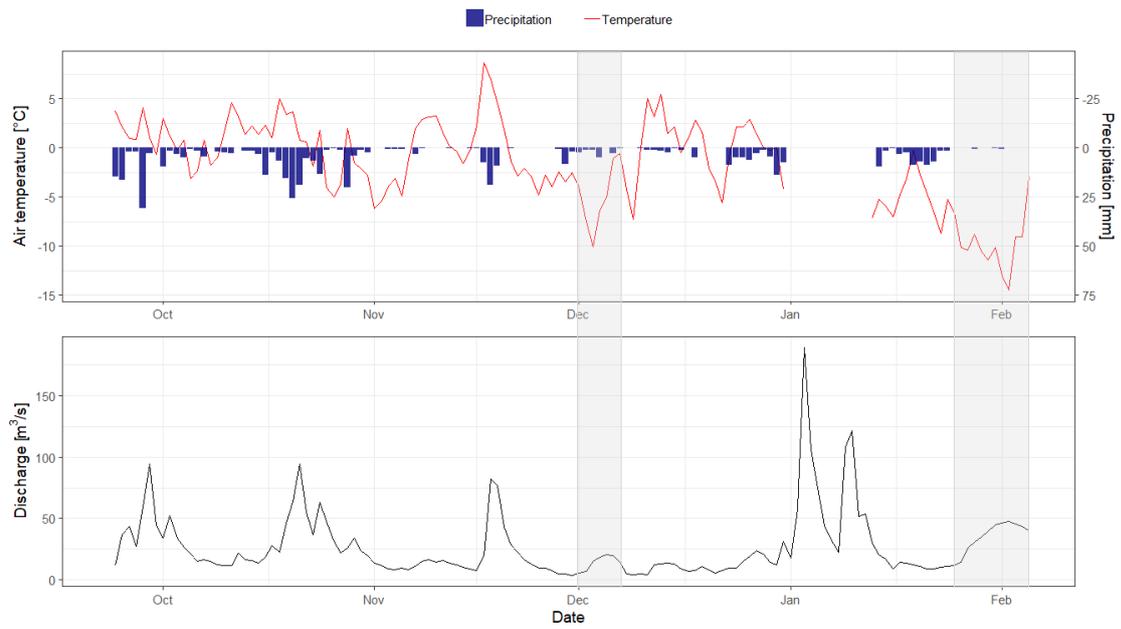
A.4. vhm 411



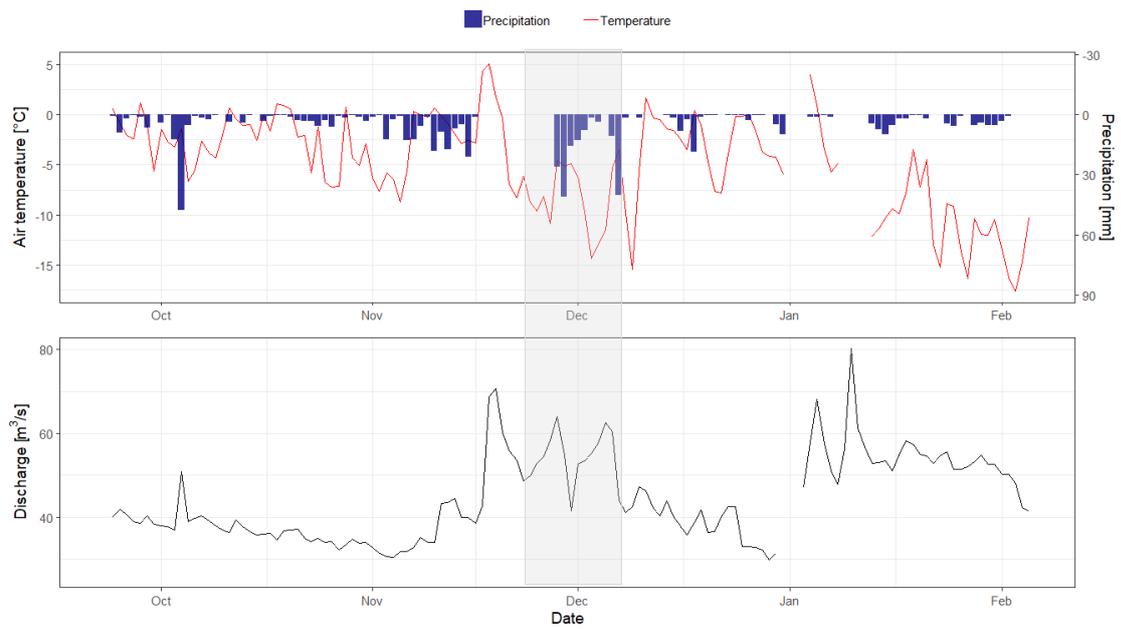
A.5. vhm 12



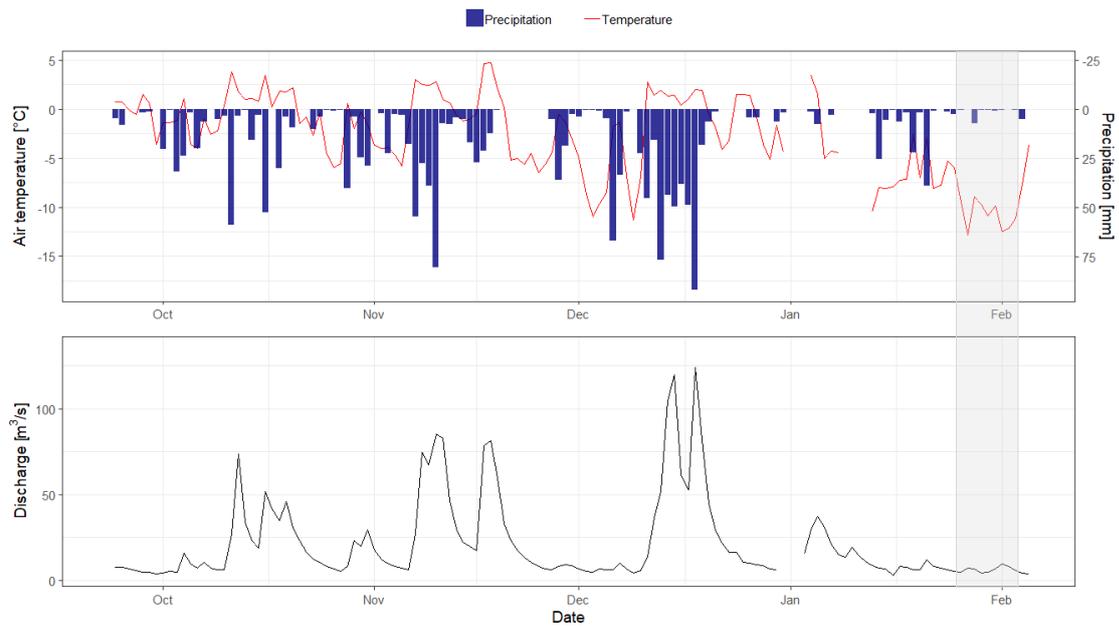
A.6. vhm 128



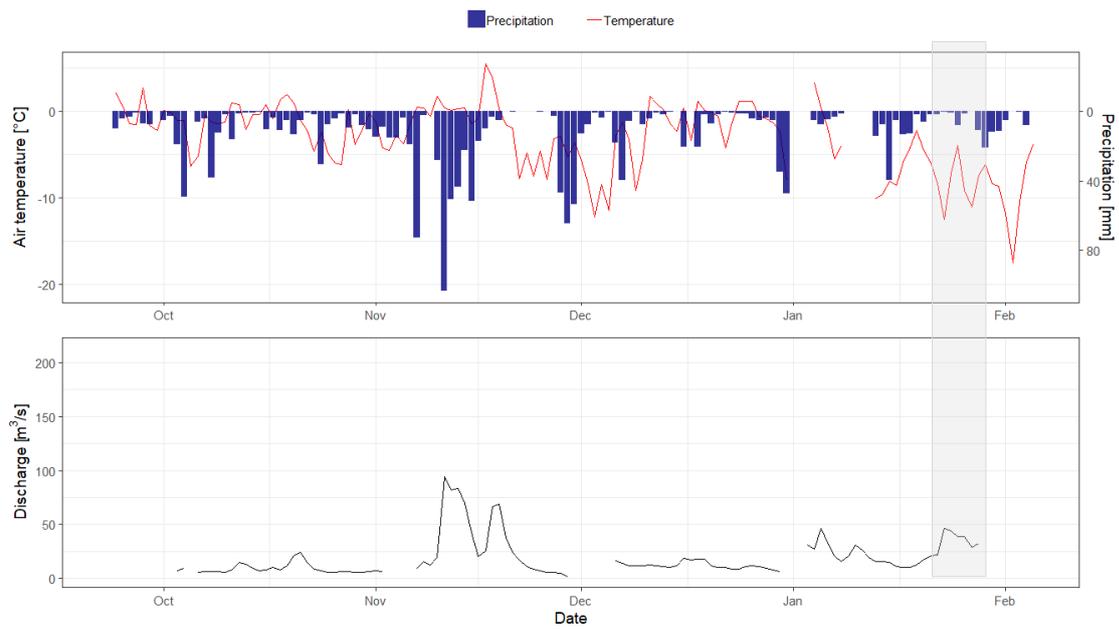
A.7. vhm 200



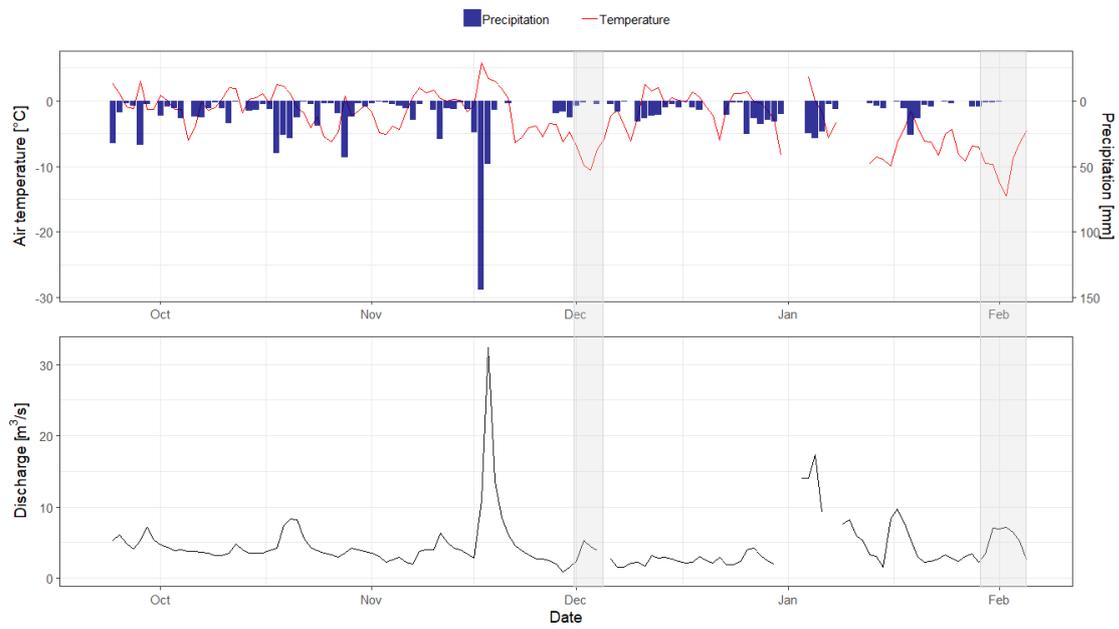
A.8. vhm 149



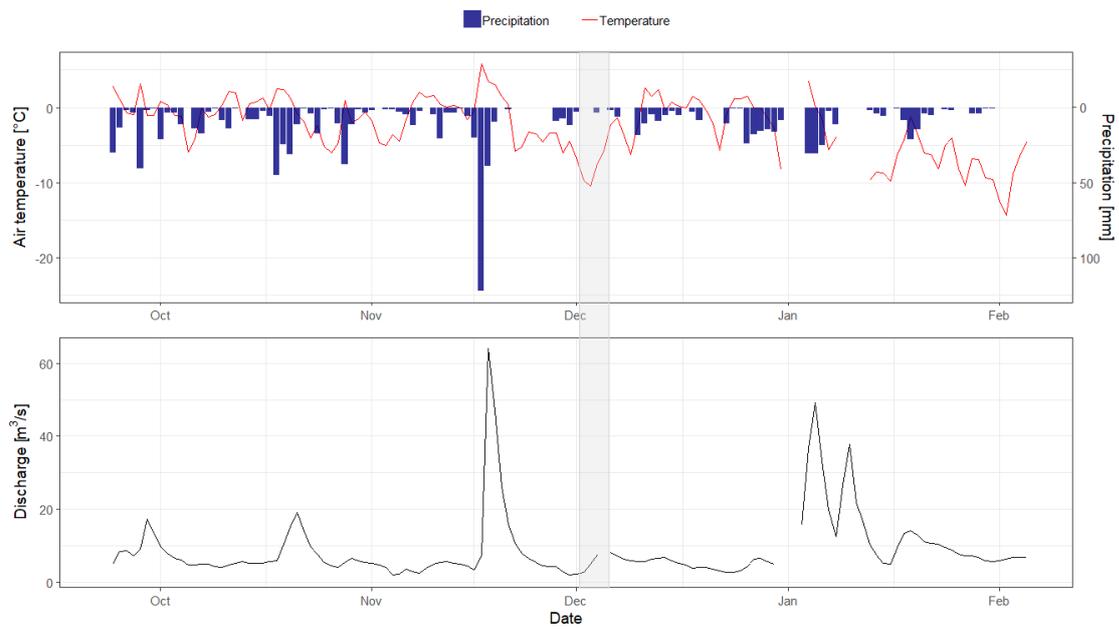
A.9. vhm 198



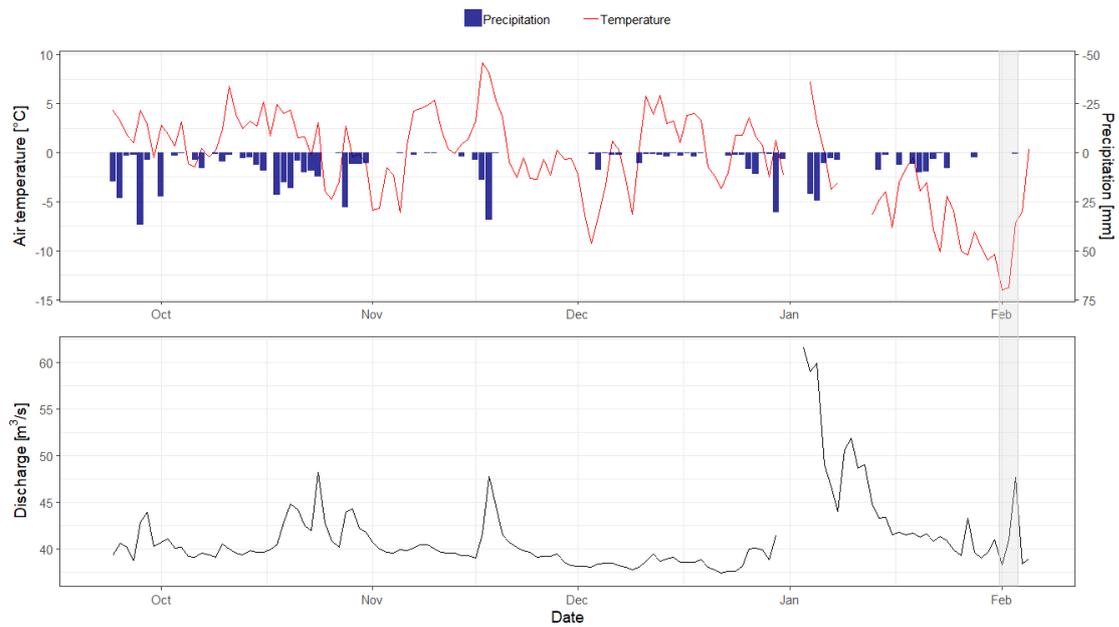
A.10. vhm 19



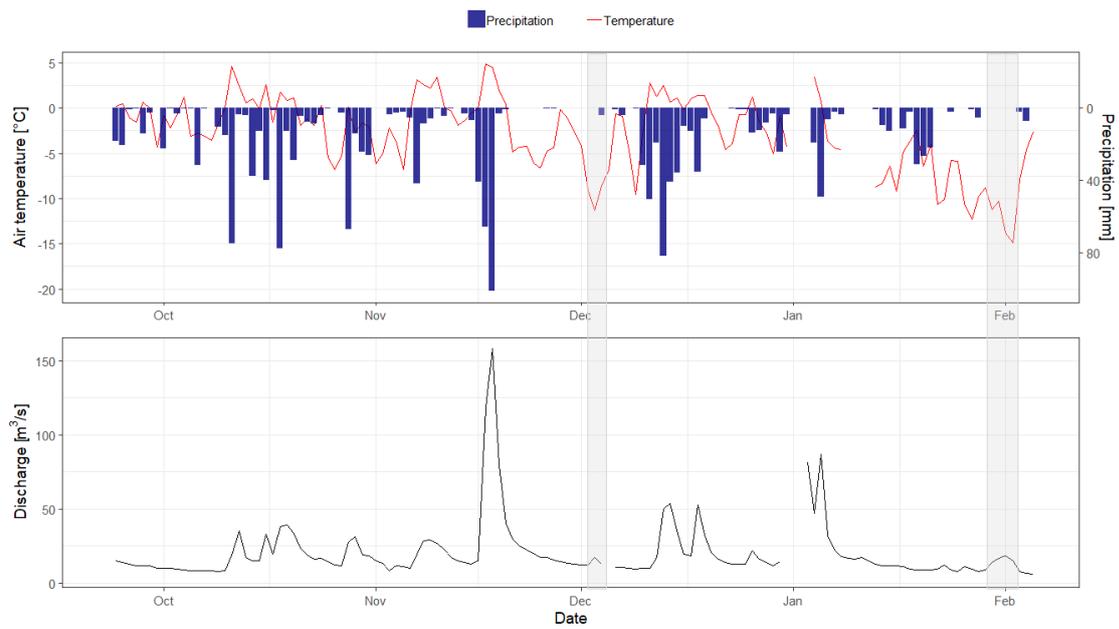
A.11. vhm 204



A.12. vhm 68



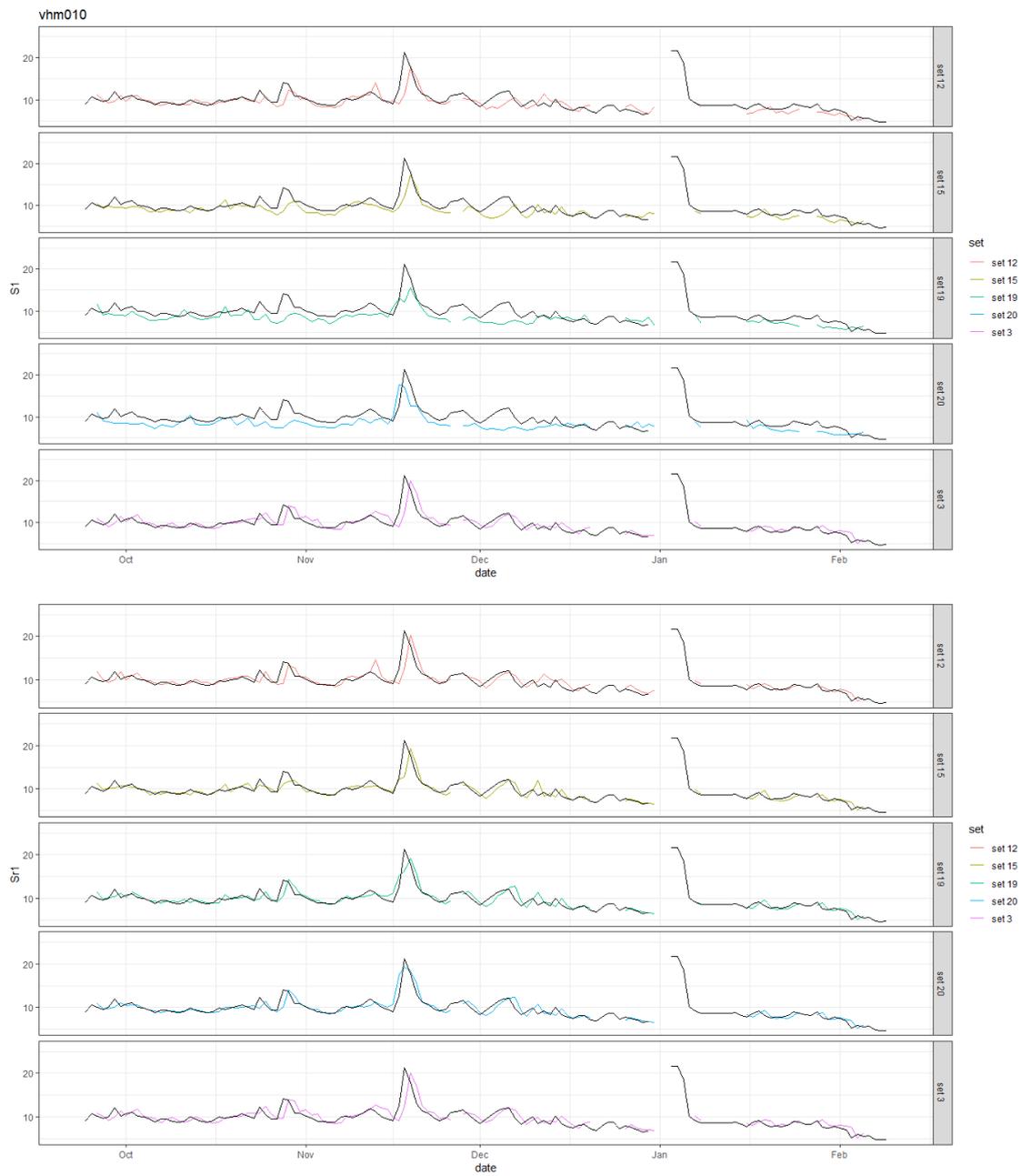
A.13. vhm 150

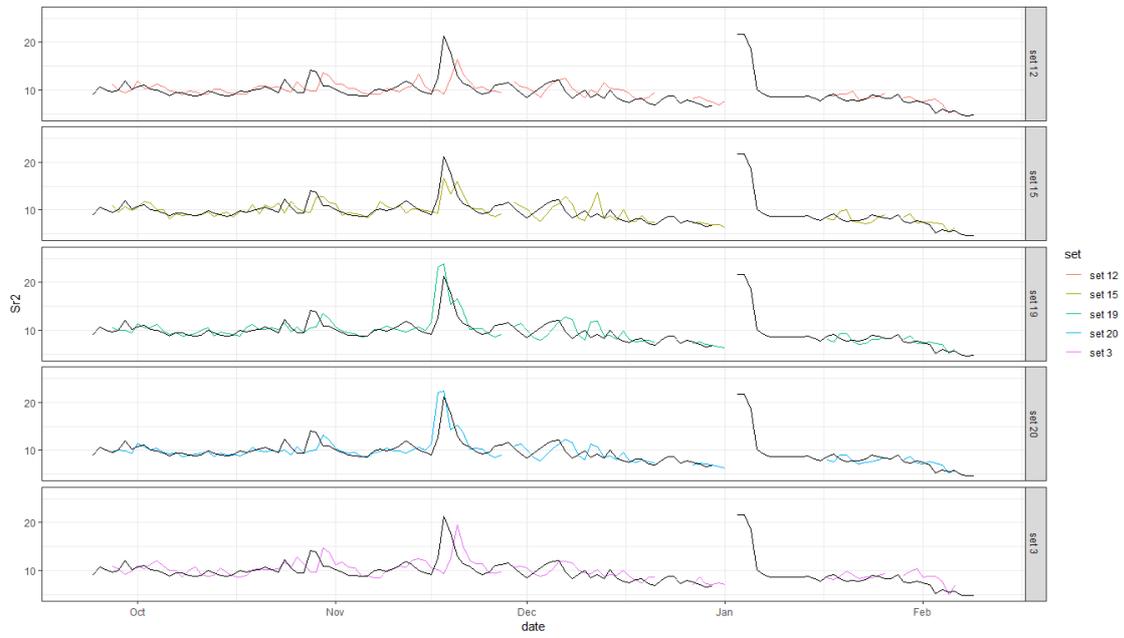
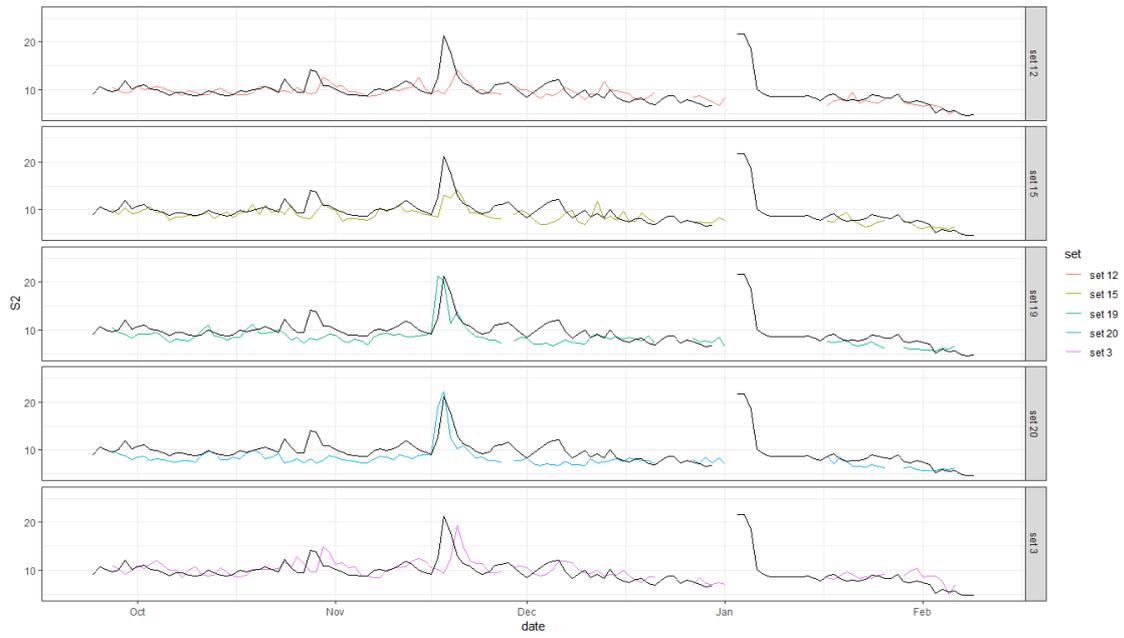


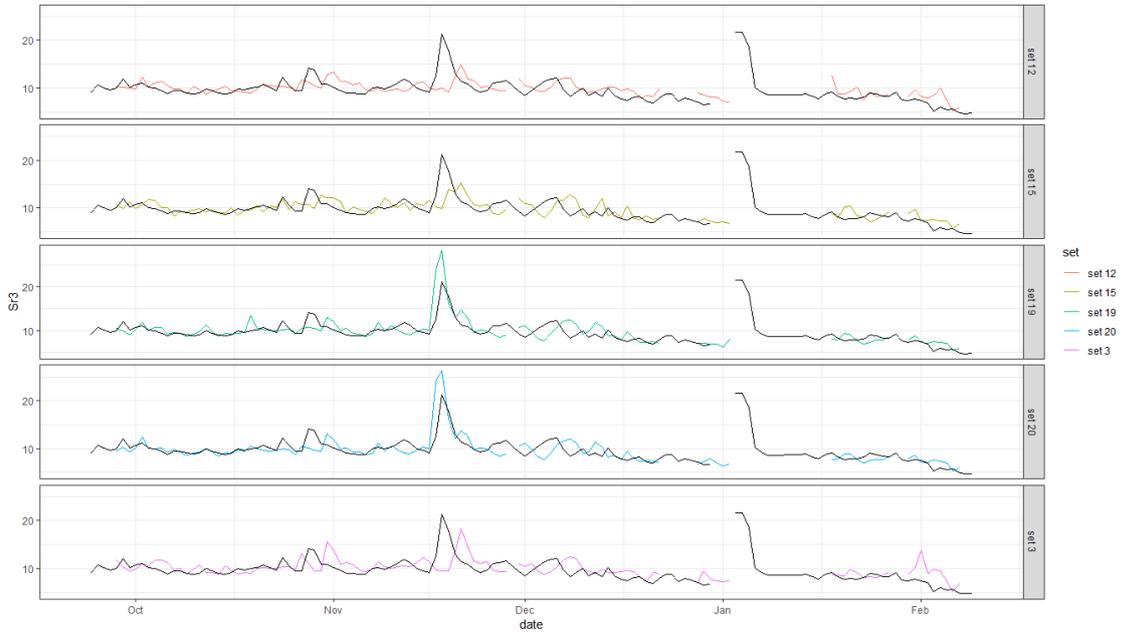
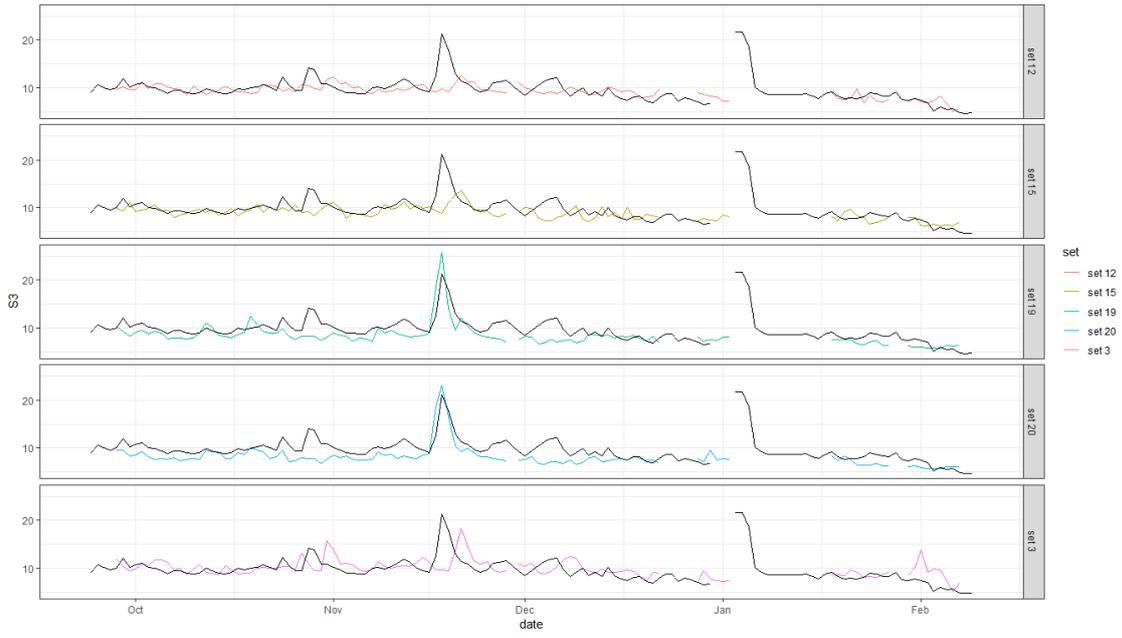
Appendix B. Best sets

