

Optimization of the Hydrologic Response Units (HRU) using gridded meteorological data and spatially varying parameters

David Poblete¹, Jorge Arévalo², Orietta Nicolis³, and Felipe Figueroa⁴

¹Universidad de Valparaíso

²Universidad de Valparaíso

³Universidad Andres Bello

⁴Plataforma de Investigación en Ecohidrología y Ecohidráulica (EcoHyd)

November 24, 2022

Abstract

Although complex hydrological models with detailed physics are every day more common, lumped and semi-distributed models are still used for many applications and offer some advantages as its reduced computational cost. Most of these semi-distributed models use the concept of Hydrological Response Unit or HRU. In its original conception, HRUs are defined as homogeneous structured elements having similar climate, land-use, soil and/or pedotransfer properties, hence a homogeneous hydrological response under equivalent meteorological forcing. This work presents a quantitative methodology to construct HRUs based on Principal Component Analysis and Hierarchical Cluster Analysis of gridded meteorological data and hydrological parameters. The methodology is tested using the Water Evaluation and Planning System (WEAP) model for the Alicahue River Basin, a small catchment in Central Andes, in Central Chile. The results show that with four HRUs it is possible to reduce up to about a 10% the relative within variance of the catchment, an indicator of homogeneity of the HRUs. Evaluation of the simulations show a good agreement with streamflow observations in the outlet of the catchment with a NSE value of 0.79 and also show the presence of small hydrological extreme areas that generally are neglected due to their relative size.

1 **Optimization of the Hydrologic Response Units (HRU) using gridded meteorological**
2 **data and spatially varying parameters**

3 **Enter authors here: D. Poblete^{1,†}, J. Arevalo^{2,3}, O. Nicolis⁴ and Felipe Figueroa⁵**

4 ¹ Escuela de Ingeniería Civil, Universidad de Valparaíso, Valparaíso, Chile.

5 ² Departamento de Meteorología, Universidad de Valparaíso, Valparaíso, Chile.

6 ³ Department of Hydrology and Atmospheric Sciences, University of Arizona, Tucson, AZ,
7 USA.

8 ⁴ Facultad de Ingeniería, Universidad Andrés Bello, Viña del Mar, Chile.

9 ⁵ Plataforma de Investigación en Ecohidrología y Ecohidráulica (EcoHyd), Santiago, Chile.

10 Corresponding author: David Poblete (david.poblete@uv.cl)

11

12 † David Poblete, email address: david.poblete@uv.cl

13

14 **Key Points:**

15 • The paper presents a methodology to optimize the definition of Hydrologic Response
16 Units (HRUs) for semi-distributed hydrologic modelling.

17 • The optimization minimizes the internal variability within an HRU and maximize the
18 variance between different HRUs.

19 • The results show different hydrological responses by each HRU, in terms of total volume,
20 timing, distribution and peak discharge.

21

22 **Abstract**

23 Although complex hydrological models with detailed physics are every day more common,
24 lumped and semi-distributed models are still used for many applications and offer some
25 advantages as its reduced computational cost. Most of these semi-distributed models use the
26 concept of Hydrological Response Unit or HRU. In its original conception, HRUs are defined as
27 homogeneous structured elements having similar climate, land-use, soil and/or pedotransfer
28 properties, hence a homogeneous hydrological response under equivalent meteorological forcing.
29 This work presents a quantitative methodology to construct HRUs based on Principal
30 Component Analysis and Hierarchical Cluster Analysis of gridded meteorological data and
31 hydrological parameters. The methodology is tested using the Water Evaluation and Planning
32 System (WEAP) model for the Alicahue River Basin, a small catchment in Central Andes, in
33 Central Chile. The results show that with four HRUs it is possible to reduce up to about a 10%
34 the relative within variance of the catchment, an indicator of homogeneity of the HRUs.
35 Evaluation of the simulations show a good agreement with streamflow observations in the outlet
36 of the catchment with a NSE value of 0.79 and also show the presence of small hydrological
37 extreme areas that generally are neglected due to their relative size.

38

39 **Plain Language Summary**

40

41 **1 Introduction**

42 Since the works of Leavesley et al. (1983) and Flügel (1995), the concept of Hydrologic
43 Response Units (HRU) has risen as one of the most common approaches for semi-distributed
44 hydrological modelling. Flügel (1995) defined an HRU as a homogeneous structured element
45 having similar climate, land-use, soil and/or pedotransfer properties, hence a homogeneous
46 hydrological response under equivalent meteorological forcing. An important assumption is that
47 the variation of the hydrological process dynamics within a single HRU is small compared with
48 the hydrologic dynamics and responses to other units defined in the model. Many authors assume
49 that HRU do not necessarily represent contiguous geographical areas so the topology of the
50 elements is simplified or just neglected (Pilz, Francke, & Bronstert, 2017; Savvidou, Efstratiadis,
51 Koussis, Koukouvinos, & Skarlatos, 2018) and the total discharge of the watershed is calculated
52 as the incremental input of every independent element and propagated to its outlet; assumption
53 that we will also consider in the rest of this study.

54 Traditionally, land use/land cover, topographic characteristics and soil types have been used as
55 proxies of many of the parameters involved in the governing equations and parameterizations of
56 the lumped, semi-distributed and even distributed models, but always with a certain degree of
57 uncertainties (Höge, Wöhling, & Nowak, 2018; Nijzink et al., 2016; Orth, Dutra, &
58 Pappenberger, 2016). Most of the methodologies to delineate HRUs rest on the expected
59 relationships between physical-ecological characteristics of the catchment and the corresponding
60 hydrological properties reflected on the hydrological model parameters. Hence, HRUs are
61 usually defined by the superposition of land use and soil type and after the classification,
62 quantitative or qualitative relations are used to estimate hydrologic parameters on each HRU.
63 One of the most common approaches has been to include the sub-basins in the process, hence the

64 intersection of the sub-basins, land use categories and soil type polygons in a GIS represents the
65 minor elements for hydrologic modelling (Dehotin & Braud, 2008; Savvidou et al., 2018).
66 A different approach is used in Savvidou et al. (2018), as they estimate the CN Curve Number
67 parameter for reference conditions using soil permeability, vegetation classes and drainage
68 capacity maps and then the HRUs are defined based on the separation of areas according to the
69 CN values. According to the authors, this delineated HRUs can be used in any hydrological
70 model as the SCS-CN model, which is widely used and understood.
71 Even though the importance on defining properly the HRU for a good representation of the
72 hydrological processes and dynamics, methods and tools for identifying an appropriate scale, are
73 often missing. The challenge is to identify a proper method for discretization of the basins, losing
74 the least information possible and maximizing the model reliability and utility that in turn plays a
75 crucial role in the accuracy of the models (Haghnegahdar, Tolson, Craig, & Paya, 2015; Han,
76 Huang, Zhang, Li, & Li, 2014; Haverkamp, Fohrer, & Frede, 2005). If over simplification of the
77 basin characteristics is done, small areas of extreme hydrologic behavior can be neglected by a
78 lack of representation in the aggregation procedures (Haverkamp et al., 2005). On the other hand,
79 if the used data is highly detailed and fragmented, it can lead to an excessive number of HRUs,
80 making the modelling impracticable.
81 Although meteorological variables are inputs to every model, none of the methodologies use that
82 information directly in the construction process of the HRU. Flügel (1995) suggested more than
83 two decades ago that the use of meteorological information to construct HRU is advisable, but it
84 has not been explored in depth probably due to the lack of good quality spatial meteorological
85 information. Today, this idea is more plausible and can be considered because one of the basic
86 assumptions on HRU is that meteorological forcing is homogeneously spatialized over the
87 domain of the HRU. Therefore, the spatial heterogeneity of the precipitation and other variables
88 can be incorporated in the delineation of HRUs. An indirect approach to include climate
89 information is used by Young et al. (2009), where 15 watersheds of the Sierra Nevada in
90 California are discretized in HRU by the intersection of sub-basins, soils type, vegetation cover
91 and elevation bands in the Water Evaluation And Planning System model (WEAP; Yates, Sieber,
92 Purkey & Huber-Lee (2005)). They calculate fractional areas for each sub-basin using a
93 vegetation cover/soil type combination in 250 meters elevation bands ranging from 500 to 4000
94 meters above sea level, in order to provide a finer discretization for snow accumulation and melt
95 modelling. This has been a common practice in the use of this model in semi-arid basins in Chile
96 (for instance, Bonelli, Vicuña, Meza, Gironás, & Barton (2014) and Vicuña, Garreaud, &
97 McPhee (2010)).
98 Given all these issues, some questions arise: How to use the detailed information available on
99 land use, geomorphologic properties and climatic behavior for the separation of a manageable
100 number of independent modelling units? Which criteria must be used to simplify the complexity
101 of hydrologic dynamics of a watershed into the smallest number of homogeneous units as
102 possible without losing valuable information? Does the use of these independent modeling units
103 ensure heterogeneity of hydrologic response between them?
104 This paper presents a quantitative methodology for the determination of unstructured HRU based
105 on homogeneity of the hydrological parameters used by any specific hydrological model and its
106 meteorological inputs. A Principal Component Analysis (PCA) is performed in order to get an
107 independent set of vectors to be used in a Hierarchical Clustering (HC) algorithm to obtain the
108 desired HRUs. The result minimizes the internal variability of hydrologic properties in each

109 HRU and simultaneously maximizes the variability between different HRUs, subsequently of the
110 hydrologic responses of each element.

