Veðurstofa Íslands

Design-flood estimates from daily runoff simulations using the Icelandic Reanalysis (ICRA): a first step towards estimating extremes in ungauged catchments

Andréa-Giorgio R Massad
Tinna Þórarinsdóttir
Matthew J. Roberts

Greinargerð

Veðurstofa Íslands
Bústaðavegur 7–9
IS 108 Reykjavík

+354 522 60 00
vedur@vedur.is

AGM/ofl/2022-01

Veðurstofa Íslands

LYKILSÍÐA

Greinargerð nr. AGM/ofl/2022-01	**Dags.** Mars 2022	**Dreifing:** **Opin** ☒ **Lokuð** ☐ **Skilmálar:**
Heiti greinargerðar: Design-flood estimates from daily runoff simulations using the Icelandic Reanalysis (ICRA): a first step towards estimating extremes in ungauged catchments		**Upplag:** Rafræn útgáfa **Fjöldi síðna:** 41 **Framkvæmdastjóri sviðs:** Jórunn Harðardóttir
Höfundar: Andréa-Giorgio R Massad, Tinna Þórarinsdóttir og Matthew J. Roberts		**Verkefnisstjóri:** Tinna Þórarinsdóttir / Matthew J. Roberts **Verknúmer:** 4812-0-0003
Gerð greinargerðar/verkstig:		**Málsnúmer:** 2021-0060
Unnið fyrir: Rannsóknasjóð Vegagerðarinnar		
Samvinnuaðilar:		

Útdráttur:

Extreme flood estimates are important in the design of hydrological infrastructure, including highways, stormwater drains, bridges, and culverts. In this research, a first attempt to estimate extreme values based on simulated runoff from the ICRA reanalysis data is investigated. Firstly, runoff is converted into discharge and compared to measurements from 44 stations around Iceland. Hierarchical cluster analysis is then used to identify groups of stations that cluster similarly whether the analysis is based on observed or simulated discharge. Cluster-based corrections are then investigated to correct systematic overestimation in the simulated dataset. An Extreme Value Analysis is then performed and showed closer results between return level values based on observations and simulations after applying the correction in a majority of cases. Overall, these result show that extreme discharge values based on catchment-accumulated runoff from the ICRA dataset is able to simulate the observed high discharge after correction. The findings of this study represent an initial methodology that could successfully assess design-flood values for ungauged catchments throughout the country.

Lykilorð: Flood analysis, hierarchical clustering, ungauged rivers	**Undirskrift framkvæmdastjóra sviðs:** **Undirskrift verkefnisstjóra:** **Yfirfarið af:** SG

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.

Table of Contents

1 Introduction

In Iceland, periods of extreme rainfall have led to numerous damaging floods, including widespread flooding in southeast Iceland in September 2017, as well as recent rainfall-induced landslides in Seyðisfjörður in December 2020, and 50-year flooding in the north of the country in June 2021. The landslides in Seyðisfjörður were caused by record-breaking rainfall, amounting to almost 570 mm over five days. Extreme flood estimates are important in the design of hydrological infrastructure, including highways, stormwater drains, bridges, and culverts.

A recently published Icelandic study by Massad *et al.* (2020) reassessed precipitation return levels in Iceland, resulting in a new national map of 24-hour precipitation thresholds for a 5-year event. The study also presented intensity-duration-frequency curves for numerous locations in Iceland, describing the relationship between rainfall intensity, rainfall duration, and return periods.

The 2020 study was based on hourly precipitation data made recently available by the Icelandic reanalysis of atmospheric conditions, known as the ICRA dataset (Nawri *et al.*, 2017). The ICRA dataset was derived from the HARMONIE numerical weather prediction model, providing access to various atmospheric parameters from over 11,000 grid-points at 2.5 km horizontal resolution. The dataset begins in 1979, providing over 38 years of hourly data.

In this project, the runoff parameter from the ICRA dataset is investigated using the same extreme-value approach by Massad *et al.* (2020) and Þórarinsdóttir *et al.* (2021). Comprising liquid precipitation and snowmelt, the runoff parameter represents a new means for design-flood estimates at any non-glaciated location in Iceland. In the future, this dataset will allow flood return-periods to be estimated for ungauged catchments, enabling small-scale engineering assessments of runoff extremes at virtually any location.

Extreme runoff estimates from ungauged catchments are challenging. In fact, such estimates represent one of the leading problems in flood hydrology. In several recent studies, Veðurstofan has investigated flood forecasting in ungauged catchments, including simulations using the WaSIM hydrological model in the Westfjords and Tröllaskagi regions (Crochet and Þórarinsdóttir, 2014). An index-flood method was also tested in the Eastfjords, leading to promising initial results (Crochet and Þórarinsdóttir, 2015). With the increasing dependence on Iceland's road infrastructure, combined with the uncertainties of rapid climate change, there is a need to develop updated design-flood methods for rapid and widespread assessments. This project is a first step towards delivering such a methodology.

Building on previously funded Vegagerðin research projects, the goal of this study is to investigate how reliable the ICRA runoff is to estimate flood extremes. The project will follow several steps:

1. Firstly, daily runoff from the ICRA will be extracted for 44 catchments where discharge measurements have been recorded for more than 20 years.
2. In a second step, two hierarchical clustering will be presented: one based on the ICRA discharge, the other on measurements with the aim of determining which rivers cluster similarly in both analyses. These rivers will constitute a group of control stations.
3. Flood extremes will then be calculated based on both datasets using the Block Maxima method, and a cluster-based correction will be proposed to improve the extreme values results from the simulated discharge.
4. Finally, these corrections will be applied to the other catchments, based on the clustering results from the simulation, as would be for ungauged rivers. Values will then be validated by the observations, offering a step towards the estimation of flood extremes in ungauged areas.

2 Data

2.1 Measurements from the gauging station network

Since the first gauging stations were set up in Iceland, the gauging network has expanded to record most of the major rivers in the country, allowing for high resolution measurements down to 10-minute intervals. River discharge is not measured directly: the gauges measure the water level, which is then converted into a discharge using flow rating curves. The rating curves are measured at the gauge location through cross-section of the river and establish the correspondence between water level and discharge. The rating curves are updated regularly, as river path and characteristics change over time.

For this study, daily-averaged discharge measurements from a total of 44 gauging stations are used (Figure 1). Those stations were previously used for testing and calibrating the hydrologic model AirGR (Atlason *et al.*, 2021) as well as for the analogue forecast set up for Vegagerðin (Priet-Mahéo *et al.*, 2020 and 2021). On Figure 1, the gauging stations and their associated catchments are shown on a map of Iceland, with a colour code indicating their river type. According to that classification, four kinds of rivers exist in Iceland; although, in reality, they are often a combination of two or three different types. In the North, East and in the Westfjords, direct-runoff rivers (18 catchments, in green on the figure) lie on old, rather impermeable bedrock. On newer bedrock, spring-fed rivers (16 catchments, in blue on the figure) are dominant. Mostly fed by Vatnajökull, seven rivers are classified as glacial rivers (in grey). Finally, two catchments are primarily considered as lake rivers (in orange). Three gauging stations associated with the river Skaftá (VHM 70, VHM 183, and VHM 328) are qualified as jökulhlaup rivers, and are therefore particularly complicated to forecast because of the unpredictability of those events. They will be discarded later in this study.

Table 1 enlists all the rivers used for this study, and indicates the beginning and end year of the observed timeseries along with the number of missing days. Note that most stations are still recording as of today, but only data until 2017 were needed for this study, to match the reanalysis.

2.2 Simulated runoff from the Icelandic Reanalysis

2.2.1 The Icelandic Reanalysis (ICRA), and extraction of the relevant variables

The operational numerical weather prediction (NWP) system used by the Icelandic Meteorological Office (IMO) is the non-hydrostatic HARMONIE–AROME model, with a horizontal resolution of 2.5 × 2.5 km and 65 vertical levels (Bengtsson *et al.*, 2017). The fine-scale gridding gives 66,181 terrestrial points over Iceland. The model has been used to reanalyse atmospheric conditions in Iceland at hourly time-steps between September 1979 and August 2017, resulting in the Icelandic Reanalysis (ICRA) dataset (Nawri *et al.*, 2017).

As in most NWP systems, runoff (*ro*) is not a direct output from the model, but it is a combination of three variables: the rainfall rate (*rf*), the rate of evaporation (*evap*) and the melting (*mlt*). Hourly runoff can therefore be calculated as:

$$ro = rf + mlt - evap$$

It should be noted that the melting variable is also an undirect product of the model resulting from the combination of sleet- and snowfall rates, sublimation, and snow water equivalent. Therefore, in total, six variables need to be extracted from the reanalysis in order to estimate the daily runoff.

Based on the 2.5 km horizontal resolution of the dataset, timeseries were extracted for the watersheds by summing the runoff from all grid-points within the catchment outlines. All those variables being cumulative, they were summed in order to have for each catchment one daily runoff timeseries covering nearly 40 years of reanalysis.

2.2.2 Conversion of runoff into discharge

To compare with the daily discharge timeseries from the gauges, the simulated daily runoff needs to be converted into a simulated discharge for each catchment. This is done with the following formula:

$$Q(m^3 s^{-1}) = \frac{runoff(mm) * 0.001 * cell\ area(m^2)}{60 * 60 * 24}$$

The main assumption is that all the simulated runoff reaches the river within the day and no infiltration occurs. This approximation is not expected to work similarly in all the watersheds: it is assumed to give good results for small, direct-runoff catchments, but lead to larger errors for catchments with a strong groundwater component, or with water reservoirs such as lakes or meres.

In this study, the focus is on extreme discharge values. Hence, even if a time lag exists between observed and simulated discharge (as a result from the fact that the runoff does not reach the river within the day), it is not expected to affect the flow analysis as the focus is on peak values, and not on the time of occurrence.

Table 1 – Station list, timeseries available, a number of missing days among that period.

River	Time-period	Missing days	River	Time-period	Missing days
10 - Svartá	1932 – 2017	0	121 - Ormarsá	2005 – 2016	0
12 - Haukadalsá	1950 – 2017	3165	128 - Norðurá	1970 – 2017	1360
19 - Dynjandisá	1956 – 2017	789	144 - Austari-Jökulsá	1971 – 2017	0
26 - Sandá	1965 – 2017	460	148 - Fossá	1968 – 2017	224
30 - Þjórsá	1947 – 2017	2564	149 - Geithellnaá	1971 – 2017	5457
38 - Þverá	1980 – 2017	0	150 - Djúpá	1968 – 2017	1
43 - Brúará	1948 – 2017	0	162 - Jökulsá á Fjöllum	1984 – 2017	0
45 - Vatnsdalsá	1948 – 2017	2088	183 - Skaftá	1980 – 2017	0
48 - Selá	1982 – 2017	0	185 - Hólmsá	1980 – 2017	0
51 - Hjaltadalsá	1980 – 2017	0	198 - Hvalá	1976 – 2017	1
59 - Ytri-Rangá	1961 – 2015	0	200 - Fnjóská	1976 – 2017	0
60 - Eystri-Rangá	2005 – 2017	0	204 - Vatnsdalsá	1976 – 2017	4261
64 - Ölfusá	1980 – 2017	0	205 - Kelduá	1977 – 2017	2745
66 - Hvítá	1980 – 2017	0	206 - Fellsá	1976 – 2017	2956
68 - Tungufljót	1951 – 2017	2023	218 - Markarfljót	1982 – 2001	0
70 - Skaftá	1951 – 2017	1755	233 - Kreppa	1985 – 2017	0
81 - Úlfarsá	1956 – 2017	0	238 - Skjálfandafljót	1987 – 2017	0
83 - Fjarðará	1958 – 2017	2558	271 - Sog	1972 – 2017	0
92 - Bægisá	1980 – 2017	0	328 - Eldvatn	1993 – 2017	0
102 - Jökulsá á Fjöllum	1980 – 2017	0	400 - Vattardalsá	1980 – 2017	0
110 - Jökulsá á Dal	1963 – 2017	6194	408 - Sandá	1999 – 2017	0
116 - Svartá	1985 – 2017	0	411 - Stóra-Laxá	2000 – 2017	0

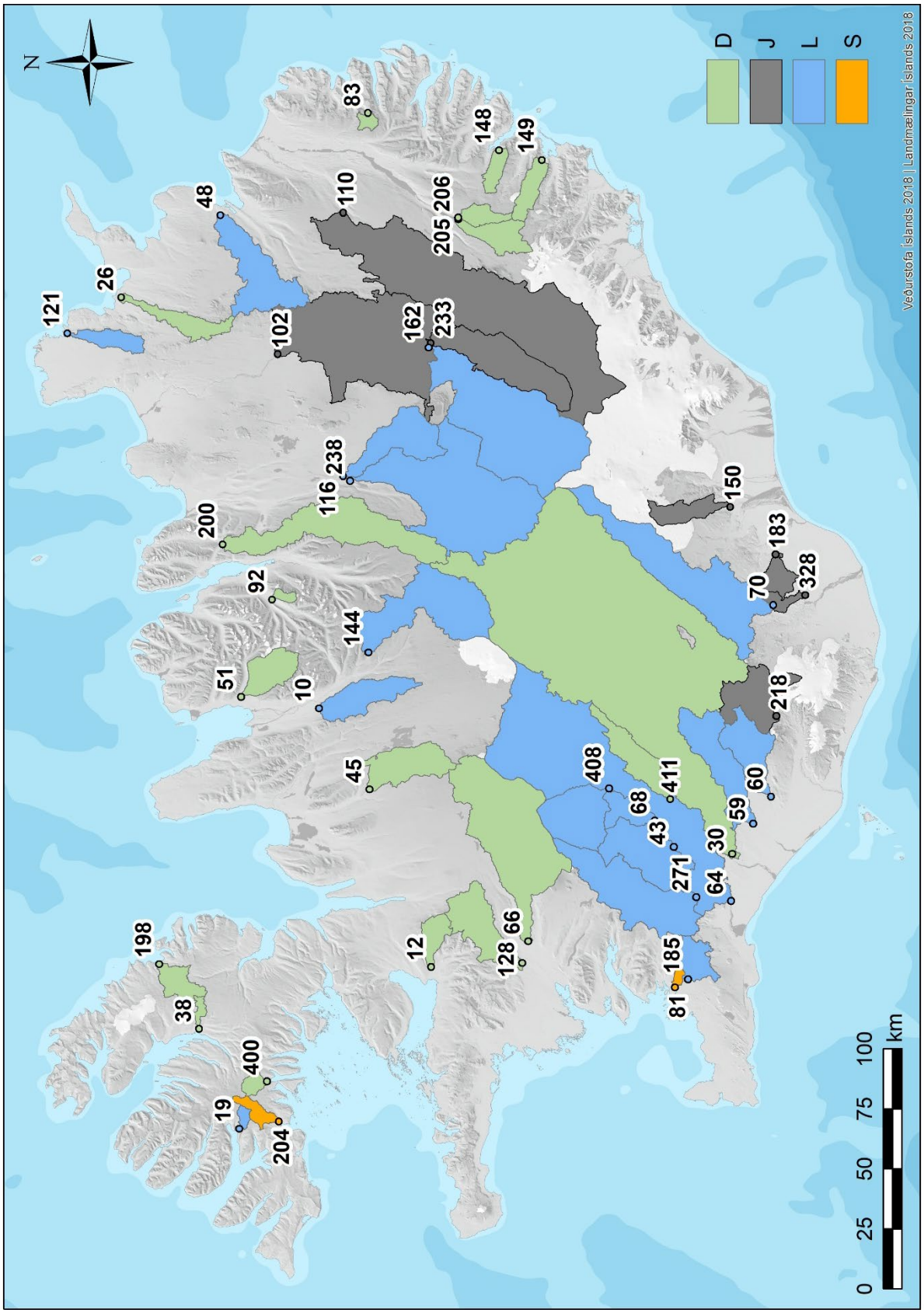

Figure 1 – Location of the gauging stations used in this study, and outlines of the associated catchments. Colour of the catchment areas depends on the type of river: direct-runoff (green), spring-fed rivers (blue), glacial rivers (grey), lake rivers (orange). Map from Atlason et al. (2021), based on the classifications of Rist (1990), Hróðmarsson et al. (2009, 2020), Hróðmarsson and Þórarinsdóttir (2018).