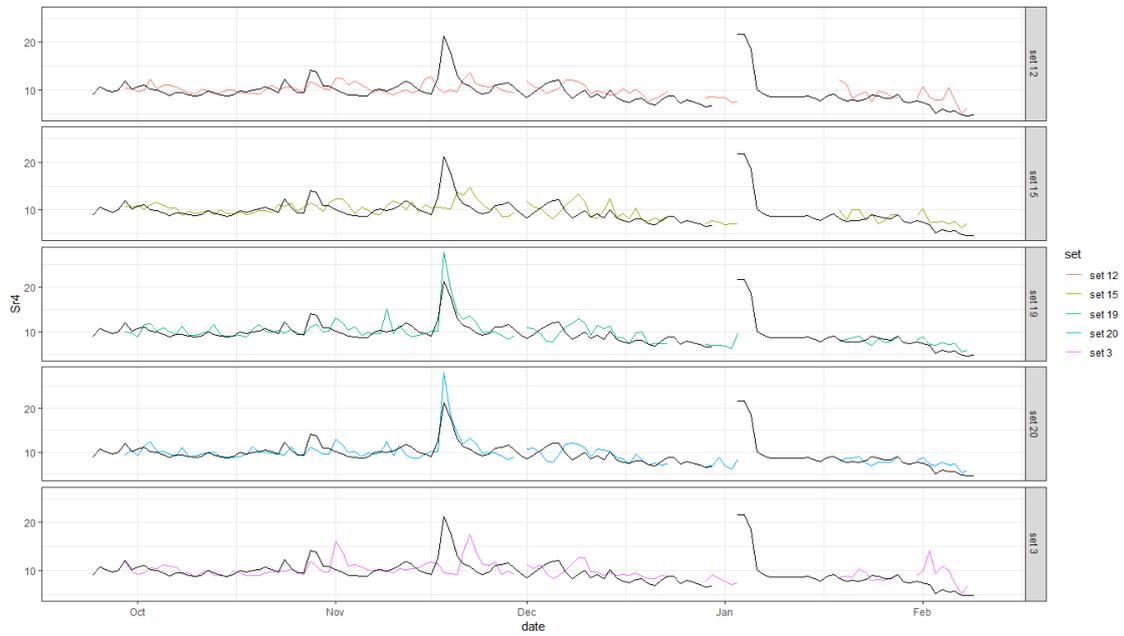
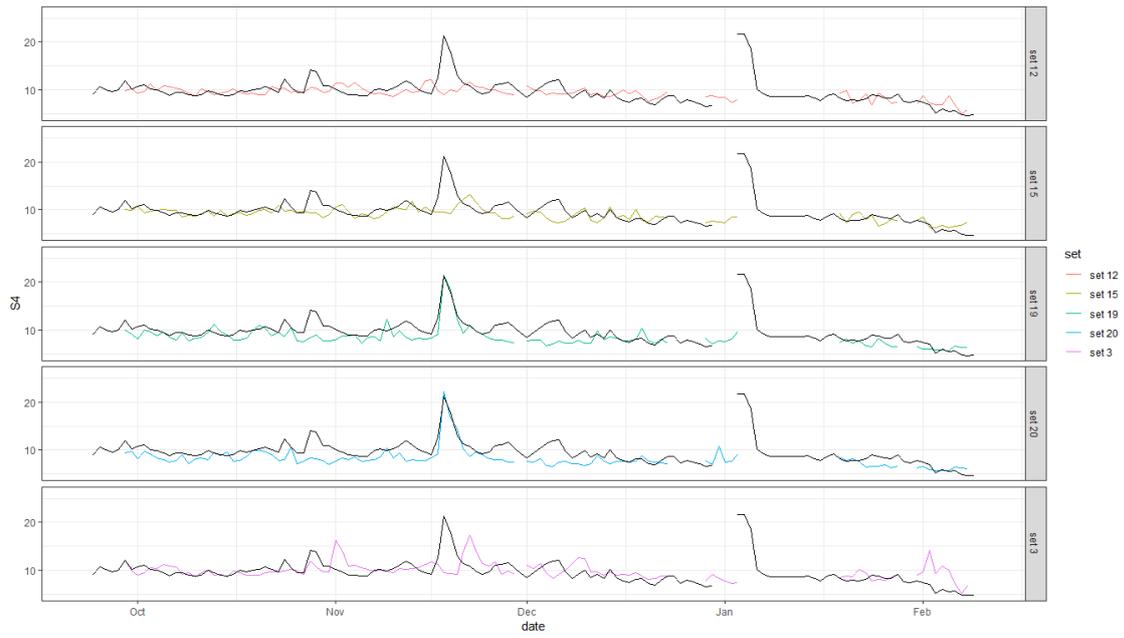
B.1. Class A

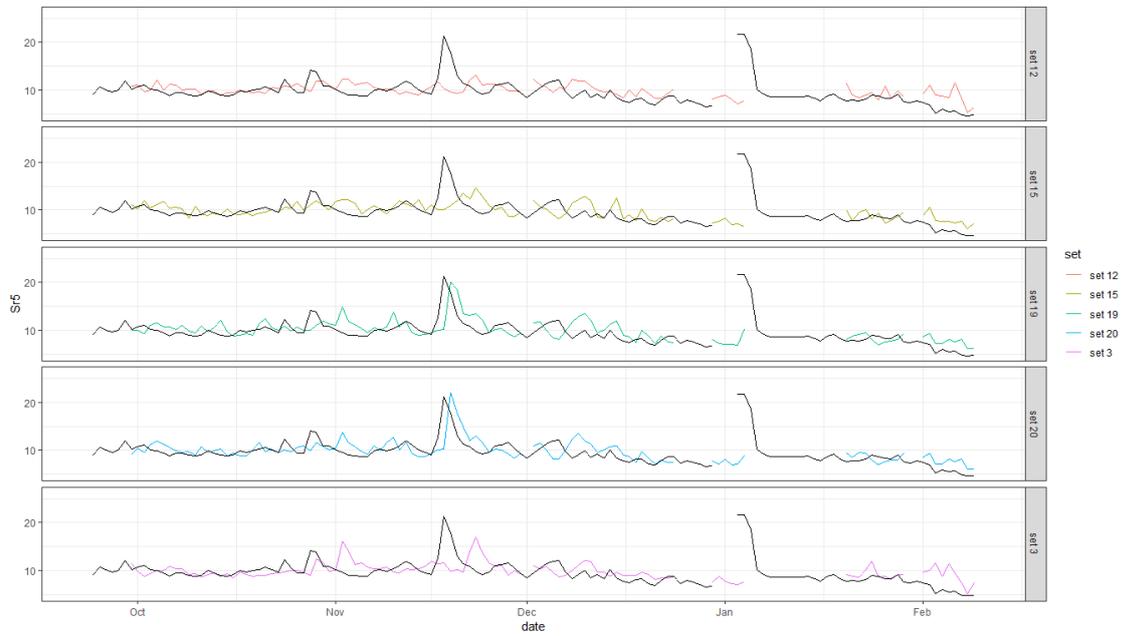
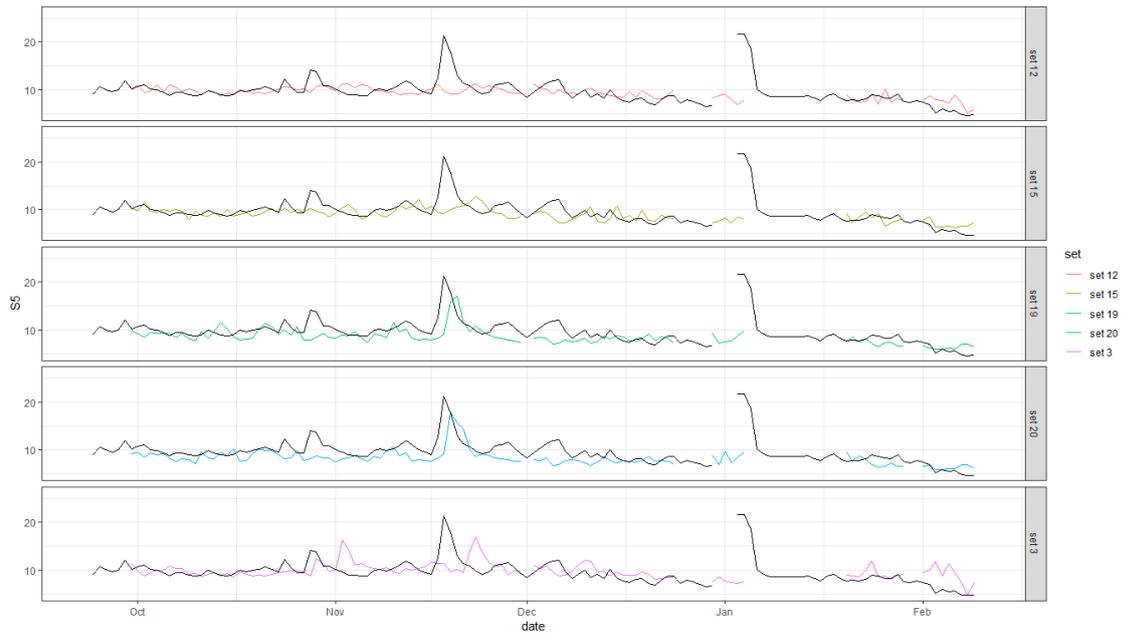
B.1.a. vhm 10



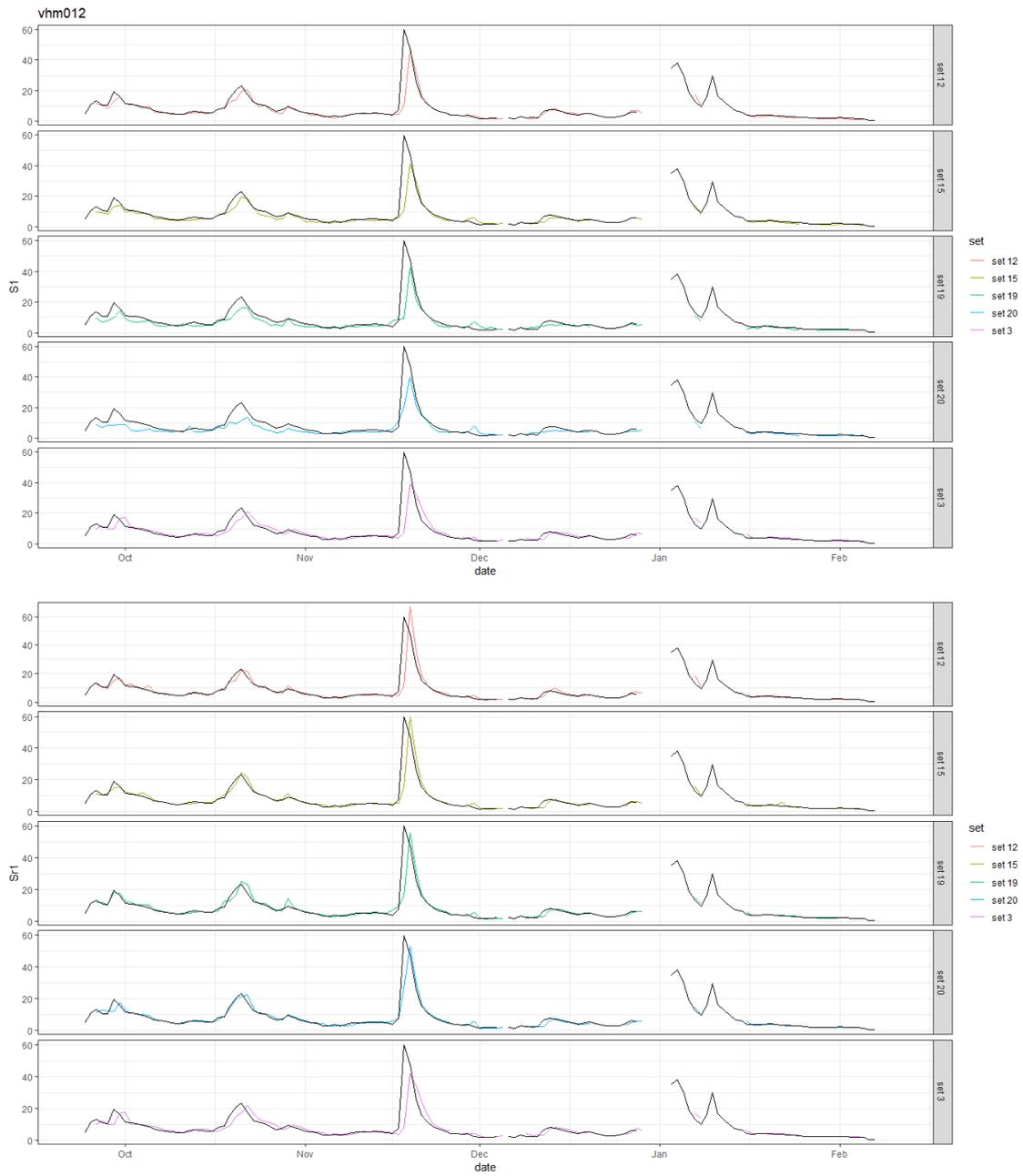


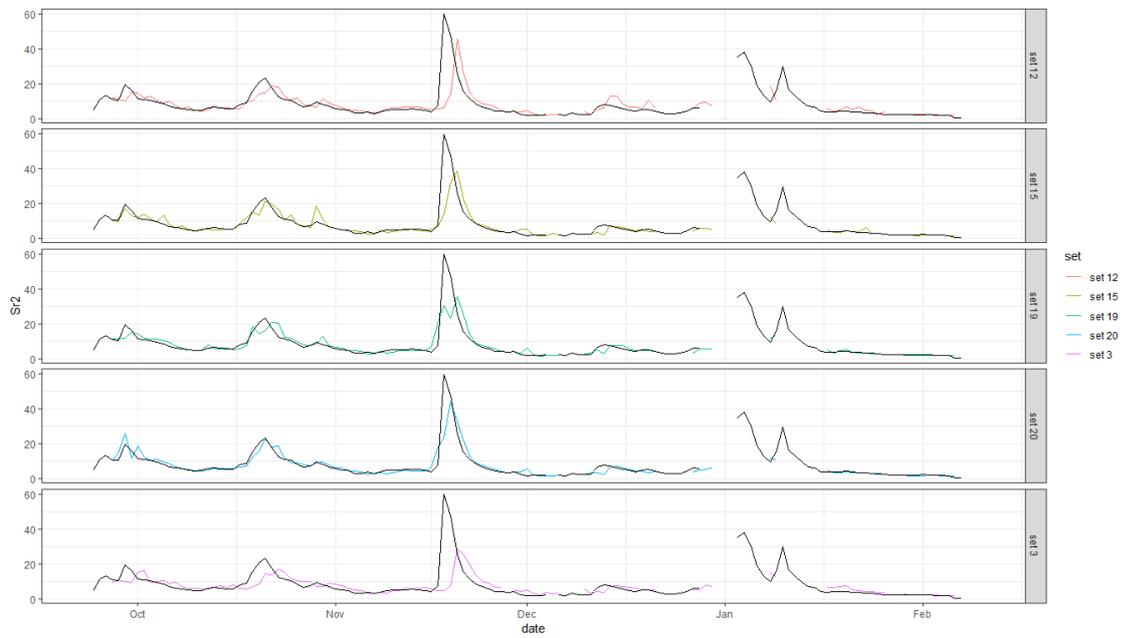
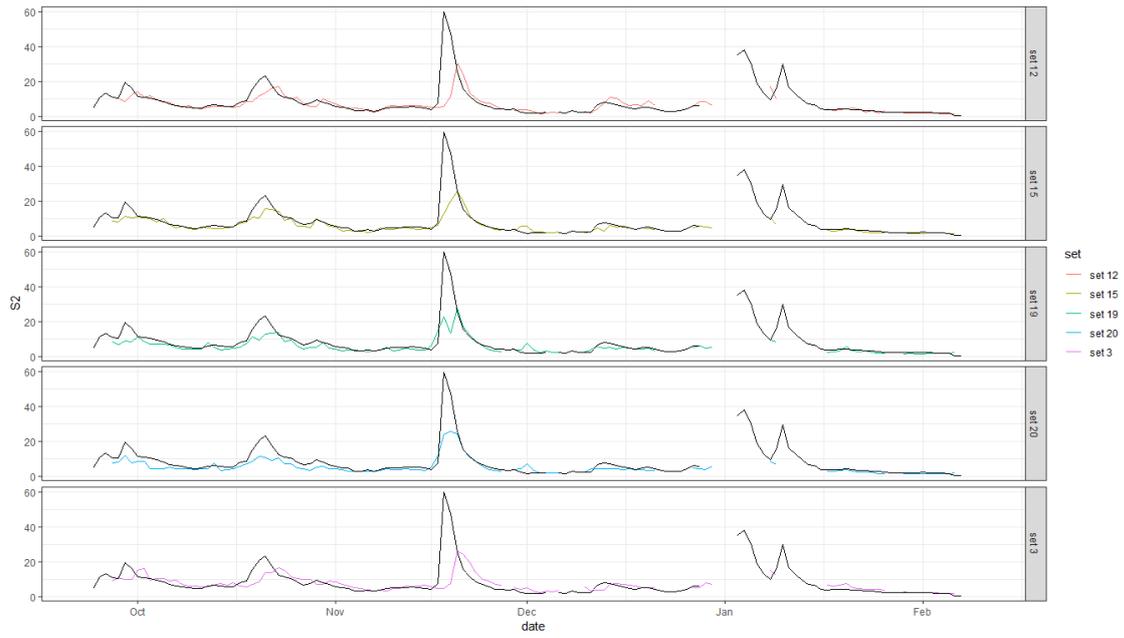


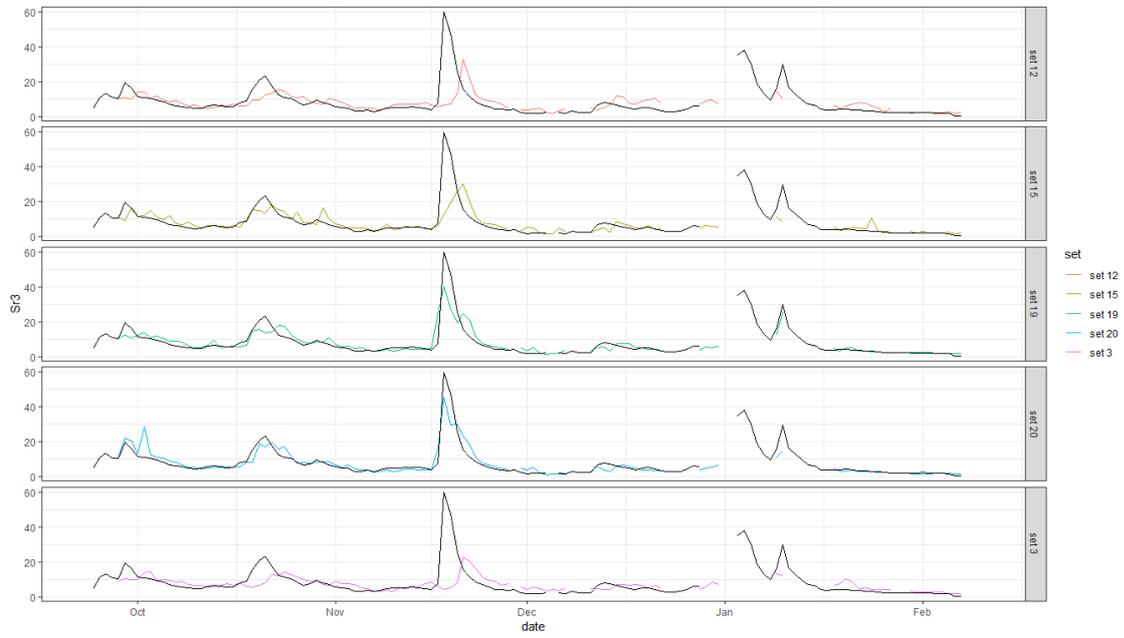
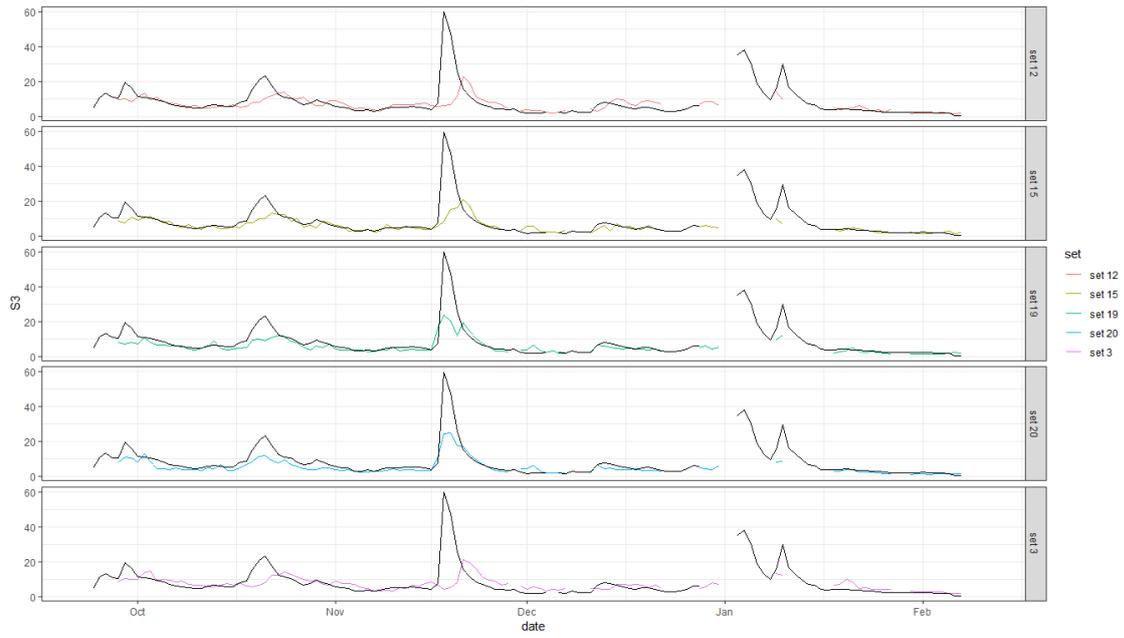


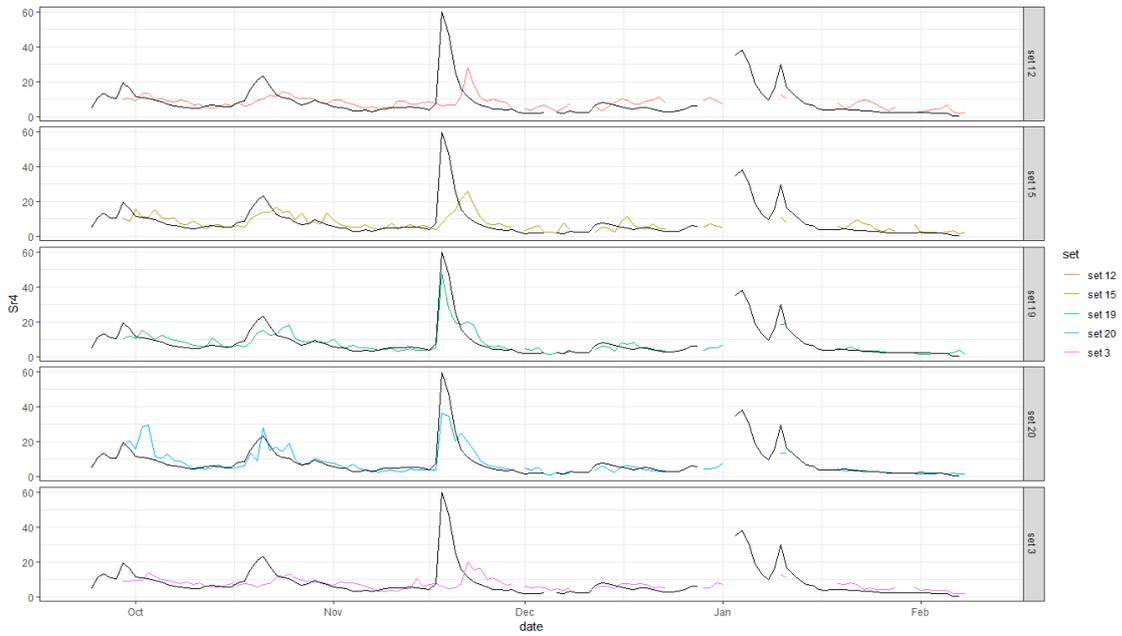
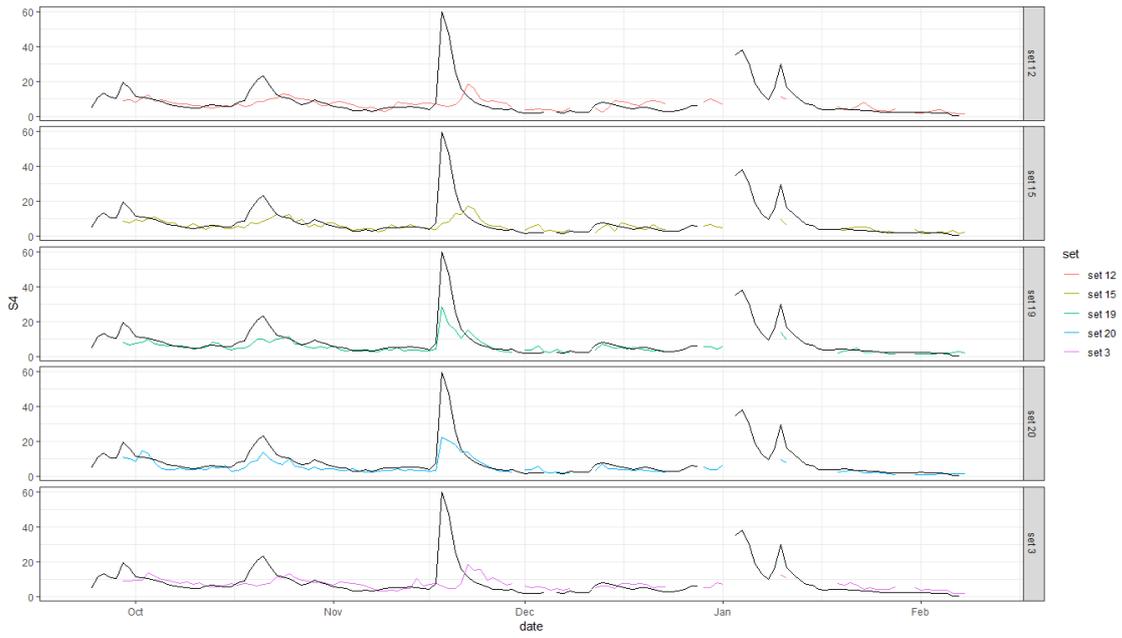


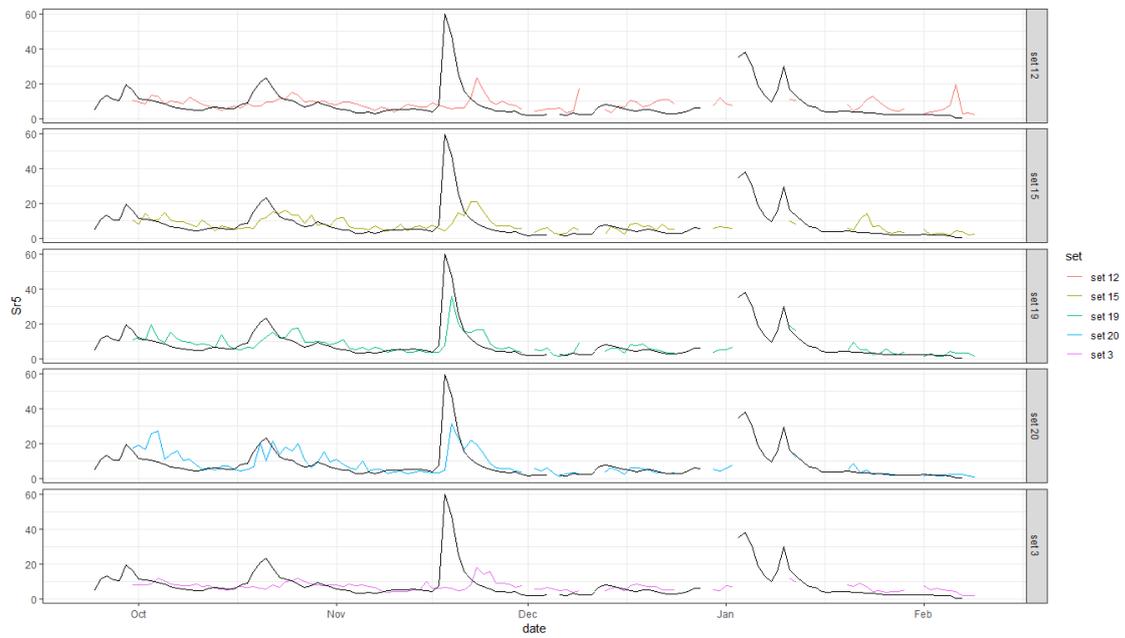
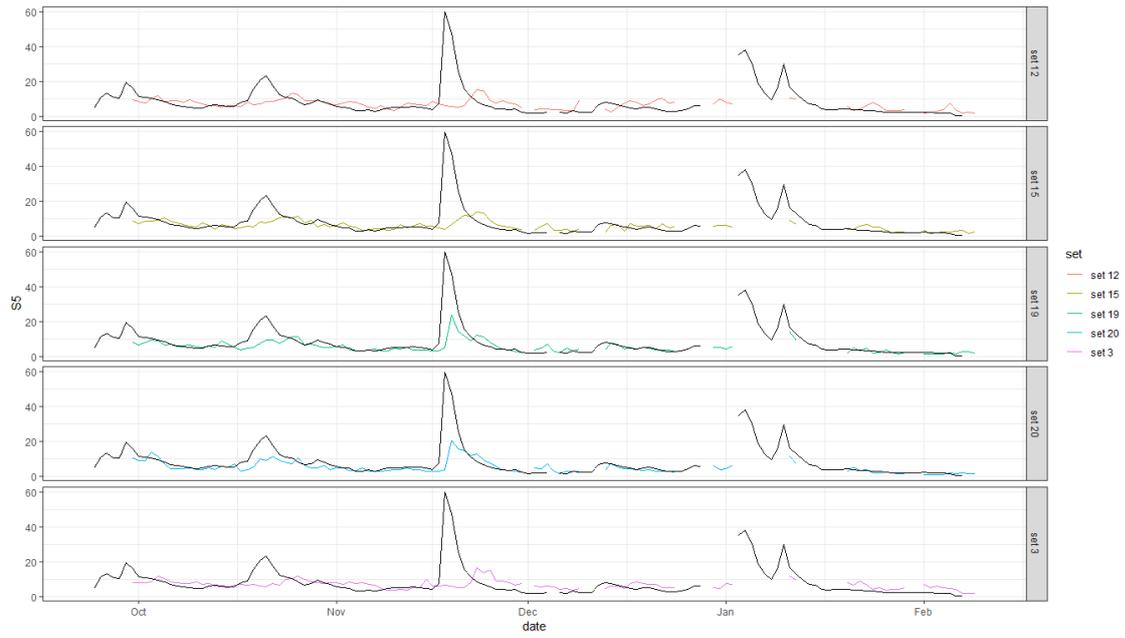
B.1.b. vhm 12