111 To test the proposed methodology, HRU delineation is performed for the Alicahue river basin, an
112 Andean semi-arid basin located in Central Chile. Hydrologic parameters and climate averaged
113 values used by the semi-distributed WEAP model (Water Evaluation and Planning System) are
114 calculated over a regular grid, that in turn are used to classify each cell in the mentioned HRUs.
115 Climate variables are based on a 1km resolution bias-corrected model output for three periods of
116 12-month using the WRF model (Skamarock et al., 2008) and the hydrologic parameters are
117 estimated by topographic characteristics derived from 30m ASTER DEM (Tachikawa et al.,
118 2011) and Land Use data from Natural Resources Research Center of Chile (Martínez, Flores,
119 Retamal, Ahumada, & Brito, 2013). Finally, the performance, accuracy and skill of the model
120 using ten different configurations of HRU are analyzed using common modelling indicators.
121

122 **2 Materials and Methods**

123 The proposed methodology to delineate HRUs consists in the creation of a dataset of raster files
124 comprised of hydrologic parameters and meteorological variables used by the target hydrological
125 model. Then, through principal component and hierarchical cluster analyses, every cell of the
126 raster files is classified into a specific cluster to form the different HRUs.
127

128 **2.1 Area of the study case**

129 Central Chile has a landscape with a very complex topography. It is surrounded by the high
130 peaks of the Andes Mountains usually above 4,000 m.a.s.l. at the East and the Pacific Ocean at
131 only about 150 km West of the mountains. Most of the river basins in this area have a latitudinal
132 preferential path, downstream of the Andes up to the Ocean. Its climate corresponds to
133 Mediterranean with dry summers, temperatures are usually mild ranging from about 0°C as
134 minimum during winter up to 35°C as maximum during summer, except for the high elevation
135 lands where below freezing temperatures are usual during winter. Mean annual precipitation is
136 about 400 mm for the valleys and coastal areas, which is mostly due to winter frontal storms,
137 hence the spatial variability is mostly modulated by the orographic effects.

138 The domain of the area of study corresponds to the Alicahue river basin, which is located
139 between geographical coordinates 32.39°S to 32.21°S in latitude and 70.76°E to 70.41°E in
140 longitude, in the province of Petorca, Valparaiso Region, Central Chile (Figure 1). This is a sub-
141 catchment of the La Ligua river basin, that receives water from other minor streams and flows
142 into the sea, with a total length of nearly 200km. The Alicahue river has a length of just 30 km
143 and its drainage area is just 354 km², but its topography ranges from 780 m.a.s.l up to 3985
144 m.a.s.l., with almost half of its area located above 2500 m.a.s.l. In winter and during rainfall, the
145 0°C isotherm in central Chile is typically located at about 2500 m.a.s.l. (Garreaud, 1992),
146 allowing snow accumulation in most part of the Andes mountains. Hence, at the outlet of the
147 basin there is an important flow between mid-spring and beginning of summer in southern
148 hemisphere (from October to January) due to snow melting. Agriculture uses the waters from La
149 Ligua river, but most of the discharge of this river during the dry season comes from upper
150 basins as the Alicahue river, where snow accumulation is possible during the cold and wet
151 winter.

152 The Alichahue river basin is a small catchment and has limited human intervention, which
 153 simplifies the analysis of the results. Also, its behavior is comparable to several other high-
 154 altitude catchments in Central Chile, where complex topography is dominated for the traversal
 155 valleys downstream of the Andes Mountains and snowmelt is one of the dominant hydrologic
 156 drivers.

157



158

159 *Figure 1. Area of Study. Relative location in South America and WRF Domains (left), topography of the Area of Study indicating*
 160 *the limits of Alichahue and La Ligua river basins (red and purple polygons, respectively) and the location of the weather stations*
 161 *(streamflow station co-located with Alichahue Hacienda weather station).*

162 The only available stream gauge station is located in the outlet of the Alichahue river basin
 163 (32.20°S, 70.45°W), from station BN 05200001-7 “*Rio Alichahue en Colliguay*”, from the General
 164 Dictatorate of Water (DGA in Spanish), with a recording period starting the year 1963.

165 Although only the station Alichahue Hacienda is located inside the basin, the frontal nature of the
 166 precipitation makes reasonably to correlate near observations. The values recorded at these
 167 stations are 1.16 m3s⁻¹ for mean annual streamflow, 267mm for total annual precipitation and
 168 15.1°C for mean annual temperature.

169 2.2 Hydrologic Parameters and meteorological datasets

170 The methodology uses raster maps of the hydrologic parameters and mean annual values of the
 171 meteorological variables used by the chosen hydrologic model. These maps need to be
 172 constructed or generated previously by any methodology.

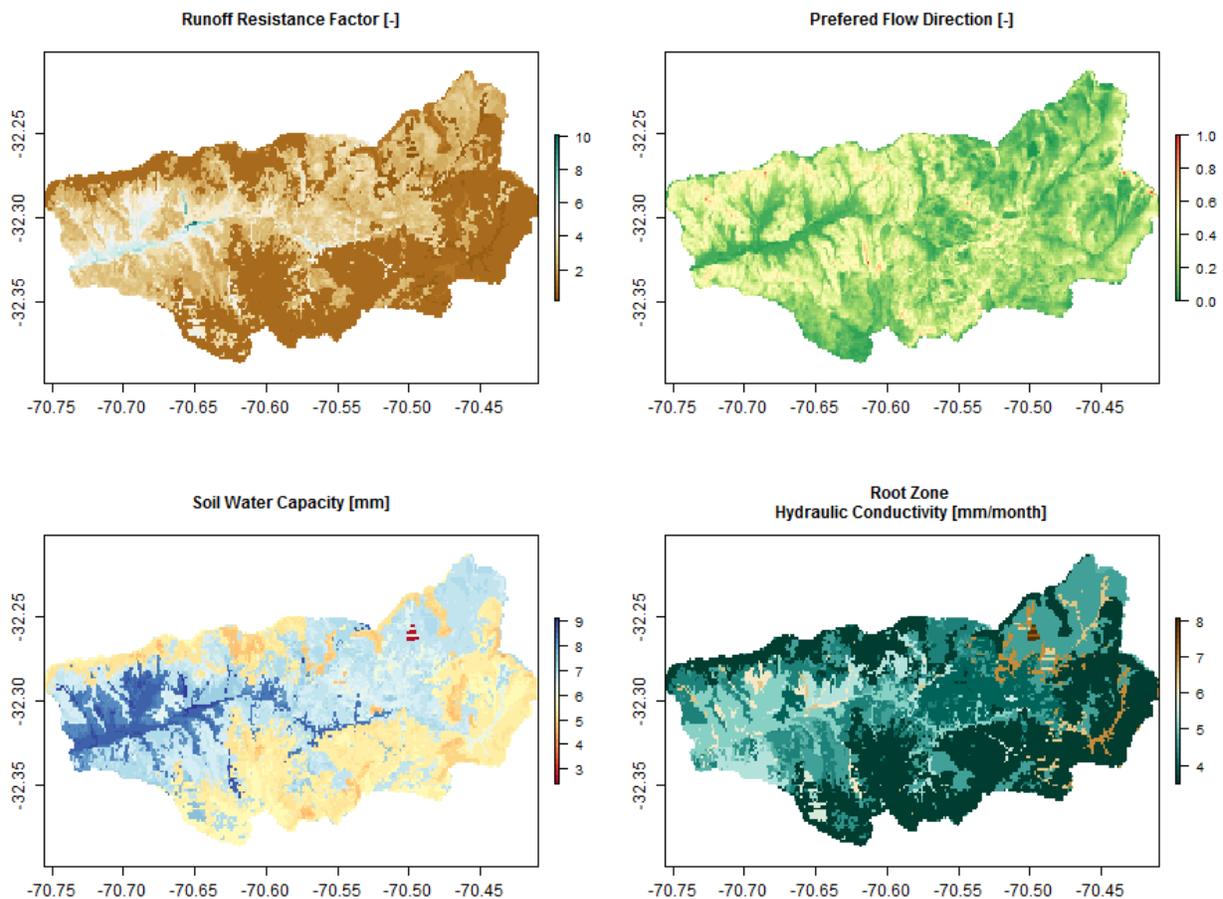
173

174 In the case of this study, the WEAP model (Yates et al., 2005) is used to test the methodology.
 175 WEAP is a water allocation model, that has been used for water resources management in
 176 several studies over several catchments around the world (for instance Purkey et al., 2007;
 177 Young et al., 2009) and particularly in Chile (Bonelli et al., 2014; Vicuña et al., 2010). It has
 178 incorporated a hydrology module that represents the mass balance in elements called *catchments*,
 179 in which simplified hydrological fluxes and storages are modelled using a one dimensional and
 180 2-layers storage system. Although a WEAP catchment can be used as a single HRU, the
 181 *catchment* element can be internally divided in more separate units, each of them as a single
 182 HRU. The methodology most widely used, divides the catchment in elements by land use cover
 183 within a given elevation band and sub-basin; with all the HRU having the same meteorological
 184 condition on each elevation band.