3 Cluster Analysis

3.1 Methodology

Over the years, several types of classifications have been developed with the aim of grouping rivers together according to their type. In 2014, rivers were classified based on the geology of the catchments and the presence of lakes and meres (Stefánsdóttir *et al.,* 2014), while Rist (1990), and Hróðmarsson and Þórarinsdóttir (2018) based their classification on observations made over more than 50 years of field measurements. More recently, a hierarchical cluster analysis has been used to categorize rivers in groups that share more similarities than with any other rivers from other groups. This analytic method was previously used by Crochet (2012) and was adapted for Icelandic rivers in two previous Vegagerðin-funded projects (Priet-Mahéo *et al.*, 2019 and 2021). According to Demirel and Kahya (2007), the Ward's method based on Euclidean distances is better suited when performing a cluster analysis for hydrological data.

In this research, the dataset used for clustering is comprised of both discharge timeseries and several independent catchment characteristics. Only values between 2007 and 2017 were used in order to work with a homogeneous set of data. Discharge data were then combined in three different ways, each method reflecting a different behaviour of the river:

- *Seasonality.* Discharge is averaged over the whole timeseries by Julian day, emphasizing the seasonal pattern of each river. For each catchment, only monthly-averaged discharge is kept so that only the general trend is kept in the analysis, as weekly variations are unrelated to the type of river and rather reflective of punctual weather conditions.
- *Flow-duration curves*. Discharge is ranked decreasingly, and then plotted against 10% exceedance steps to create flow-duration curves. Those graphs express how often a discharge level is exceeded, providing a good indication of the river's power potential.
- *Mass curves*. Discharge is averaged over the whole timeseries by Julian day, and then summed cumulatively over day of year. Monthly differences are then computed between cumulated discharge and the values obtained if the discharge remained constant all year long.

An example of each discharge plot is shown on Figure 2 for Dynjandisá catchment (VHM 19), in the Westfjords. On the top panel, the seasonality plot is shown based on daily-averaged values (grey line) and monthly-averaged values (black line). For the cluster analysis, as mentioned previously, only the monthly values were used. Contrarily to those figures, it should also be noted that discharge timeseries were normalised before being processed, in order to facilitate the comparison between rivers with very different baseflows.

Additionally, several catchment characteristics (Table 2) were added to complete the analysis, including the area, aspect ratio, longest flow-path, mean elevation, and geological properties.

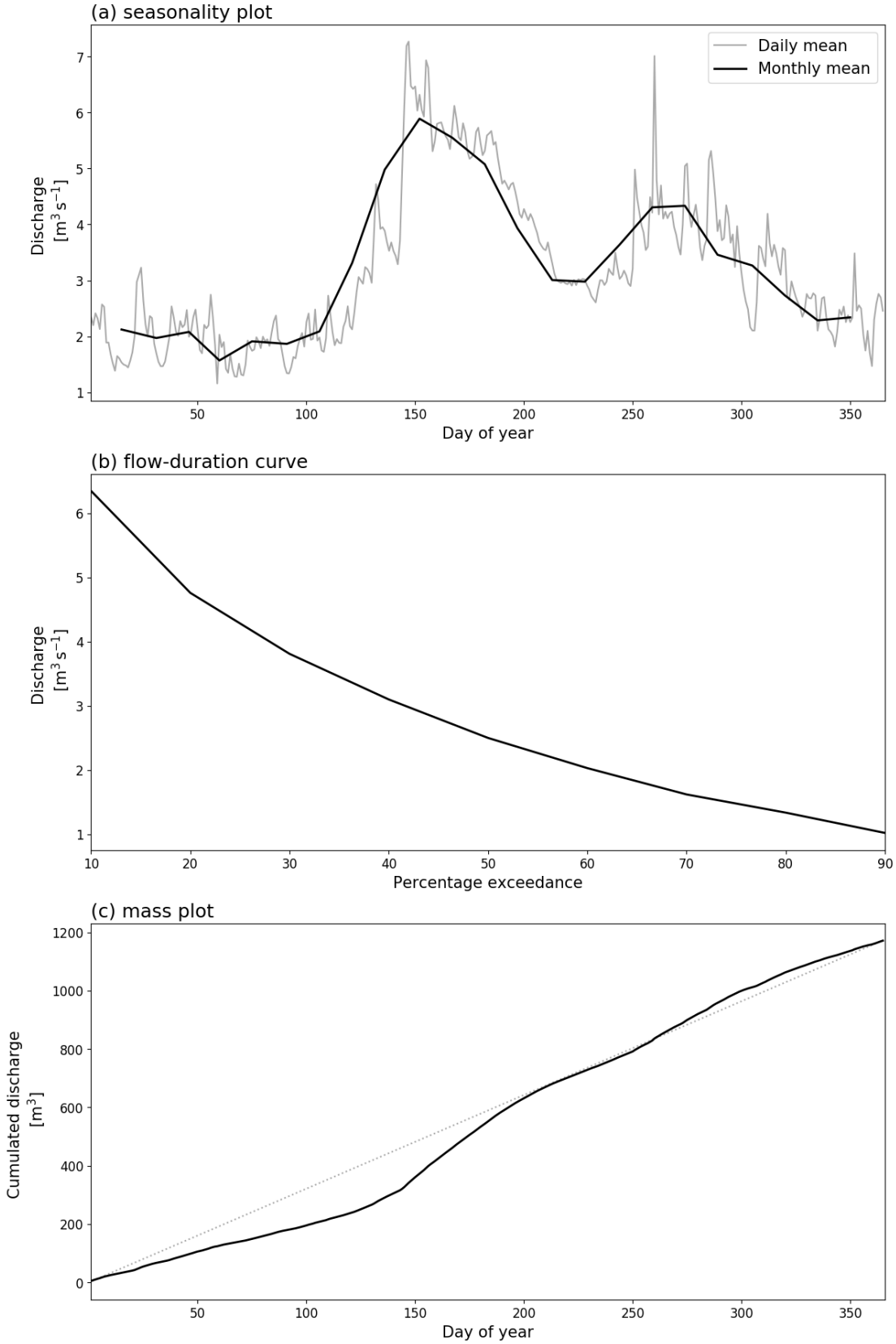

Figure 2 – Discharge timeseries for station VHM 19 as used in the cluster analysis: (a) seasonality plots, shown for both daily-averaged discharge (grey line), and monthly-averaged discharge (black line). (b) flow-duration curve with 10% exceedance steps. (c) mass curve (black line) showing daily-averaged cumulated discharge over the year, with the identity line shown with the grey-dashed line.

Table 2 – Main characteristics of the river catchments used for the cluster analysis.

VHM	Area *km²*	Aspect ratio	Longest flowpath *m a.s.l.*	Average Elevation *m a.s.l.*	Glacial cover *%*	Old bedrock *%*	Young bedrock *%*	Bedrock *%*
10	396.1	2.96	55,167	527	0	99.2	0.5	99.7
12	164.7	1.74	31,960	408	0	96.8	0	96.8
19	38.4	1.7	15,040	510	0	100	0	100
26	266.3	3.36	64,555	387	0	61.3	38.7	100
30	7313.5	2.57	247,279	702	13.22	15.5	66	81.5
38	42.8	2.39	20,971	428	0	100	0	100
43	640.7	1.73	50,958	307	0	3.1	96.9	99.9
45	458.3	2.37	58,072	547	0	67.1	32.9	100
48	701.4	1.74	74,306	543	0	48.9	51.1	100
51	299.6	1.87	34,990	723	2.96	97	0	97
59	621.9	2.9	84,104	354	0	0	98.6	98.7
60	419.9	1.92	60,336	572	2.02	0	97.4	97.4
64	5661.9	2.41	169,493	304	11.84	22	62.9	84.9
66	1574.4	2.24	123,017	650	20.3	22.7	53.7	76.4
68	201.1	1.33	35,345	245	0	6.2	93	99.2
70	1409.1	3.93	129,717	771	30	2	68	70
81	41.9	2.46	20,560	171	0	38.2	58.5	96.7
83	47.5	1.1	11,878	683	0	100	0	100
92	37.4	1.93	13,904	900	0	77.8	0	77.8
102	5097.1	2.6	189,195	538	28.64	0	71.3	71.3
110	3283	3.31	167,744	878	42.33	44.8	12.6	57.4
116	527.1	1.87	62,858	645	0	1.2	98.8	100
121	183.3	3.6	48,916	209	0	0	100	100
128	513	2.01	58,289	338	0	93.7	1.7	95.4
144	1085.2	1.69	93,866	960	12.88	55.9	28.7	84.6
148	115.1	2.47	28,963	577	0	99.8	0	99.8
149	189.4	3.27	37,033	609	4.83	91	0	91
150	225.9	3.03	45,563	767	40.23	47	12.8	59.8
162	2023.1	2.01	110,507	1,195	56.92	0	43.1	43.1
183	1627.2	2.36	133,710	249	26.31	10.9	62.1	73
185	216.8	1.42	31,153	294	0	0	100	100
198	192.9	1.31	31,543	399	0	100	0	100
200	1102.2	3.51	131,238	723	0	97.1	0.4	97.6
204	102.3	2.47	28,106	466	0	100	0	100
205	264.6	2.11	42,434	731	2	82	6	88
206	126.3	2.29	28,303	865	0	100	0	100
218	516.9	1.14	53,731	737	12.22	0	71.7	71.7
233	818.1	3.42	81,106	1,130	37.65	0	62.3	62.3
238	2163	1.54	118,032	822	4.51	26.5	68.7	95.2
271	1027	2.64	97,404	394	3.35	5	83.3	88.3
328	1496	2.99	146,594	155	28.62	5.8	64.7	70.6
400	73.2	1.41	16,634	435	0	100	0	100
408	581.3	1.13	58,363	756	49.27	0	50.7	50.7
411	387.1	3.71	73,405	559	0	97.7	2.3	100

3.2 Results

The cluster analysis was performed both based on measurements, and on simulated discharge in order to analyse the difference between clustering within both datasets. Results are presented on two dendrograms, shown on Figure 3. For both datasets, the cophenetic distances are close, with a value around 0.8. The cophenetic distance is an indicator of the correlation between distance and cophenetic matrices resulting from the cluster analysis. As it approaches the value of 1, it can be concluded that in the two analyses, data were clustered successfully.

In the figure, it is decided to keep five clusters, and vertical bars are drawn on the dendrograms at a distance value of 2.8. The following stations belong to the same cluster in both analyses, and will be referred as control stations later on:

- *Cluster A:* VHM 43, VHM 59, VHM 68, VHM 81, VHM 185, VHM 271
- *Cluster B:* VHM 19, VHM 38, VHM 51, VHM 83, VHM 92, VHM 148, VHM 149, VHM 198, VHM 200, VHM 204, VHM 205, VHM 206, VHM 400
- *Cluster C:* VHM 10, VHM 12, VHM 45, VHM 128, VHM 411
- *Cluster D:* VHM 102, VHM 110, VHM 150, VHM 162, VHM 233, VHM 408, VHM 183, VHM 70
- *Cluster E:* VHM 48, VHM 116, VHM 121, VHM 238

Overall, out of the 44 stations, 36 were clustered similarly in both analyses. Those catchments are shown on Figure 4, with a colour code for each cluster. In the rest of the study, three stations associated to river Skaftá are discarded because of the effects of jökulhlaup on the flood analysis: VHM 70 and VHM 183 (Cluster D in both analyses), and VHM 328 (Cluster A or Cluster D, depending on the cluster analysis). This leaves this study with 34 rivers used as control stations, and seven that will be studied later on.

Results from the cluster analysis indicate that the control stations were generally classified according to river types, and weather conditions. Cluster A mostly gathers spring-fed rivers, some of them originating from glacial rivers and located on the southwestern part of Iceland. In Cluster B, most of the rivers are direct runoff, influenced by snowmelt, and located in the northern half of Iceland. Cluster C is more difficult to describe and quite mixed, with rivers located in the western part of the country, sometimes controlled by small ponds and lakes. Cluster D comprises glacial rivers, and all watersheds are partially covered by glaciers. Finally, in Cluster E, rivers are mainly groundwater-fed, which accounts for a large part of the baseflow.

For further insights into the clustering process, normalised seasonality plots (similar to Figure 2.a) are shown for each cluster, based on measured (Figure 5) and simulated (Figure 6) discharge. Within each cluster, mean monthly values are drawn with a solid line, and the minimum-maximum interval is shown with shaded area. For each cluster, the seasonality plots show very distinctive trends, indicating that river with the same behaviour were successfully clustered together by the analyses. For instance, in Cluster A, maximum discharge for all stations reaches its peak during the winter months, and a low point in August. The trend is completely different for Cluster D, with maximum discharge reached at the end of the summer. This is typical of glacial rivers: melting accumulates over the summer months and slowly increases the river flow. For Cluster C and D, the trends are similar in both analyses, with two peaks being reached: one in springtime and the other one in the fall. The general pattern for stations belonging to Cluster B differ between both analyses, which might be attributed to the fact that it is the largest cluster with 13 rivers.