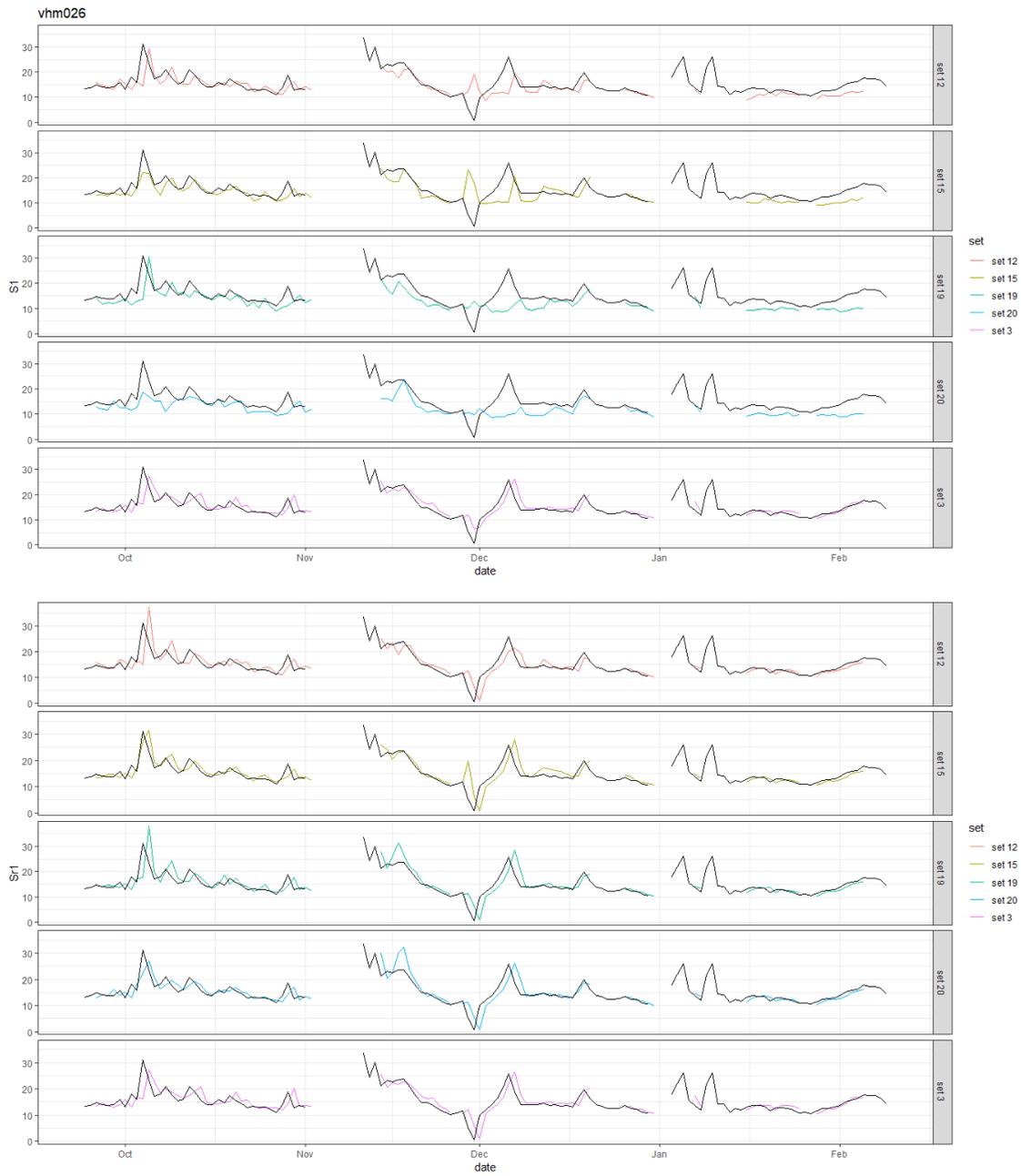


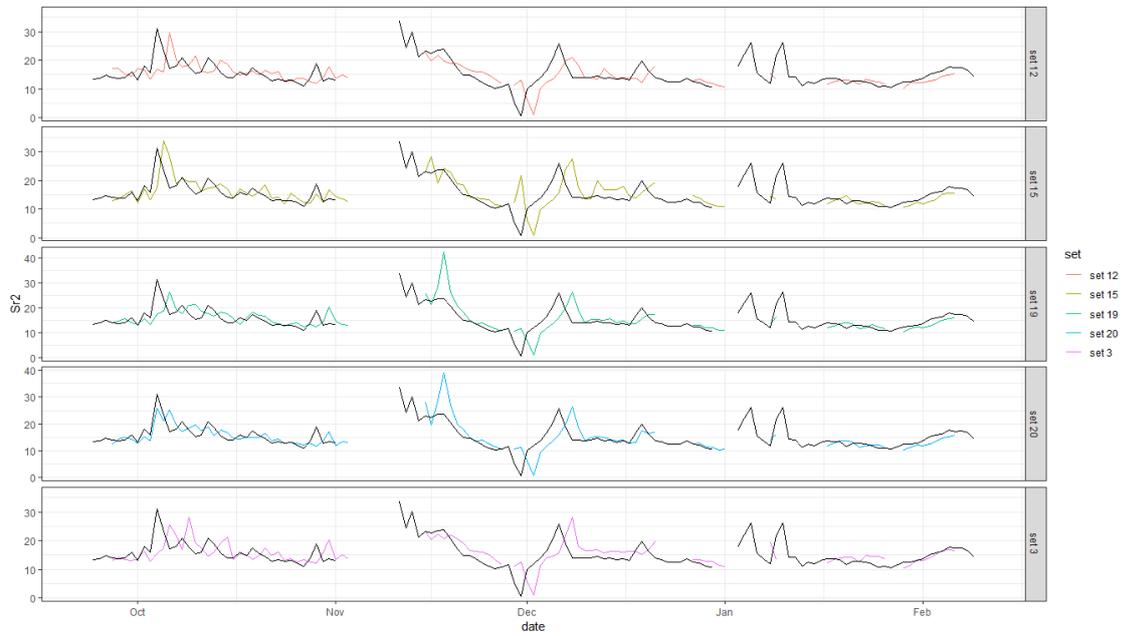
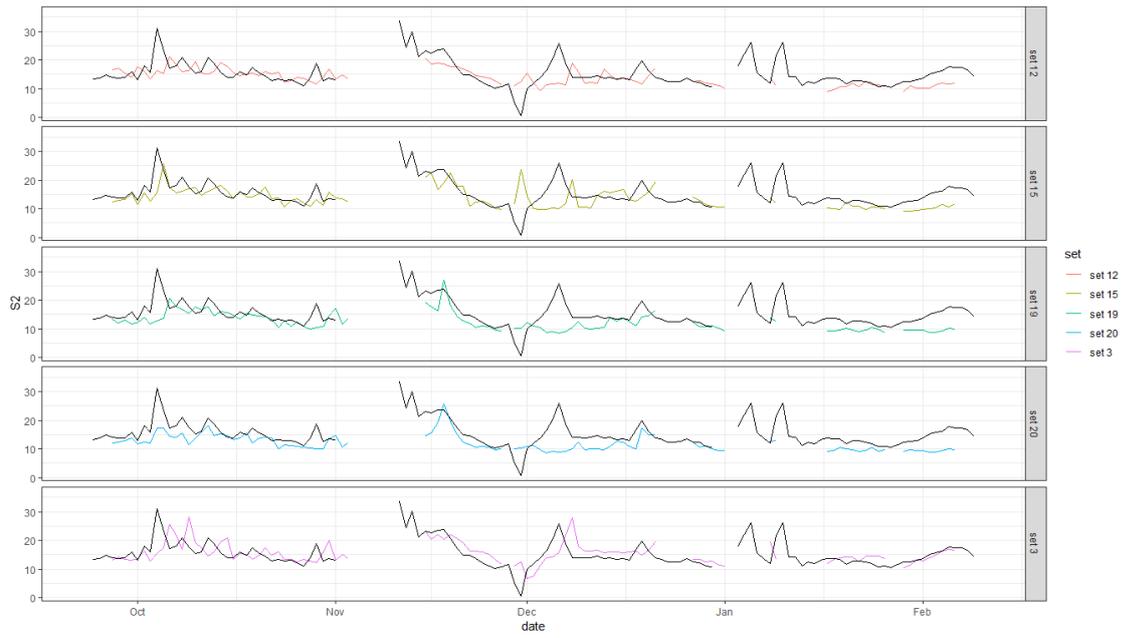


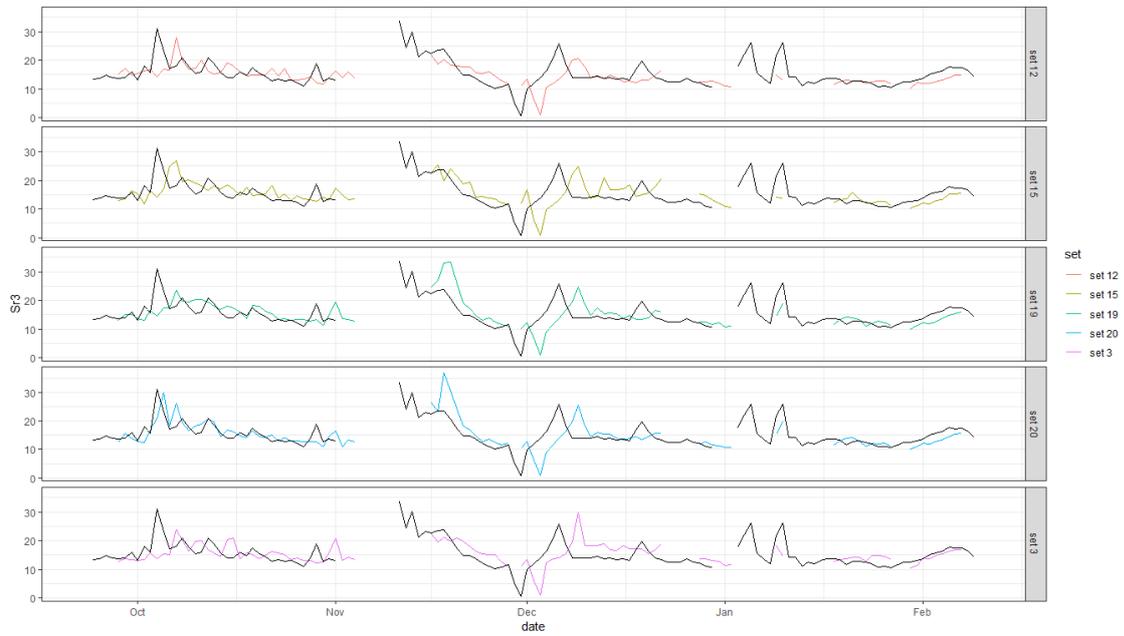
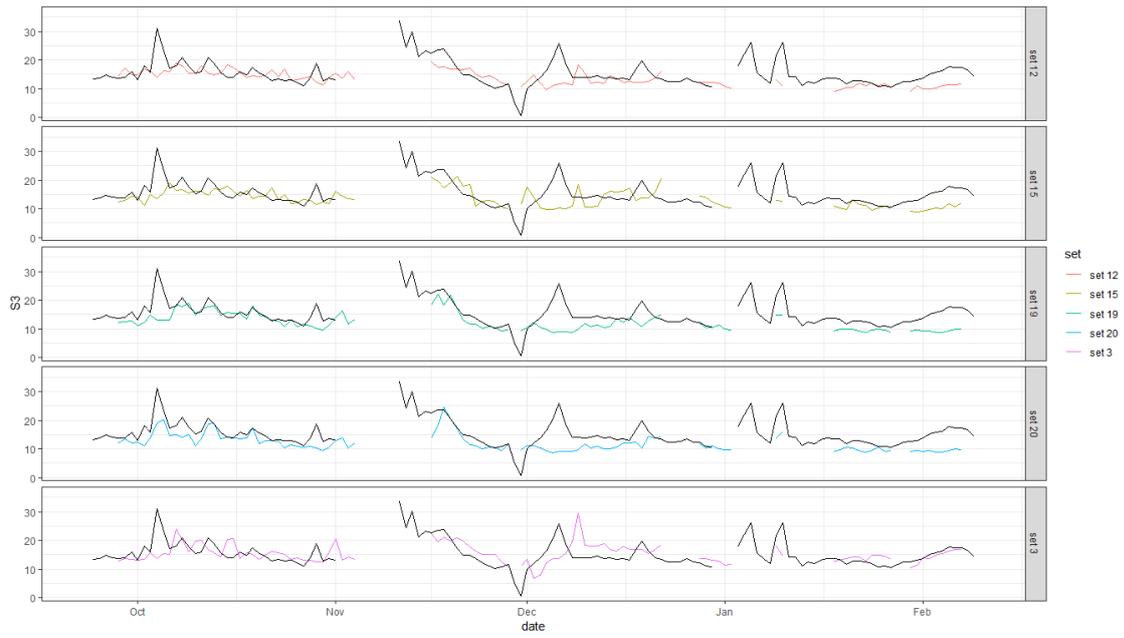


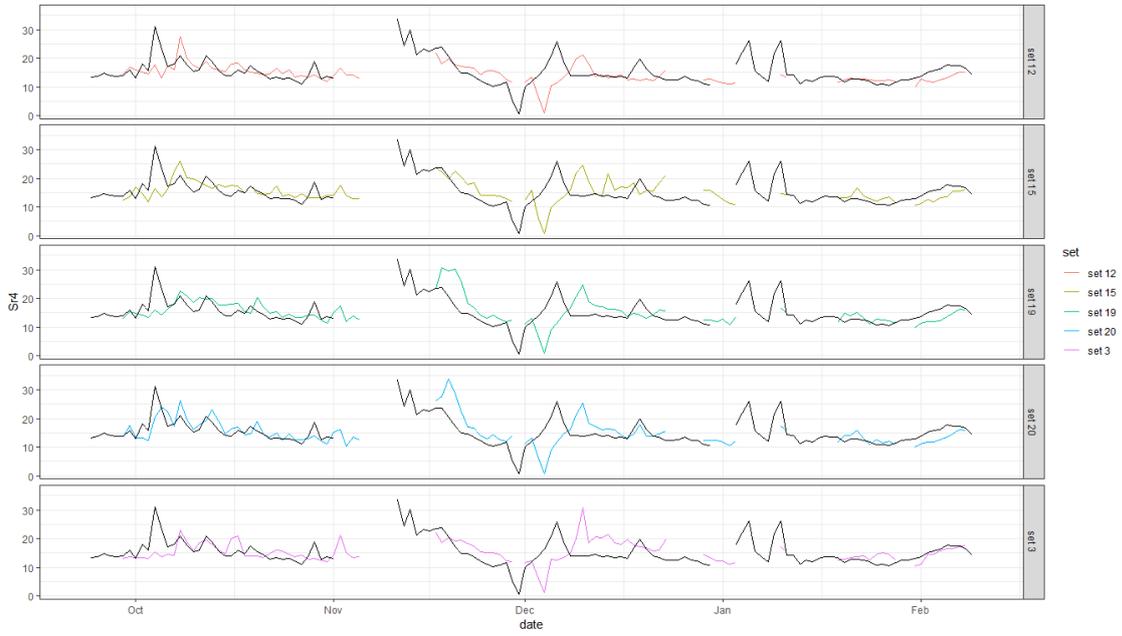
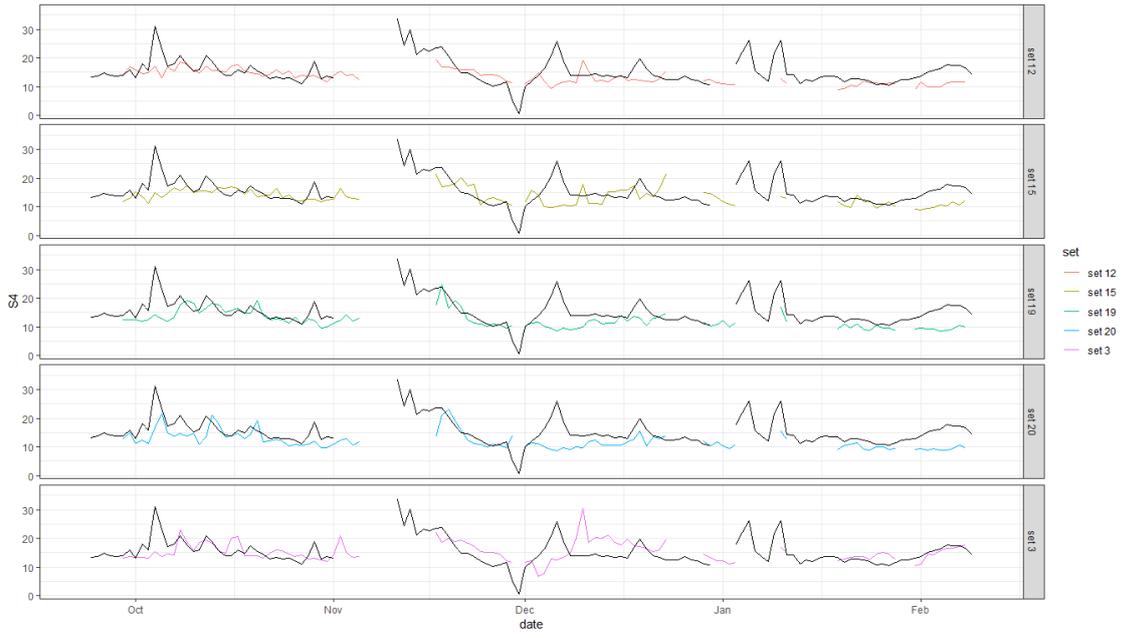


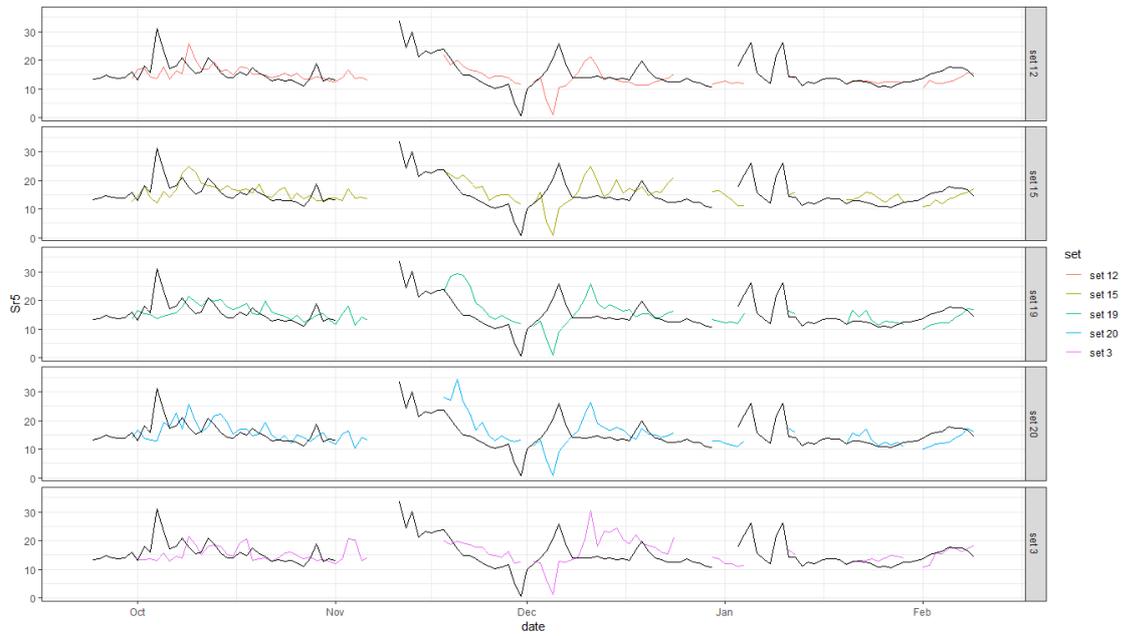
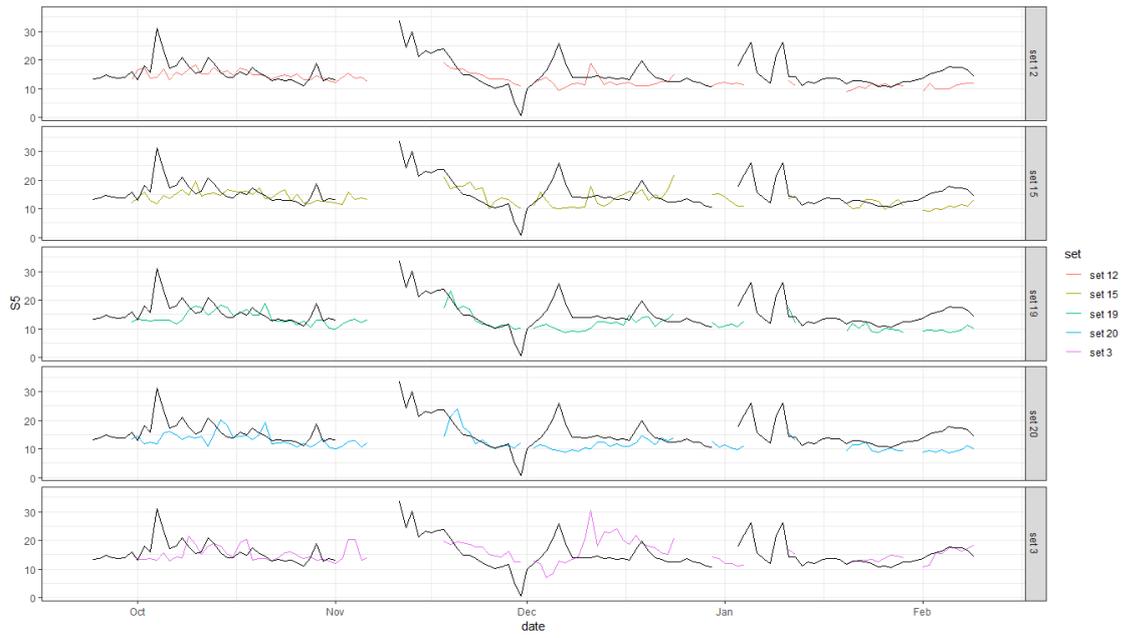
B.1.c. vhm 26

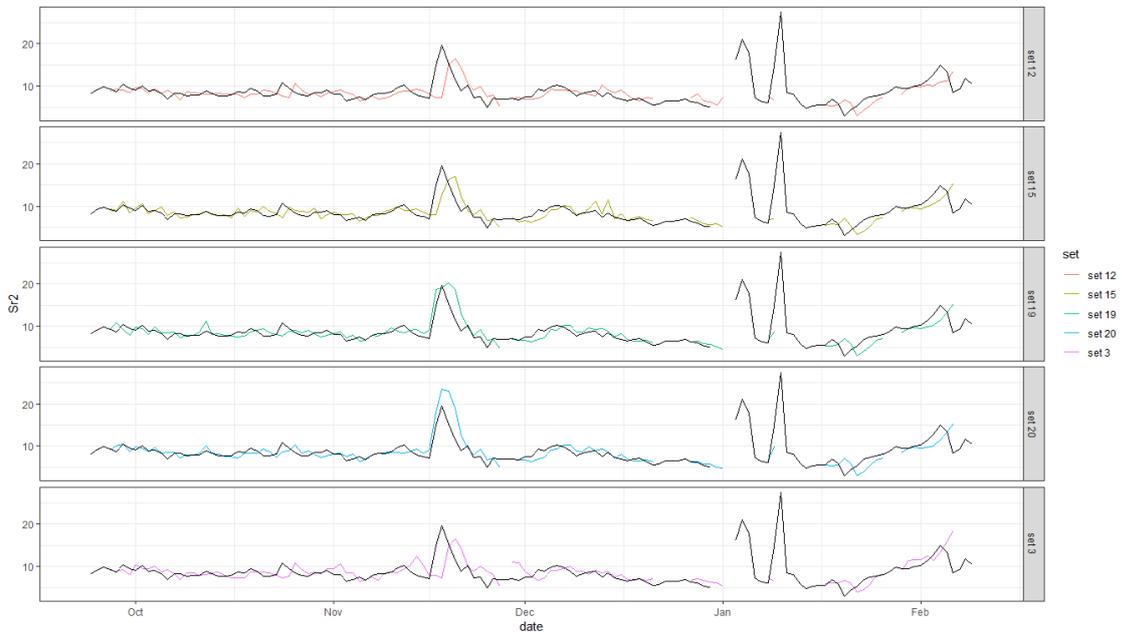
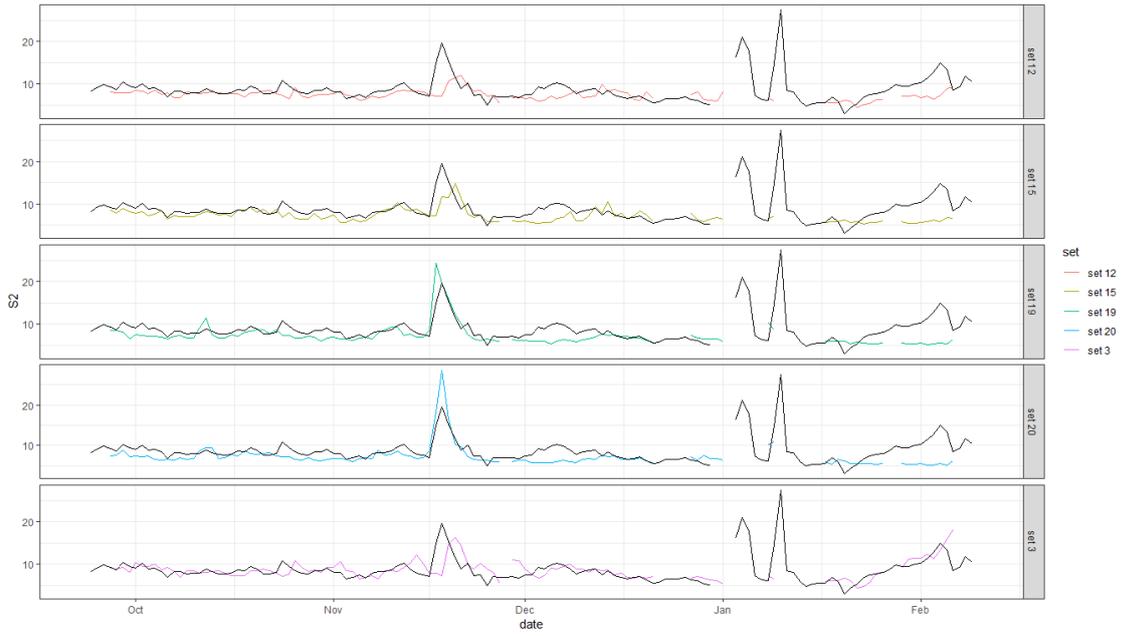


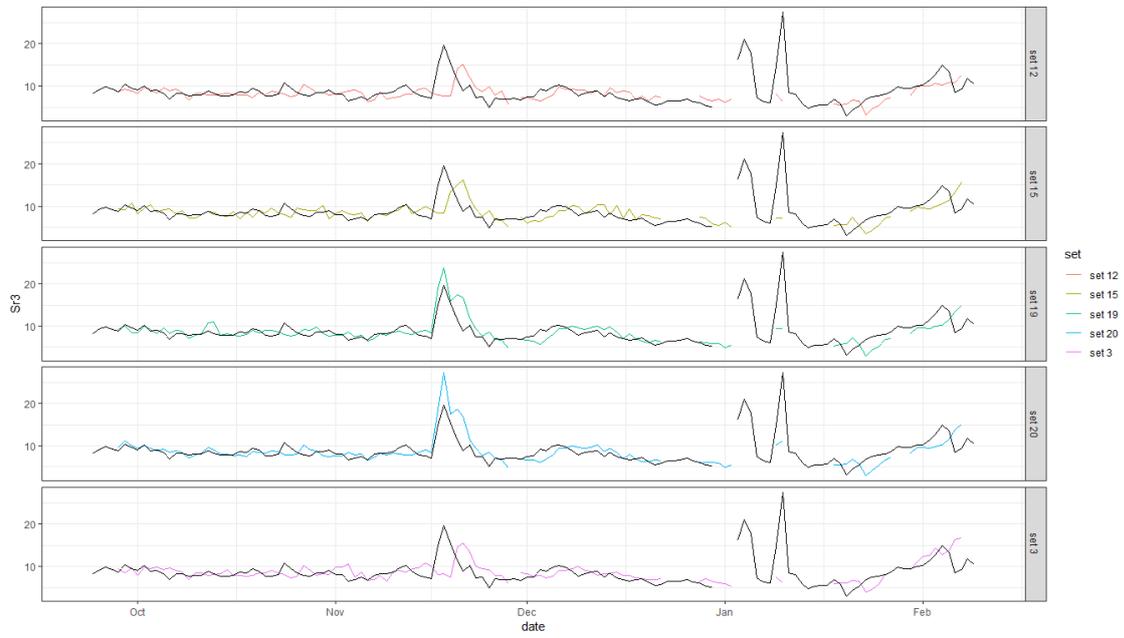
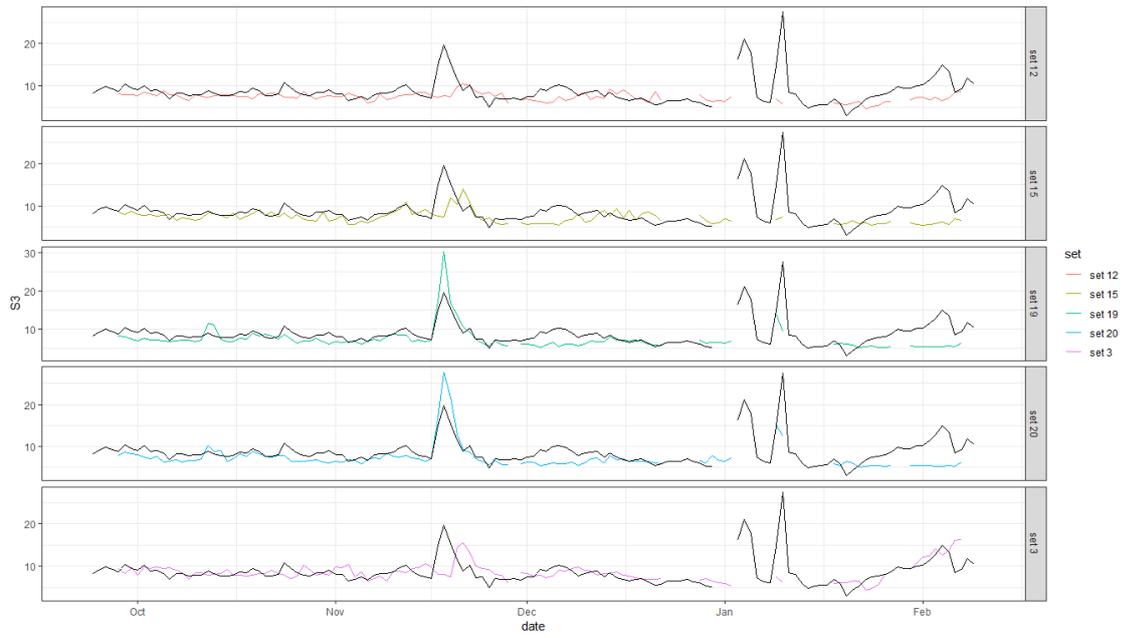


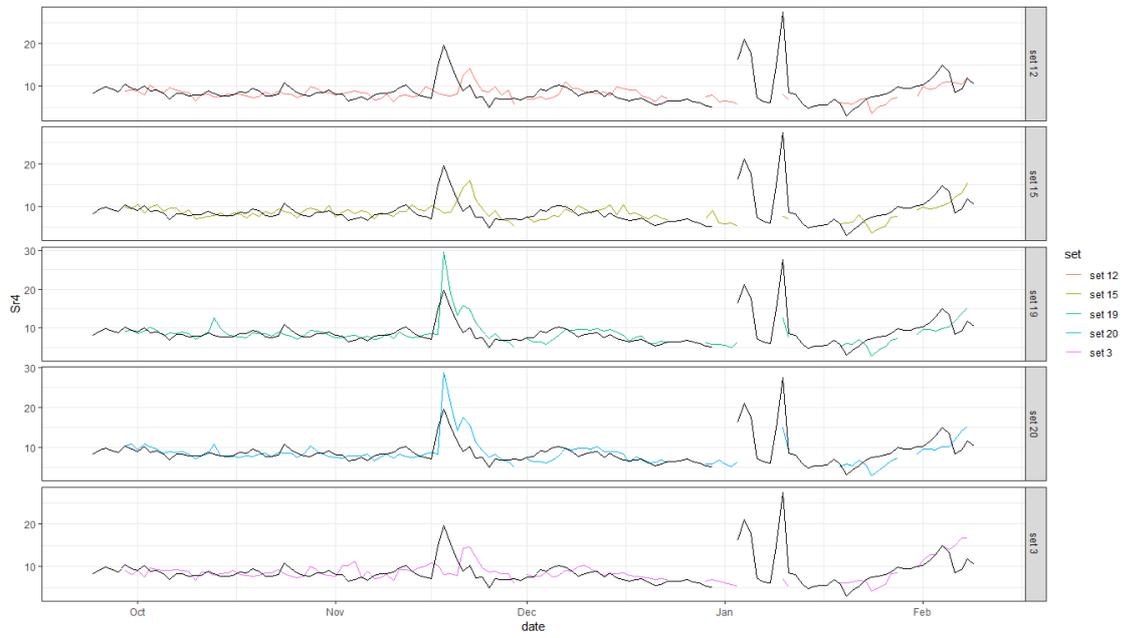
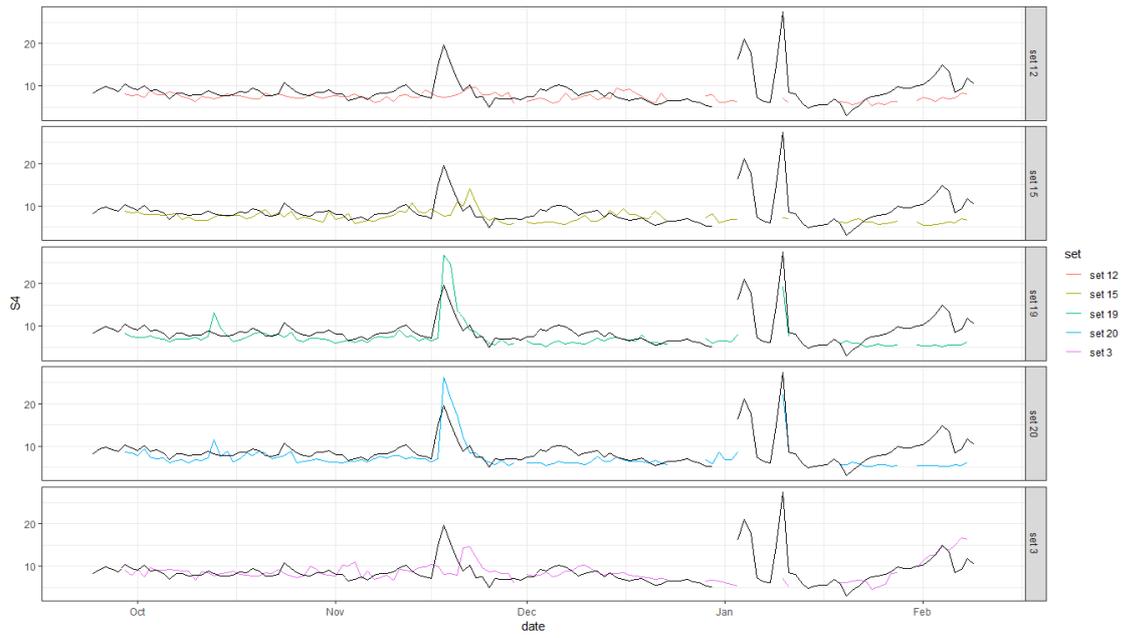