185 The upper layer of the *catchment* element has four hydrologic parameters: Sw (Soil Water
 186 Capacity in mm) represents the soil layer depth; RRF (Runoff Resistance Factor) is equivalent to
 187 the run-off coefficient in the rational equation; Ks (Root Zone Conductivity in mm/month)
 188 corresponds to the saturated hydraulic conductivity in the soil layer; and kc to the crop
 189 coefficient of the vegetation. The parameter f (Preferred Flow Direction) controls the water
 190 flowing from the upper layer to the lower layer as Interflow or Deep Percolation ($f=1$ for total
 191 horizontal flow and $f=0$ for total vertical flow). Dw and kd represent the depth and the saturated
 192 hydraulic conductivity of the deeper layer of the *catchment* element, respectively. Finally, the
 193 simple snow model uses two temperature thresholds, for melting and freezing (T_1 and T_s ,
 194 respectively), totalizing nine parameters (for detailed information on the water balance equations
 195 see Yates et al., 2005).

196 Figure 2 shows the maps of the WEAP parameters f , RRF and the log values of Sw and Ks ,
 197 respectively for the Alicahue river basin. RRF (top-left) is related to milder slopes and vegetated
 198 terrain; f (top-right) depends on the terrain slope and soil properties. Sw (bottom-left) and Ks
 199 (bottom-right) are shown in a log scale for better visualization. Sw present the deepest soils in the
 200 flat lands near the main water course, in contrast to the sides of the hillslope. The latter is also
 201 related to the vegetation land cover and slope and dominated by very permeable areas in high
 202 altitude wetlands.

203

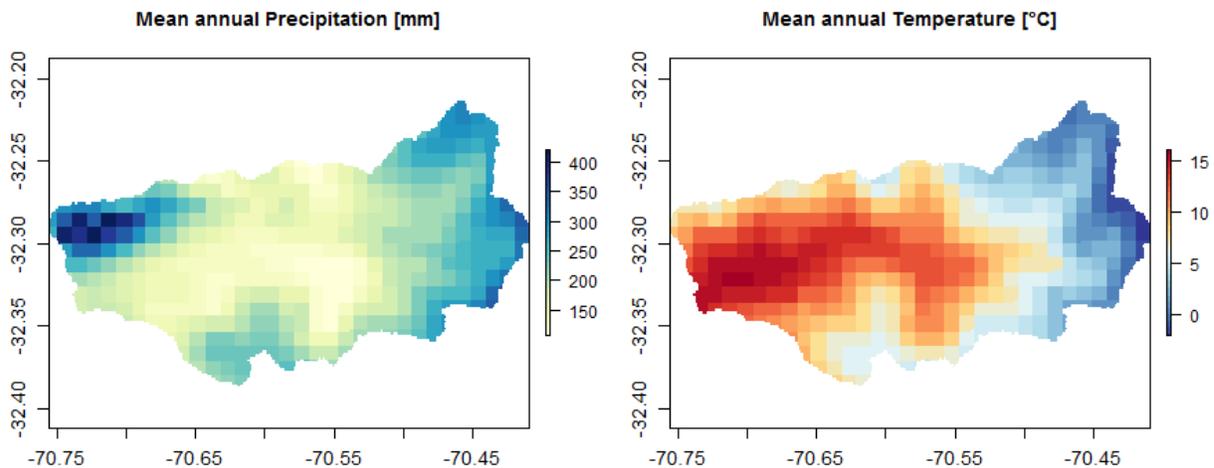


204

205 Figure 2. Parameters map covering the Alicahue River basin for: (top left) f , (top right) RRF, (bottom left) S_w and (bottom right)
 206 K_s . S_w and K_s are plotted in log scale.

207 As the spatial representativeness of meteorological stations is small in complex terrain and
 208 observations are usually scarce, the WRF model version 3.4.1 was used to simulate three periods
 209 of 12 consecutive months each: (1) 1 July 2003 to 30 June 2004, (2) 1 June 2009 to 31 May
 210 2010, and (3) 1 Jun 2010 to 31 May 2011. Those simulations were performed in four nested
 211 domains with 27, 9, 3 and 1 km of horizontal resolution and 50 terrain-following vertical levels,
 212 the innermost domain was used for the analysis and cover entirely the study area, as could be
 213 seen in Figure 1. As the WRF simulation period covered only 36 non continuous months, it was
 214 not suitable to drive the long-term hydrological modeling. Hence, a simpler relation between the
 215 available observed precipitation time series near the basin and thus in each HRU was used.
 216 Long-term temperature time series were extrapolated to the HRUs using a simple linear model
 217 between the mean annual temperature modeled in WRF and “Alicahue Hacienda” station records
 218 using variables as elevation, aspect, mean longitude and mean latitude for the HRUs in each of
 219 the simulations.

220 Figure 3 shows the mean annual precipitation and the mean temperature for the 36 months of
 221 WRF simulations. Other climatological variables as relative humidity, net radiation, albedo,
 222 evapotranspiration and wind speed are not shown.



223 Figure 3. Maps covering the Alicahue River basin for: (left) mean annual precipitation, (right) mean temperature.
 224

225
 226 As both types of datasets have different cell sizes and extension, to join both datasets the
 227 meteorological raster maps are resampled to a common grid system into the parameters base grid
 228 by the nearest neighbor method using the Vincenty (ellipsoid) great circle distance from the
 229 *distm* function of the *geosphere* package in R (Hijmans, 2017).

230

231 2.3 Clustering processes and HRU delineation

232 In this section, the core of the HRU delineation process is detailed. The gridded model-specific
 233 parameters and the climatological information are used in the Principal Component Analysis to
 234 later use its firsts components in the Hierarchical Clustering.

235 The Principal Component Analysis (PCA) technique consists on describing a multidimensional
 236 data set using a smaller number of uncorrelated variables (principal components) that incorporate
 237 as much information as possible. If the data set X is composed by i individuals and p variables,
 238 the PCA finds the first k principal components (with $k < p$) with the maximum variance. The
 239 Principal Component Analysis is performed using the function `PCA` from the `FactoMineR`
 240 package (Lê, Josse, & Husson, 2008) for multivariate data analysis in R (R Core Team, 2019) by
 241 assigning greater weights to the most important variables, in order to capture more variance of
 242 these variables. In this study, precipitation and temperature variables were weighted by a factor
 243 of two given its importance in the water balance equation, while the rest of the variables had
 244 weights equal to one. This allows the use of the expertise of the modeler in assigning more
 245 importance to specific variables.

246 Working with principal components instead on the original data, allows to obtain more stable
 247 results in the clustering process. Since the first dimensions (or components) extract the most
 248 information from data and the last ones represent the noise (Husson, Julie, & Jérôme, 2010), the
 249 first components accounting at least 90% of variance are used in the Hierarchical Cluster
 250 Analysis function `HCPC` from the same `FactoMineR` package. The objective is to capture most
 251 of the variability of the most important variables and simultaneously not to capture the variability
 252 of the least important variables or represent a minor proportion of the main variables.

253 The Hierarchical Clustering used in this work has been implemented using the Ward's criterion
 254 Husson et al. (2010). The Ward's method is based on an agglomerative approach or "bottom-
 255 up", where the clustering starts considering each observation as a cluster, and pairs of clusters
 256 are merged as one moves up in the hierarchy. The initial cluster distances in Ward's method may
 257 be defined by the squared Euclidean distance between the individuals' values and their averages.

258 By considering a multivariate database composed by i spatial individuals (cells) and K variables
 259 (both hydrologic parameters and meteorological variables), the total variance of Q clusters (with
 260 $Q < i$) is evaluated according to its decomposition in the between and within variances given by:

$$261 \sum_{k=1}^K \sum_{q=1}^Q \sum_{i=1}^{N_q} (x_{iqk} - \bar{x}_k)^2 = \sum_{k=1}^K \sum_{q=1}^Q N_q (\bar{x}_{qk} - \bar{x}_k)^2 + \sum_{k=1}^K \sum_{q=1}^Q \sum_{i=1}^{N_q} (x_{iqk} - \bar{x}_{qk})^2, \quad (1)$$

263 where x_{iqk} is the normalized value of the variable k for the individual i of the cluster q , \bar{x}_{qk} is
 264 the mean of the variable k for cluster q , \bar{x}_k is the overall mean of variable k (equal to zero if
 265 normalized) and N_q is the number of spatial points in cluster q . The first member at the right
 266 side of the equation represents the between inertia (or between variance) and the second member,
 267 the within inertia (within variance).

268 The importance on this equation is that the total variance of the system remains constant and as
 269 the within variance decreases (the clusters become more homogeneous), the between variance
 270 increases (the clusters become more and more different between each other).

271 At each step of the aggregating procedures algorithm the increase of within variance is
 272 minimized (or the increase of the between variance is maximized). This analysis detects groups
 273 of individuals with similar characteristics and hydrologic behavior based on parameters and
 274 meteorological similitude between cells. Each group of cells belonging to a cluster represents a
 275 single HRU to be used in the hydrological model. It is not necessary for cells to be contiguous to

276 belong to the same cluster as proximity in the space of attributes does not ensure proximity in the
 277 geographical space Fouedjio (2016), although this may be desirable if contiguous HRU are to be
 278 delineated.