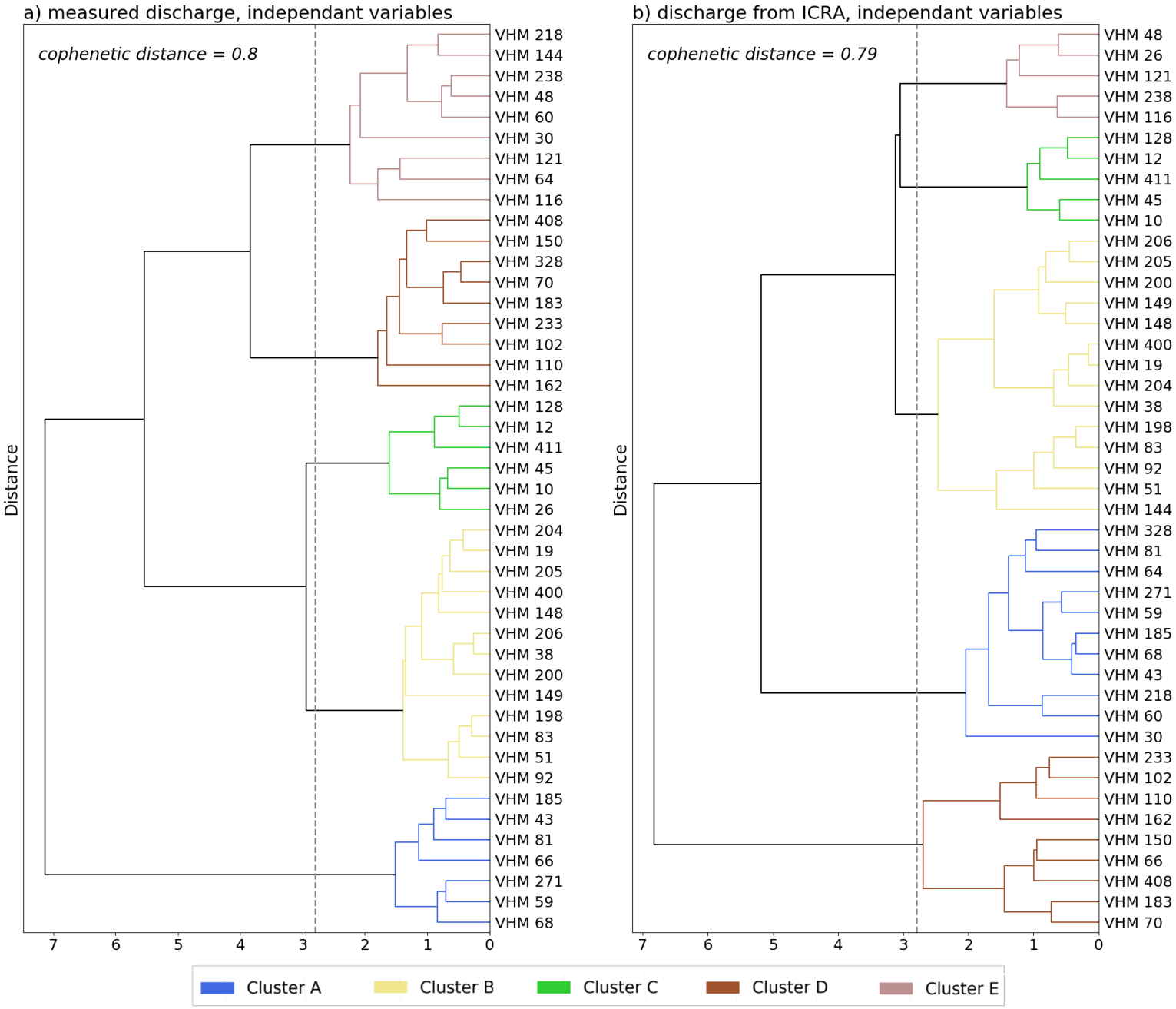

Figure 3 – Dendrograms resulting from the cluster analysis on measured (left) and simulated (right) discharge timeseries.

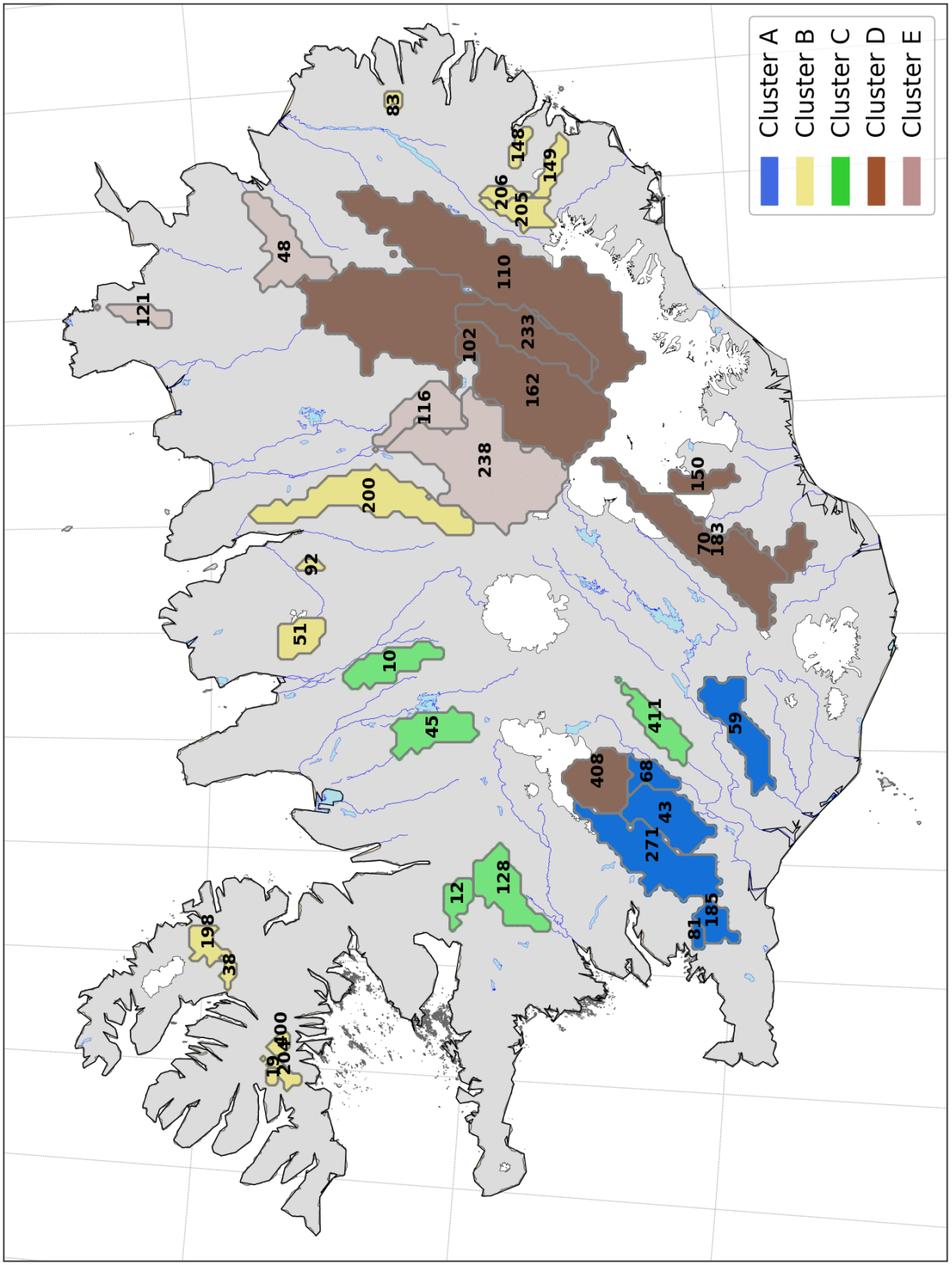

Figure 4 – Map of Iceland including catchments that clustered similarly after analysis on measured and simulated discharge.

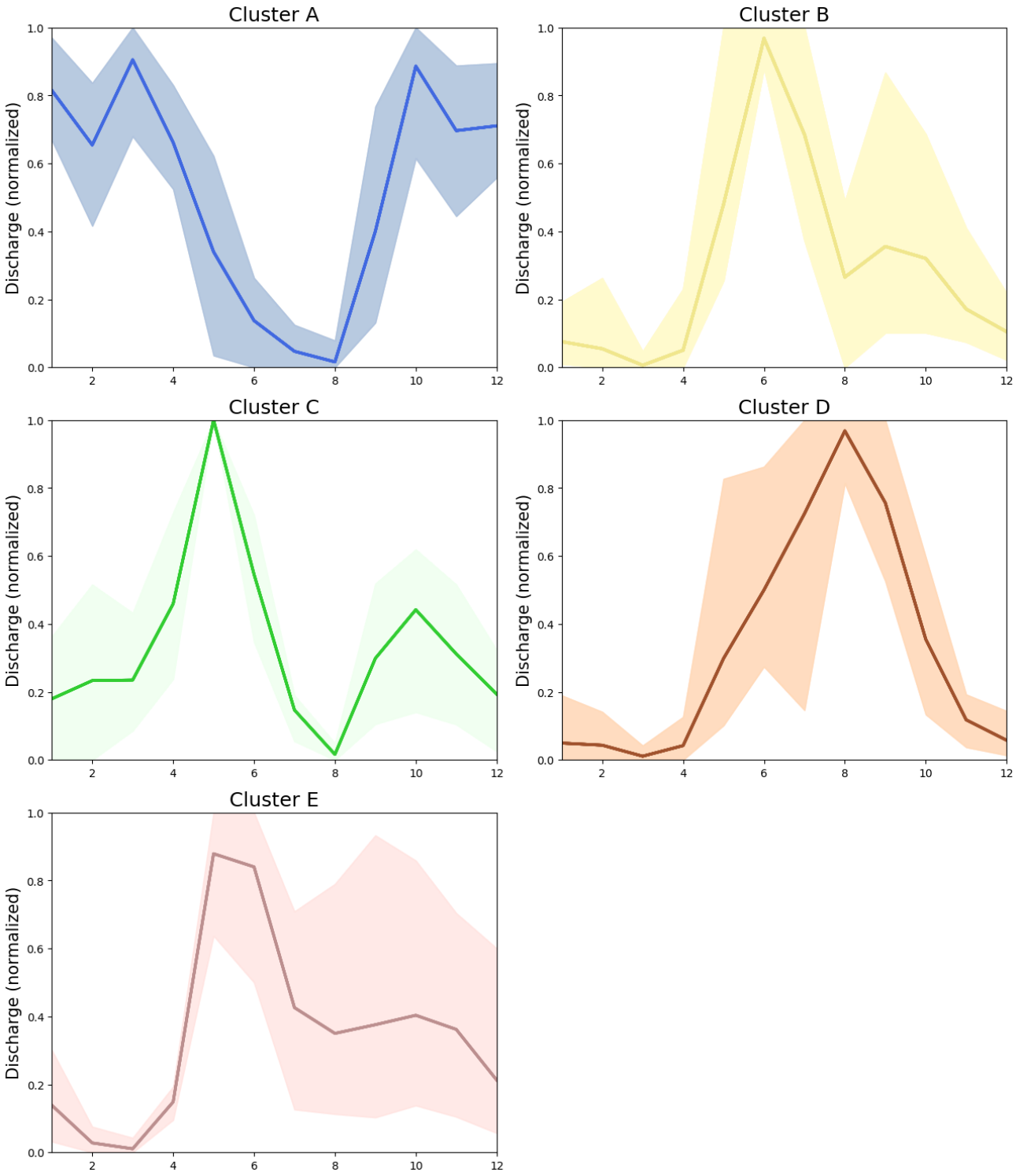

Figure 5 – Seasonality plots for each cluster, based on discharge timeseries measurement. For each cluster, mean monthly values among all the stations belonging to the same cluster are shown with the solid lines, and the minimum-maximum intervals are shown with the shaded area.

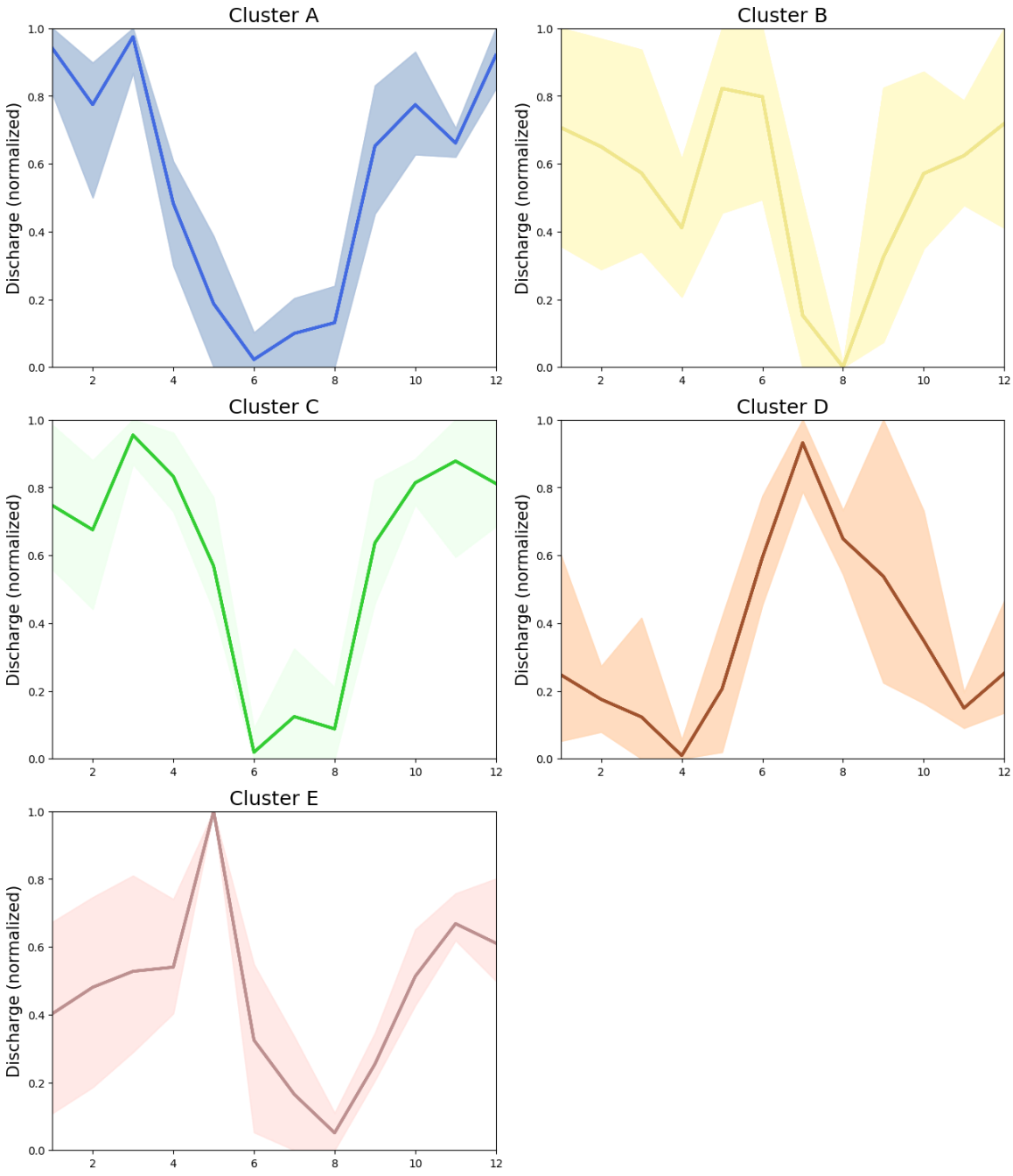

Figure 6 – Seasonality plots for each cluster, based on discharge timeseries simulation from the ICRA dataset. For each cluster, mean monthly values among all the stations belonging to the same cluster are shown with the solid lines, and the minimum-maximum intervals are shown with the shaded area.