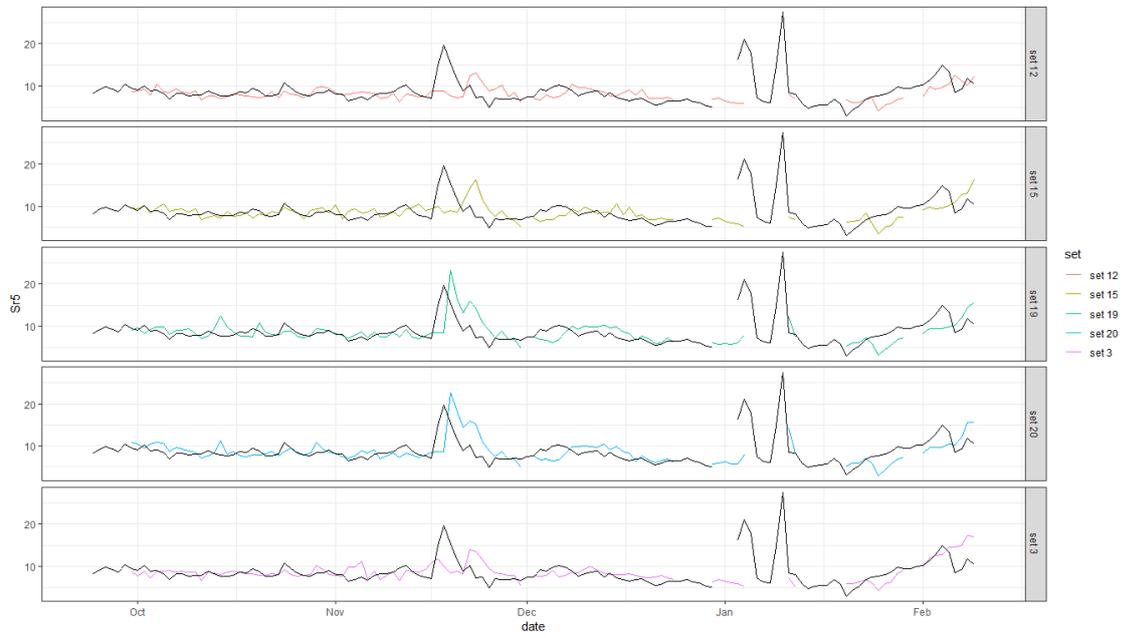
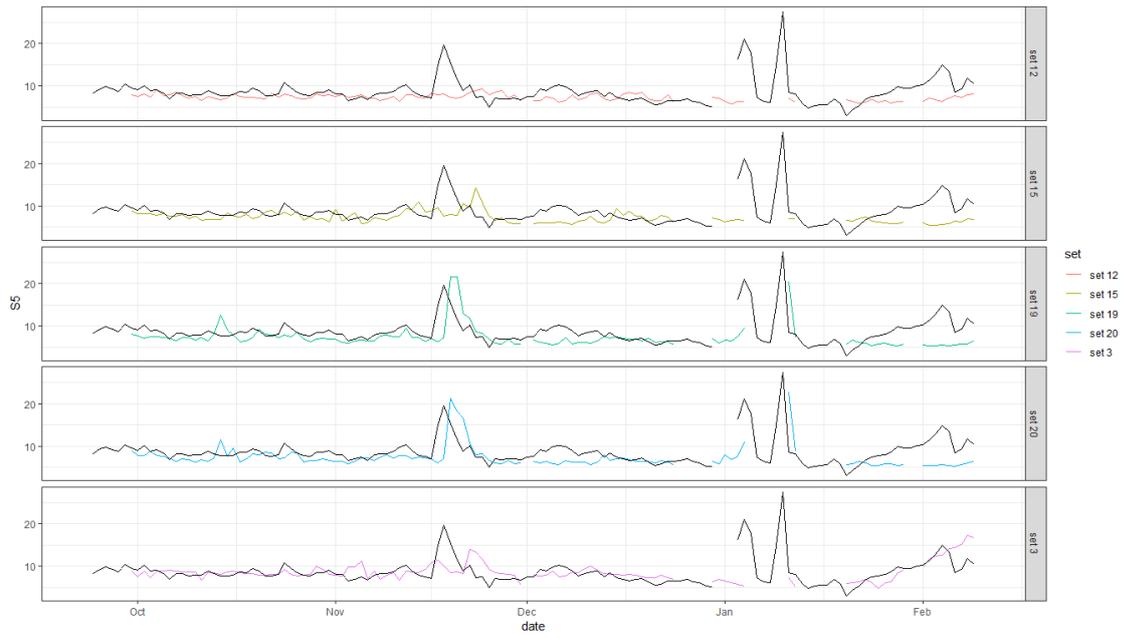




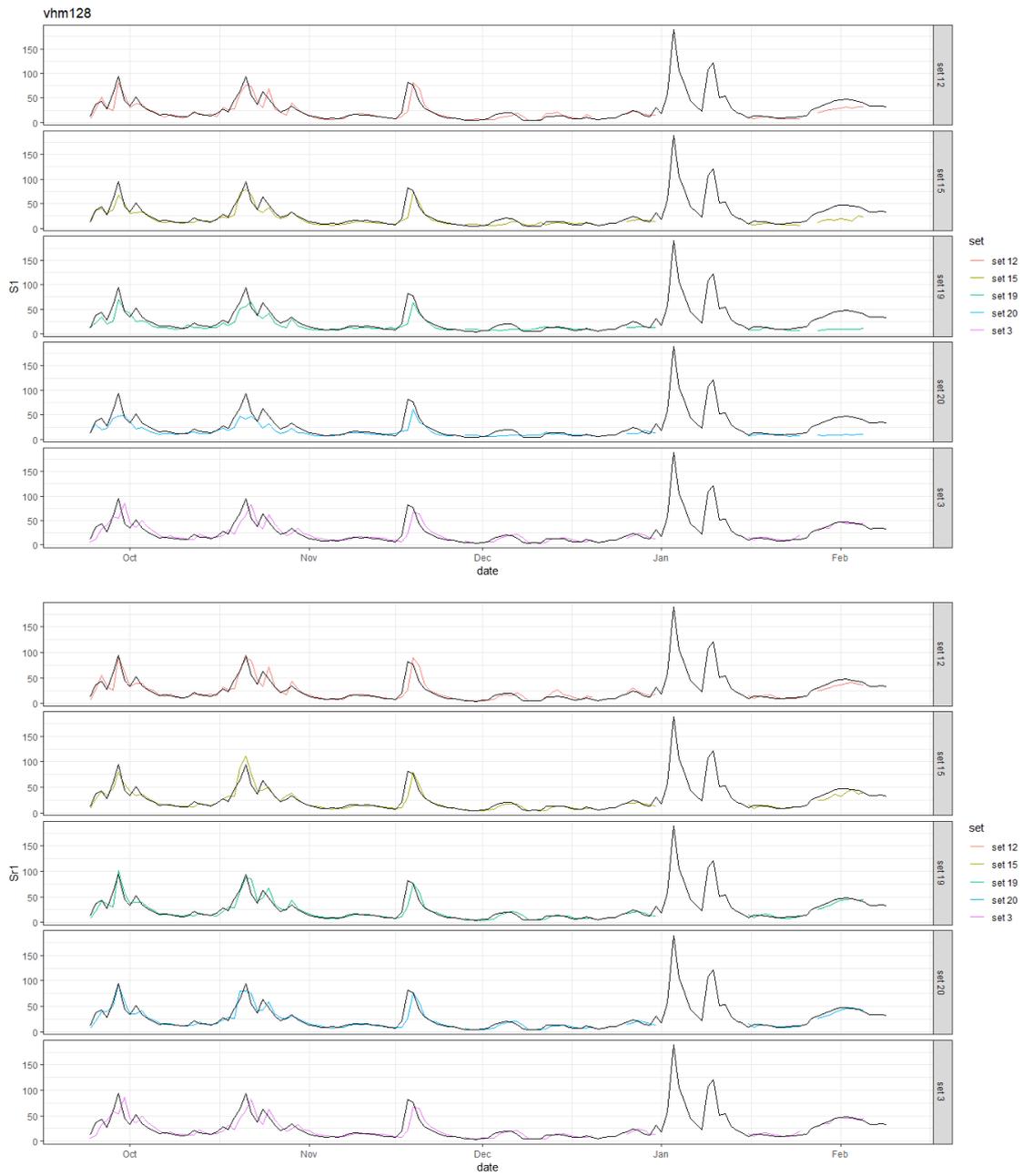


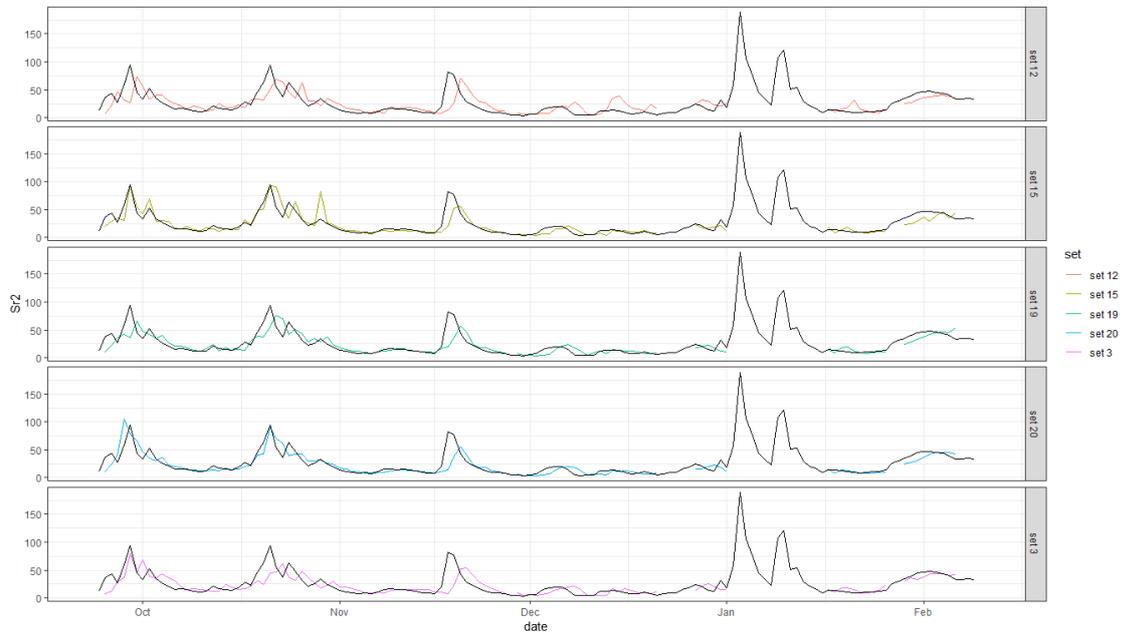
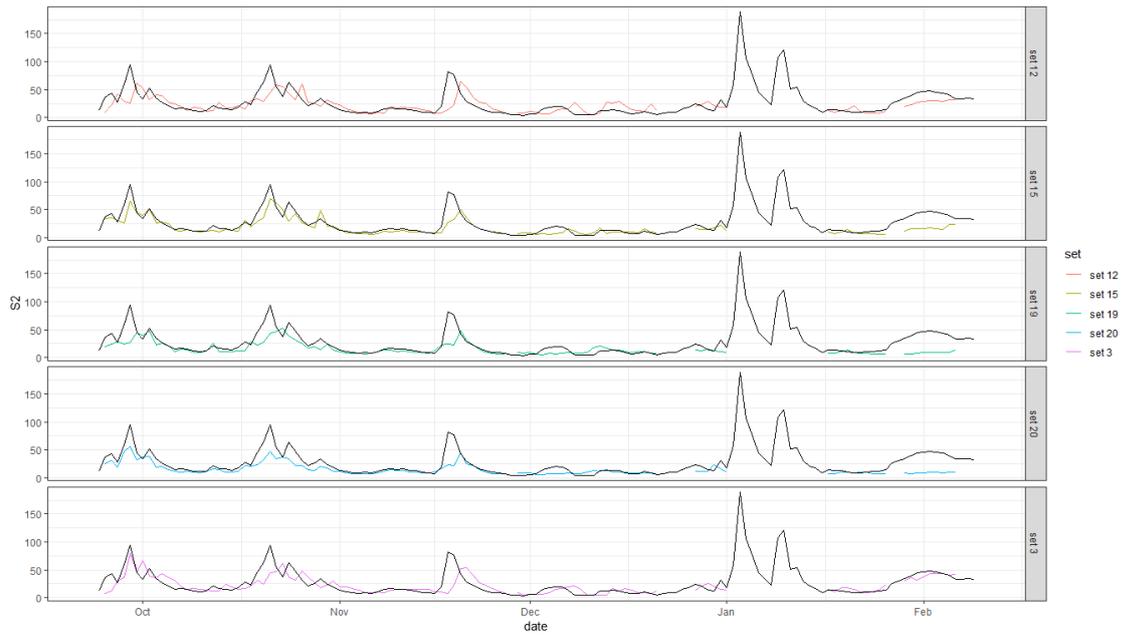


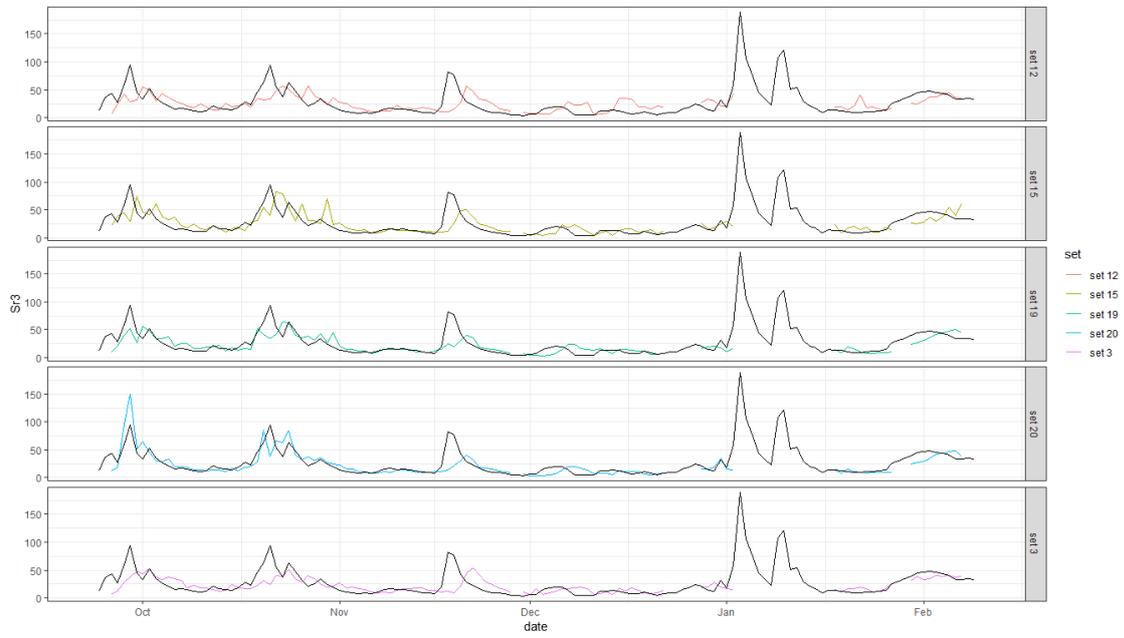
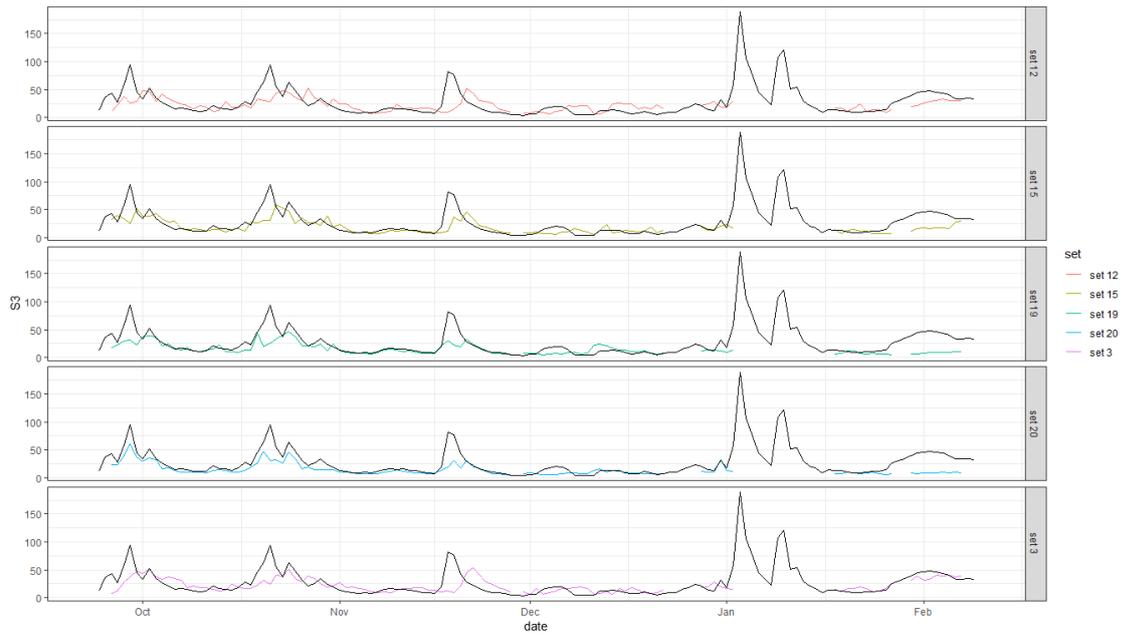


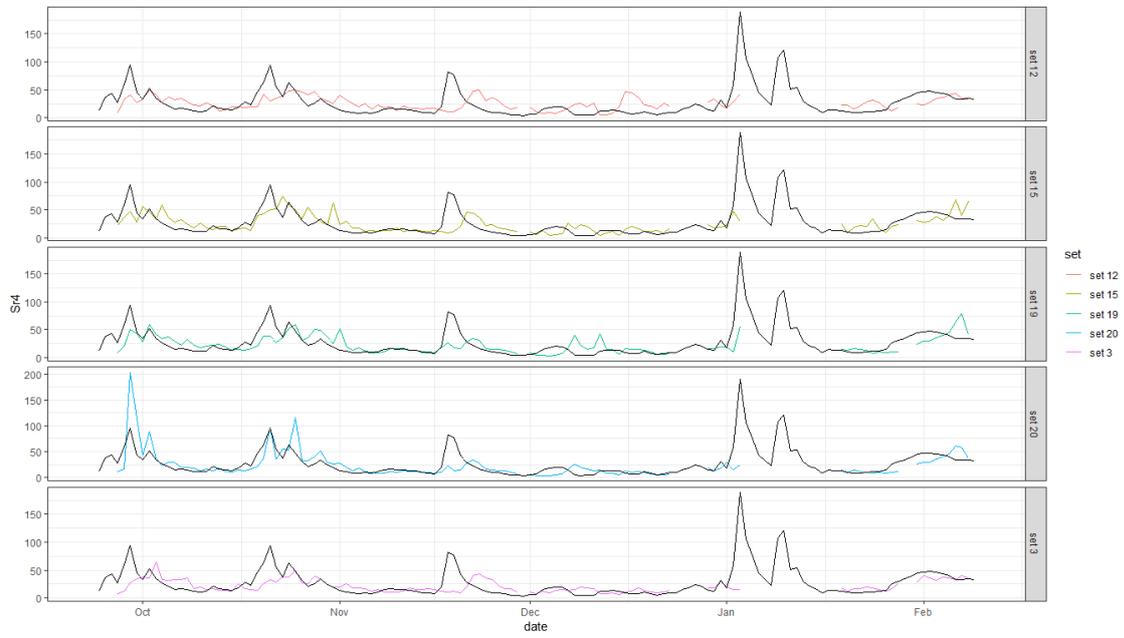
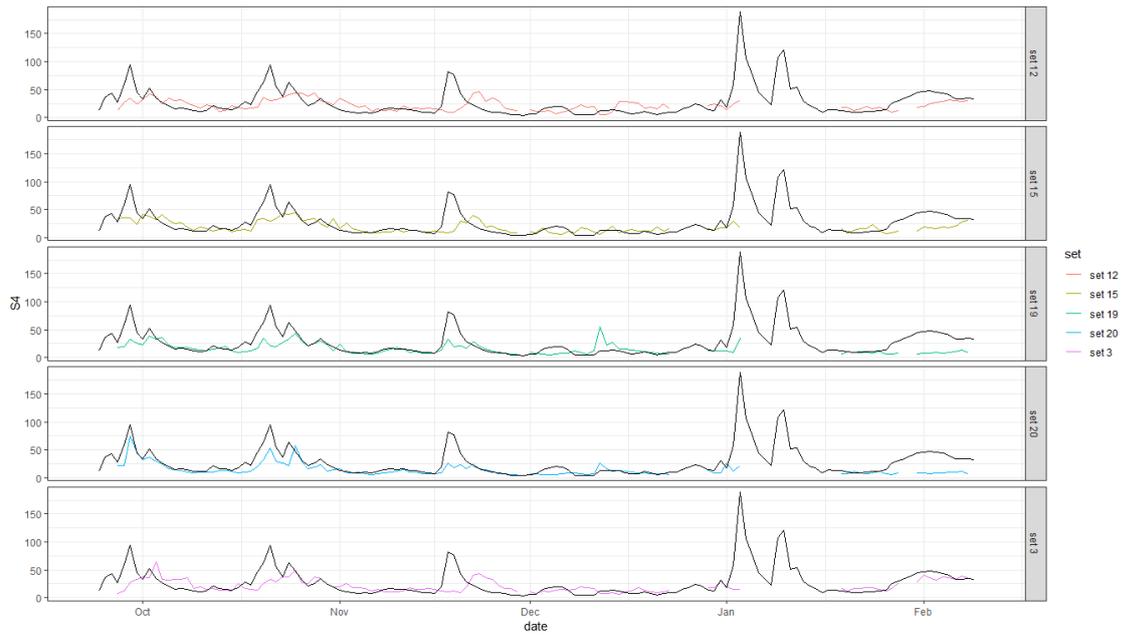


B.1.e. vhm 128

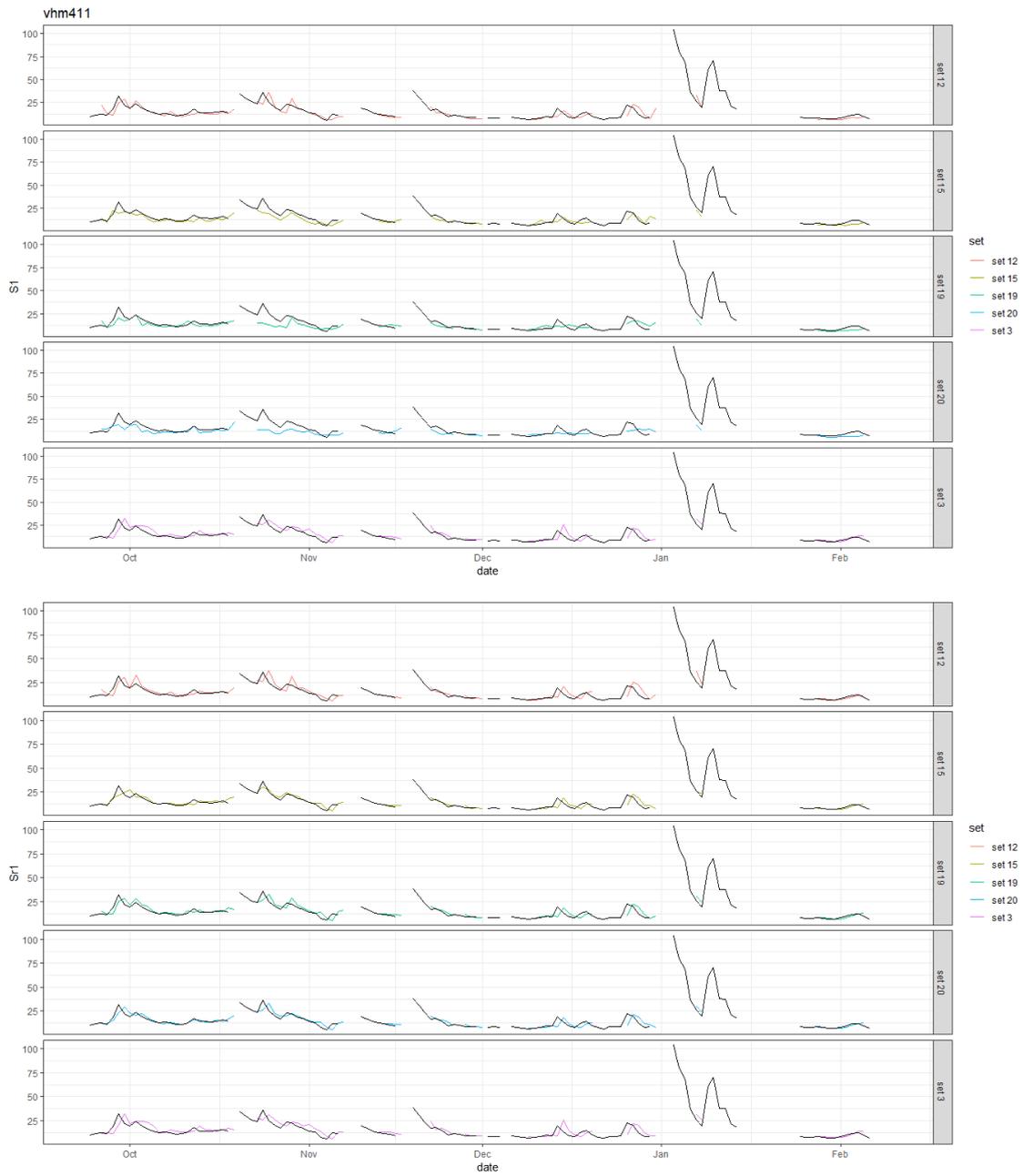


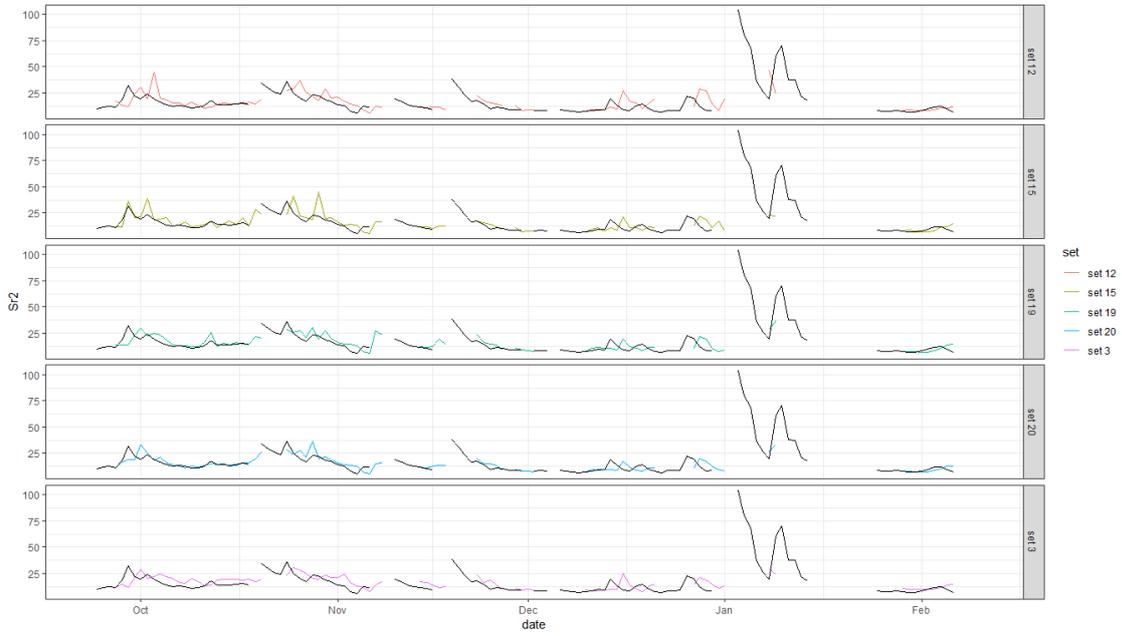
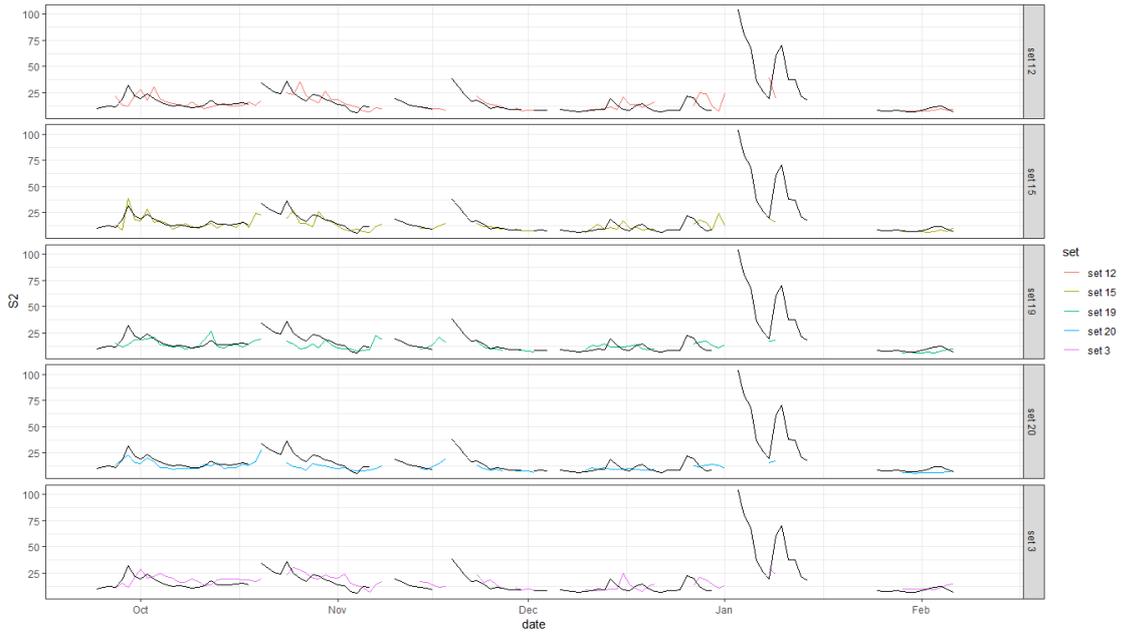