279 The optimal number of clusters in the data is selected using the clustering tree and is calculated
 280 automatically by the function when the within variance reaches a minimum plateau, using the
 281 least number of clusters. In the method described by Husson et al. (2010), if $\Delta(Q)$ is the between
 282 inertia increase when moving from $Q-1$ to Q clusters, the optimal number of clusters Q is the
 283 one which minimize the relation $\Delta(Q)/\Delta(Q+1)$. Other indexes to assess the optimal number of
 284 clusters are described in Fouedjio (2016).

285 To test the present methodology, we calculate 10 scenarios in which each scenario has a number
 286 of s clusters (s from 1 to 10). For each scenario, the method stores for each cell the HRU it
 287 belongs to. This is done to evaluate the sensitivity of the hydrological model to the number of
 288 HRUs, ranging from a single HRU (a completely lumped model) to a more semi-distributed
 289 scheme of the basin with as many HRUs as clusters generated.

290 2.4 Hydrological model setup and simulations

291 The WEAP model is ran using the ten different scenarios described previously. Every
 292 configuration of the model uses a different number of HRU. The lumped configuration was
 293 called HRU_01 and uses just one HRU to model the basin. The second scenario uses two HRU
 294 and is called HRU_02. The rest of the configurations are called similarly depending on the
 295 number of HRU used.

296 The time series of monthly precipitation and mean monthly temperatures were derived from the
 297 meteorological dataset and the observed values recorded in the meteorological stations.

298 The values assigned for the hydrological parameters in each HRU are calculated as the average
 299 value for all the cells belonging to such HRU defined in the previous step. Also the values of
 300 Wind Speed, Relative Humidity and Albedo were set constant for every simulation for
 301 simplicity. These values were obtained by intersecting the area for each cluster defined in the
 302 previous section with the raster corresponding to the annual mean of each variable obtained from
 303 WRF outputs.

304 For calibration purposes, WEAP model has spatially-constant calibration factors for each of the
 305 four parameters assessed in the PCA/HPCP methodology. They are assumed initially as one but
 306 can be adjusted in the calibration process to adjust the results of the modelling by mean of
 307 automated or manual techniques.

308 Finally, the results of the hydrological modelling are analyzed by some common hydrological
 309 indicators as the Nash–Sutcliffe model efficiency coefficient (NSE) and the Root Mean Squared
 310 Error (RMSE) standardized by the mean discharge.

$$311 \quad NSE = 1 - \frac{\sum_{i=1}^N (Q_{oi} - Q_{si})^2}{\sum_{i=1}^N (Q_{oi} - \bar{Q}_o)^2}, \quad (2)$$

$$312 \quad RMSE = \frac{1}{\bar{Q}_o} \cdot \sqrt{\frac{\sum_{i=1}^N (Q_{oi} - Q_{si})^2}{N}}, \quad (3)$$

313 Where:

314 Q_{oi} : Observed discharge at time step i .

315 Q_{si} : Simulated discharge at time step i .

316 \bar{Q}_o : Average of the observed discharge over the simulation period.

318 3 Results

319 This section shows the main results in each part of the methodology: (i) the results of the
 320 Principal Component Analysis and the Hierarchical Clustering and (ii) the results of the
 321 hydrological modelling using the different schemes of the HRU.

322 3.1 PCA and cluster analysis

323 This section shows the results of the Principal Component Analysis and the Hierarchical
 324 Clustering Analysis, the core of the HRU delineation. The PCA was performed over the set of
 325 the meteorological variables (Precipitation, Temperature, Relative Humidity, Wind Velocity,
 326 Albedo and Evapotranspiration) and hydrologic parameters (S_w , f , RRF and K_s). The first result
 327 to highlight is that the two first dimensions resulting from the PCA account for 66.8% of the total
 328 variance. Adding the following 3rd, 4th and 5th dimensions, they account for the 78.7%, 84.6%
 329 and 89.6% respectively of the total variance of the master dataset. Table 1 shows the variance
 330 explained by each consecutive eigenvector or dimension and the contribution of each variable to
 331 the dimensions. The first dimension (more than 50% of the total variance) is composed mainly
 332 on meteorological variables, being Temperature, Albedo, Wind Speed and Rainfall the ones with
 333 more contribution. The second-dimension accounts for more than 16.1% of the total variance and
 334 is composed mainly by rainfall and the hydrological parameters.

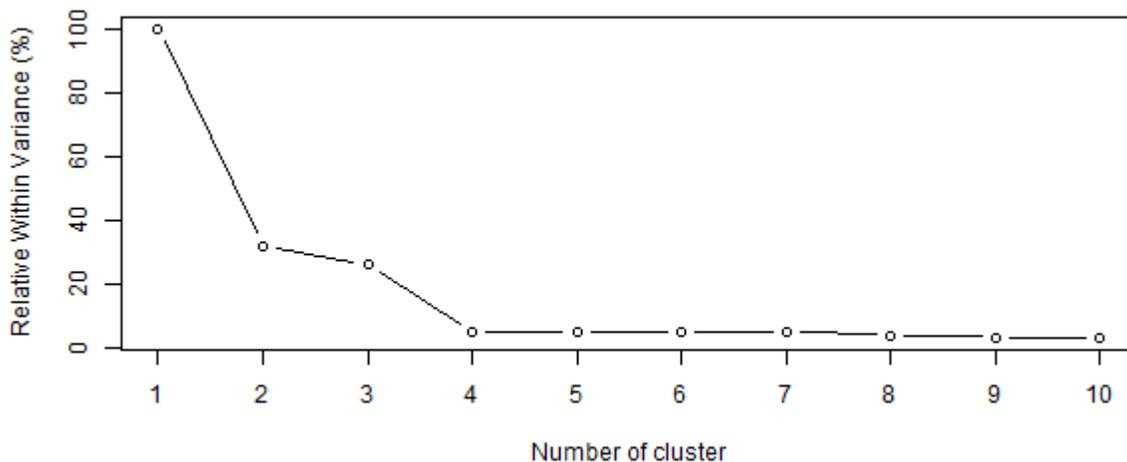
335
 336
 337 *Table 1. Summary from the PCA analysis results for the first five dimension. Where the upper part shows the total variance*
 338 *explained by each dimension and lower, the contribution of each variable to that dimension.*

| | Dim.1 | Dim.2 | Dim.3 | Dim.4 | Dim.5 |
|-------------------------------|------------------------------------|-------|-------|-------|-------|
| Variance Explained (%) | 50.7 | 16.1 | 11.9 | 5.9 | 5.0 |
| <i>Variables</i> | Contribution to each dimension (%) | | | | |
| Temp | 28.2 | 0.1 | 0.7 | 0.2 | 5.1 |
| Albedo | 11.5 | 1.1 | 5.6 | 1.6 | 10.0 |
| WS | 10.8 | 1.3 | 0.6 | 0.0 | 8.5 |
| Pp | 10.7 | 41.5 | 22.6 | 0.0 | 0.6 |
| EVPM | 10.2 | 2.2 | 1.5 | 0.5 | 12.3 |
| RNet | 9.7 | 3.5 | 9.6 | 1.4 | 10.7 |
| HR | 9.5 | 5.6 | 1.4 | 1.0 | 0.0 |
| RRF | 4.9 | 13.9 | 10.5 | 12.1 | 2.5 |
| Sw | 4.1 | 20.3 | 6.4 | 4.4 | 9.7 |
| F | 0.3 | 4.7 | 32.1 | 0.4 | 35.6 |
| Ks | 0.0 | 5.7 | 9.0 | 78.3 | 5.0 |

339

340 Table 1 suggests that the firsts five dimensions are carrying most of the information, cleaning the
 341 statistical noise and hence these first five components will be used in the HCPC function for the
 342 cluster analysis. Based on the new dataset composed by only these principal components, the
 343 total variance of the system is fixed for the cluster analysis, therefore the within variance will be
 344 expressed as relative to such total hereinafter.

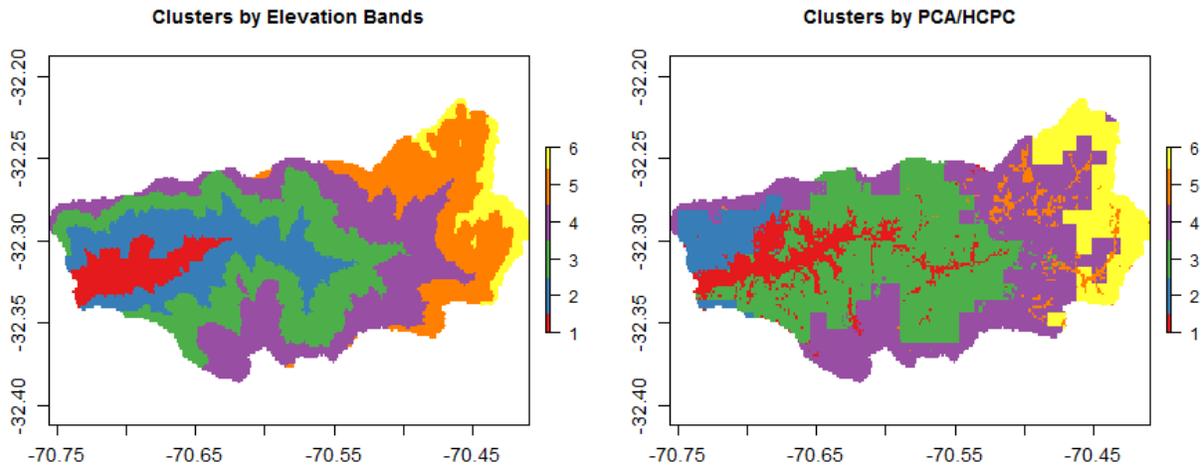
345 Figure 4 shows the proportion of the within variance relative to the total variance, where the
 346 major decrement of the variance occurs up to the case with four clusters and decreases until the
 347 case with seven or eight clusters. A decrease in the within variance means that the internal
 348 variability of each cluster decreases and hence the variance between clusters increases (Equation
 349 1). As this happens, the hydrologic behavior between HRUs is also expected to be more
 350 heterogeneous and simultaneously more homogeneous within each individual HRU, which is
 351 expected to lead in a better hydrologic modelling. Hence, it was expected that the optimal
 352 number of HRU for hydrological modeling is four. As described in the methodology, the
 353 Hierarchical Cluster Analysis was performed to produce ten scenarios with different number of
 354 HRUs partitioning the basin, varying between one (lumped model; HRU_1) to ten (HRU_10), to
 355 be tested in the hydrological model.