4 Extreme Value Analysis

4.1 Methodology

Extreme Value Analysis (EVA) is a statistical discipline used to predict the occurrence of rare events by assessing their frequency from the most extreme values of a dataset. EVA allows the calculation of return levels associated with periods that can be much longer than the length of the timeseries available for the analysis. Two approaches exist: the Peak-over-Threshold method and the Block Maxima method. In this study, only the latter method is used, as in recent hydrological projects at IMO (Pagneux *et al.,* 2017, 2018 and 2019; Þórarinsdóttir *et al.*, 2021).

The Block Maxima approach consists of dividing the timeseries into non-overlapping periods of equal size and retaining only the maximum values within each period. When dealing with hydrological data, it is common to use the maximum daily values from each calendar year. A new timeseries that includes only the maxima is thus generated and referred to as an Annual Maxima Series (AMS). Under extreme value conditions, the AMS follows a General Extreme Value (GEV) family of distribution:

$$G(z) = exp\left\{\left[1 + \xi\left(\frac{z-\mu}{\sigma}\right)\right]^{-1/\xi}\right\}$$

where z is the extreme value and μ, σ and ξ are the three parameters of the GEV model $G(z)$, defining location, scale and shape parameters, respectively. Three types of GEV distribution exist, depending on the value of the shape parameter ξ. In this study, for consistency with previous work for flood analysis, a GEV distribution of type I (Gumbel) is used to fit the AMS, with ξ set to zero:

$$G(z) = exp\left\{-exp\left[-\left(\frac{z-\mu}{\sigma}\right)\right]\right\},$$

The return level r associated with the return period $1/p$ can finally be estimated with the formula:

$$r = \mu - \sigma log\{-log(1-p)\}$$

and r is defined as the value expected to be exceeded on average once every $1/p$ year.

For more details, see Coles (2001).

4.2 Comparison between simulated and observed discharge

4.2.1 Flow-duration curves

In order to compare the discharge measured at the gauges with the discharge derived from the ICRA runoff, highest daily values are ranked and plotted decreasingly for each control station, using only values above the 95th percentile. For comparison purposes, the timeseries are identical for each station and days with no observations were also discarded in the reanalysis. Examples for each cluster are shown on Figure 7, on plots that look very similar to flow-duration curves, even though the x-axis is a number of days instead of a percentage exceedance. Only 5% of the data are actually plotted since values below the 95th percentile are discarded in this comparison. Results show that the simulated discharge is much higher, although the general distribution of the flow-duration curves look very similar, as if the simulated discharge could easily be scaled by a correcting factor to match the observations. These results can be extended to most of the stations, as the simulation show an overestimation in the vast majority of cases (32 out of the 34 control stations). This is not surprising, as no infiltration is considered when converting the runoff into discharge in the simulations, leading to a larger amount of runoff pouring straight into the river flows.

More flow-duration curves are shown on Figure 8, for five other rivers that display different results. In two cases, VHM 205 and VHM 206 (Figure 8), a very good match between simulated and measured discharge can be observed. They are both very small catchments associated to direct-runoff rivers, where the assumption that all of the runoff goes straight to the river within a day is valid. For the stations VHM 68, VHM 185 and VHM 408, the simulated discharge is overestimated, but the duration curves show different trends. In the first two cases, the flow-duration curves based on measurements are more incurvated than the one based on simulations. While for VHM 408, the flow-duration curve based on measurements is almost flat which the simulated discharge fails to reproduce.

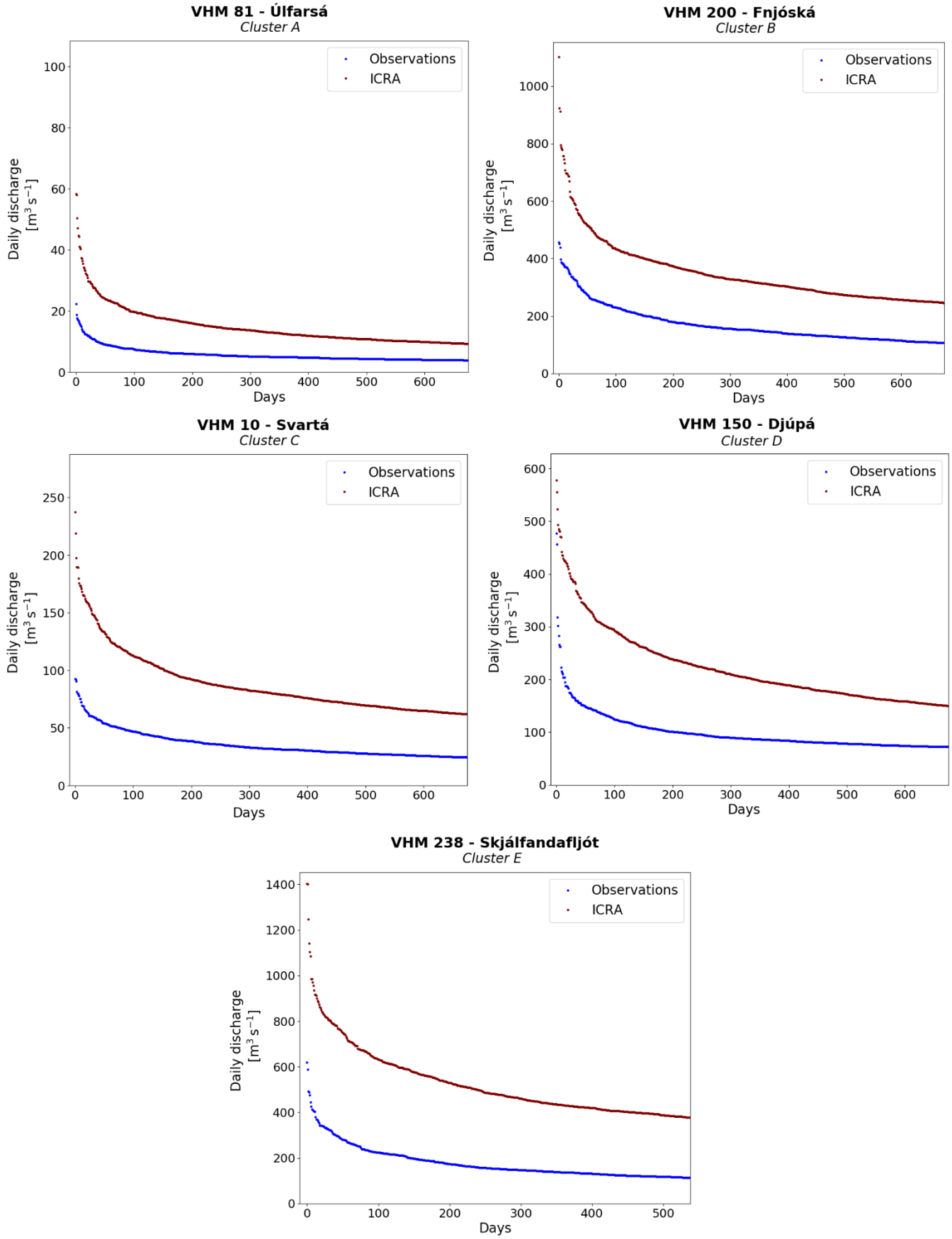

Figure 7 – Flow-duration curves for five stations including the 5% highest values comparing discharge from observations (blue lines) to simulations (brown lines).

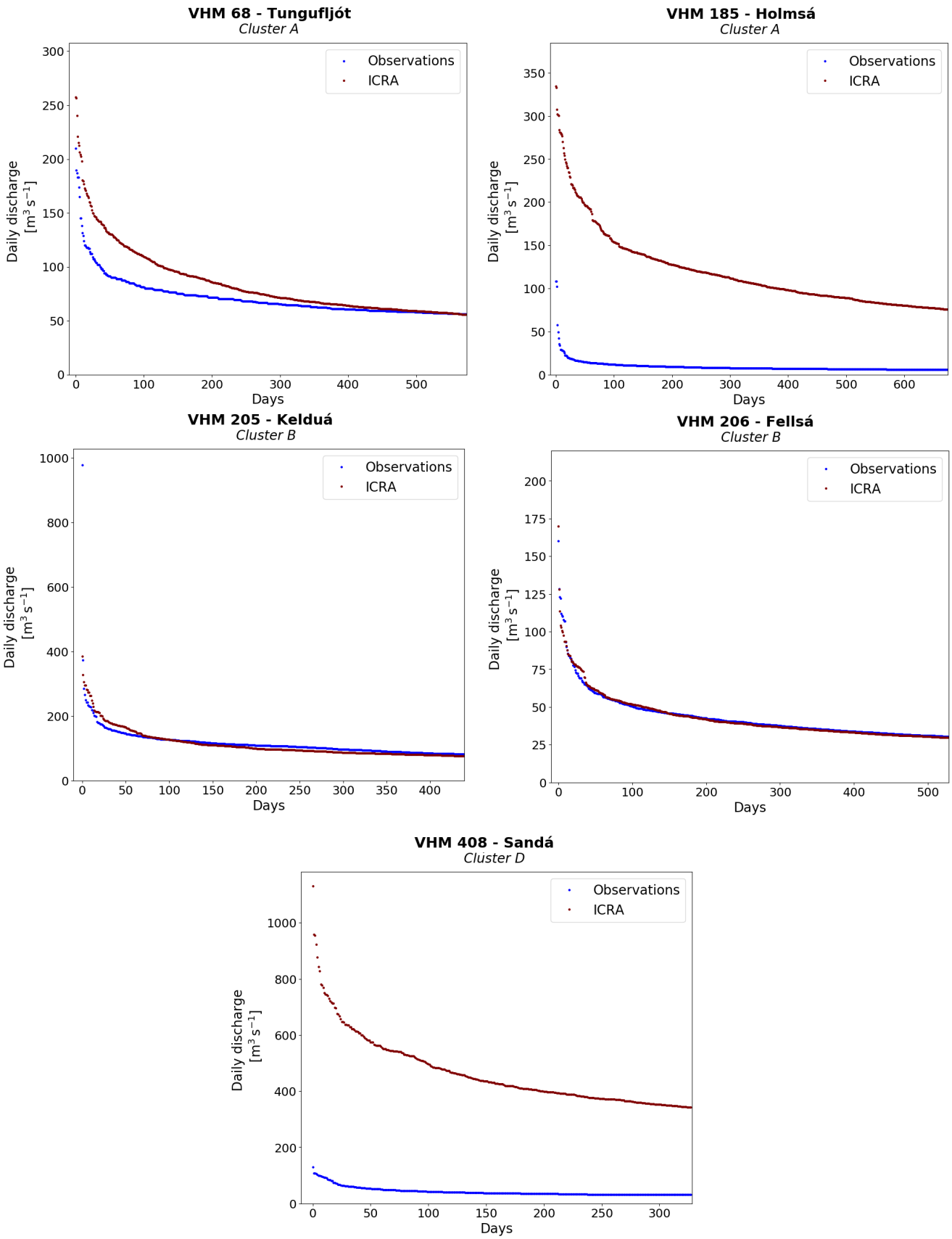

Figure 8 - Flow-duration curves for five stations including the 5% highest values comparing discharge from observations (blue lines) to simulations (brown lines).

4.2.2 Correction of the simulated discharge

In order to correct the discharge overestimation noted in most simulated flow-duration curves of the 5% highest values (Figure 7 and Figure 8), the difference between simulated and observed discharge is assessed using a coefficient of proportionality. This coefficient will then be used as a correcting factor to improve the match between measured and simulated discharge.

For each station, the mean coefficient of proportionality is calculated by comparing daily ranked discharge above the 95th, 90th and 75th percentiles. For the different percentiles, each coefficient is then averaged over all the stations that cluster together. Results are shown in Figure 9, in histograms, each cluster is represented by a panel on the figure. Individual stations are shown for each cluster, with a colour according to the amount of data the correcting coefficient was calculated on: light orange when averaged over the 5% highest discharge data (above 95th percentile), dark orange when averaged over the 10% highest data (above 90th percentile), and brown when averaged over the 25% highest data (above 75th percentile). The dashed lines show the average coefficient of proportionality for each cluster, with the colour matching the percentile the correction was based on. A coefficient of proportionality equal to 1 means there is no difference between simulated and observed discharge. Above 1, the mean observed values are higher than the simulations; under 1, the mean simulated values are higher than the measurements.

Results do not vary much between the coefficients based on the 95th, 90th and 75th percentiles. Stations within Cluster C and E show very similar results with their nearest-neighbours, with the coefficients varying within a 0.15 range among each station of those clusters. In the other clusters, results are overall good, although a few stations stand out: VHM 68 and 185 in Cluster A, VHM 205 and 206 in Cluster B, and VHM 233 and 408 in Cluster D, VHM 233 and VHM 408. As expected from Figure 8, the coefficients of proportionality for stations VHM 205 and 206 are close to 1, indicating the good match between the two datasets. Both stations VHM 185 in Cluster A and VHM 408 in Cluster D, show very low coefficients, around 0.1, which is expected as the trends of the simulated and observed flow-duration curves were very different (see chapter 4.2.1).

Results of the calculated average among the coefficient of proportionality (Figure 9, dashed lines) are extremely close , with variation within a 0.1 range between all percentiles, except in Cluster A. Stations belonging to Cluster B have a relatively high mean coefficient of proportionality (0.58 for the 95th and 75th percentiles, 0.59 for the 90th percentile), which can be attributed to the presence of VHM 205 and 206 among its members. The largest difference between observed and simulated discharge values can be found in Cluster E, where the average coefficient of proportionality is around 0.35.

In order to adjust the simulated high discharge for each station to better fit the observed one, daily discharge values calculated from the ICRA runoff are multiplied by the mean coefficient of proportionality from the belonging cluster. For instance, to obtain the light orange curves on the figure, simulated data were multiplied by 0.36 for the stations belonging to Cluster A, 0.58 for stations from Cluster B, and so on. Same for the orange and red lines, where data were multiplied by the mean coefficients calculated from 90th and 75th percentiles, respectively.

Results are shown on Figure 10 for the same stations that can be seen on Figure 7, one from each cluster. Daily values are represented in blue for the observed discharge, in brown for the simulated discharge before correction, in light orange after correction based on data above the 95th percentile, in orange after correction based on data above the 90th percentile, and in red after correction based on data above the 75th percentile. For all the stations shown, results are significantly improved after applying the corrections. In the figure, results based on the different corrections do not vary much, with the corrected lines almost overlaying one another. Except for Cluster C, corrections derived from data above the 95th percentile give closer discharge to the measurements. For station VHM 200, only the red line is visible, as the correcting factors

are approximately the same, no matter if they were based on the 75th, 90th or 95th percentiles. Overall, the correcting factors improve the results from the reanalysis, making the highest simulated discharge values more in line with the measurements.