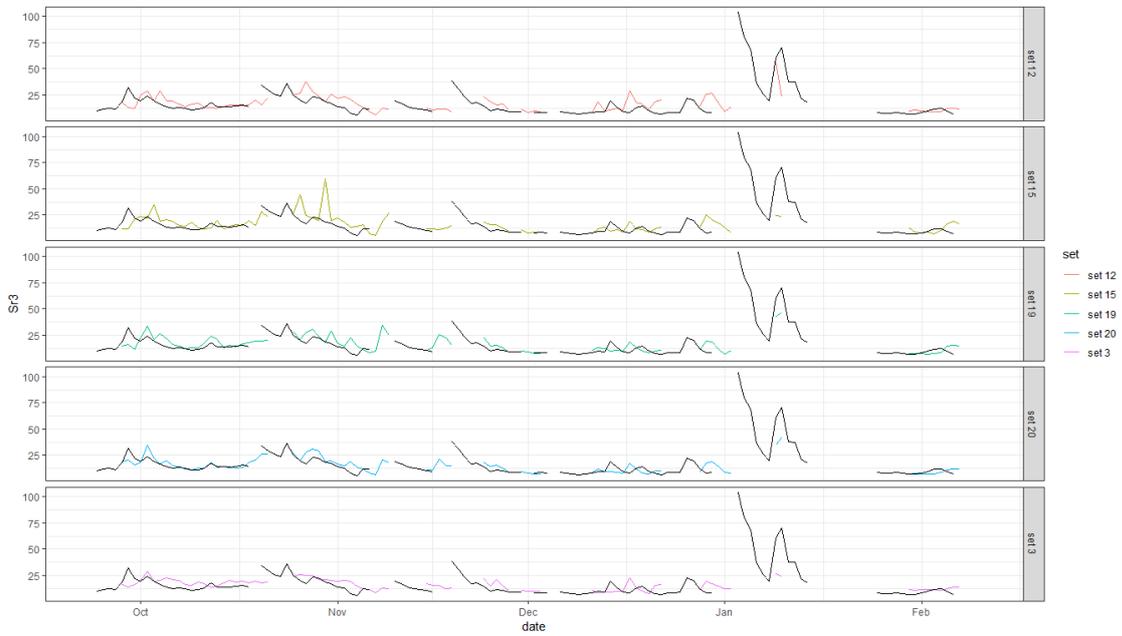
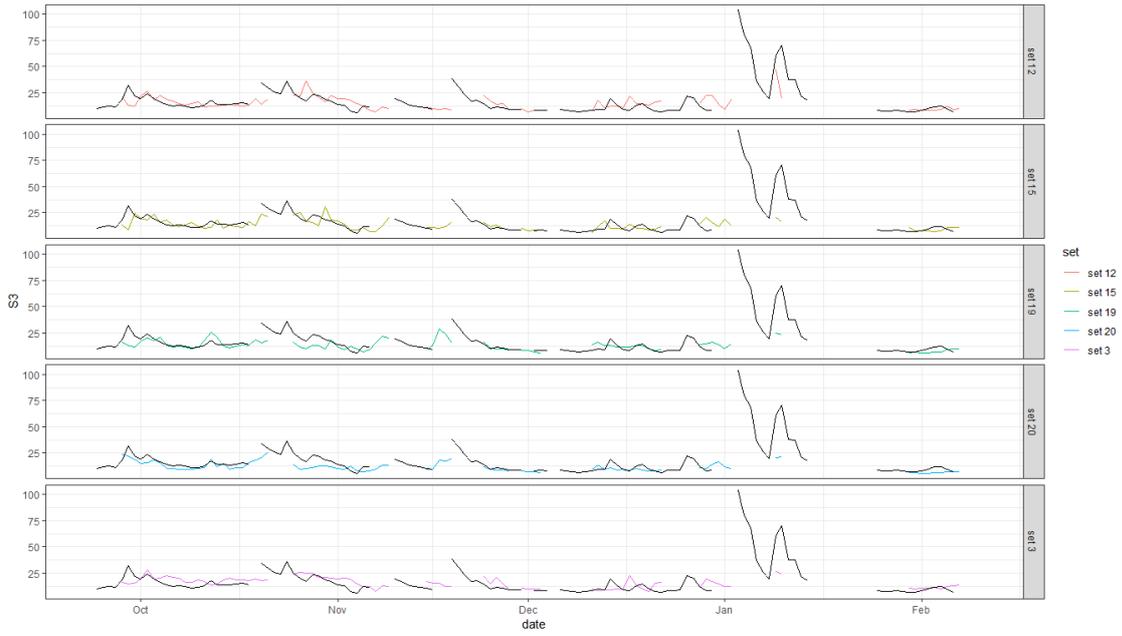


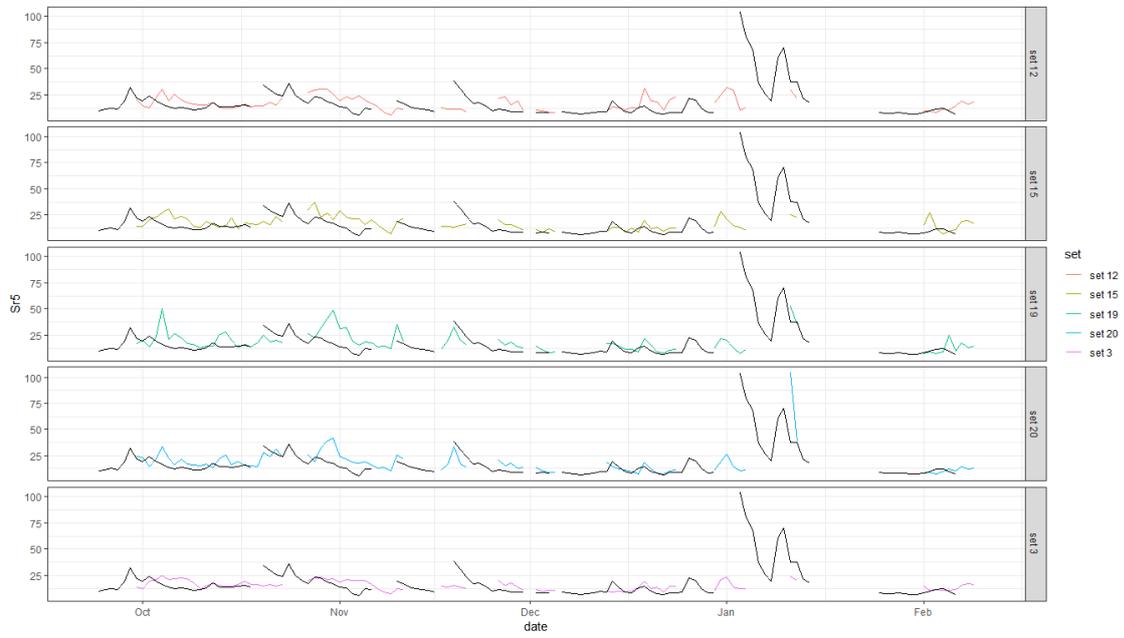
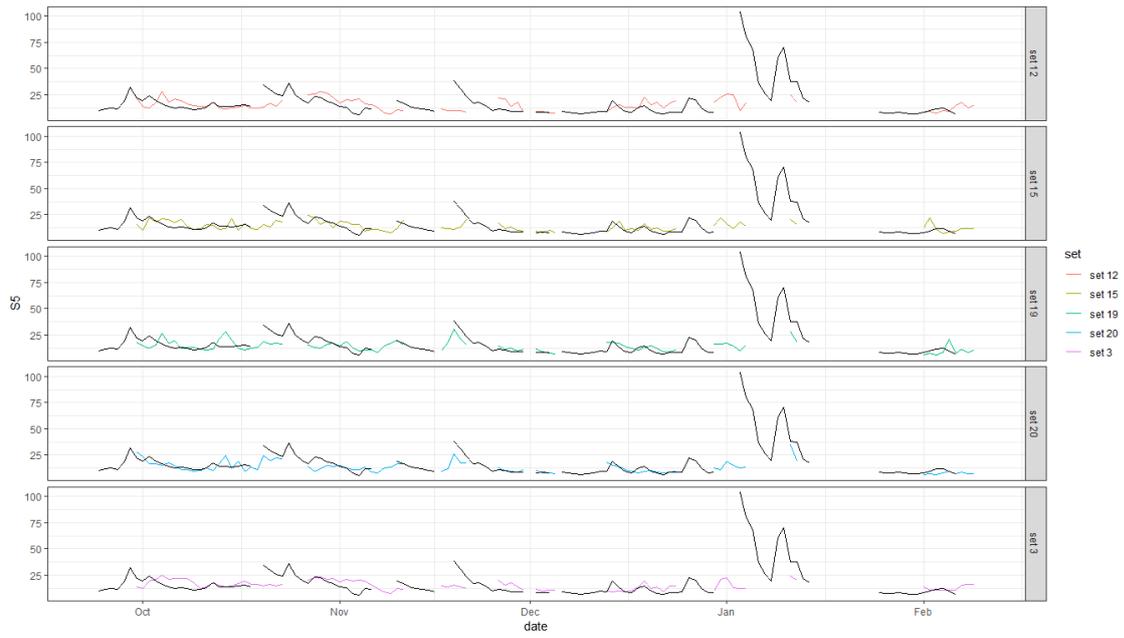


B.1.f. vhm 411



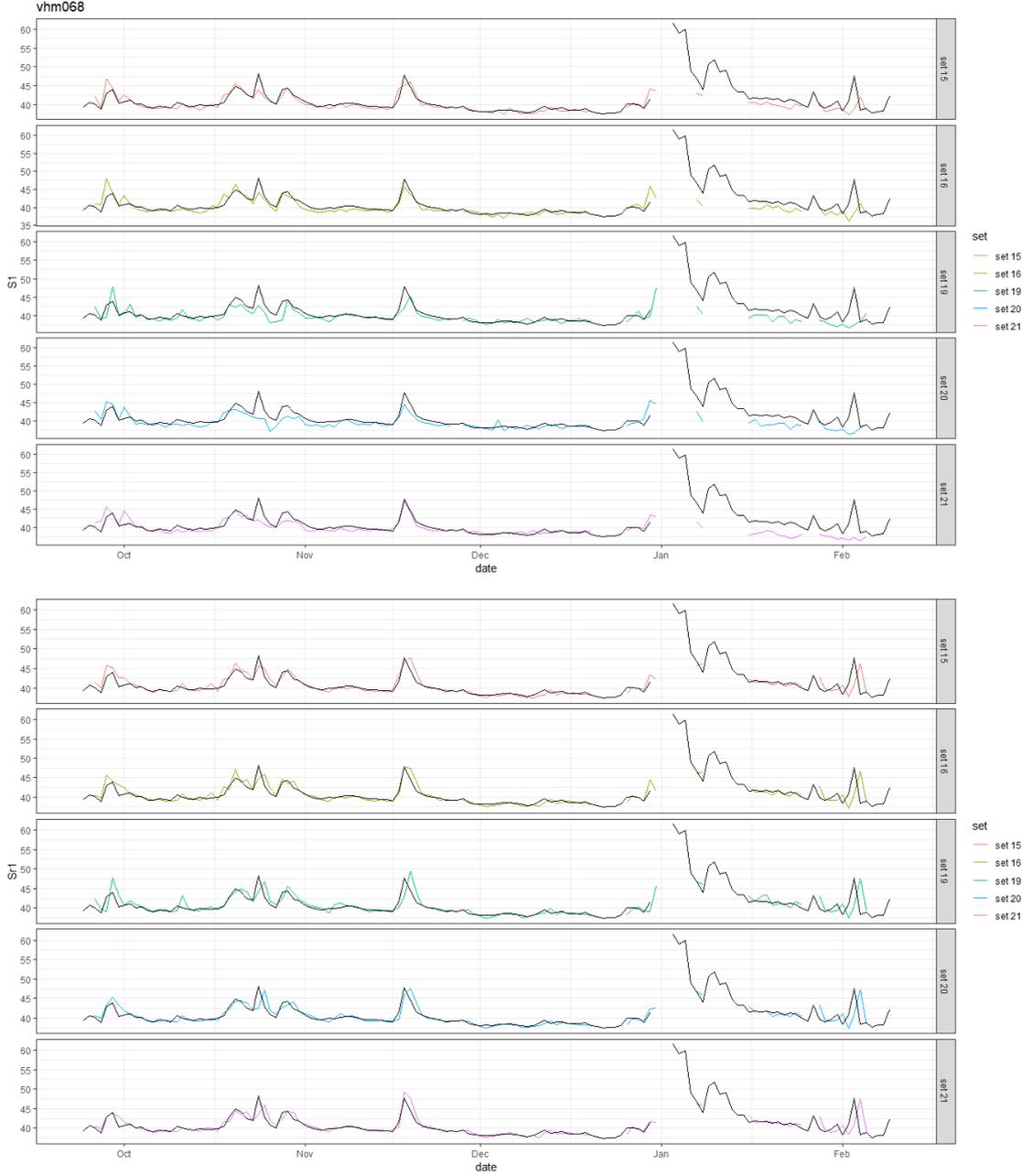


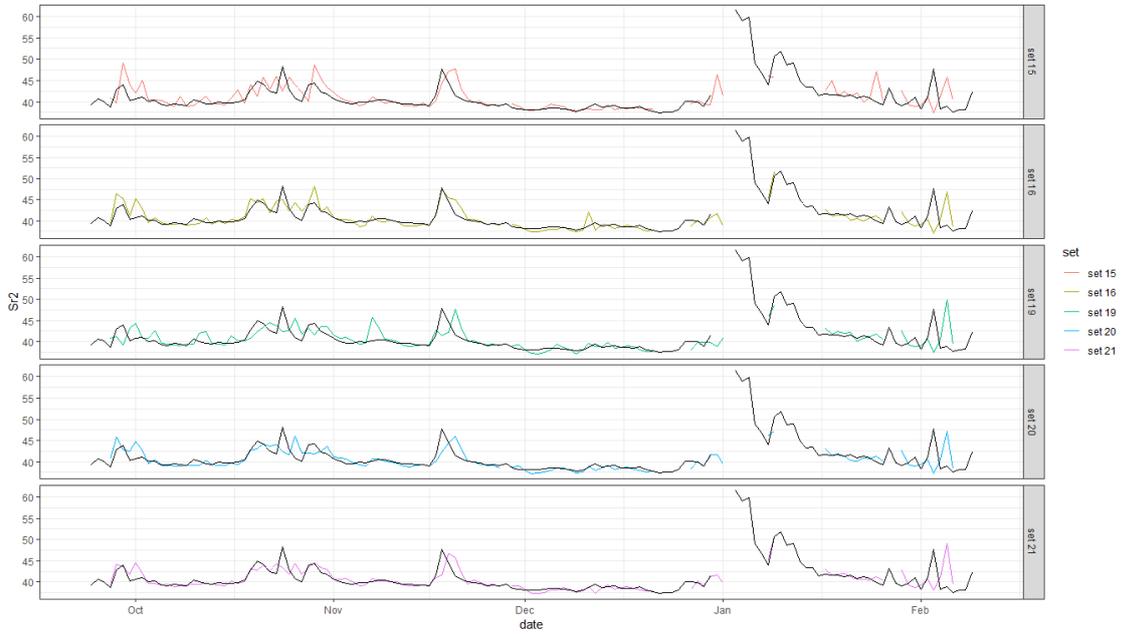
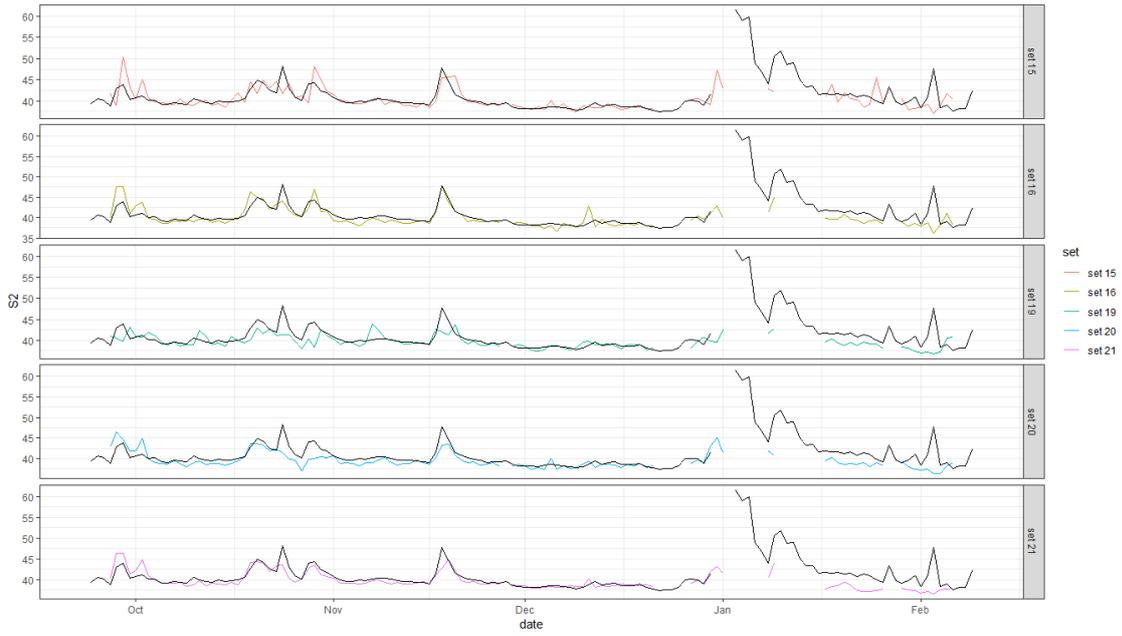


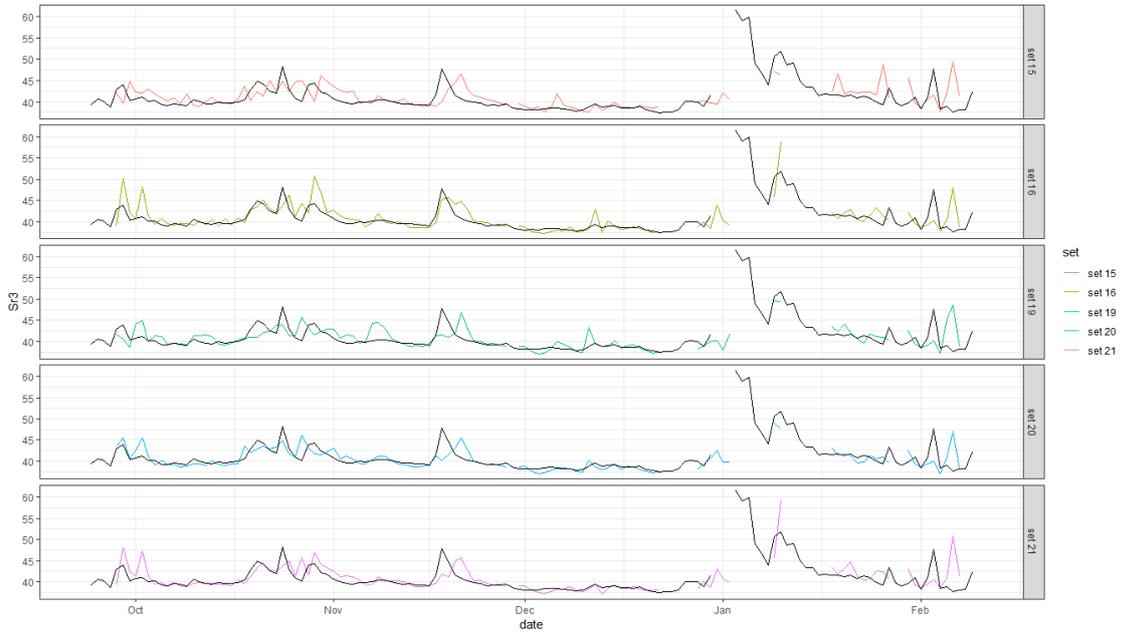
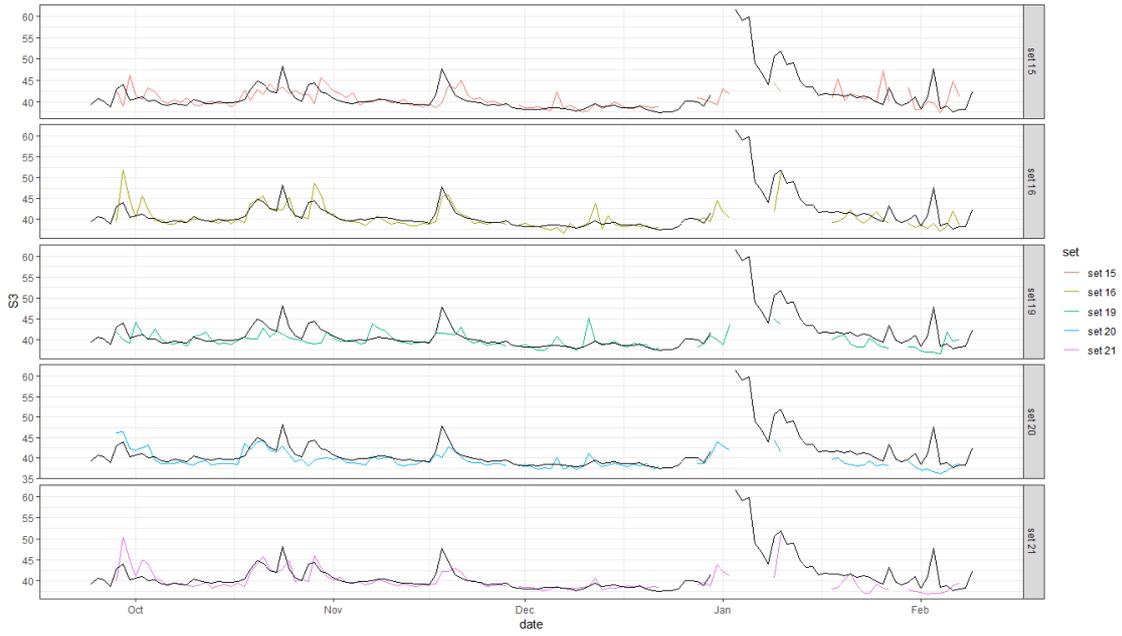


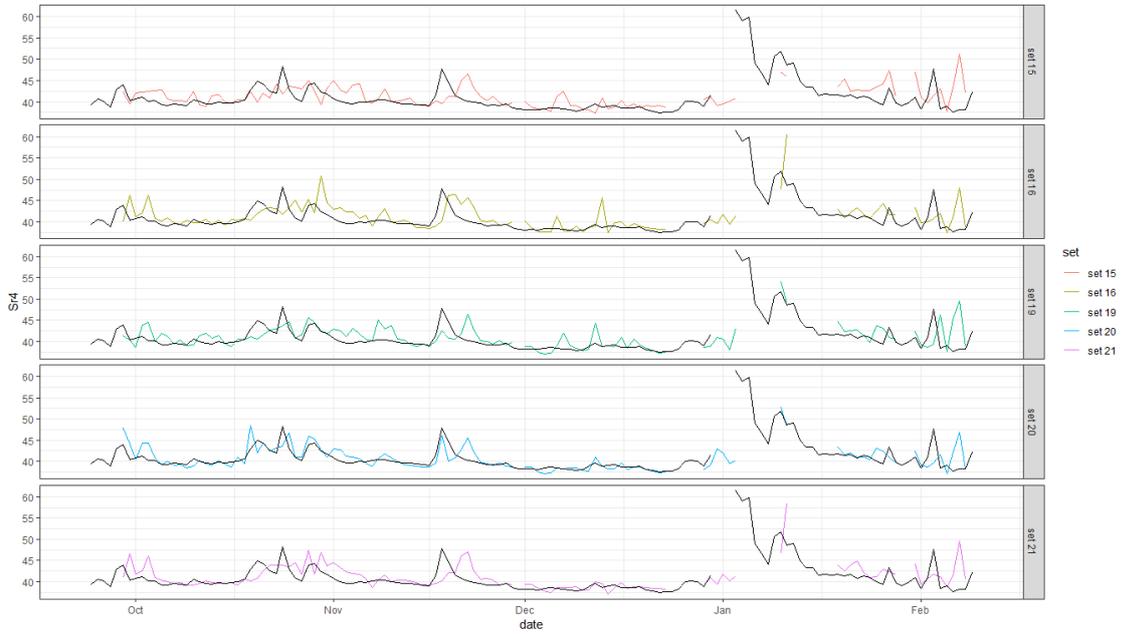
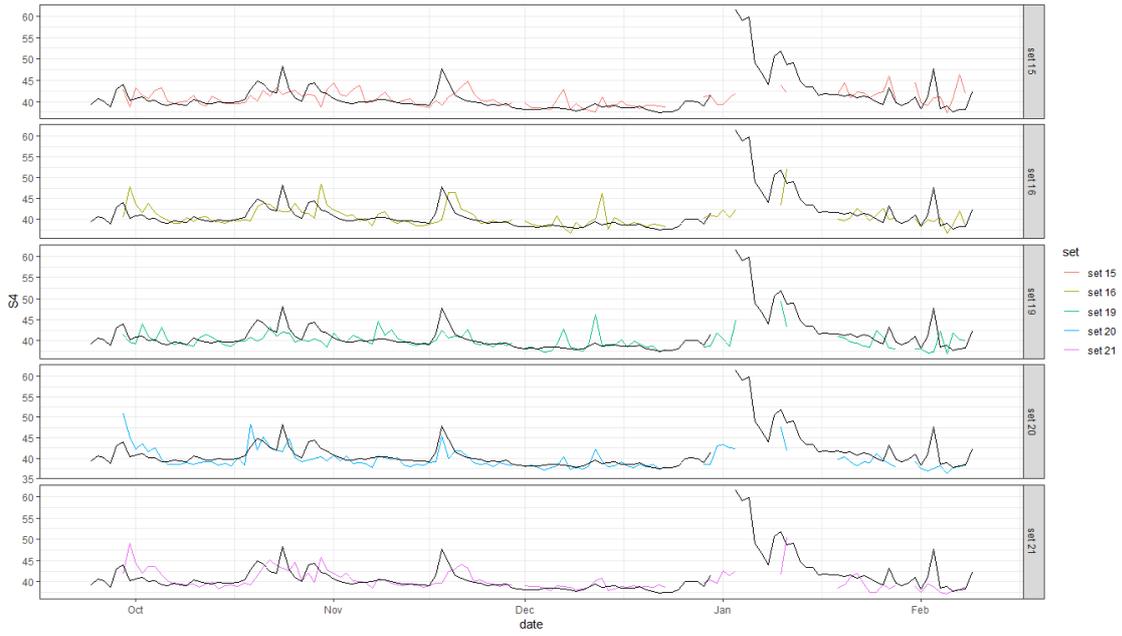
B.2. Class B

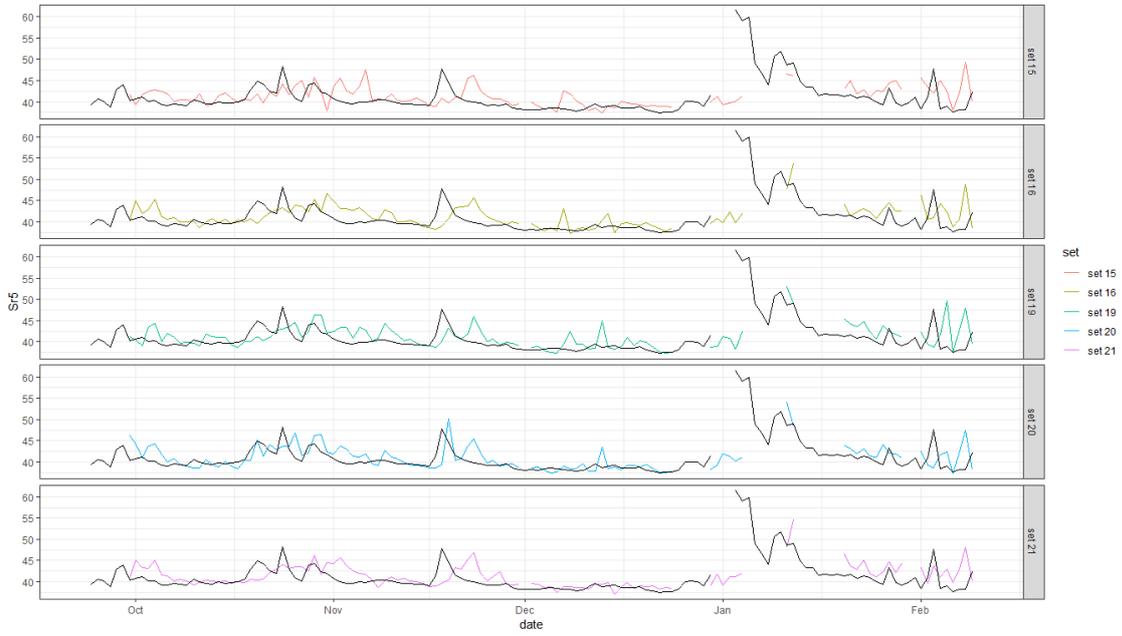
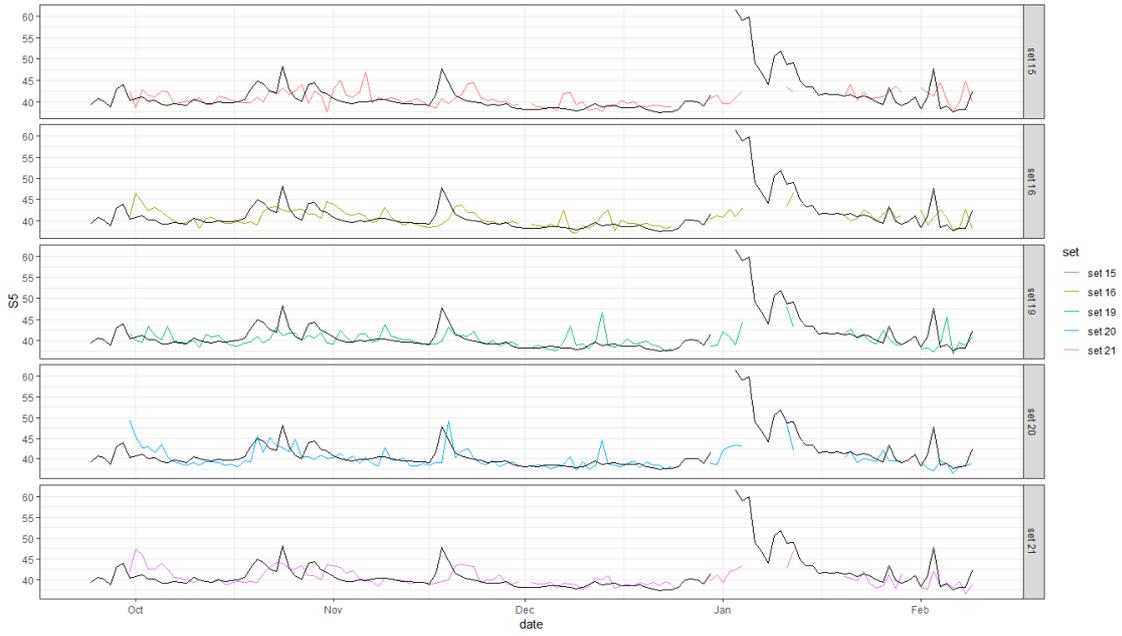
B.2.a. vhm 68