356
 357 *Figure 4. Relative within variance for each number of clusters.*

358 Figure 5 shows the distribution of cells in six equal elevation bands, roughly of 550m each (left
 359 plot), as generally used in WEAP as first step for separation of HRUs (Vicuña et al., 2010;
 360 Young et al., 2009) and six clusters following the methodology proposed in this work, named
 361 PCA/HCPC (right). This number of clusters was chosen because of the best hydrologic results
 362 (section 3.3). Cluster 6 is similar in shape with the highest elevation band as they are
 363 concentrated in the eastern part of the basin where the highest elevations are located. For other
 364 clusters, the figure shows a clear difference; for instance, cluster 2 in the PCA/HCPC
 365 methodology is concentrated in the northwestern part of the catchment, consistent with the high
 366 precipitation area identified in Figure 3a, which is not identified in the traditional elevation
 367 bands. The cluster with the lowest elevations (cluster 1) is not as regular as its corresponding
 368 elevation band, as this cluster seems to follow the riverbed and the flat riparian zone. Cluster 5
 369 seems to be concentrated in higher and colder areas with Andean vegetation and vegas,
 370 characterized by their high-water content or retention capacity, compared to the surroundings
 371 composed mainly by bare soil and disperse and small shrubs (cluster 4). Clusters still follow a
 372 tendency by elevation, as Mean Annual Temperature is the main variable composing the first
 373 dimension, but other variables tend to get importance as the number of clusters increase.

374



375

376

377

Figure 5. Elevations bands every 550m (left), as used in the traditional methodology, and HRU delimitations (right) for the simulation with six HRUs by the PCA/HPCPC methodology.

378

379

380

381

382

383

384

385

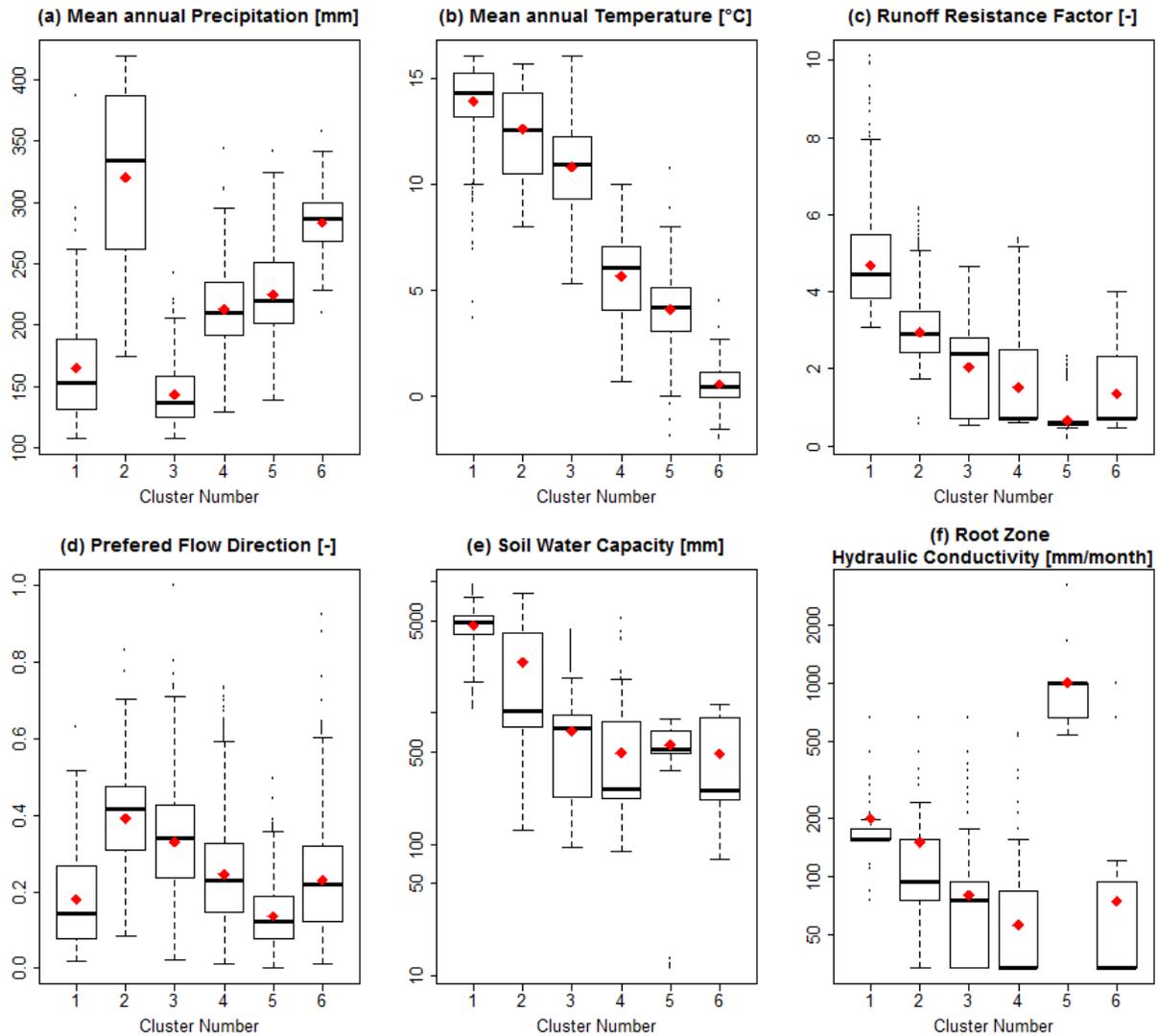
386

387

388

Figure 6 shows the boxplot for the values of six selected variables on each cell grouped by cluster to highlight the differences between them. It is possible to observe distinct characteristics between clusters: cluster 6 is the coldest cluster and has one of the highest precipitation rates; cluster 1 is the warmest, the lowest in altitude, very dry and the one having the deepest soil capacity, probably due to its location in the deepest part of the valley, where soils tend to be deeper and with higher runoff resistance factor, due to its flat terrain and vegetation. Cluster 2 is the one with highest precipitation. Cluster 5 is the one with highest hydraulic conductivity, due to the presence of marshes and wetlands. Clusters 3 and 4 are similar although cluster 4 has a mean value for precipitation of almost 50% more than cluster 3, and also have differences in the variables not shown. Clusters 6 and 4 have similar hydrologic parameters, but cluster 6 is colder and receives more rainfall.

389



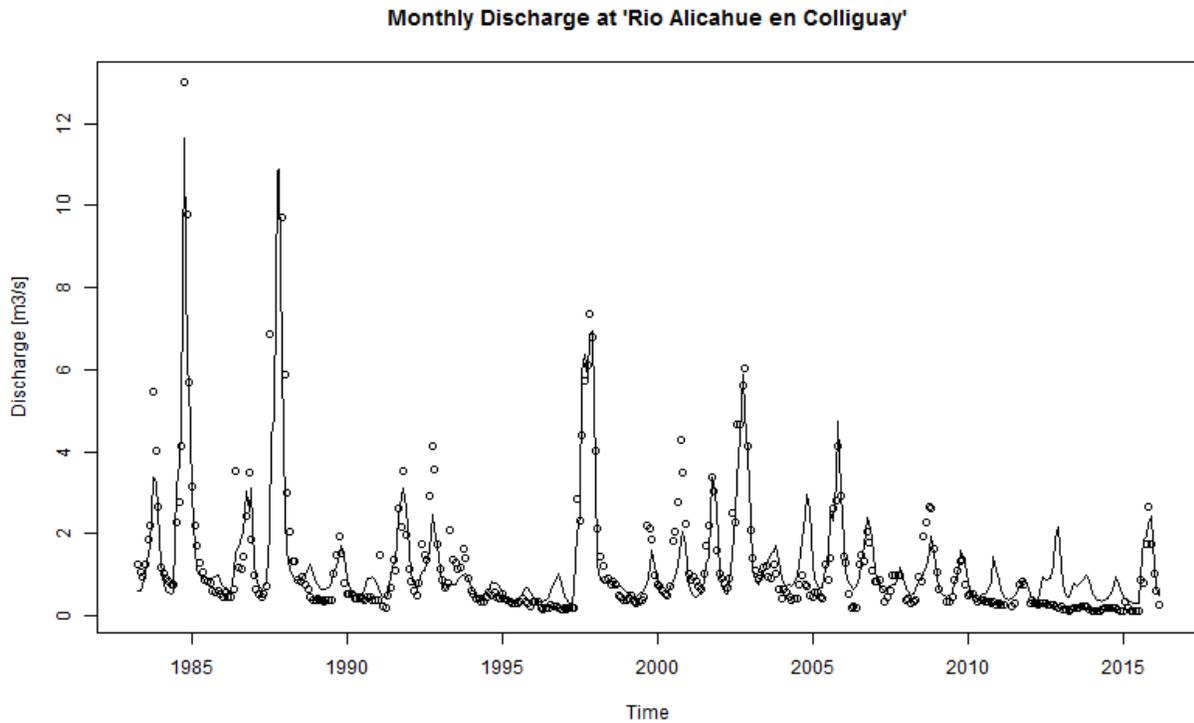
390
 391 *Figure 6. Boxplots per cluster in the simulation HRU_06 for: (a) Mean annual precipitation, (b) Mean temperature, (c) RRF, (d)*
 392 *f, (e) Sw and (f) Ks. Red circles represent the mean value of each cluster. Sw and Ks are plotted in log scale. Means are shown as*
 393 *red dots.*