Results for the rivers that showed the largest difference of coefficients of proportionality compared to the rest of stations within their cluster, are presented in Figure 11. The same stations were presented in Figure 8 and the results are not as successful for those five stations. For stations VHM 68, VHM 205 and VHM 206, the corrected timeseries are now too low. For station VHM 185, after corrections, simulated timeseries still overestimate the measurements. However, the two highest measured discharge values (above 100 $m^3 s^{-1}$) are better simulated than before applying the corrections. For station VHM 408, the corrected simulated discharge is quite far from the observed one, although it is closer than the uncorrected one.

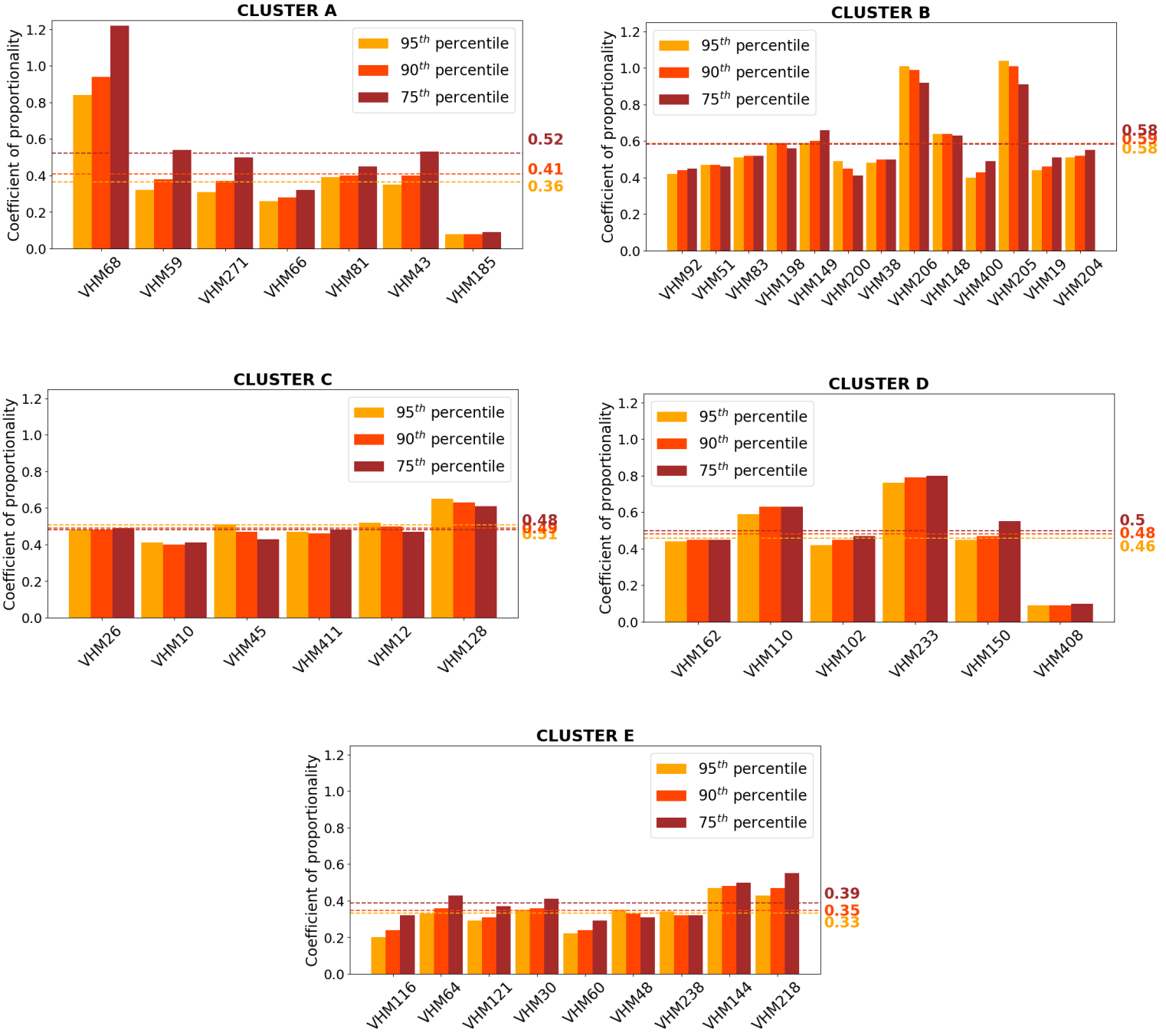

Figure 9 – Histograms showing coefficients of proportionality for each control stations and based on values above the 75th (brown), 90th (orange) and 95th (light orange) percentiles. Stations are shown by cluster, and mean coefficients averaged among all stations are represented by the dashed lines.

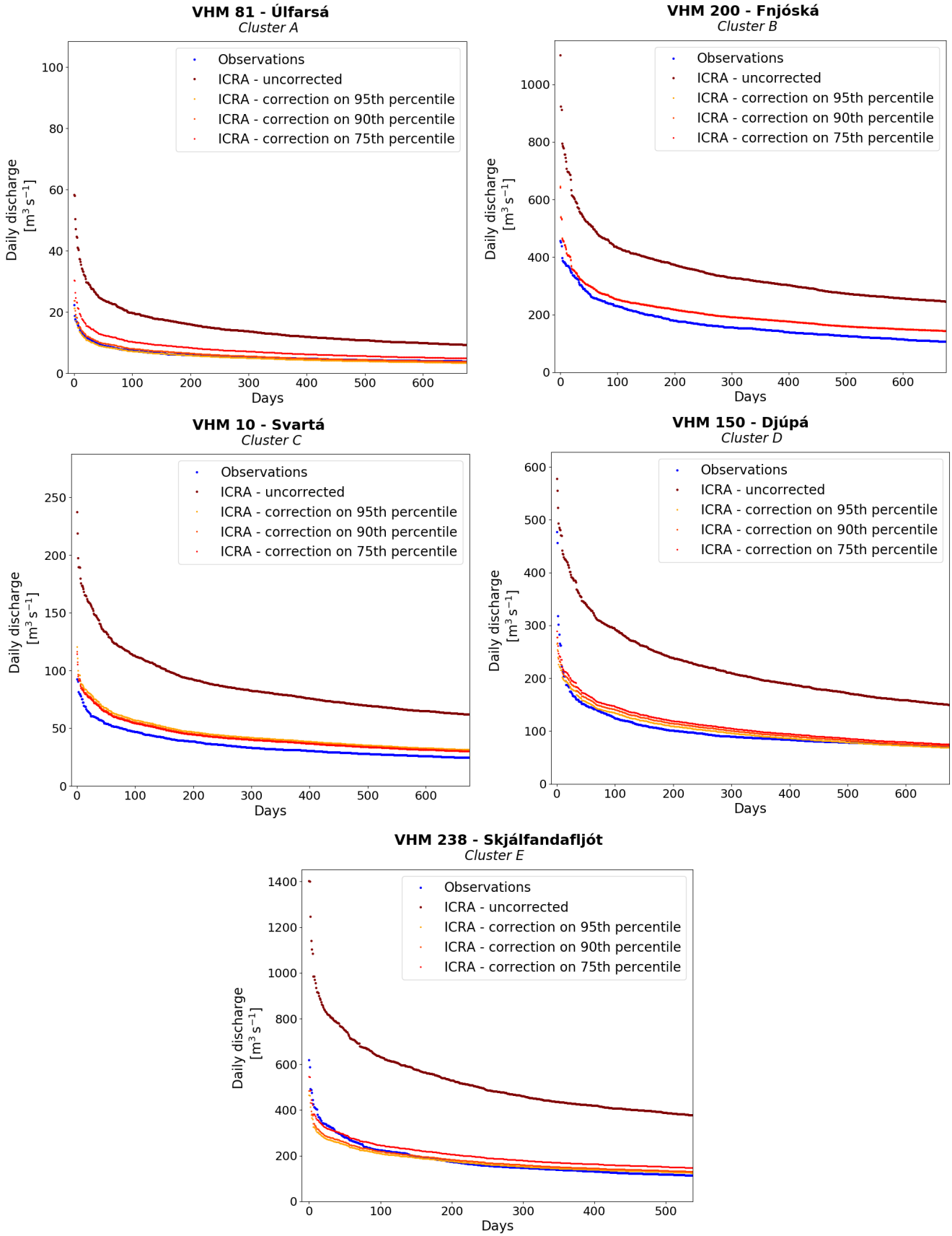

Figure 10 - Flow-duration curves for five stations including the 5% highest values. Discharge values are based on observations (blue), and based on the ICRA dataset before (brown) and after applying the coefficients of proportionality (yellow, orange, and red).

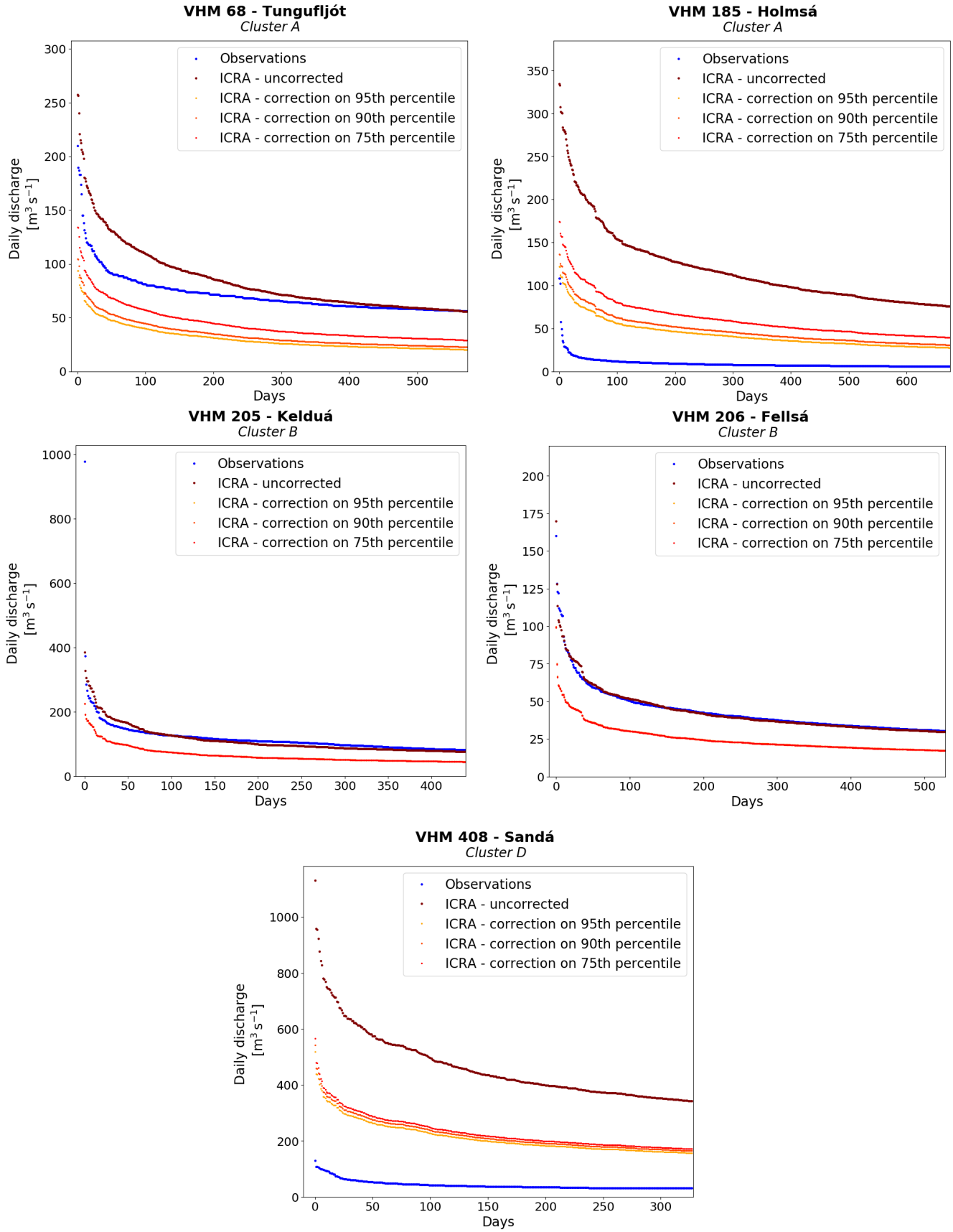

Figure 11 - Flow-duration curves for five stations including the 5% highest values. Discharge values are based on observations (blue), and based on the ICRA dataset before (brown) and after applying the coefficients of proportionality (yellow, orange, and red).

4.3 Flood analysis for control stations

4.3.1 Flood return levels

In the rest of this study, only mean corrections based on data above the 95th percentiles are used. The Block Maxima method is applied to all control stations based on measured discharge and simulated timeseries derived from the ICRA dataset, before and after applying the correction factor.

Return levels are calculated based on observations, and simulations before and after correction. The correction coefficient depends on the cluster in question, and is applied to the timeseries before the EVA. Daily flood return levels with a 10-, 25-, 50-, 100-, 200-, and 500-year return period are calculated for all stations. Three examples are shown in Table 3, for stations VHM 200, VHM 110 and VHM 206. Here, stations VHM 200 and 206 belong to Cluster B, and the ICRA simulations are multiplied by the correction factor 0.58; while for VHM 110 (Cluster D), the simulations are multiplied by 0.46. For stations VHM 200 and VHM 110, results before correction show a large overestimation of the return levels compared to the measurements, with simulated values being more than twice the observed discharge for all return periods. For instance, the 200-year return level calculated from observations is 671 m^3 s^{-1}, and 1299 m^3 s^{-1} based on the simulations. After applying the correction, results based on the ICRA dataset are considerably lowered, and very close to the ones based on observations. After correction, the 200-year return level for station VHM 200 is equal to 758 m^3 s^{-1}.

These results are further illustrated on the return level plots presented in Figure 12 and 13. In those figures, discharge values are plotted against the return periods on a logarithmic scale. Here, values from the measured AMS between years 1980 and 2016 are represented by the blue dots. A straight line shows the fit between those data and return periods and horizontal dashed lines indicates the values for the 25-year flood. The same is done for discharge derived from the ICRA dataset on the top plots in red, and for simulated discharge after correction on the lower plot in orange. In both cases, results are significantly improved by the scaling, with simulated AMS much closer to the observation AMS after applying the correction. For instance, the 25-year return levels percentage difference between simulation and observation for station VHM 200 is 70.1 % before correction, and only 19.4 % after correction. For VHM 110, it decreases from 82.3% before correction to 9.5%.

A counterexample is given for VHM 206 in Table 3 (bottom) and Figure 14. As explained before, this station showed a very good match between measured and simulated discharge right away, presumably because of the small size of the catchment and the fact that it is a direct-runoff river, with minimal infiltration. As expected, the return level estimation suffers from applying the correcting factor, as is illustrated on the return level plot. Before correction, the 25-year flood analysis only gives a 2 m^3 s^{-1} difference between simulated and measured discharge, showing a very close fit between the two datasets. Once the analysis is redone after multiplying the simulated timeseries by the correction factor, the updated return level drops to 84 m^3 s^{-1} w, indicating a percentage difference of 54.5%, much larger than before the correction.