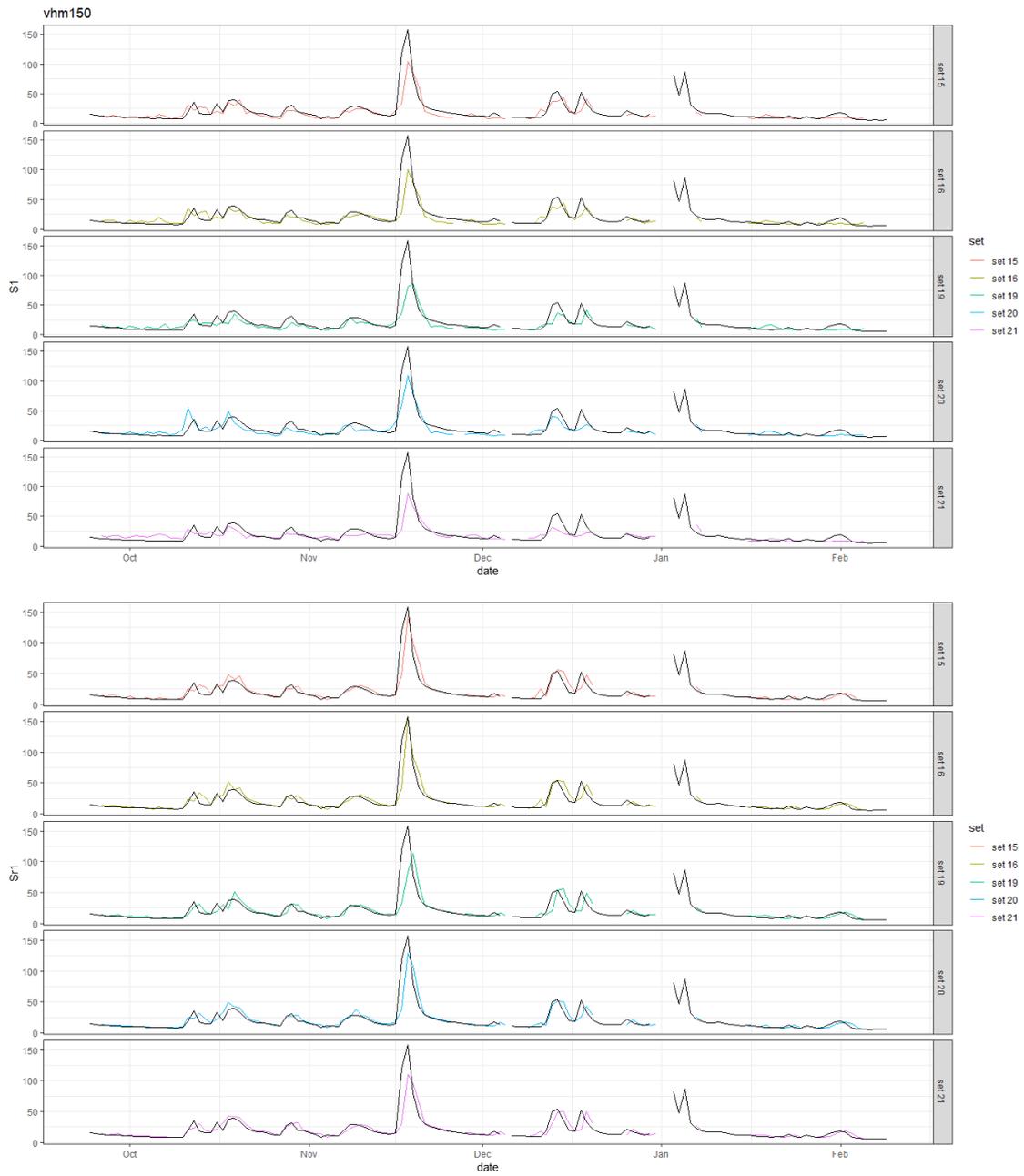


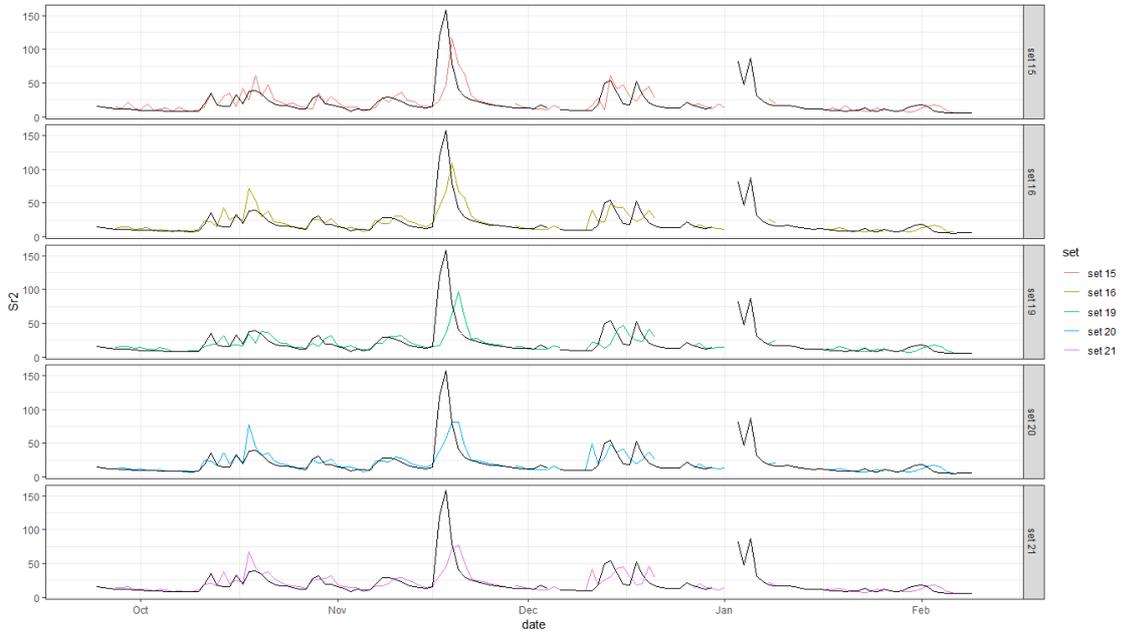
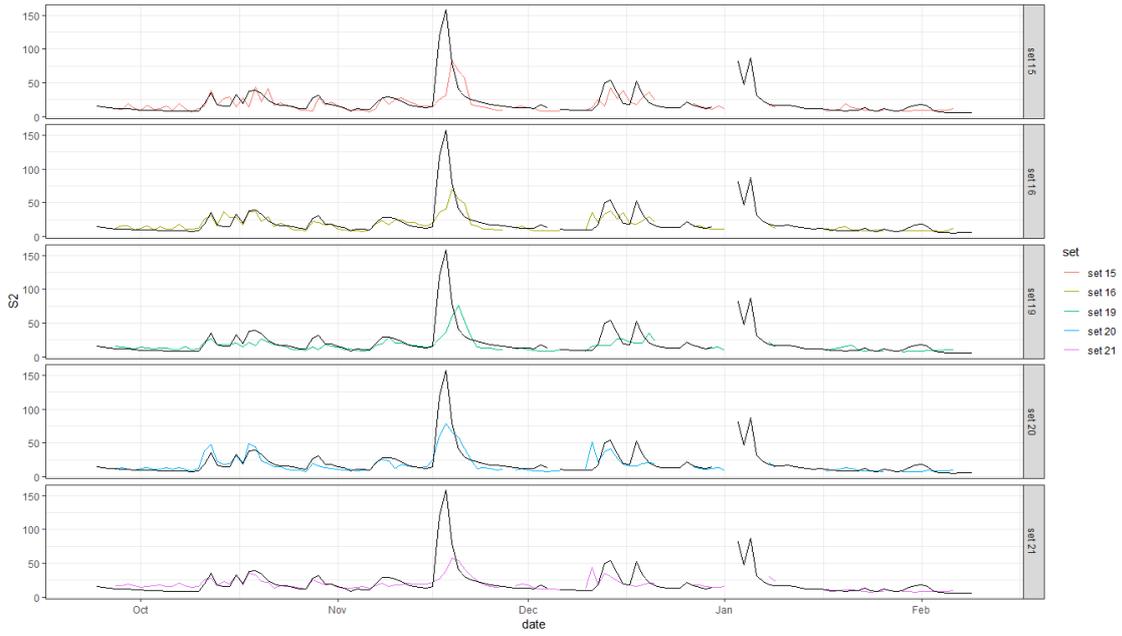


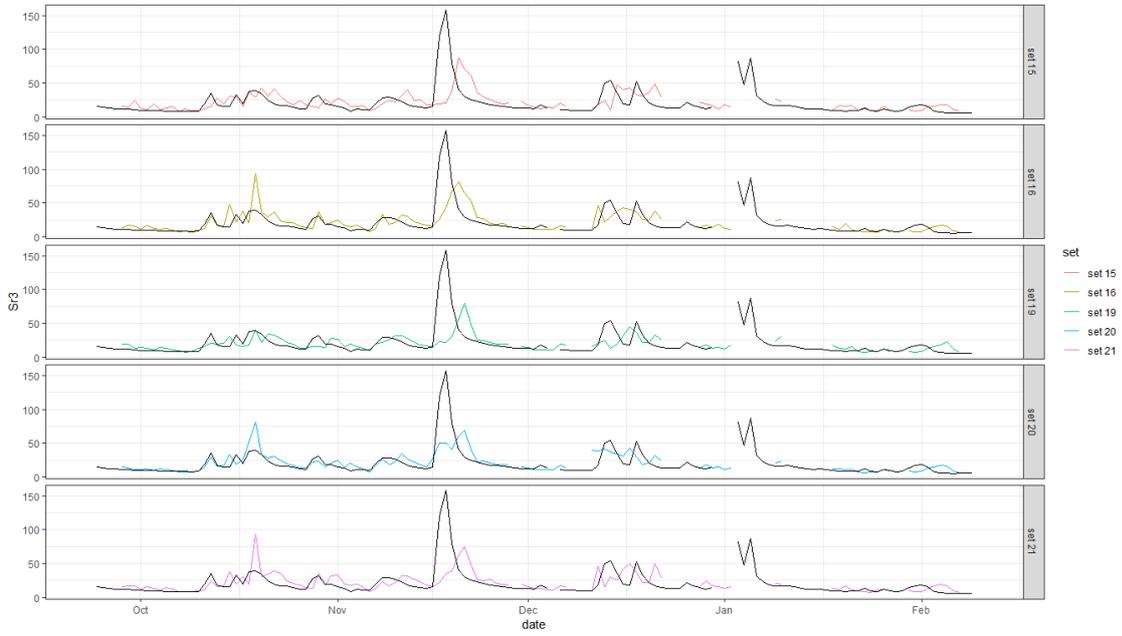
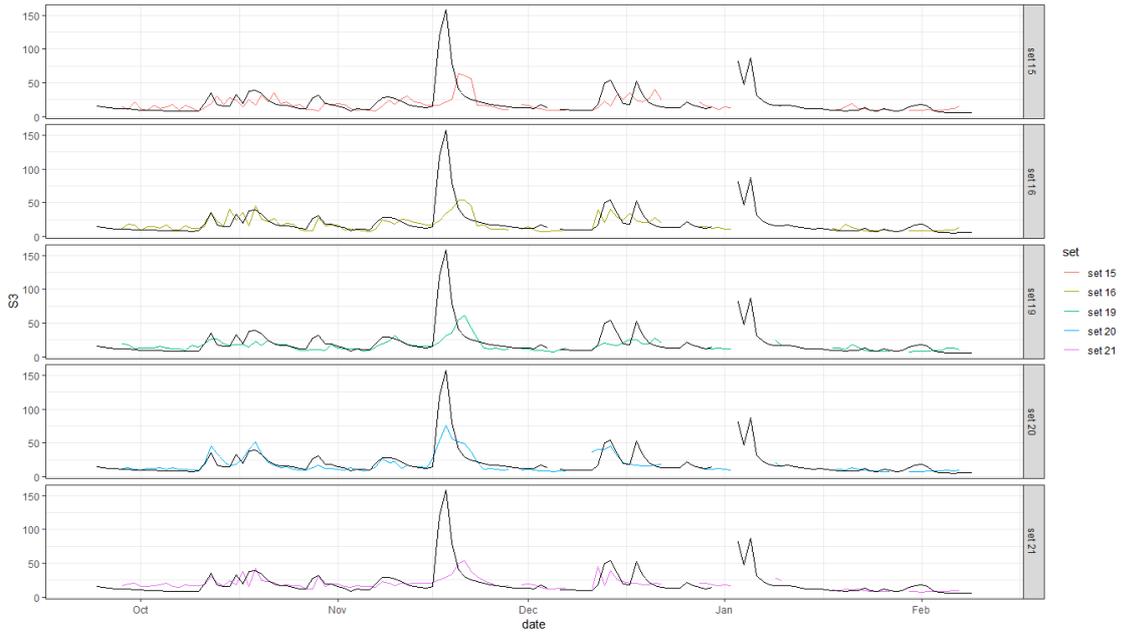


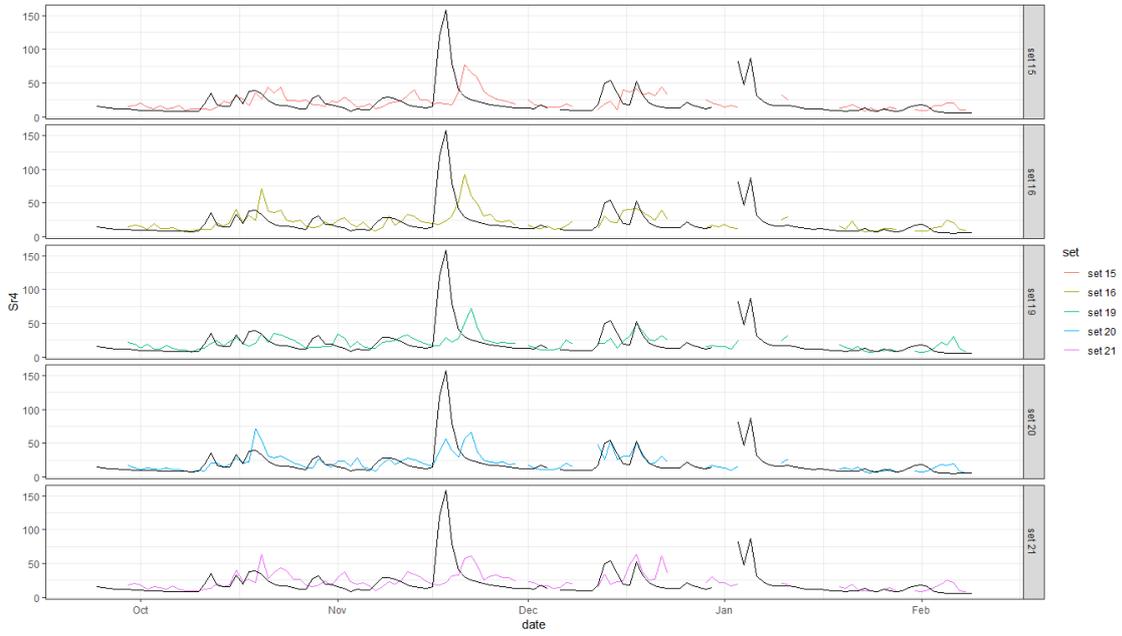
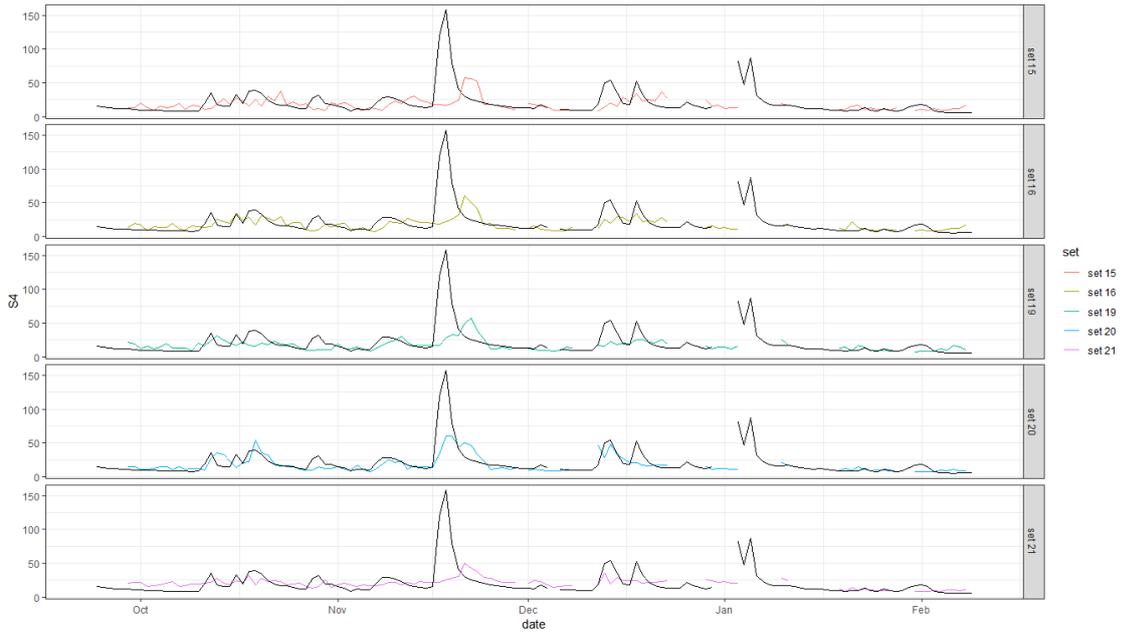


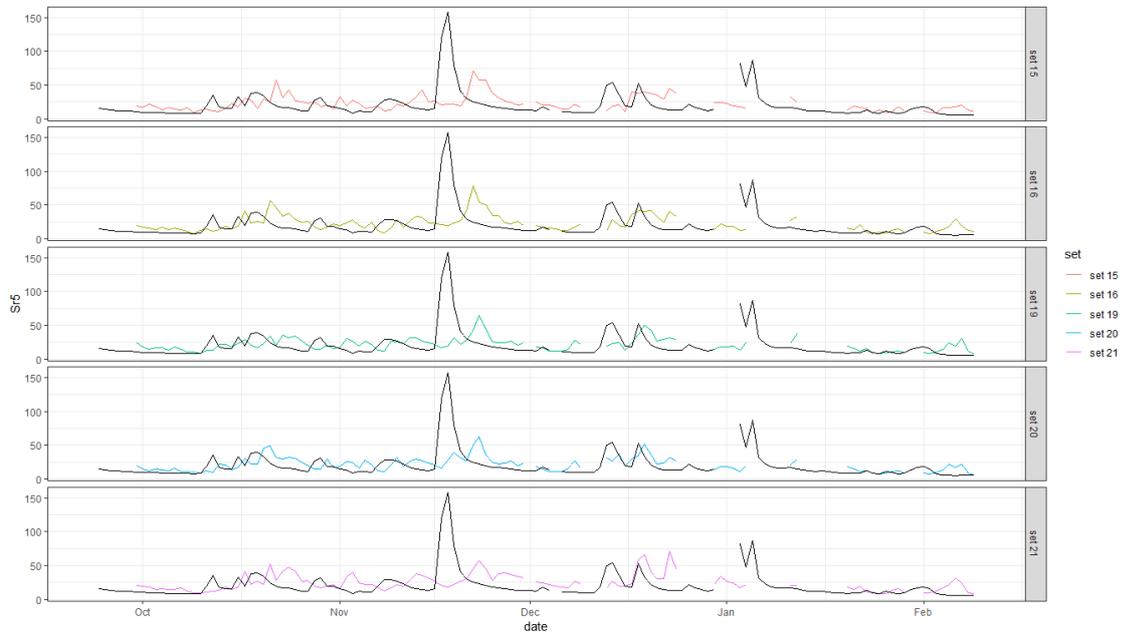
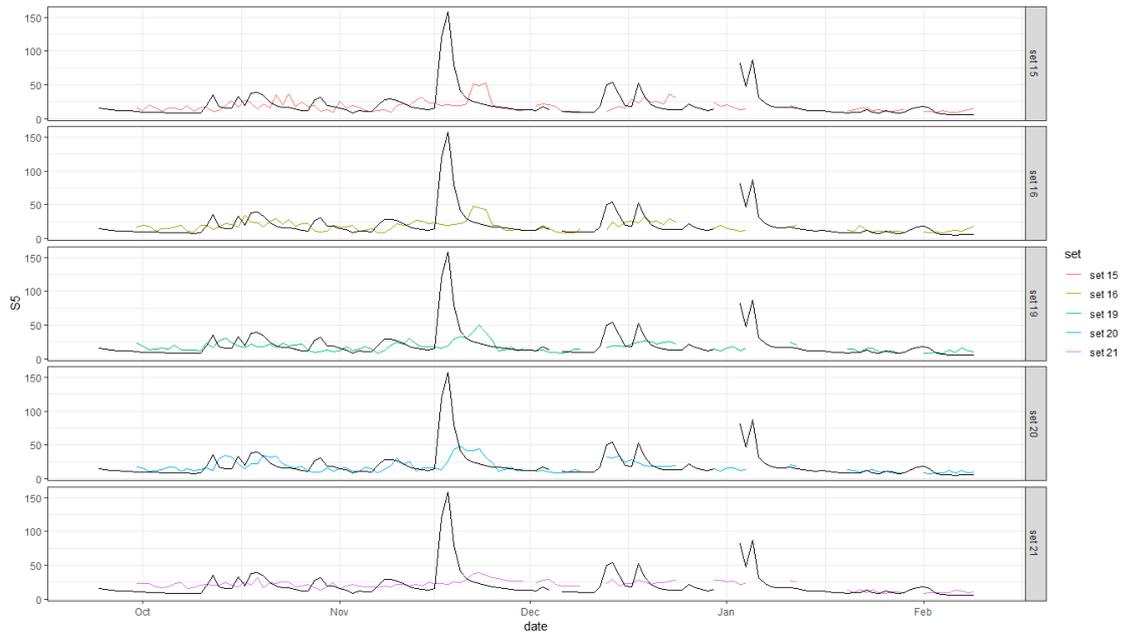
B.2.b. vhm 150