394 3.2 Hydrological Modelling and HRUs contribution

395 The WEAP Model was ran using ten scenarios, each one with a different number of HRUs
 396 ranging from 1 to 10 (labeled as HRU_01 to HRU_10). Each simulation was run in a monthly
 397 time step starting from April 1979 until March 2016, following the Water Year commonly used
 398 in Chile, although the first four years were dismissed due to the warming of the model.

399 As the number of HRU increases and the level of spatial discretization is more detailed, also the
 400 model efficiency increases. For the first simulation with a lumped scheme, the NSE and RMSE
 401 (Equation 2 and Equation 3) values were 0.58 and 4.1% respectively and both indexes improved
 402 as more HRU were used. But for simulations with more than six HRUs, the extra clusters or
 403 HRU are not making any considerable improvement in the results, consistent with what was
 404 described in Haverkamp et al. (2005) and the model efficiency fluctuates in a plateau of 0.76-

405 0.79 while the RMSE near 3.1-3.2%. Figure 7 shows the observed and simulated monthly
 406 hydrograph for HRU_06.



407
 408 *Figure 7. Hydrograph of observed and modeled streamflow in 'Rio Alicahue en Colliguay' station for the simulation HRU_06.*
 409 *Observed discharge is shown in dots and the simulated discharge in a continuous line.*

410 Table 2 presents the differences between clusters in terms of inputs and responses and it is used
 411 to assess the different hydrological processes that each HRU represent. It shows the mean annual
 412 temperature, rainfall and elevation, the mean annual discharge and its standard deviation, the
 413 variation coefficient and the centroid of the annual flow volume as an index to measures timing
 414 of peak discharge in the season, calculated as a weighted average of the month and the discharge
 415 associated to each month (Young et al., 2009):

$$416 \text{Hydro}_{\text{Centroid}} = \frac{\sum_{i=1}^N \text{month_index}_i \cdot Q_{si}}{\sum_{i=1}^N Q_{si}}, \quad (4)$$

417
 418 , where January has index 1 and December, 12.

419 Although clusters 6 and 3 have similarities in total discharge, its annual variation, cluster 6
 420 groups most of the coldest cells in the basin where the snow melts late during the season and its
 421 discharge center of mass is in the middle of November and peaks in January, while cluster 3
 422 peaks in the middle of September, coinciding with its center of mass of the hydrograph. Both
 423 clusters are controlled by very different hydrological process and parameters.

424 Interest is on cluster 5, as it has a relatively small amount of area but its proportional contribution
 425 to the total discharge doubles its relative area. It peaks at the beginning of summer, has the lower
 426 Evaporation/Precipitation ratio and it is an example of an extreme hydrologic behavior that must
 427 be characterized and not dismissed. It is possible to argue that clusters 5 and 4 can be merged, as

428 their peak at the same time and they have the same elevation, but that decision depends on to
429 what extent it is possible to aggregate.

430 Clusters 1 and 2 are the lower in elevation, but the relative contribution of cluster 1 compared to
431 its relative area indicates that is the less important and its discharge is comparable in absolute
432 terms to cluster 2, although their areas are 34.0 and 22.1 km².

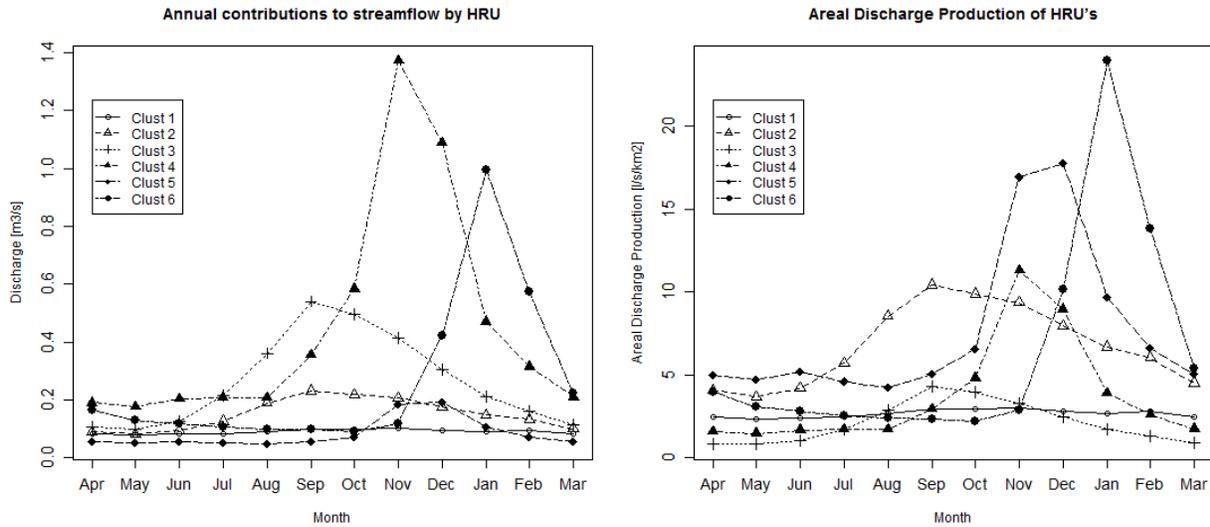
433

434 *Table 2. Summary of mean annual variables for the six clusters used in the hydrological simulation. The values correspond to the*
435 *mean values for the simulation period of 1984-2016.*

| | Cluster 1 | Cluster 2 | Cluster 3 | Cluster 4 | Cluster 5 | Cluster 6 |
|--|-----------|-----------|-----------|-----------|-----------|-----------|
| Area [Km²] | 34.0 | 22.1 | 125.2 | 121.7 | 10.9 | 41.5 |
| % over total area | 9.6% | 6.2% | 35.2% | 34.2% | 3.1% | 11.7% |
| Elevation [m] | 1483 | 1460 | 2063 | 2080 | 2837 | 2951 |
| Precipitation [mm] | 204 | 413 | 175 | 269 | 287 | 357 |
| Evapotranspiration [mm] | 126 | 203 | 110 | 154 | 50 | 161 |
| E/P [-] | 0.62 | 0.49 | 0.63 | 0.57 | 0.17 | 0.45 |
| Discharge | | | | | | |
| Mean [m3/s] | 0.09 | 0.15 | 0.26 | 0.45 | 0.08 | 0.26 |
| % over total discharge | 7% | 12% | 20% | 35% | 6% | 20% |
| Standard Deviation [m3/s] | 0.03 | 0.11 | 0.30 | 0.77 | 0.08 | 0.47 |
| Coefficient of Variation | 0.34 | 0.76 | 1.16 | 1.72 | 0.99 | 1.79 |
| Hydrograph Centroid [month index] | 9.65 | 9.85 | 9.79 | 10.57 | 10.35 | 11.49 |

436

437 Finally, Figure 8 presents the mean monthly discharge (left) and the mean monthly areal
438 discharge production (right) for each of the six clusters. Each cluster shows different
439 hydrographs in terms of total volume and peak timing, consistent with the goal of maximization
440 of the between-variability; this behavior is also seen in other scenarios with different number of
441 clusters, although not shown here. Cluster 4 is the main contributor to the annual discharge, and
442 it is clear its nival hydrologic regime with its peak in November (half Spring in Southern
443 Hemisphere), coinciding with the basin peak due to snowmelt. Clusters 1, 2 and 3 present
444 hydrographs with peaks in or near September, two months later of the precipitation peaks for this
445 region during austral winter, probably due to the firsts snowmelts but also from interflow
446 produced from rainfall that reacts slower than direct runoff. Cluster 1 is also the more stable in
447 terms of discharge, mainly due to the availability to hold water because of its larger soil water
448 capacity and cluster 3 presents a more distinct peak in the end of the winter probably due to the
449 first snowmelt but also rainfall and humidity leaving the upper part of the soil. Cluster 6 presents
450 the most retarded hydrograph peak in the season, explained by late melt of snow due to its
451 relatively higher mean elevation compared to the other HRUs, hence, the lowest values of mean
452 temperatures. Clusters 4 and 5 also present a nival regime as their peaks match with the snow
453 melting season, but its total volume is completely different as cluster 5 is explained by a
454 concentration in a relatively small area of marshes and Andes wetlands while cluster 4 shows the
455 biggest contribution to the total streamflow, mainly given by its high portion of area (34.2%).