Table 3 – Return levels ($m^3 s^{-1}$) for stations VHM 200 (top), VHM 110 (middle), and VHM 206 (bottom). Results are based on the measured discharge, simulated discharge from the ICRA runoff, and simulated discharge from the ICRA runoff after correction based on the values above the 95th percentile. Values are given for a 10-, 25-, 50-, 100-, 200-, and 500-year return period.

VHM 200 - Fnjóská

Return-period	Return levels ($m^3 s^{-1}$)		
	Observations	**ICRA**	**ICRA, corrected**
10 years	392	859	502
25 years	479	996	582
50 years	543	1098	641
100 years	607	1199	700
200 years	671	1299	758
500 years	755	1432	836

VHM 110 – Jökulsá á Dal

Return-period	Return levels ($m^3 s^{-1}$)		
	Observations	**ICRA**	**ICRA, corrected**
10 years	870	2125	974
25 years	1026	2461	1128
50 years	1142	2711	1242
100 years	1257	2958	1356
200 years	1371	3205	1469
500 years	1522	3530	1618

VHM 206 - Fellsá

Return-period	Return levels ($m^3 s^{-1}$)		
	Observations	**ICRA**	**ICRA, corrected**
10 years	120	122	71
25 years	147	145	84
50 years	166	161	94
100 years	186	178	104
200 years	206	195	114
500 years	232	217	127

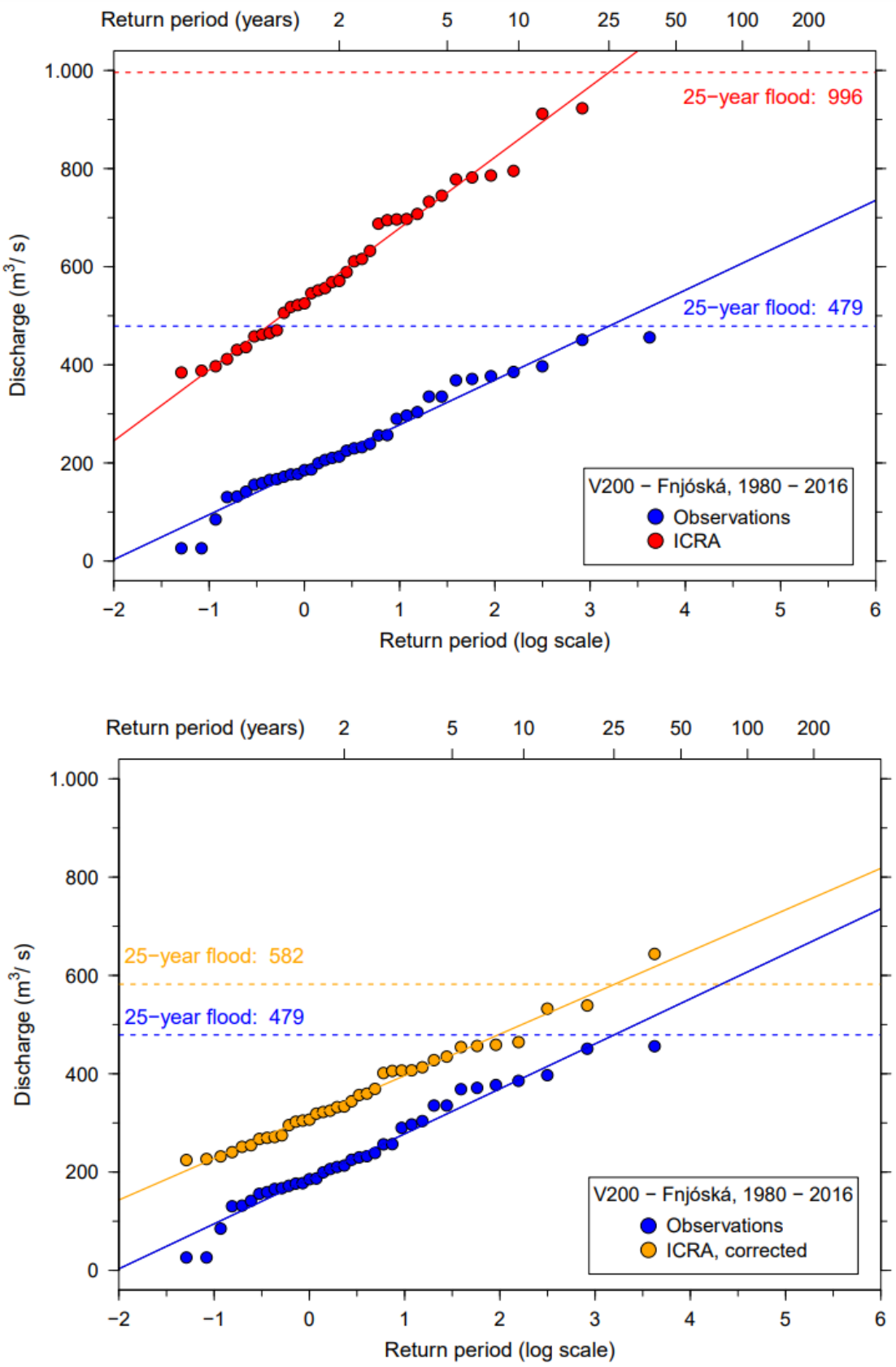

Figure 12 – Return level plot for station VHM 200, based on observations (blue), simulations before (red) and after correction (orange). Dashed-lines show the 25-year return level for the different dataset.

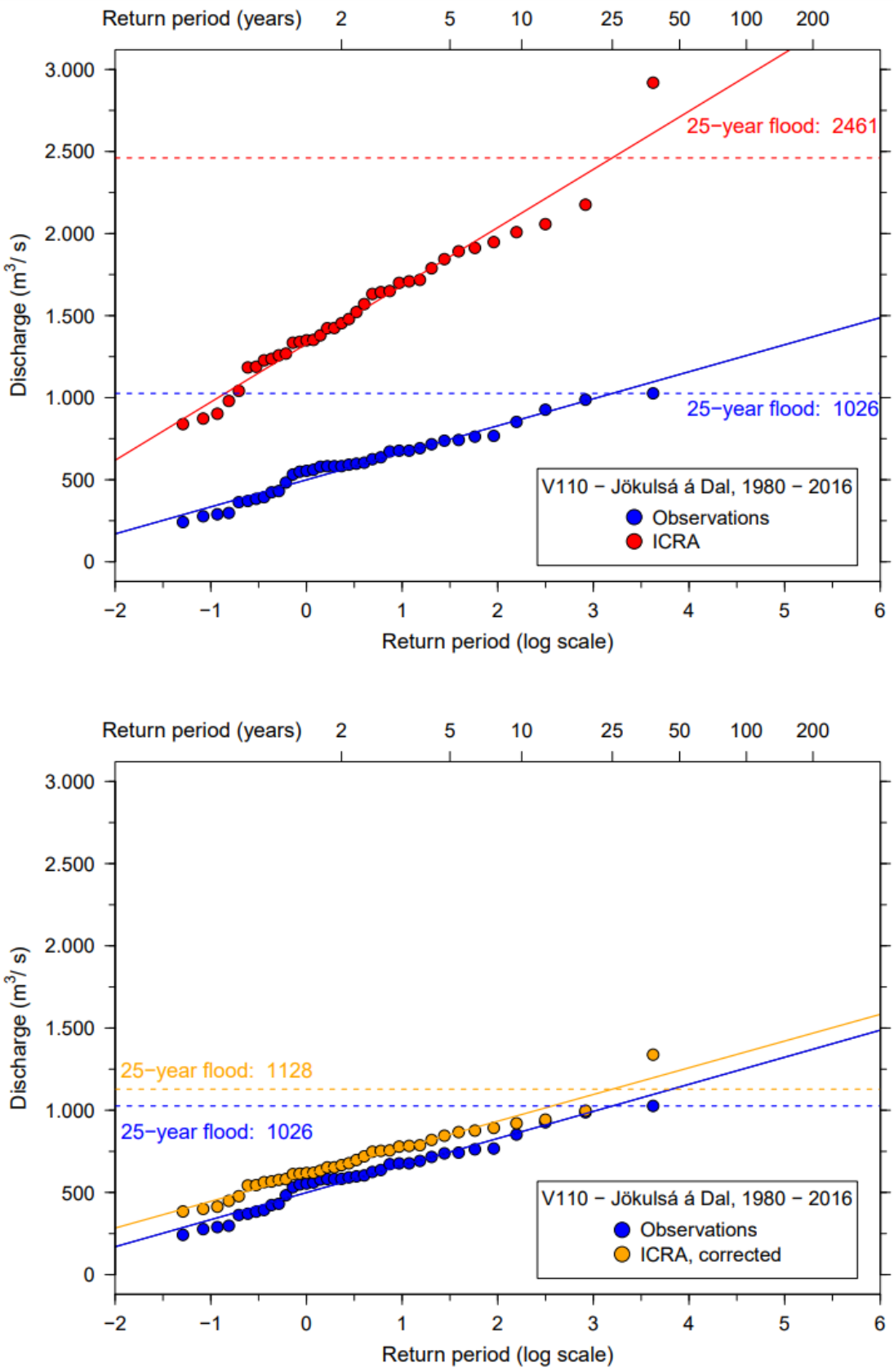

Figure 13 – Return level plot for station VHM 110, based on observations (blue), simulations before (red) and after correction (orange). Dashed-lines show the 25-year return level for the different dataset.

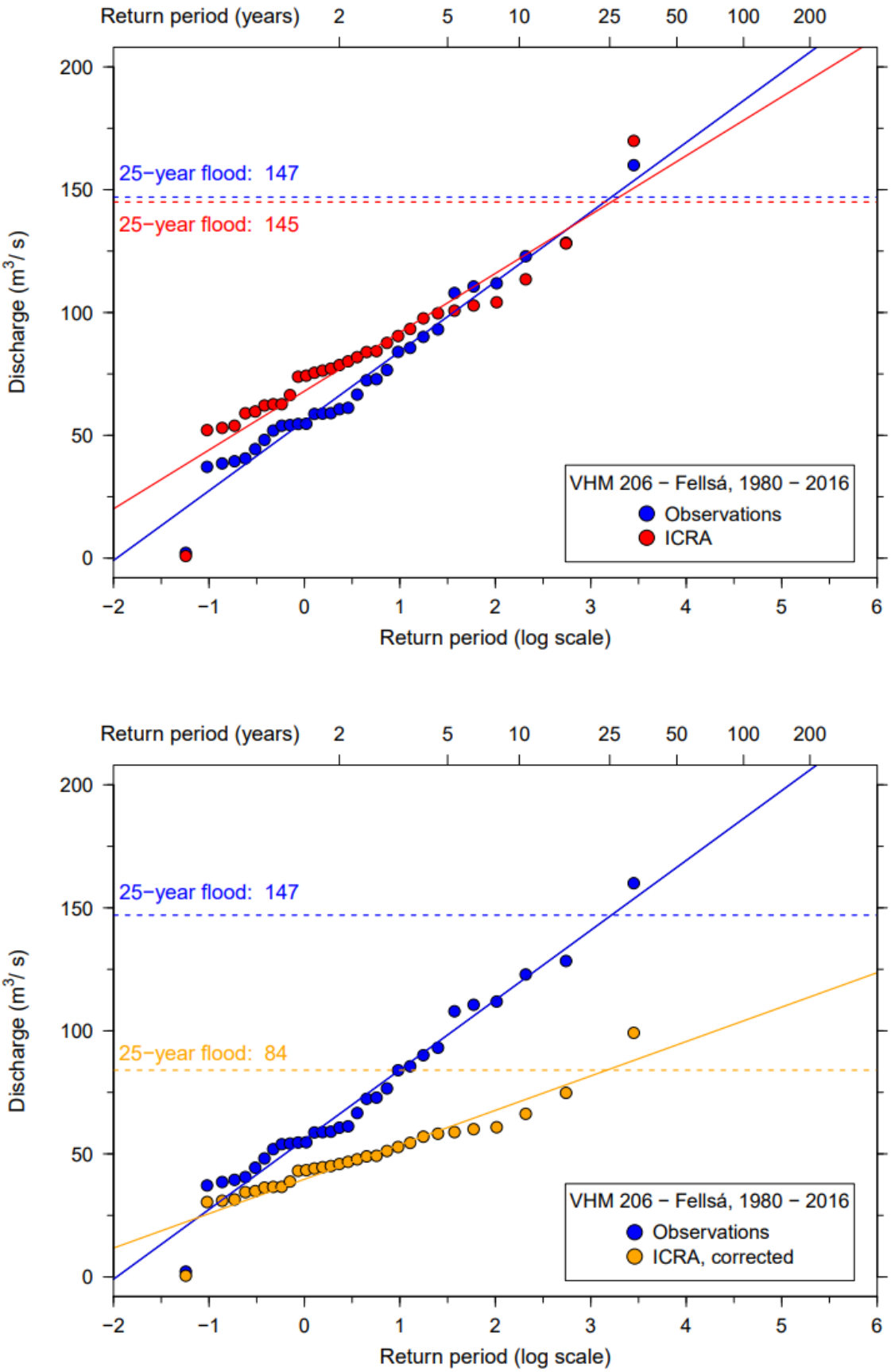

Figure 14 – Return level plot for station VHM 206, based on observations (blue), simulations before (red) and after correction (orange). Dashed-lines show the 25-year return level for the different dataset.

4.3.2 Closeness Coefficient values

To simplify the comparison between observed and simulated extreme discharge values, a closeness coefficient (*CC*) is introduced to determine how well the simulated values match the measurements:

$$CC = \frac{\min(obs, sim)}{\max(obs, sim)} \times 100$$

This coefficient quantifies simply how close the simulated value is to the observed one, independently of whether the value is higher or lower than the observation. In that sense, *CC* can be used as a percentage match between two values of a same event.

Coefficients were calculated for all the control stations for the 25-, and 200-year return levels, before and after applying the correction. Results are shown in map form, on Figure 15 and 16. On the maps, the *CC* are given a colour code according to their values: if the percentage match is below 50%, the value appears on a red circle. If the percentage match is between 50 and 75%, the circle is orange; and if it is higher than 75%, the circle is green. As expected, both for the 25- and the 200-year floods, the values are greatly improved by scaling the simulated timeseries. For the 25-year flood, before applying the correction, 20 *CC* values are inferior to 50%, 7 stations between 50 and 75%, and 7 stations are superior to 75%. After applying the correction, only 7 stations have a *CC* inferior to 50%, 15 stations have a *CC* between 50 and 75%, and 12 stations have a *CC* superior to 75%. Maps on Figure 16 show similar results for the 200-year flood. Before correction: 17 stations under 50% (7 after correction), 11 stations between 50 and 75% (16 after correction), and 6 stations above 75% (11 after correction).