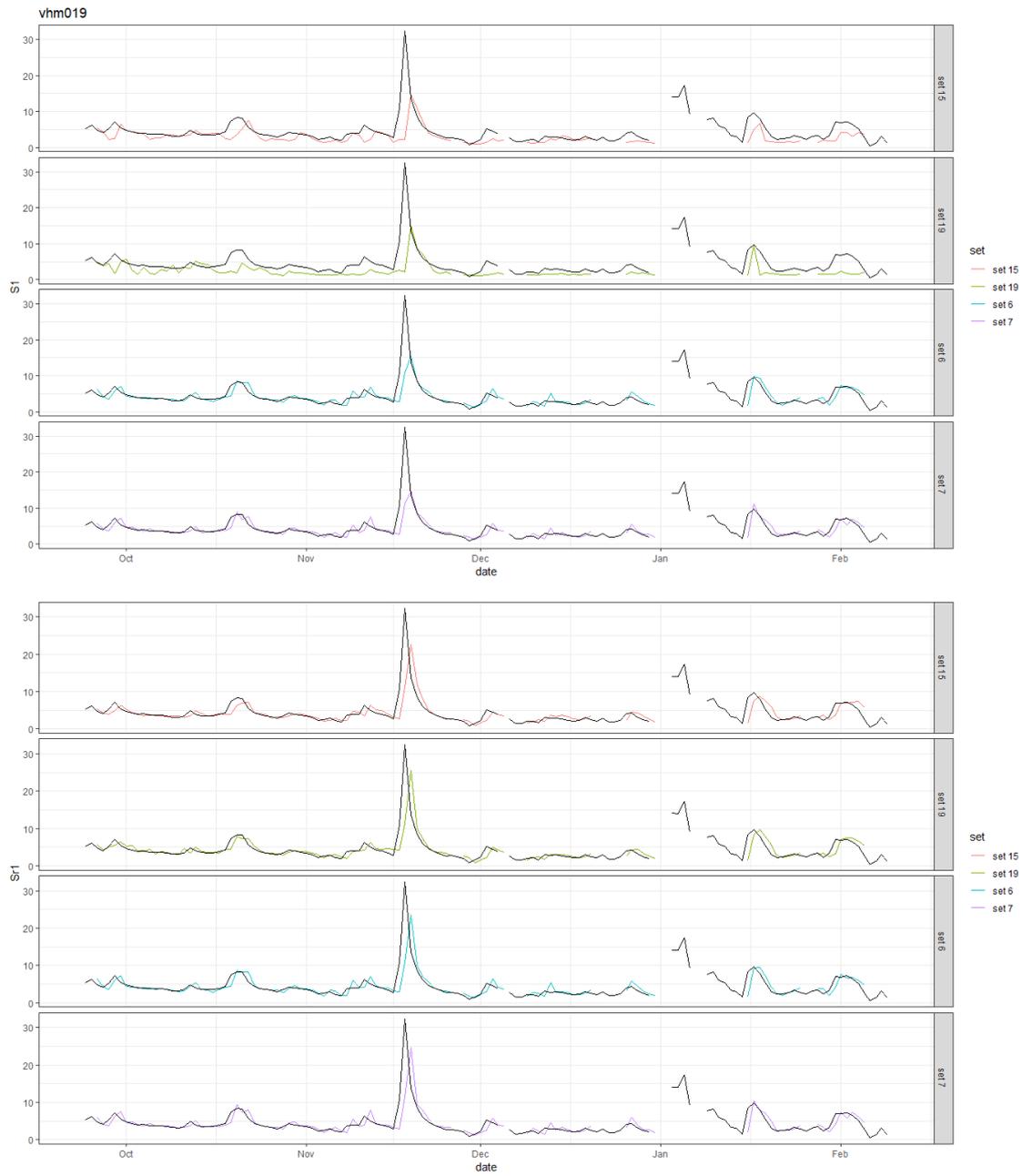


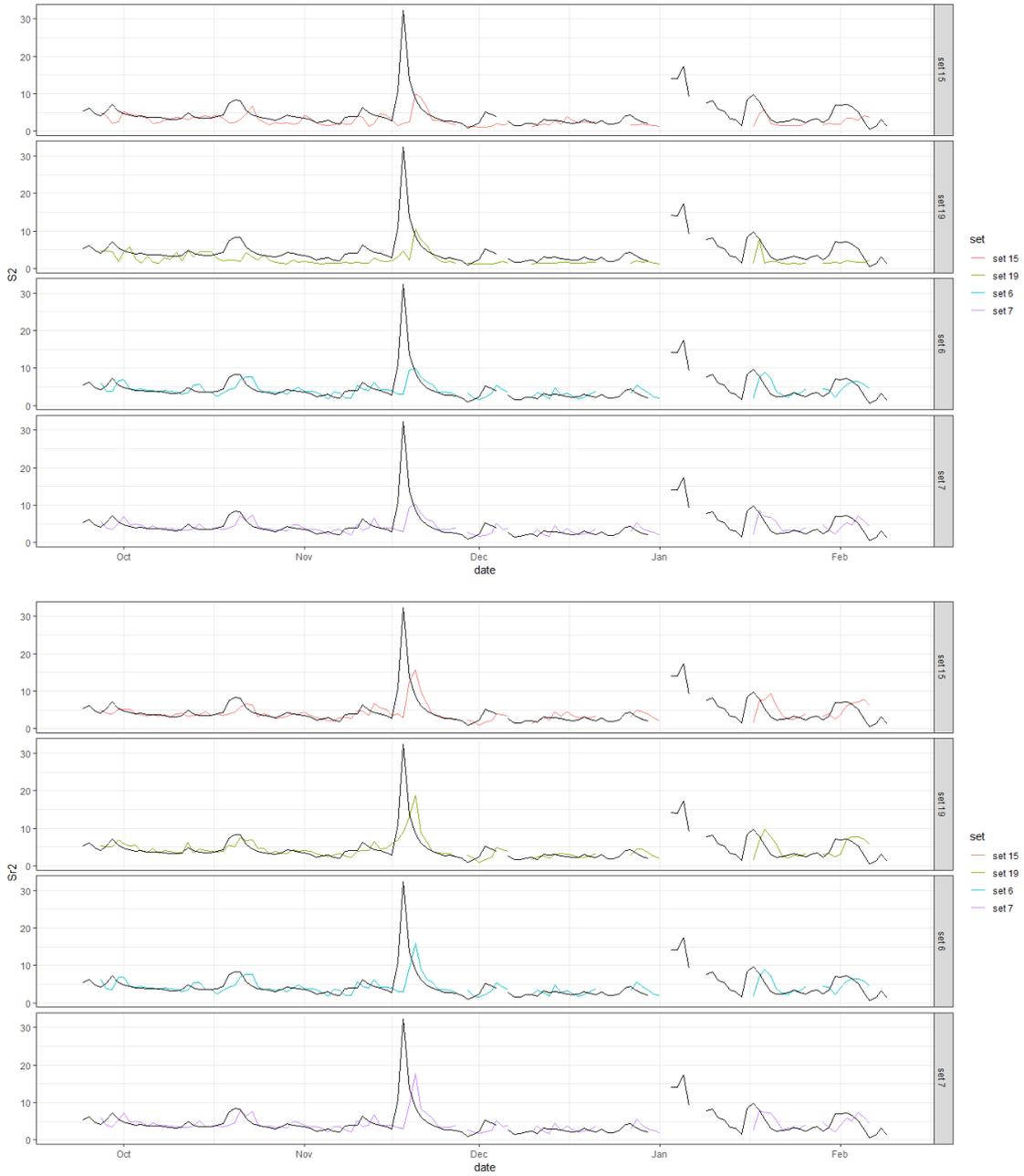


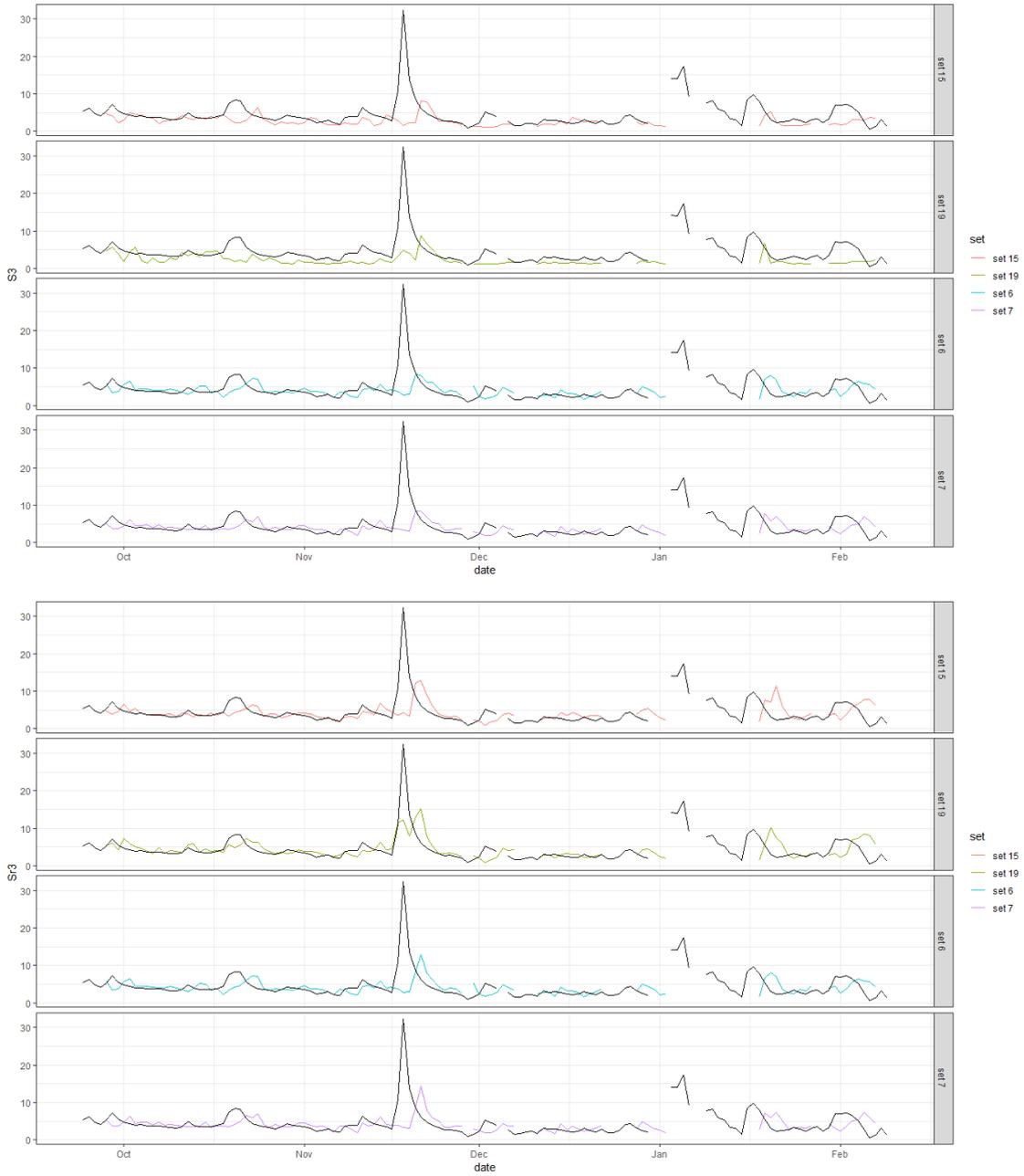


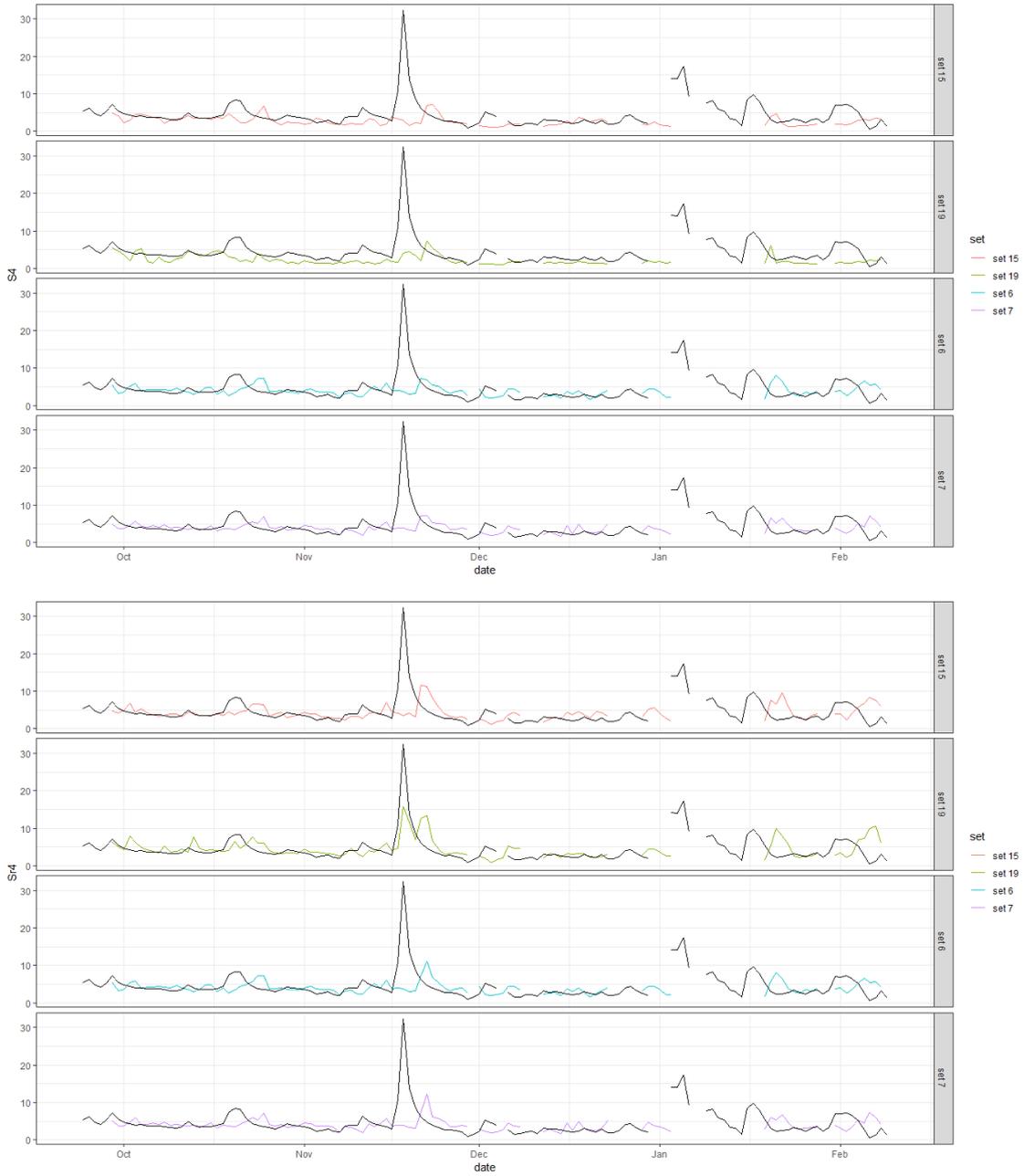
B.3. Class C

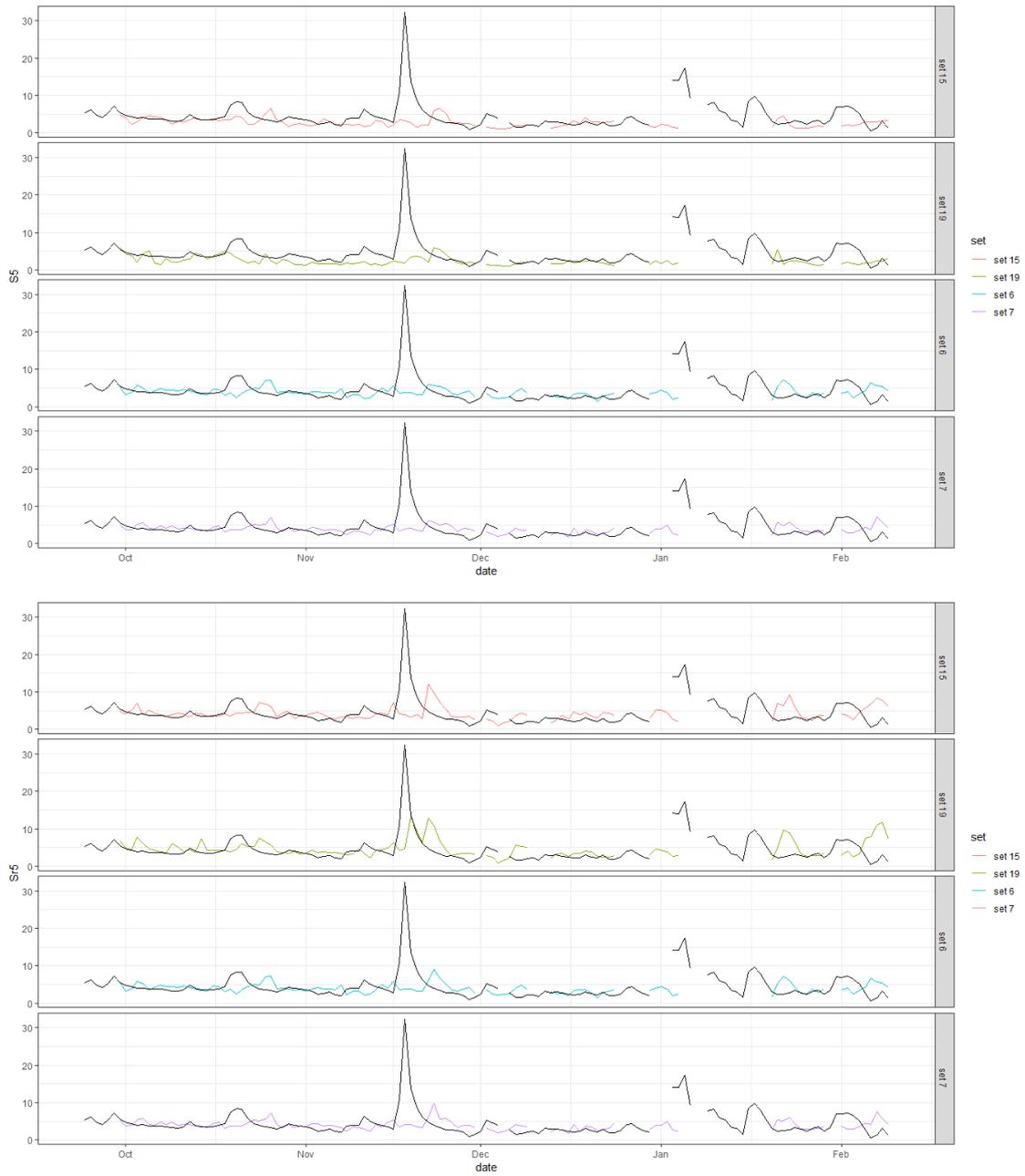
B.3.a. vhm 19



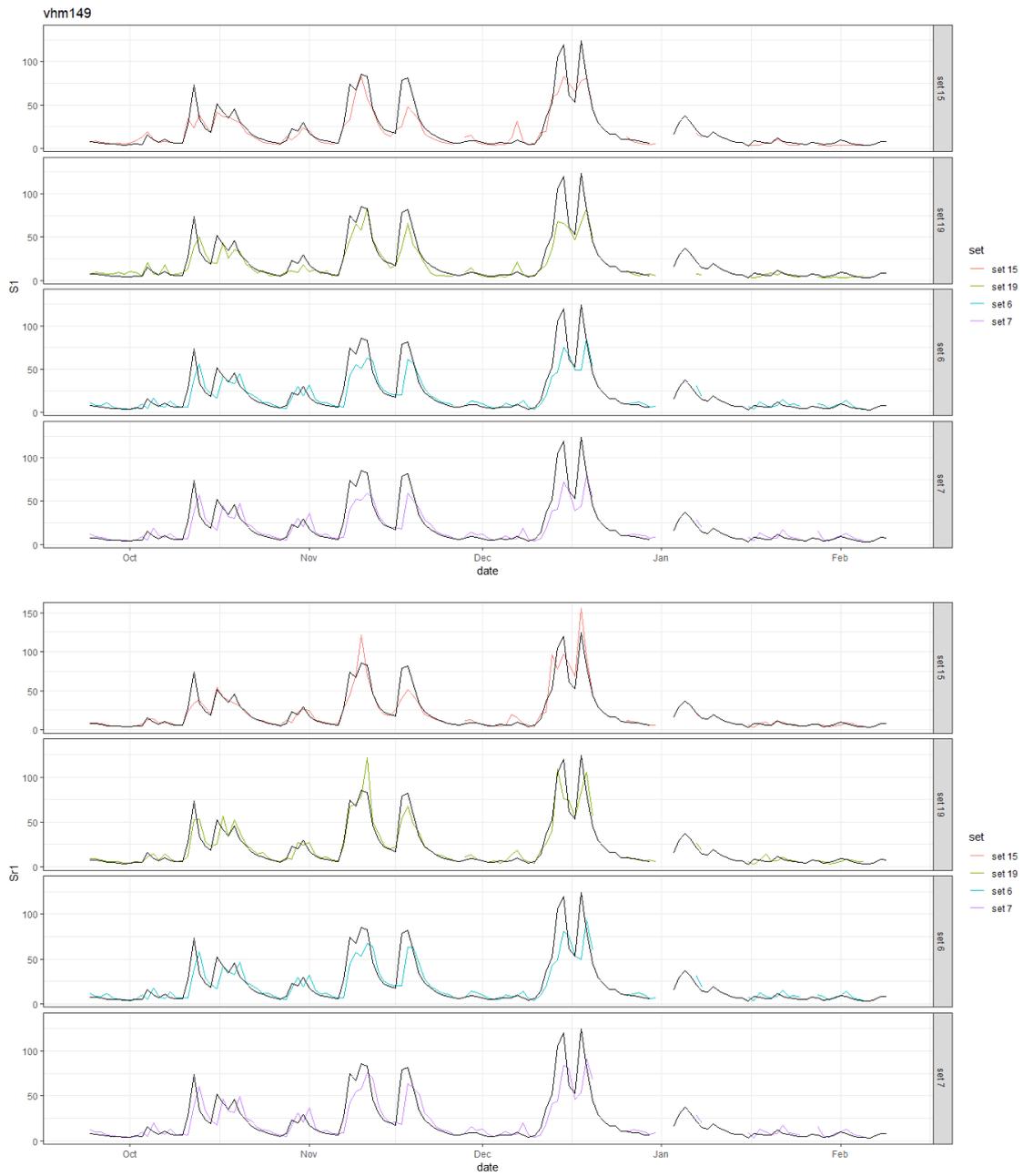


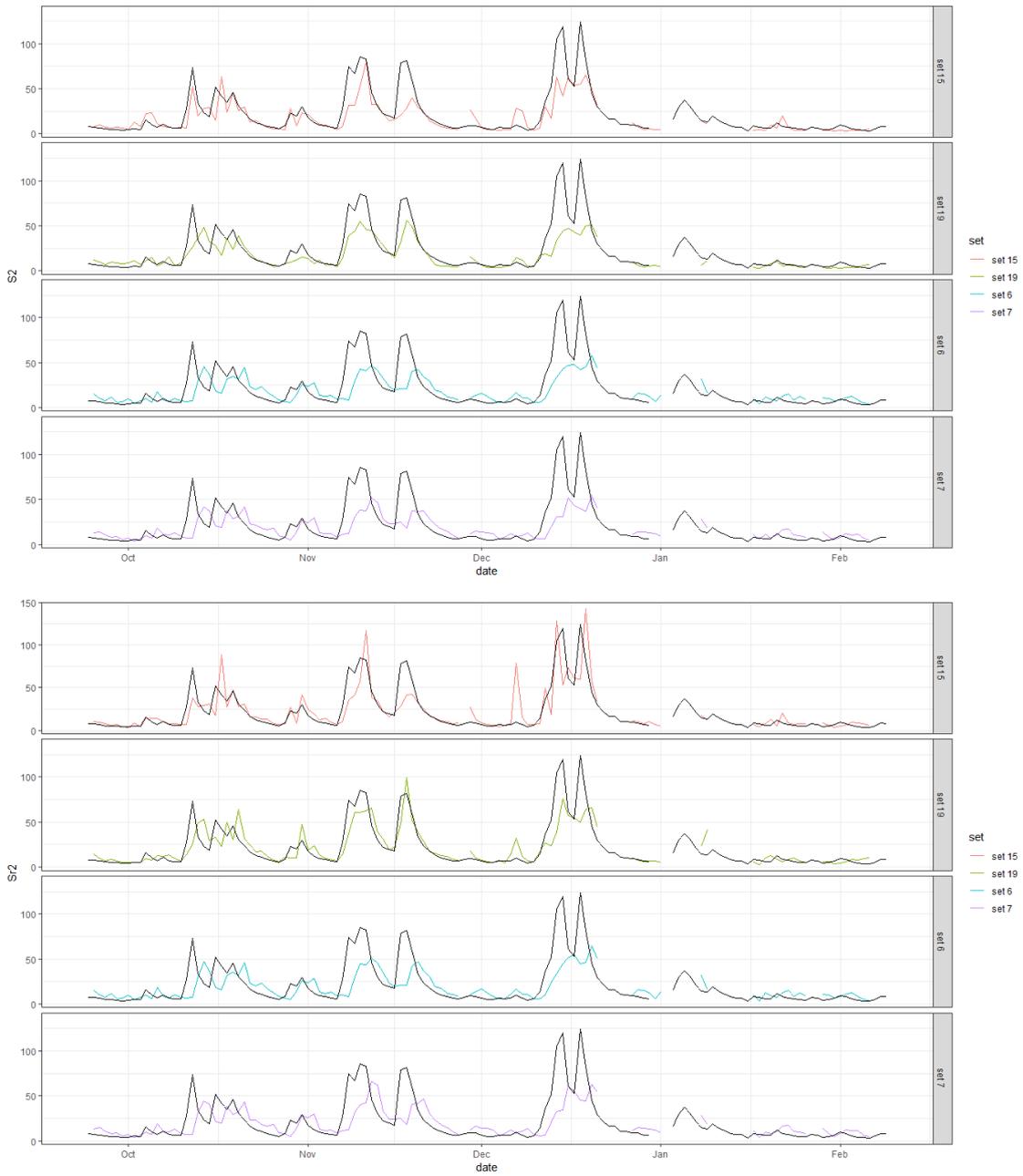


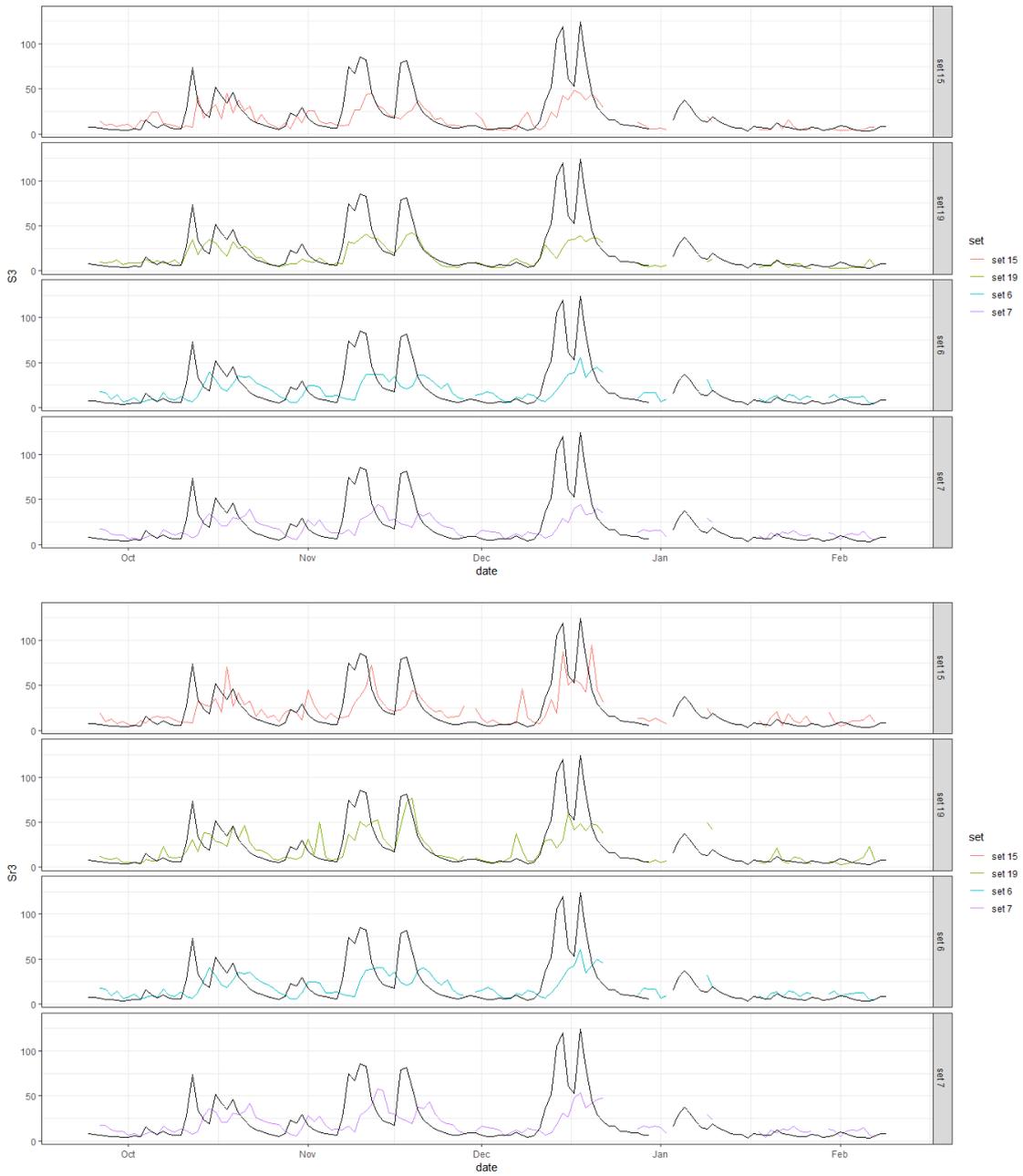


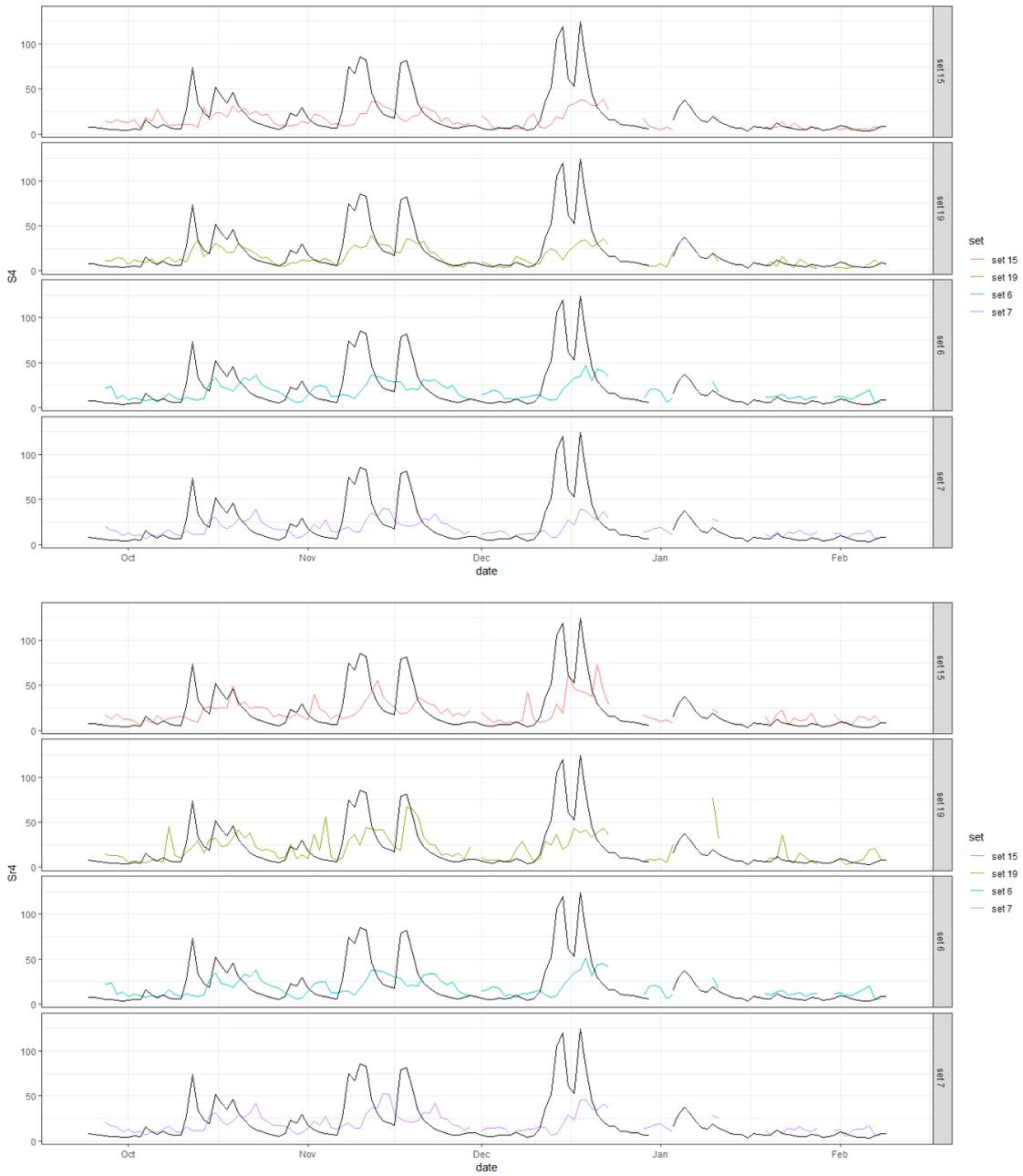


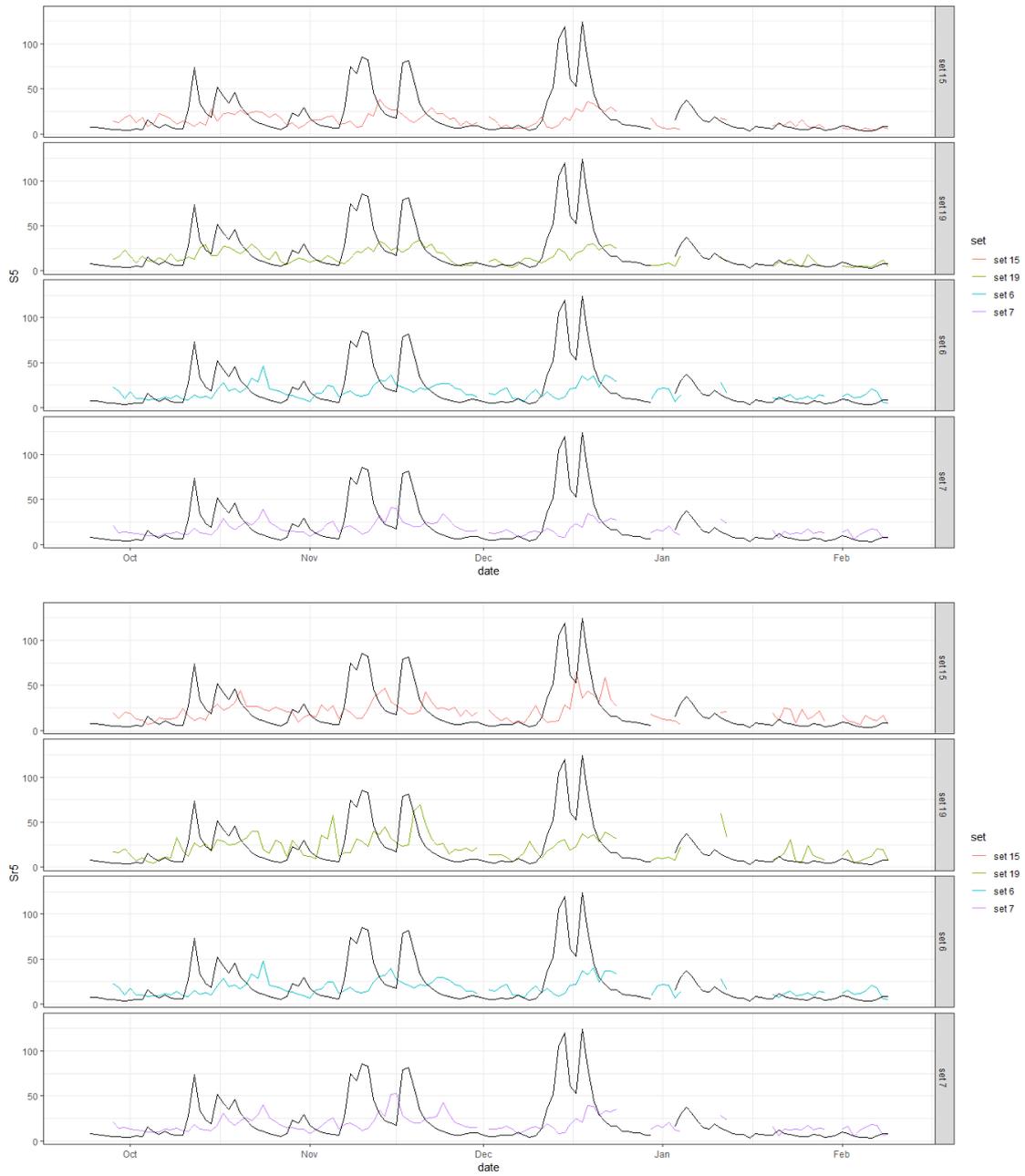
B.3.b. vhm 149



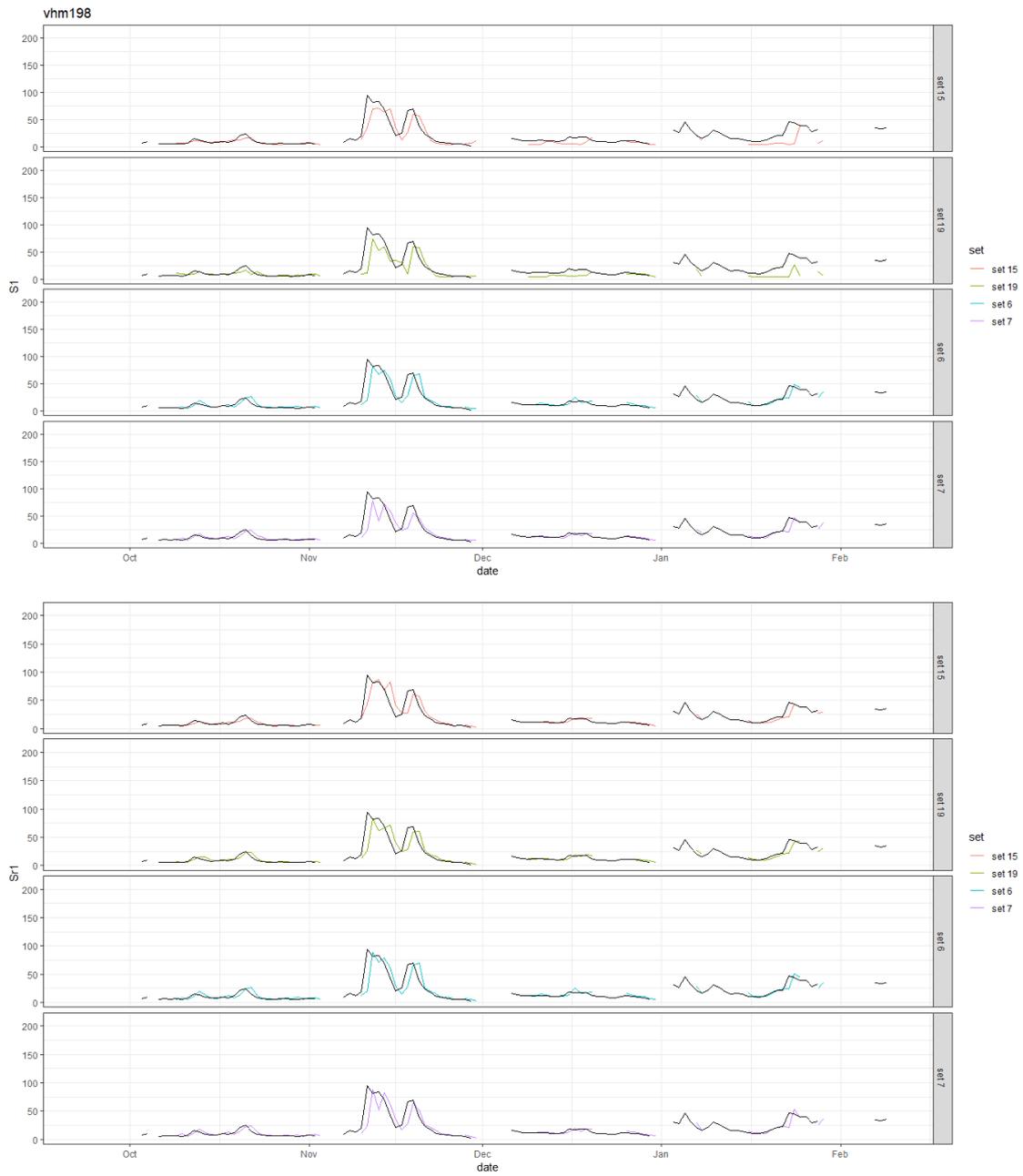


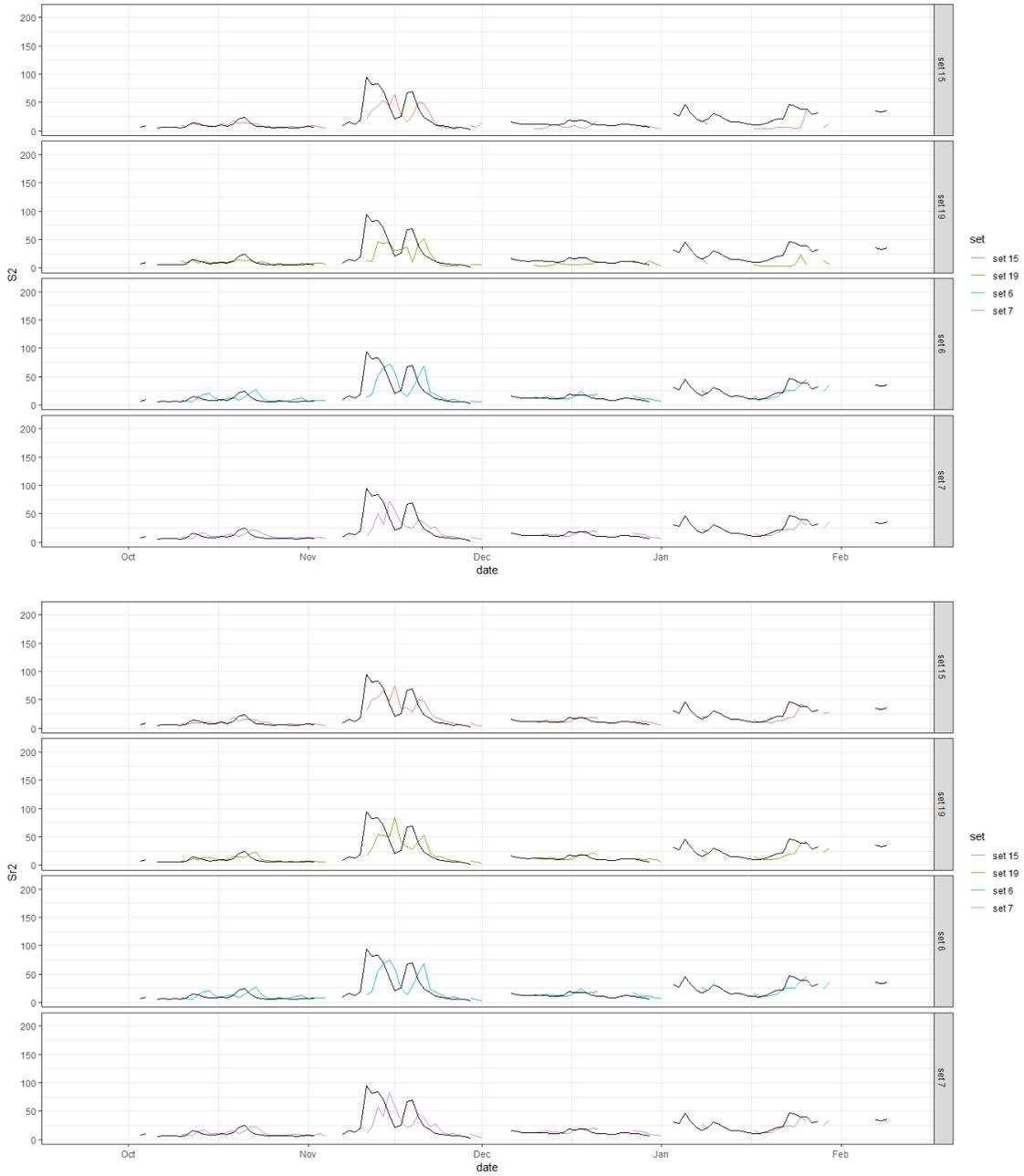