456
457 *Figure 8. Mean monthly discharge of each HRU contributing to total streamflow at the outlet, for the six HRUs scenario. Left*
458 *panel: mean annual discharge. Right panel: Mean annual discharge production by area.*

459 The image at the right shows the production of discharge relative to the area of the HRU (in
460 l/s/km²). The relative importance of each HRU changes, especially for clusters 2, 5 and 6. As
461 shown in Table 2, the relative contribution to the total discharge of those clusters doubles their
462 area relative to the total area. The hydrologic regime of cluster 2 tends to be closer to the
463 precipitation season (May to August) and it has a high areal production of water due mainly to
464 the concentration of rainfall in that area of the catchment. Cluster 5 presents the higher average
465 of areal production (7.6 l/s/km²) and even its base value of near 5.0 l/s/km² is also higher than the
466 rest of the base values. Again, this may be explained by the nature of the vegetation covering
467 most of that cluster and by the slow release of water stored in them. The highest peak of 23.9
468 l/s/km² in the month of January corresponds to cluster 6 and it is a combination of high rates of
469 precipitation during winter in a relatively small area of the catchment, accumulating a massive
470 volume of snow with the rise of temperatures in Summer, producing the highest peak of
471 discharge per unit of area due to snowmelt.

472 **4 Discussion**

473 This paper presents a new methodology for HRU delineation based on the catchment attributes,
474 explained by the model parameters and climate variables. The units generated are expected to be
475 used in lumped and semi-distributed hydrological models where the topology of the elements
476 could be neglected. The methodology present two main steps: (i) a Principal Component
477 Analysis to reduce the number of variables while most variance is kept, and (ii) a Hierarchical
478 Clustering decomposition to delineate the HRUs with minimum internal variability but
479 maximum variability among the created units.

480 The methodology was tested on the Alicahue River Basin with the WEAP model which has a
481 hydrology module. The model was run under ten scenarios with different numbers of clusters
482 (HRUs) and evaluated using the Nash-Sutcliffe Efficiency Index and the Mean Squared Error.

483 4.1 Methodology and data uncertainties in the dataset preparation

484 Although the generation of the complete dataset used to derive the HRUs was not presented, the
485 estimation of the parameters and the climate dataset can be calculated independently using any
486 methodology or information previously available. The only two main characteristics that must be
487 preserved from such methodology are: (1) climate information needs to be from a gridded dataset
488 and (2) hydrologic parameters must be specific to the target model and calculated spatially prior
489 to the HRU delineation.

490 The main reason to use a WRF simulation, instead of longer and publicly available datasets, was
491 its high resolution (1 km) which is very important in regions with very complex topography as
492 the Alichahue basin. That resolution is much higher than Reanalysis (Kalnay et al., 1996) which is
493 usually about 0.5° and even higher than the newest local dataset CR2MET available at 0.05° for
494 the continental Chilean territory (Alvarez-Garreton et al., 2018; DGA, 2017). But it is important
495 to highlight the evident limitation on the case of study due to have only 3 years of WRF
496 simulations to represent the climate of the region.

497 4.2 Clustering method and Results

498 Once the data was pre-processed, it contained more than 17 thousand cells and each cell had
499 information on four hydrologic parameters and seven meteorological variables; moreover, in
500 other implementations this number could be even bigger. That amount of data had to be
501 summarized in order to be treatable, but at the same time, it was desirable that the aggregation
502 process carried most of the variability, without losing valuable information. The Principal
503 Component Analysis was chosen, as it selects orthogonal vectors whom carry much of the
504 information gathered in the previous process.

505 The PCA function uses weights to account for the relative importance of the variables. This gives
506 to the modeler the option to assess the most important variables given the model and/or the
507 problem to solve. In this study case, weights for rainfall and temperature were equal to 2 while
508 for all other variables it was set to 1. Temperature plays a crucial role controlling
509 evapotranspiration and snow melting, both main hydrological characteristics of an Andean semi-
510 arid basin, as the Alichahue basin in the study case and Precipitation controls the water income to
511 the basins and water simulations are highly sensitive to the amount of water used to model.
512 Sensitivity analysis of PCA to the weights was not performed, but it could help gain information
513 about the robustness of the clusters given different weights.

514 The within variance decrease obtained from the clustering process could be used as a criterion
515 for the selection of the optimal number of HRUs required to capture the main hydrological
516 behaviors in the target basin. This methodology allows to highly reduce the number of HRUs
517 involved in the simulation, decreasing the required computational time. This could favor, for
518 instance, studies with ensemble simulations, more exhaustive sensitivity analysis to some
519 parameters of the models and/or much longer (or higher temporal resolutions) simulations.

520 4.3 Discharge independence in the hydrologic modelling

521 The results by the hydrological modelling show in general a good match between observations
522 and simulations, even with the lumped scheme. As expected, the simulation with one HRU, as
523 the most lumped scheme, show the poorest results in terms of efficiency (NS=0.58). As the
524 number of HRU increases and the level of spatial discretization is more detailed, also the model

525 efficiency increases. The errors in the modelling can be explained, at least to a good extent, by
526 the uncertainties in the simple meteorological models used to derive the precipitation and the
527 temperature, the proposed relations to obtain the parameters maps and a possible lack of
528 representing all the hydrologic fluxes and storages in the hydrologic cycle in the basin given by
529 lack of soil information. Also the representation of extreme hydrologic phenomena is possible
530 only if the chosen model is capable of simulate these phenomena. If not, any discretization of
531 HRU methodology would be useless or at least not useful.

532 The streamflow at the outlet of Alicahue basin is controlled by a baseflow dominated by the
533 subsurface storage which is dependent in the storage capacity of soil and evapotranspiration
534 stress, a component driven by the winter rainfall dependent in hydraulic conductivity and rain
535 intensity, and a component driven by snow melting which is highly dependent in temperature
536 and elevation. From Figure 8, such behaviors were well captured for the different HRUs, which
537 allows to the modeler a better understanding of the underlying processes controlling the outlet
538 streamflow when compared to other methodologies for HRU delimitation.

539 Finally, it is important to note that the only variable for assessing the methodology was the river
540 discharge, which simplify the water cycle and all its components into one lumped criterion. It
541 would be advisable test the methodology and the hydrologic behavior with the rest of the
542 components of the hydrological cycle (infiltration, evapotranspiration, groundwater movement,
543 leakage, etc.), which was not possible in the case of study basin due to the lack of observations.

544 **5 Conclusions**

545 Flügel presented in 1983 the concept of HRU for the hydrological modelling. HRUs are the basic
546 units in which the equations controlled by parameters are run and meteorological data is used as
547 inputs. The basic assumptions of HRU is that each of them has a particular hydrological response
548 to rainfall, temperature and other climate data. Most of the actual methodologies account only
549 partially for the spatial variability that leads to differentiated response, particularly the spatial
550 climate variability within the basin is under- or misrepresented.

551 This paper presented a methodology for the determination of HRU, more consistent with the
552 classical definition, based on hydrological parameters (specifics to the target model) and
553 meteorological inputs; using Principal Component Analysis and Hierarchical Clustering to
554 minimize the global internal variability in each HRU and that at the same time maximizes the
555 variability among HRUs. This procedure is intended to generate different responses by each unit,
556 as defined by the modeled hydrograph, minimizing the number of required HRUs to capture the
557 internal variability.

558 The application of the methodology was assessed in the Alicahue river basin, a small basin
559 located in a semi-arid and mountainous region in Central Chile, with altitudes ranging from the
560 780 to almost 4000 meters above sea level. Results of the WEAP simulations shows a good
561 agreement between modeled and observed streamflow at the outlet, with scores comparable to
562 other studies using the same model in similar basins.

563 Better hydrological parameters and meteorological datasets could still improve the model
564 efficiency. Future research is to test the methodology in other basins with different hydrologic
565 regime and using different models. WEAP is suitable for time steps longer than one day, but the
566 methodology can be used in other long-term models or even in storm models, considering other
567 parameters sensible to the basin response (i.e. concentration time, curve number, etc.).

568 In summary, the main advantages of the proposed methodology are:

569 Computational efficiency in the hydrological simulations. As the methodology is
570 designed to minimize the required numbers of HRU to account for most of the spatial variability
571 in the climate and hydrological parameters (the main controllers of the hydrological response),
572 the computational effort is highly reduced as usually it is linear in the number of HRUs. In the
573 study case only six HRUs were necessary to achieve similar scores than those from more
574 commons methodologies that use several tens of HRUs.

575 Basin heterogeneity better captured. As the PCA captures most of the variability of the
576 parameters and climate variables, heterogeneous conditions are kept even after the reduction of
577 the number of variables used as input to the cluster analysis. Also, the hierarchical clustering
578 process ensure the delineation of the HRUs is completely driven by such variability and not by
579 arbitrary choices. For instance, in the study case, one of the HRUs correspond to a small and
580 disjoint area that has a relatively large contribution to the total streamflow, which would be
581 probably neglected with most of the traditional methodologies.