Those results are also presented in Table 4, ordered by cluster. Scaling down the ICRA reanalysis leads to closer results between observation and simulation in 27 cases out of 34. Mean *CC* values indicate that overall, all the clusters benefit from the correction. In Cluster A and C, two stations (VHM 68 for Cluster A, VHM 233 for Cluster D) have a lower *CC* after correction. In Cluster B, this concerns 3 stations (VHM 205, 206 and 148); 2 stations for Cluster C (VHM 411 and 128); none for Cluster E. *CC* values were also ranked decreasingly before and after correction for the 25-year return level on histograms, shown on Figure 17. Before correcting the simulated discharge, only 1 station had a *CC* above 90% for the 25-year flood, while after correction, 6 stations have a *CC* above 90%.

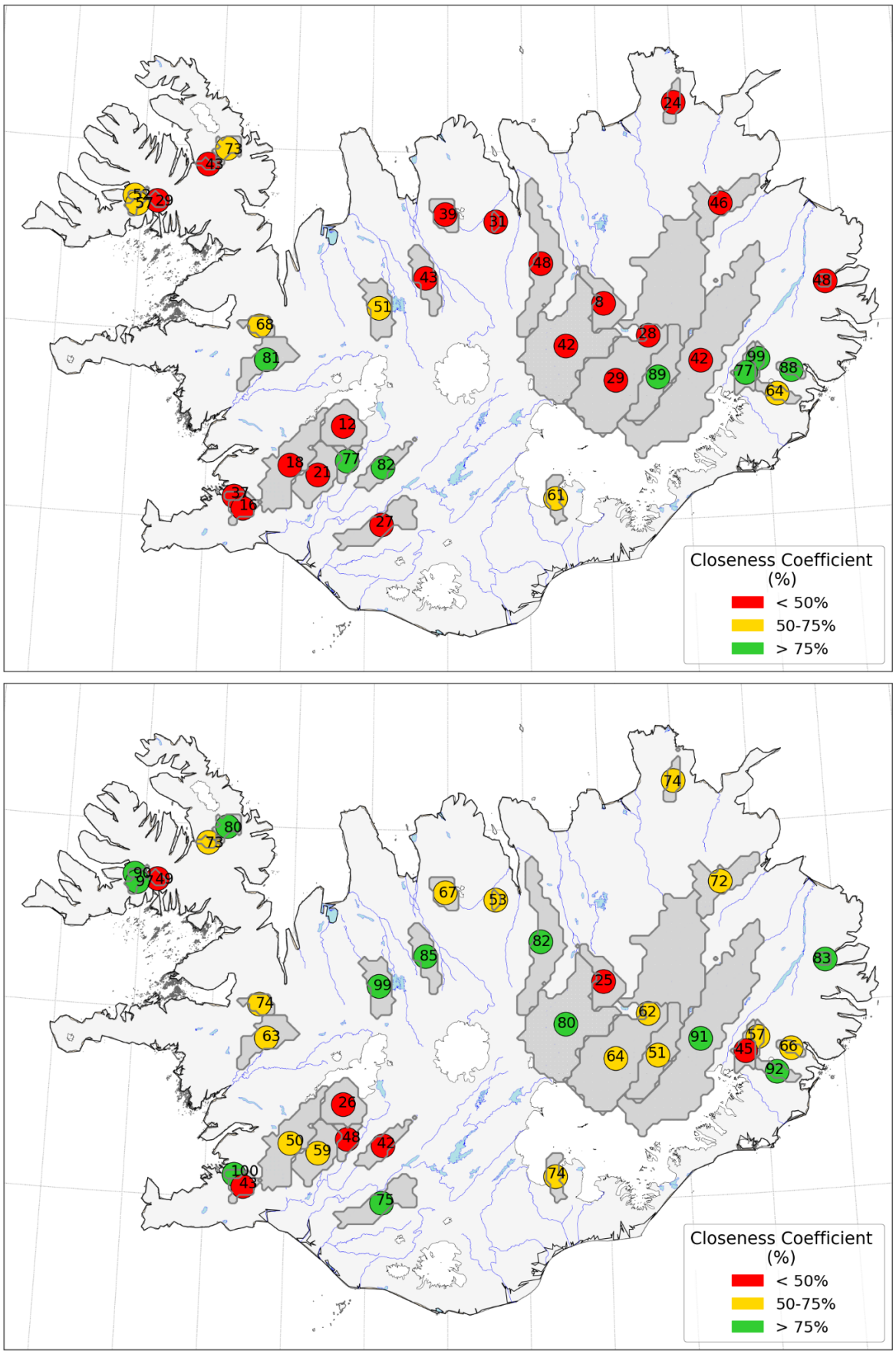

Figure 15 – Closeness Coefficient map comparing 25-year flood return level between observation and ICRA before (top) and after (bottom) applying the correcting factor.

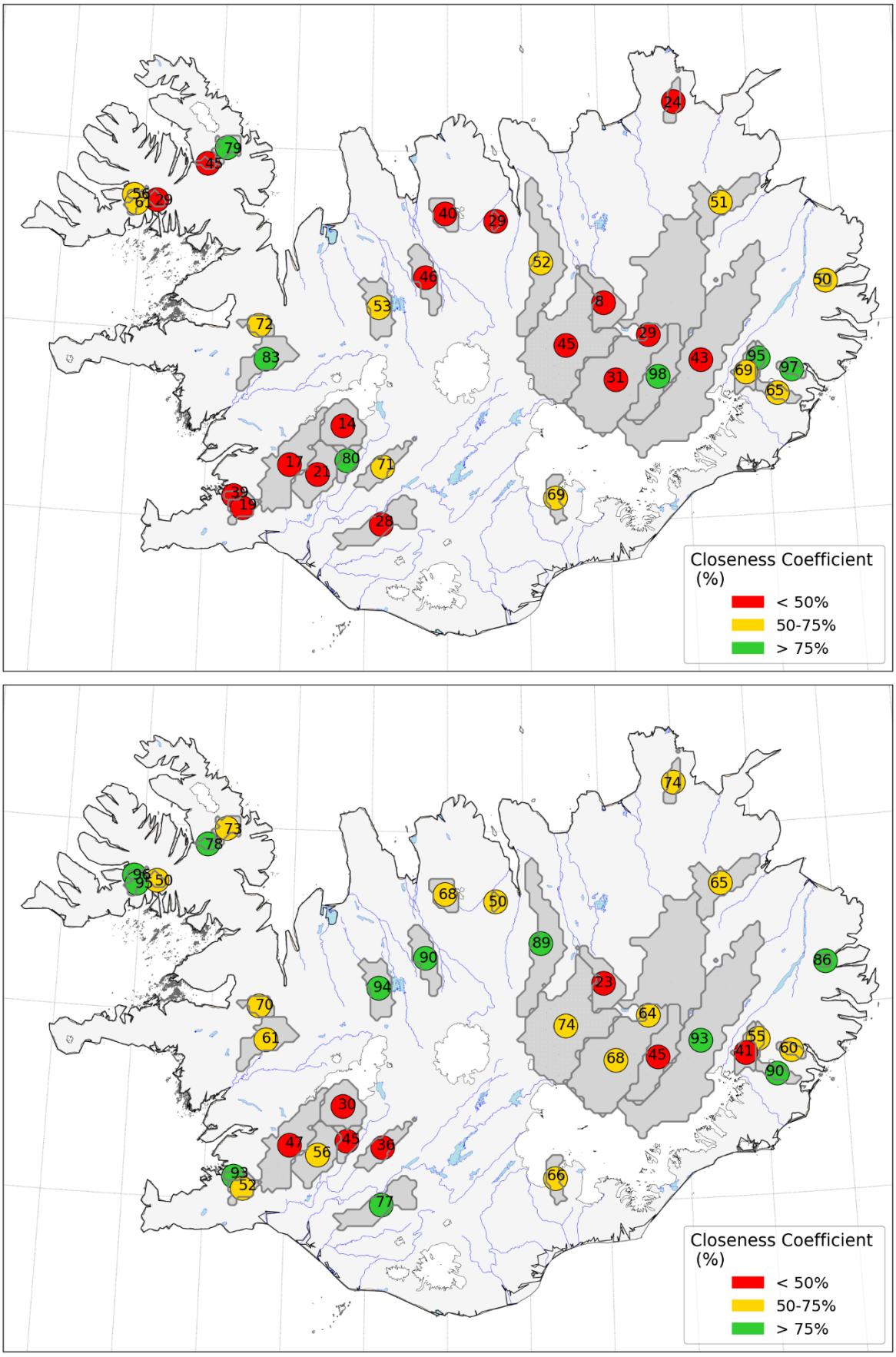

Figure 16 - Closeness Coefficient map comparing 200-year flood return level between observation and ICRA before (top) and after (bottom) applying the correcting factor.

Table 4 – Closeness Coefficient values between measurements and simulated discharge, before and after correction. Results are shown for the 25-, and 200-year return periods for all control stations. Mean CC values are given in bold for each cluster.

		Closeness Coefficient *(%)*			
		25-year return period		**200-year return period**	
		ICRA	**ICRA, corrected**	**ICRA**	**ICRA, corrected**
Cluster A	**VHM 68**	77	48	80	45
	VHM 59	27	75	28	77
	VHM 271	18	50	17	47
	VHM 81	37	100	39	93
	VHM 43	21	59	21	56
	VHM 185	16	43	19	52
	Mean CC	**33**	**63**	**34**	**62**
Cluster B	**VHM 92**	31	53	29	50
	VHM 51	39	67	40	68
	VHM 83	48	83	50	86
	VHM 198	73	80	79	73
	VHM 149	64	92	65	90
	VHM 200	48	82	52	89
	VHM 38	43	73	45	78
	VHM 206	99	57	95	55
	VHM 148	88	66	97	60
	VHM 400	29	49	29	50
	VHM 205	77	45	69	41
	VHM 19	52	90	56	96
	VHM 204	57	97	61	95
	Mean CC	**58**	**72**	**59**	**72**
Cluster C	**VHM 10**	43	85	46	90
	VHM 45	51	99	53	94
	VHM 12	68	74	72	70
	VHM 411	82	42	71	36
	VHM 128	81	63	83	61
	Mean CC	**65**	**73**	**65**	**70**
Cluster D	**VHM 162**	29	64	31	68
	VHM 110	42	91	43	93
	VHM 102	28	62	29	64
	VHM 233	89	51	98	45
	VHM 150	61	74	69	66
	VHM 408	12	26	14	30
	Mean CC	**44**	**74**	**57**	**73**
Cluster E	**VHM 116**	8	25	8	23
	VHM 121	24	74	24	74
	VHM 48	46	72	51	65
	VHM 238	42	80	45	74
	Mean CC	**30**	**63**	**32**	**59**

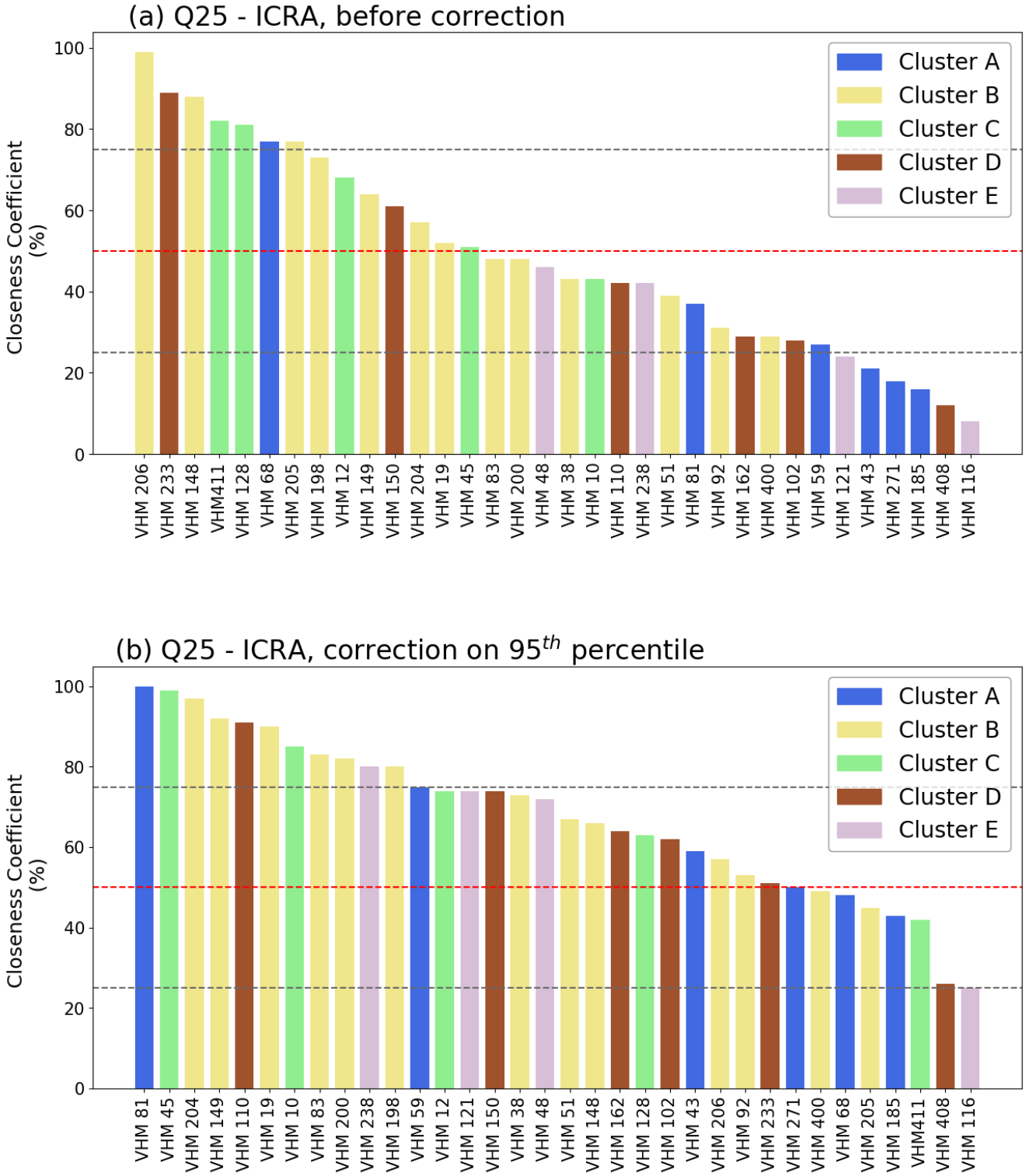

Figure 17 – Histograms presenting ranked CC values for the 25-year flood between observed and simulated discharge before (top) and after (bottom) correction. Horizontal dashed lines show the 25, 50 and 75 % CC thresholds.

4.4 Flood analysis for other stations

In the last part of this project the methodology is tested on different watersheds. In section 3, stations were classified with a cluster analysis and results were shown on two dendrograms: one based on the measured discharge, the other based on simulations derived from the ICRA runoff (REF). Out of the 44 stations used for the analysis: three were discarded because of the presence of Skaftárhlaup, 34 are clustered similarly on both dendrograms, and seven belong to different clusters.

As a test, it is decided to treat those seven rivers according to their clustering based on the simulations, as would be done if they were ungauged. This will presumably lead to larger differences between return levels based on measured and simulated datasets, than if the cluster analysis based on measurements was used.