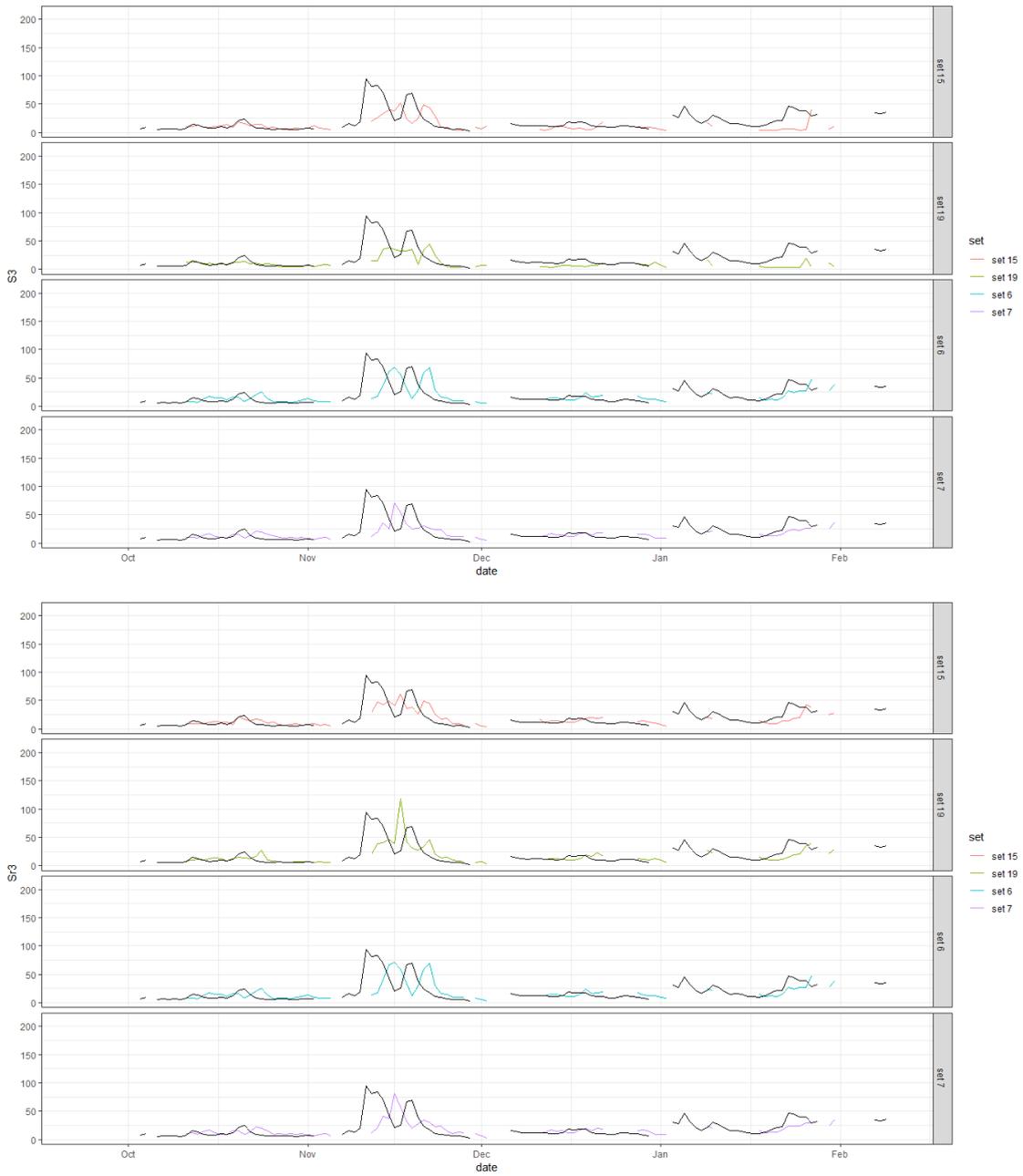


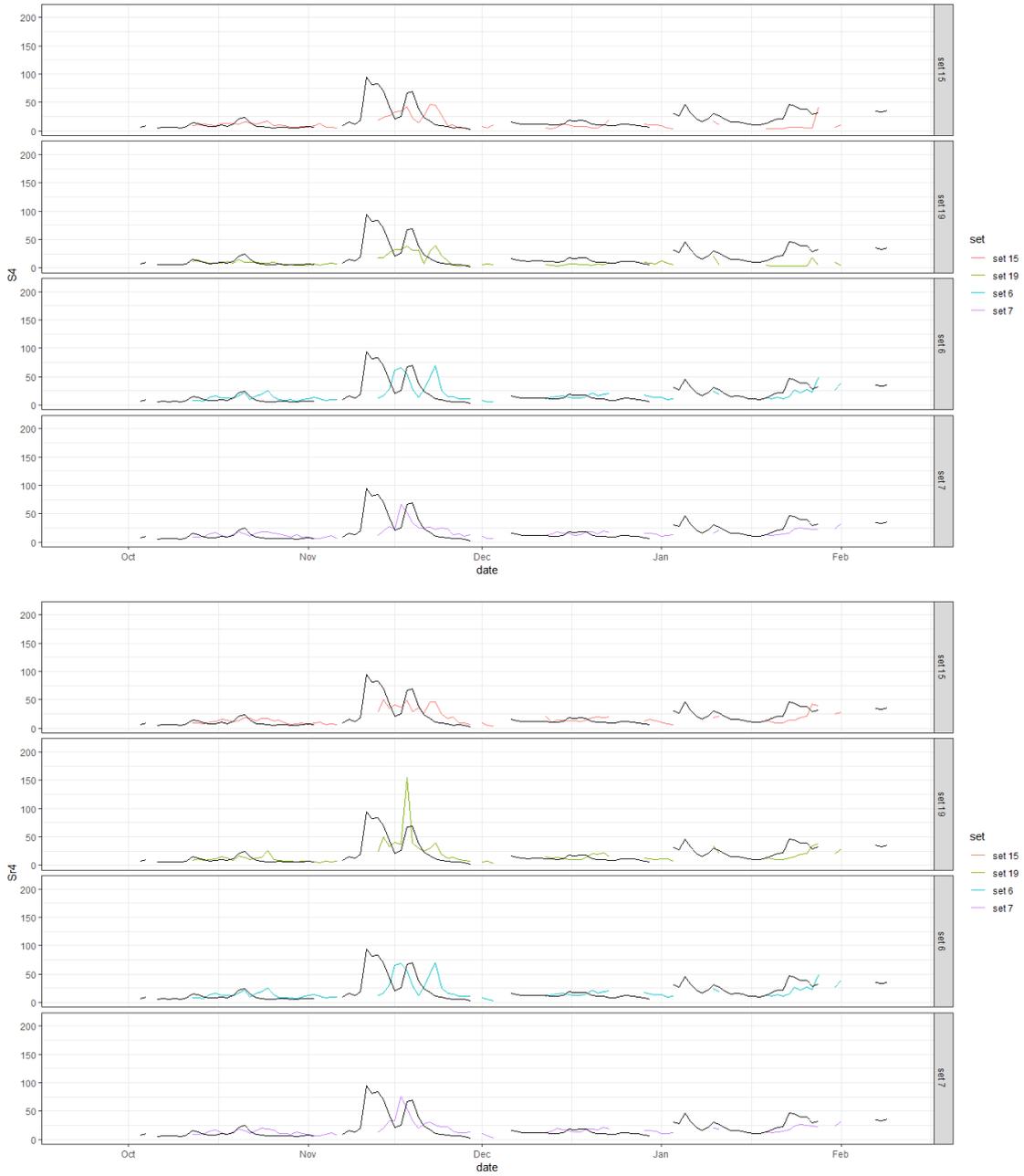


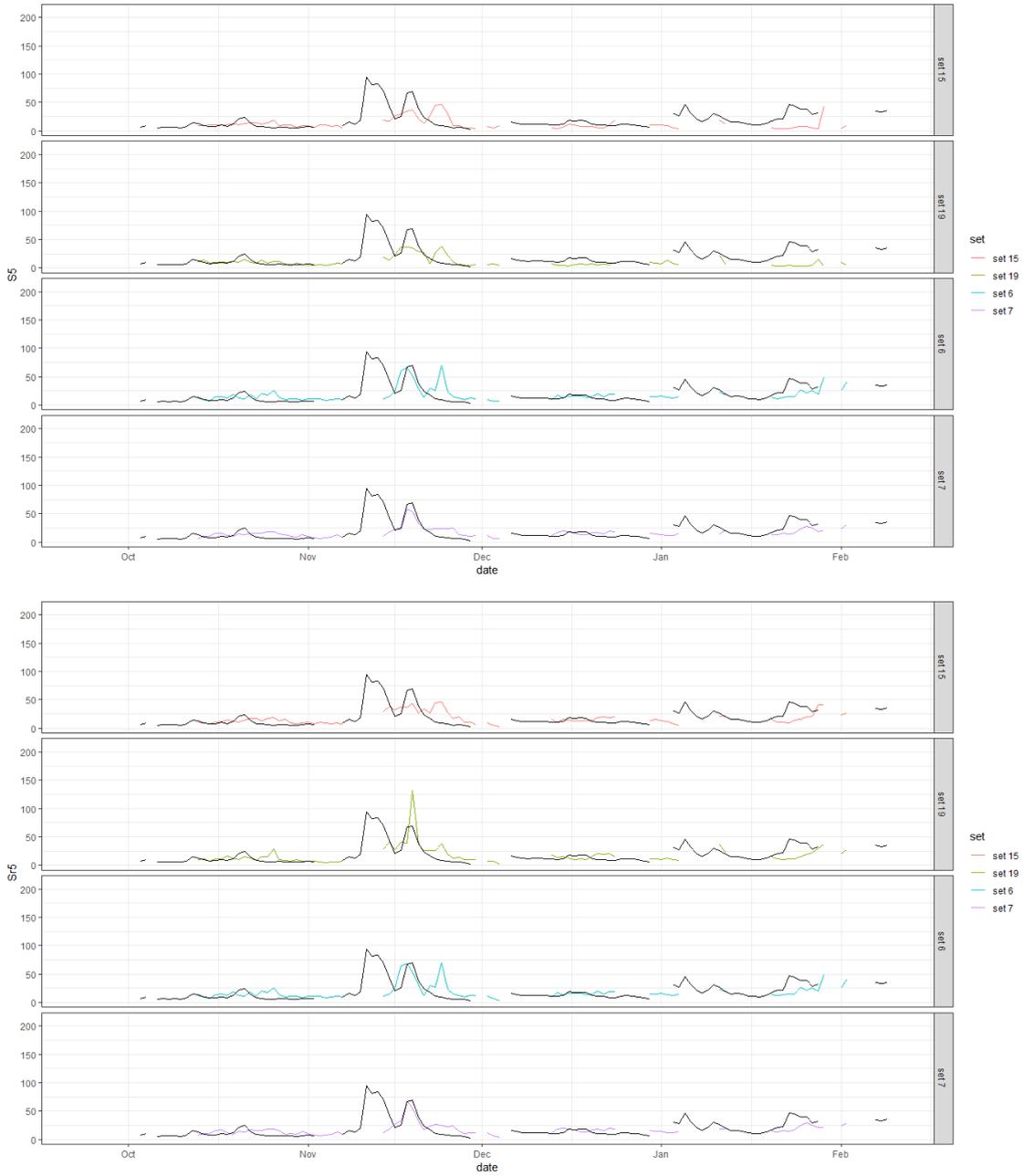
B.3.c. vhm 198



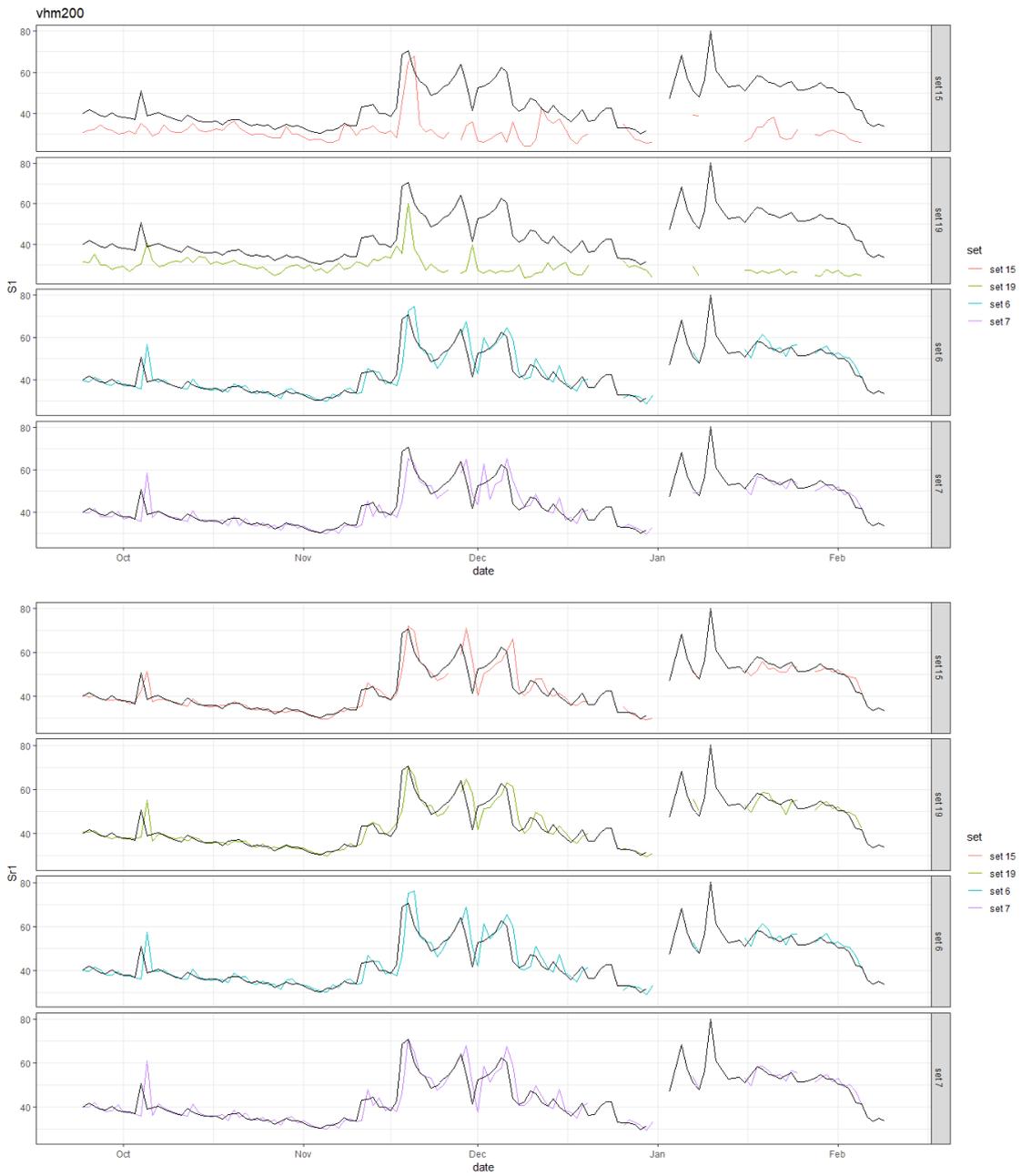


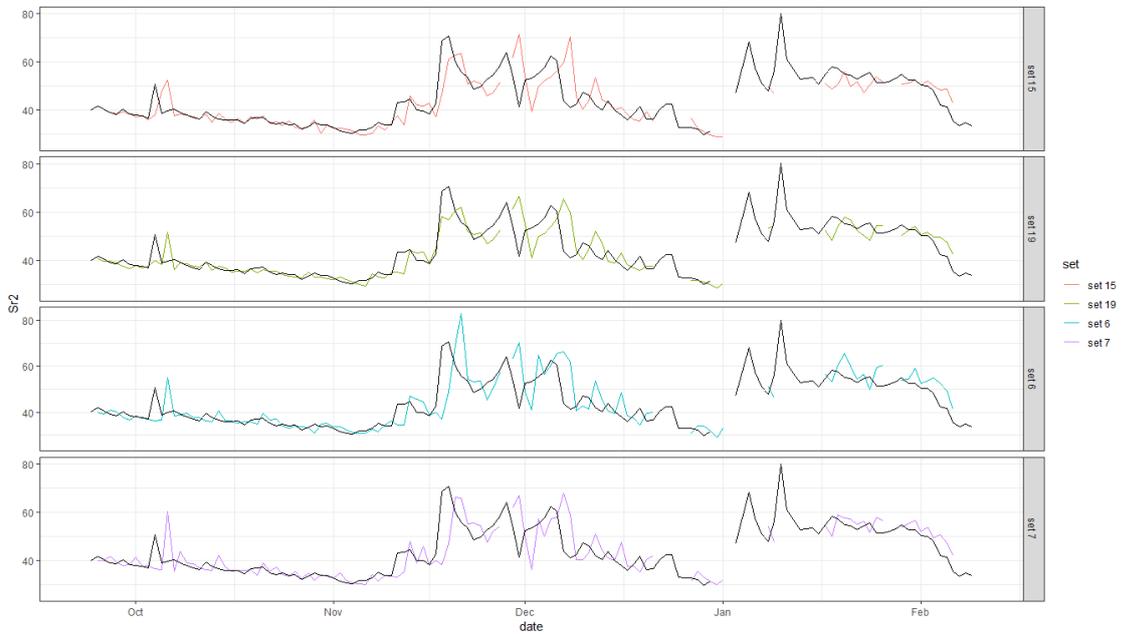
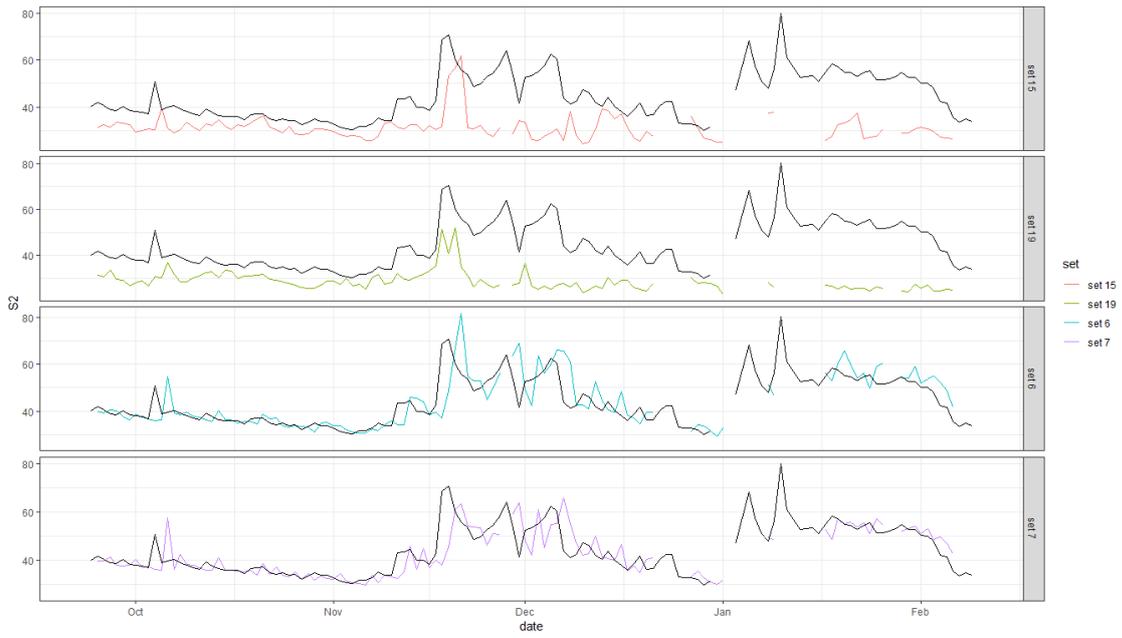


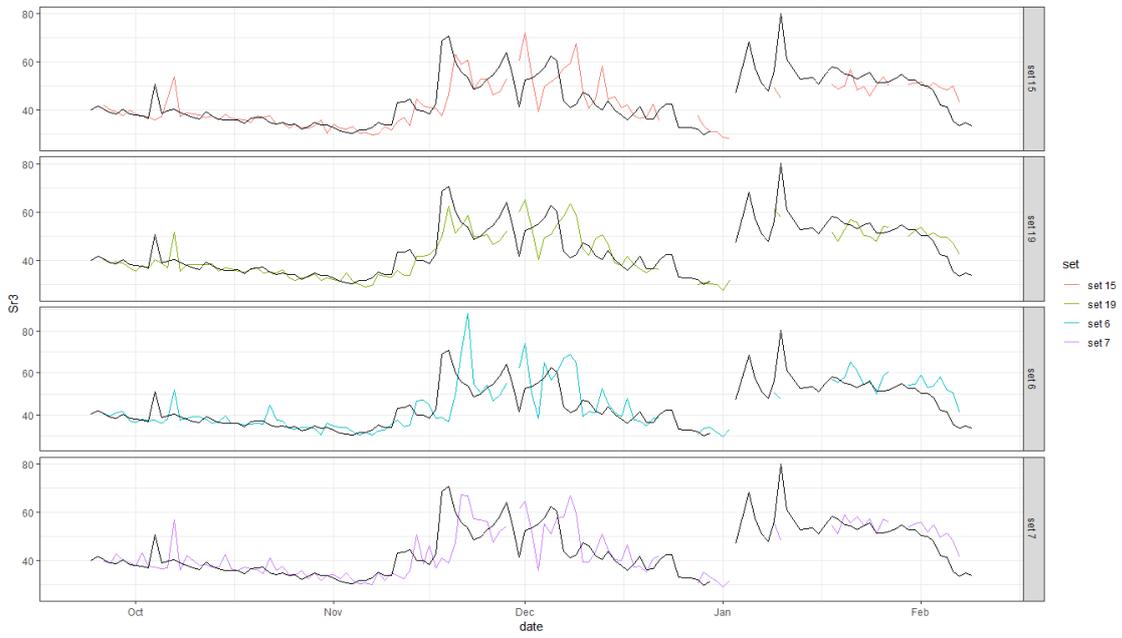
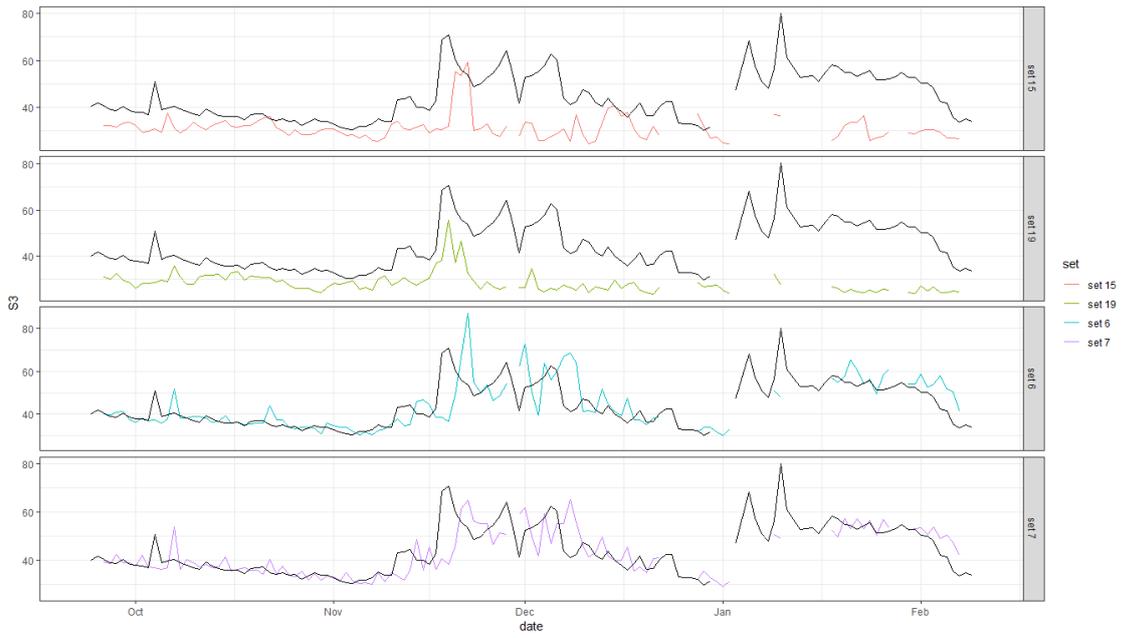


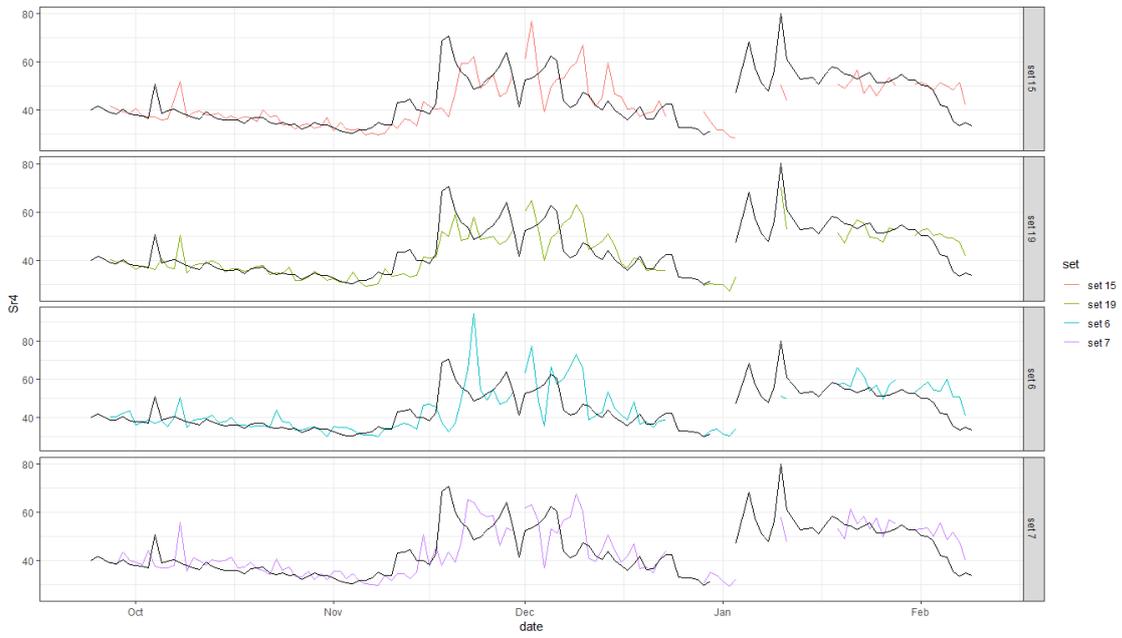
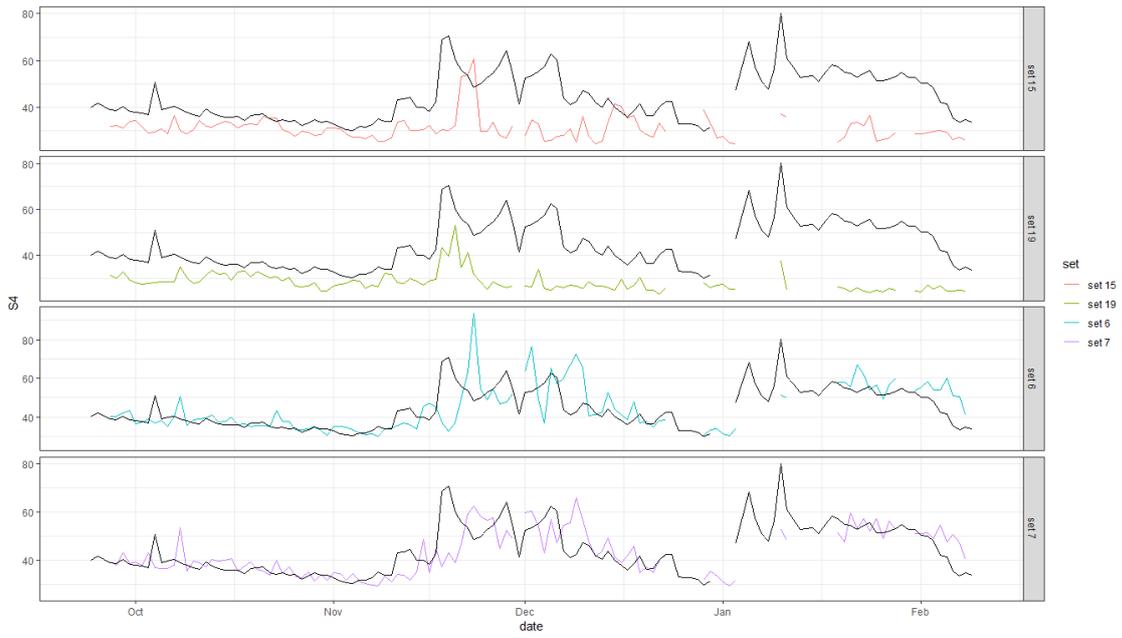


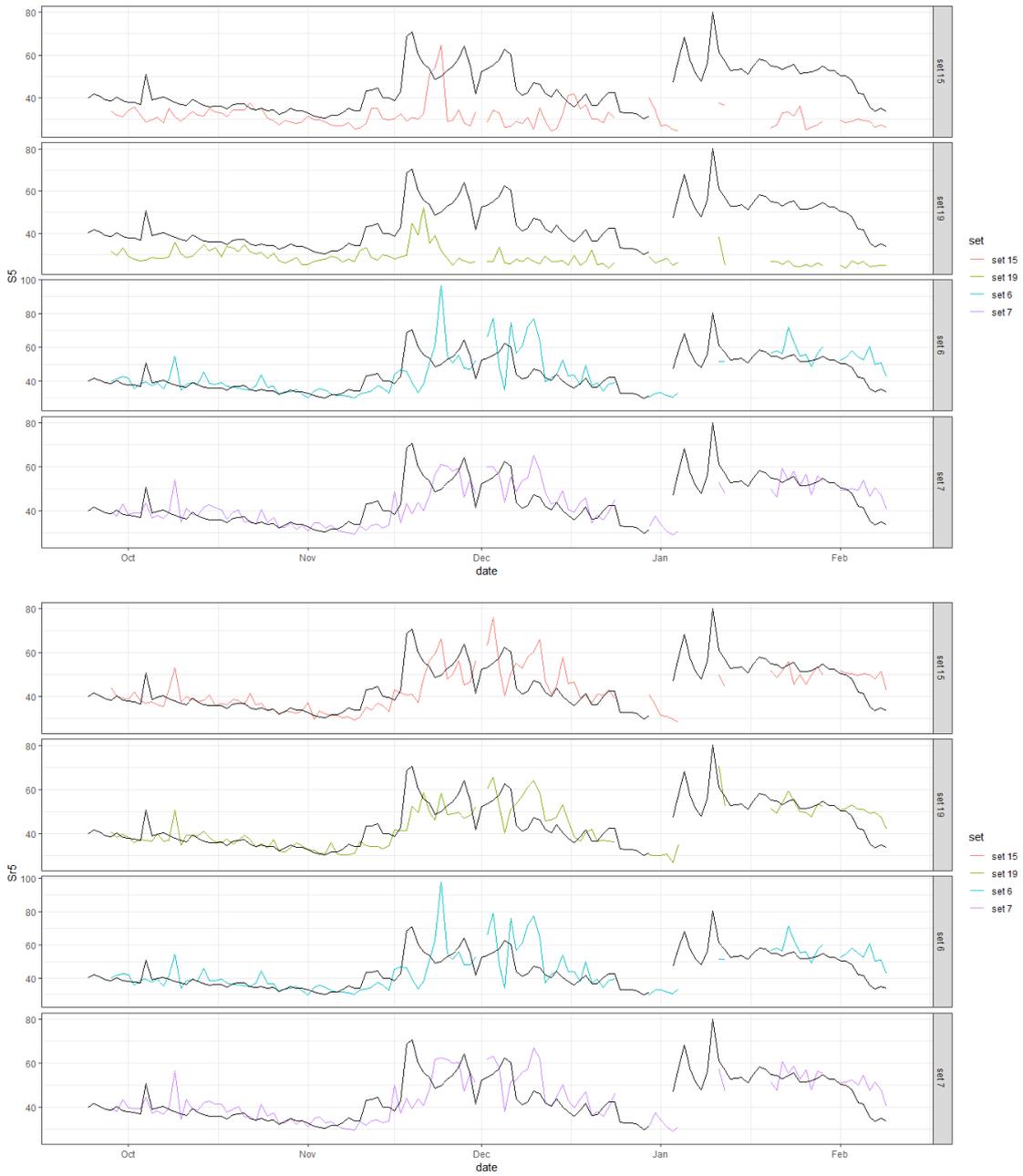
B.3.d. vhm 200











B.3.e. vhm 204