582 Better identifiability of the HRUs. As the HRU delineation was driven by the
583 minimization of the within variance in each HRU and at the same time maximization of the
584 variance between HRUs, the hydrological response is expected to be different for each HRU
585 with minimum redundancy. This will allow to the modeler to gain a better understanding of the
586 underlying hydrological behaviors that controls the response of the basin. For instance, each
587 HRU in the case of study was identified with different processes, including baseflow, quick
588 rainfall-runoff response, snow melting at different times associated with elevation and
589 temperature differences.

590 **6 Acknowledgments, Samples, and Data**

591 Supplementary data is available at DOI: 10.17632/ppgtgvyttm.2.

592 **7 References**

- 593 Alvarez-Garreton, C., Mendoza, P. A., Pablo Boisier, J., Addor, N., Galleguillos, M., Zambrano-
594 Bigiarini, M., ... Ayala, A. (2018). The CAMELS-CL dataset: catchment attributes and
595 meteorology for large sample studies-Chile dataset. *Hydrol. Earth Syst. Sci.*, 22, 5817–5846.
596 <https://doi.org/10.5194/hess-22-5817-2018>
- 597 Bonelli, S., Vicuña, S., Meza, F. J., Gironás, J., & Barton, J. (2014). Incorporating climate
598 change adaptation strategies in urban water supply planning: The case of central Chile.
599 *Journal of Water and Climate Change*, 5(3), 357–376.
600 <https://doi.org/10.2166/wcc.2014.037>
- 601 Dehotin, J., & Braud, I. (2008). Which spatial discretization for distributed hydrological models?
602 Proposition of a methodology and illustration for medium to large-scale catchments.
603 *Hydrology and Earth System Sciences*, 12(3), 769–796. <https://doi.org/10.5194/hess-12-769-2008>
- 604
- 605 DGA. (2017). *Actualización del Balance Hídrico Nacional, SIT N° 417*. Santiago. Retrieved
606 from <http://documentos.dga.cl/REH5796v1.pdf>
- 607 Flügel, W.-A. (1995). Delineating hydrological response units by geographical information
608 system analyses for regional hydrological modelling using PRMS/MMS in the drainage

- 609 basin of the River Bröl, Germany. *Hydrological Processes*, 9(3–4), 423–436.
610 <https://doi.org/10.1002/hyp.3360090313>
- 611 Fouedjio, F. (2016). A hierarchical clustering method for multivariate geostatistical data. *Spatial*
612 *Statistics*, 18, 333–351. <https://doi.org/10.1016/j.spasta.2016.07.003>
- 613 Garreaud, R. (1992). Impact of the variability of snowline in winter discharge peaks in basins of
614 mixed regime in Central Chile (in spanish). Sociedad Chilena de Ingeniería Hidráulica.
- 615 Haghnegahdar, A., Tolson, B. A., Craig, J. R., & Paya, K. T. (2015). Assessing the performance
616 of a semi-distributed hydrological model under various watershed discretization schemes.
617 *Hydrological Processes*, 29(18), 4018–4031. <https://doi.org/10.1002/hyp.10550>
- 618 Han, J.-C., Huang, G.-H., Zhang, H., Li, Z., & Li, Y.-P. (2014). Effects of watershed subdivision
619 level on semi-distributed hydrological simulations: case study of the SLURP model applied
620 to the Xiangxi River watershed, China. *Hydrological Sciences Journal*, 59(1), 108–125.
621 <https://doi.org/10.1080/02626667.2013.854368>
- 622 Haverkamp, S., Fohrer, N., & Frede, H. G. (2005). Assessment of the effect of land use patterns
623 on hydrologic landscape functions: A comprehensive GIS-based tool to minimize model
624 uncertainty resulting from spatial aggregation. *Hydrological Processes*, 19(3), 715–727.
625 <https://doi.org/10.1002/hyp.5626>
- 626 Hijmans, R. J. (2017). Geosphere: Spherical Trigonometry. R package version 1.5-7.
627 <https://CRAN.R-project.org/package=geosphere>.
- 628 Höge, M., Wöhling, T., & Nowak, W. (2018). A Primer for Model Selection: The Decisive Role
629 of Model Complexity. *Water Resources Research*, 54(3), 1688–1715.
630 <https://doi.org/10.1002/2017WR021902>
- 631 Husson, F., Julie, J., & Jérôme, P. (2010). Technical Report – Agrocampus Principal component
632 methods -hierarchical clustering -partitional clustering: why would we need to choose for
633 visualizing data? *Applied Mathematics Department Agrocampus*.
- 634 Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., ... Joseph, D.
635 (1996). The NCEP/NCAR 40-Year Reanalysis Project. *Bulletin of the American*
636 *Meteorological Society*, 77(3), 437–471. [https://doi.org/10.1175/1520-0477\(1996\)077<0437:TNYRP>2.0.CO;2](https://doi.org/10.1175/1520-0477(1996)077<0437:TNYRP>2.0.CO;2)
- 638 Lê, S., Josse, J., & Husson, F. (2008). FactoMineR: An R Package for Multivariate Analysis.
639 *Journal of Statistical Software*, 25(1).
- 640 Leavesley, G. H., Lichty, R. W., Troutman, B. M., & Saindon, L. G. (1983). Precipitation-runoff
641 modeling system; user's manual. <https://doi.org/10.3133/wri834238>
- 642 Martínez, E., Flores, J., Retamal, M., Ahumada, I., & Brito, S. (2013). *Informe técnico final*
643 *monitoreo de cambios, corrección cartográfica y actualización del catastro de bosque*
644 *nativo en las Regiones de Valparaíso, Metropolitana y Libertador Bernardo O'Higgins*.
- 645 Nijzink, R., Hutton, C., Pechlivanidis, I., Capell, R., Arheimer, B., Freer, J., ... Hrachowitz, M.
646 (2016). The evolution of root-zone moisture capacities after deforestation: a step towards
647 hydrological predictions under change? *Hydrol. Earth Syst. Sci*, 20, 4775–4799.
648 <https://doi.org/10.5194/hess-20-4775-2016>

- 649 Orth, R., Dutra, E., & Pappenberger, F. (2016). Improving weather predictability by including
650 land surface model parameter uncertainty. *Monthly Weather Review*, *144*(4), 1551–1569.
651 <https://doi.org/10.1175/MWR-D-15-0283.1>
- 652 Pilz, T., Francke, T., & Bronstert, A. (2017). LumpR 2.0.0: An R package facilitating landscape
653 discretisation for hillslope-based hydrological models. *Geoscientific Model Development*,
654 *10*(8), 3001–3023. <https://doi.org/10.5194/gmd-10-3001-2017>
- 655 Purkey, D. R., Joyce, B., Vicuña, S., Hanemann, M. W., Dale, L. L., Yates, D., & Dracup, J. a.
656 (2007). Robust analysis of future climate change impacts on water for agriculture and other
657 sectors: a case study in the Sacramento Valley. *Climatic Change*, *87*(S1), 109–122.
658 <https://doi.org/10.1007/s10584-007-9375-8>
- 659 R Core Team. (2019). R: A Language and Environment for Statistical Computing. Vienna,
660 Austria. Retrieved from <https://www.r-project.org/>
- 661 Savvidou, E., Efstratiadis, A., Koussis, A. D., Koukouvinos, A., & Skarlatos, D. (2018). The
662 curve number concept as a driver for delineating Hydrological Response Units. *Water*
663 (*Switzerland*), *10*(2). <https://doi.org/10.3390/w10020194>
- 664 Skamarock, W. C., Klemp, J. B., Dudhia, J., Gill, D. O., Barker, D. M., Duda, M. G., ... Powers,
665 J. G. (2008). *A Description of the Advanced Research WRF Version 3*. Retrieved from
666 <http://opensky.ucar.edu/islandora/object/technotes%3A500/datastream/PDF/view>
- 667 Tachikawa, T., Kaku, M., Iwasaki, A., Gesch, D. B., Oimoen, M. J., Zhang, Z., ... Carabajal, C.
668 (2011). ASTER global digital elevation model version 2 - summary of validation results.
669 NASA Land Processes Distributed Active Archive Center. Downloaded at:
670 http://www.jspacesystems.or.jp/ersdac/GDEM/ver2Validation/Summary_GDEM2_validation_report_final.pdf, (March 2015).
671
- 672 Vicuña, S., Garreaud, R. D., & McPhee, J. (2010). Climate change impacts on the hydrology of a
673 snowmelt driven basin in semiarid Chile. *Climatic Change*, *105*(3–4), 469–488.
674 <https://doi.org/10.1007/s10584-010-9888-4>
- 675 Yates, D., Sieber, J., Purkey, D., & Huber-Lee, A. (2005). 21 - A Demand-, Priority-, and
676 Preference-Driven Water Planning Model. Part 1: Model Characteristics. *Water*
677 *International* *30*, 487-500., *30*(4), 487–500.
- 678 Young, C. A., Escobar-Arias, M. I., Fernandes, M., Joyce, B., Kiparsky, M., Mount, J. F., ...
679 Yates, D. (2009). Modeling the hydrology of climate change in California's Sierra Nevada
680 for subwatershed scale adaptation. *Journal of the American Water Resources Association*
681 (*JAWRA*), *45*(6), 1409–1423. <https://doi.org/10.1111>
- 682