According to the cluster analysis based on the ICRA dataset, the seven stations belong to the following clusters:

- Cluster A: VHM 30, VHM 60, VHM 64, VHM 218
- Cluster B: VHM 144
- Cluster D: VHM 66
- Cluster E: VHM 26

Results for the 25-year flood return level are presented for the seven stations in Table 5 and *CC* values comparing simulation to observations are shown on Figure 18. Before scaling down the discharge timeseries, *CC* values for all stations appear in red, indicating lower than 50% match between simulated and observed discharge, ranging from 21% (VHM 60 – Eystri-Rangá) to 42% (VHM 26 – Sandá). After applying the correction, flood return levels based on measurements and simulations show a better match: six stations have a *CC* value in the range 50-75%, and one station (VHM 30 – Þjórsá) has a *CC* value of 83%, appearing in green on the map. Therefore, all those stations benefit from being scaled by a correcting factor.

Table 5 – 25-year flood return level for 7 stations based on the ICRA dataset before and after scaling by the correcting factor.

		25-year return level *(m^3 s^{-1})*		
		Observations	**ICRA**	**ICRA, corrected**
Cluster A	**VHM 30**	1981	5844	2239
	VHM 60	118	567	216
	VHM 64	1599	5784	2208
	VHM 218	215	756	288
Cluster B	**VHM 144**	252	686	401
Cluster D	**VHM 66**	410	1652	757
Cluster E	**VHM 26**	126	302	89

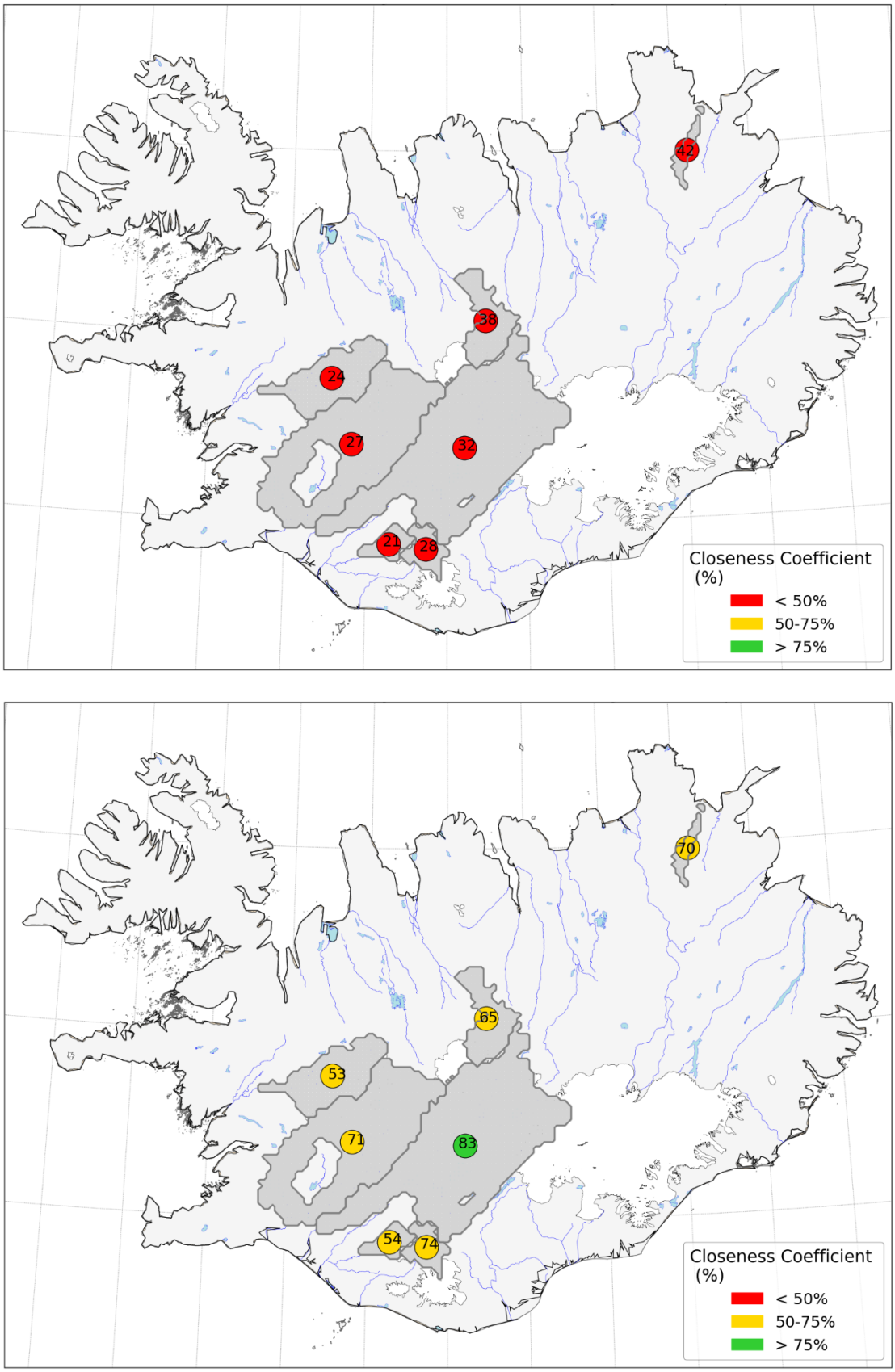

Figure 18 – Closeness Coefficient map comparing 25-year flood return level between observation and ICRA before (top) and after (bottom) applying the correcting factor, for seven stations.

5 Discussion

The goal of this study was to investigate if the ICRA runoff can be used to estimate flood extremes. This proved quite successful. Simulated runoff was accumulated on a daily-basis, which made comparison with daily-averaged observed discharge at several gauging stations, possible.

Commonly used for flood analysis is the instantaneous daily peak discharge. This has been applied before in Iceland (Crochet, 2012; Hróðmarsson *et al*., 2009, 2018) and extracting the instantaneous peak value from daily discharge has also been the subject of hydrological research over the years (Fill and Steiner, 2003; Chen *et al*., 2017). Flood analysis based on instantaneous data is of major interest: return levels reach much higher values, and such flood estimates are therefore more likely to give more realistic view of possible effects of flood extremes. Extracting the instantaneous peak from daily discharge has not been treated here and should be investigated in future works.

In section 3, a hierarchical cluster analysis was used to determine which rivers showed the most similarities and grouped together. Five clusters were defined from the dendrograms, which then led to a group of 32 control stations (after discarding jökulhlaup-affected rivers), that clustered similarly in both analyses. The coefficients later used to correct the simulated timeseries were cluster-dependent, and a different number of clusters would have changed the values of the averaged coefficients. Even though the coefficients of proportionality did not oscillate much between the individual stations, a few stations stood out (such as VHM 205 and 206) and a different interpretation of the clustering could be investigated to refine the values of the corrections applied.

For comparison purposes, the flood analysis presented in section 4 was based on observed and simulated datasets that were identical for each station, and days with no observations were also discarded in the reanalysis. Therefore, the AMS used for the calculation of the return levels was for some stations shorter than the 40 years available, and in the future a fuller use of this dataset could be appreciated. Moreover, for consistency purposes with previous studies, the Block Maxima method was used to determine the return levels. A sensitivity test could be performed to see how different those results would be if another method, such as the Peak-over-Threshold, was used.

Section 4 showed that applying a correction to the simulated discharge improved the results in 27 out of 34 cases. The seven remaining stations suffered from the correction and gave better results before applying the correction factor. Some of them are small direct-runoff rivers that were simulated correctly before applying the coefficients, while others are more problematic as the reanalysis failed to reproduce the trend of their duration curve. Expanding the analysis to a larger number of stations in the future would help understand what are the limiting factors in which cases the correction should not be used.

6 Conclusions

In this research, a first attempt to estimate extreme flood values based on simulated data from the ICRA model has been proposed. Firstly, 44 gauging stations that have been recording discharge of the main rivers in Iceland were selected. Runoff for 38-years of reanalysis was extracted and summed daily for all the catchments associated with the selected rivers. The runoff was then converted into average discharge. By the end of the first step of the project, two sets of daily discharge series were available, one built on simulated runoff and the other one on observed discharge.

In the second part of the project, it was tested if the dataset built on the ICRA data would give similar results of hierarchical clusters as the observed dataset. In both cases, discharge timeseries were normalized and combined in different ways to reflect the seasonality, duration curves and mass curves of the rivers. Various catchment characteristics were also added to the analysis and results were presented on two dendrograms. Rivers clustered in five groups, according to river types and geographical location. Most of the catchments (36 out of 44 gauging stations) clustered similarly between the two datasets. Since the data was normalized before performing the cluster analysis, the results do not give any insight into the closeness of the discharge values, but showed that the conversion of the runoff successfully kept the general behaviour of the rivers. Three stations associated to river Skaftá and subject to frequent jökulhlaup were discarded, and in the end, 34 stations were used as control stations, and seven left to be used later in this study.

Thirdly, discharge values from both datasets were compared with a focus on values above the 95^{th} percentile in order to limit the data only to the extremes. Discharge for all control stations were presented on duration curves. In the vast majority of cases (32 out of 34 rivers), the discharge was overestimated in the simulations, which reflects the lack of infiltration in the runoff conversion model. However, the general trend of the discharge was well simulated. It was therefore decided to apply a correcting factor to scale down the simulations and obtain values closer to the observed data. This was done by calculating the coefficient of proportionality for each individual station, based on highest discharge data over the 75^{th}, 90^{th} and 95^{th} percentiles. Those coefficients were then averaged by cluster and applied to the simulated dataset, which greatly improved the results. The correcting coefficient based on the 95^{th} percentile gave the closest match with the observations.

An EVA was then performed using the Block Maxima method on both the observed and simulated timeseries, before and after correction. Closeness Coefficients were calculated and showed a great improvement of the return levels after applying the correction for 27 out of 34 rivers. Eventually, the seven stations left from the control station group were studied, applying a methodology that would be applied on ungauged catchments in the future. 25-year flood values were then calculated and compared to the observed values, showing that in all cases, the *CC* values were significantly improved.

Overall, these result show that extreme discharge values based on catchment-accumulated runoff from the ICRA dataset is able to simulate the observed high discharge after correction. The findings of this study represent an initial methodology that could successfully assess design-flood values for ungauged catchments throughout the country.

References

Atlason, H., Þórarinsdóttir, T., Roberts, M. J., Massad, A.-G. R., Priet-Mahéo, M., Björnsson, B. B. (2021). Notkun einfalds vatnafarslíkans á vatnafarslega ólíkum svæðum á Íslandi. *Veðurstofa Íslands*

Bengtsson, L., Andrae, U., Aspelien, T., Batrak, Y., Calvo, J., de Rooy, W. & Køltzow, M.Ø. (2017). The HARMONIE-AROME model configuration in the ALADIN-HIRLAM NWP system. *Mon. Wea. Rev.,* **145**, 1919–1935

Bo Chen, Witold F. Krajewski, Fan Liu, Weihua Fang, Zongxue Xu (2017). Estimating instantaneous peak flow from mean daily flow. *Hydrology Research*, 1 December 2017; 48 (6): 1474–1488. doi.org/10.2166/nh.2017.200

Coles, S. (2001). An introduction to Statistical Modeling of Extreme Values. London; *Springer.*

Crochet, P. (2012). Flood-Duration-Frequency modeling. Application to ten catchments in Northern Iceland. *Veðurstofa Íslands.* VÍ 2012-006

Crochet, P. & Þórarinsdóttir, T. (2014). Flood frequency estimation for ungauged catchments inIceland by combined hydrological modelling and regional frequency analysis. *Veðurstofa Íslands.* VÍ 2014–001

Crochet, P. & Þórarinsdóttir, T. (2015). Regional flood frequency analysis: A case study in Eastern Iceland. *Veðurstofa Íslands.* VÍ 2015-007

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

Fill, H. D. & Steiner, A.A. (2003). Estimating instantaneous peak flow from mean daily flow data. *Journal of Hydrologic Engineering,* **8**, 365–369

Hilmar B. Hróðmarsson, Njáll Fannar Reynisson & Ólafur F. Gíslason (2009). Flóð íslenskra vatnsfalla–flóðagreining rennslisraða. *Veðurstofa Íslands*. VÍ 2009-001.

Hróðmarsson, H. B., & Þórarinsdóttir, T. (2018). Flóð íslenskra vatnsfalla, flóðagreining rennslisraða. *Veðurstofa Íslands*. VÍ 2018-003.

Massad, A.-G. R., Petersen, G. N., Þórarinsdóttir, T., Roberts, M.J. (2020). Reassessment of return levels in Iceland. *Veðurstofa Íslands*. VÍ 2020-008.

Nawri, N., Pálmason, B., Petersen, G.N., Björnsson, H. & Þorsteinsson, S. (2017). The ICRA atmospheric reanalysis project for Iceland. *Veðurstofa Íslands*. VÍ 2017-005

Pagneux, E., Jóhannsdóttir, G.E., Þórarinsdóttir, T., Hróðmarsson, H. & Egilson, D. (2017). Flóð á vatnasviðum Eyjafjarðarár, Héraðsvatna, Hvítár í Borgarfirði, Lagarfljóts og Skjálfandafljóts: I. Yfirlit yfir orsakir, stærð og afleiðingar sögulegra atburða. *Veðurstofa Íslands.* VÍ 2017-006

Pagneux, E., Jónsson, M.A., Þórarinsdóttir, T., Björnsson, B.B., Egilson, D. & Roberts, M.J. (2018). Hættumat vegna jökulhlaupa í Skaftá: Hermun flóðasviðsmynda. *Veðurstofa Íslands*. VÍ 2018-008

Pagneux, E., Jónsson, M. Á., Björnsson, B. B., Pétursdóttir, S., Reynisson, N. F., Hróðmarsson, H. B., Einarsson, B. & Roberts, M. J. (2019). Hættumat vegna vatnsflóða í Ölfusá. *Veðurstofa Íslands*. VÍ 2019-013

Priet-Mahéo, M., Massad, A.-G. R., Pétursdóttir, S., Þórarinsdóttir, T. & Egilson, D. (2019). Daglegar rennslisspár með notkun hliðstæðrar greiningar Harmonieveðurgagna. *Veðurstofa Íslands*

Priet-Mahéo, M., Massad, A.-G. R., Þórarinsdóttir, T., Roberts, M.J. (2020). Veflausn með daglegum rennslisspám sem byggist á hliðstæðri greiningu veðurgagna. *Veðurstofa Íslands*

Rist, S. (1990). *Vatns er þörf.* Reykjavík: Bókaútgáfa Menningarsjóðs.

Stefánsdóttir, G., Björnsson, B.B., Magnússon, S., & Egilson, D. (2014). Verklokaskýrsla vegna stjórnar vatnamála - Vinna ársins 2013. *Veðurstofa Íslands*

Þórarinsdóttir, T., Roberts, M.J., Wallevik, J.E., Björnsson, B. B., Massad, A.-G., R. (2021). Hættumat vatnasviða: Eyjafjarðará, Héraðsvötn, Hvítá í Borgarfirði, Lagarfljót og Skjálfandafljót. *Veðurstofa Íslands*