



HAL
open science

Decomposition of 2D polygons and its effect in hydrological models

P. Sanzana, J. Gironas, Isabelle Braud, N. Hitschfeld, F. Branger, F. Rodriguez, M. Fuamba, J. Romero, X. Vargas, J.F. Munoz, et al.

► **To cite this version:**

P. Sanzana, J. Gironas, Isabelle Braud, N. Hitschfeld, F. Branger, et al.. Decomposition of 2D polygons and its effect in hydrological models. *Journal of Hydroinformatics*, 2019, 1, pp.104-122. 10.2166/hydro.2018.031 . hal-02608054

HAL Id: hal-02608054

<https://hal.inrae.fr/hal-02608054v1>

Submitted on 16 May 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

1 **Decomposition of 2D polygons and its effect in hydrological models**

2 [Short title: Decomposition of 2D polygons and its effect in hydrological
3 models]

4 P. Sanzana, J. Gironás, I. Braud, N. Hitschfeld, F. Branger, F. Rodriguez, M.
5 Fuamba, J. Romero, X. Vargas, J. F. Muñoz, S. Vicuña, and A. Mejía

6 **P. Sanzana** (corresponding author)

7 **J. Gironás**

8 **J. F. Muñoz**

9 **S. Vicuña**

10 Departamento de Ingeniería Hidráulica y Ambiental, Pontificia Universidad
11 Católica de Chile, Avenida Vicuña Mackenna 4860, Santiago, Chile

12 E-mail: ppsanzana@ing.puc.cl

13 **P. Sanzana**

14 **J. Gironás**

15 Centro de Desarrollo Urbano Sustentable CONICYT/FONDAP/15110020,
16 Avenida Vicuña Mackenna 4860, Santiago, Chile

17 **J. Gironás**

18 **S. Vicuña**

19 Centro de Investigación para la Gestión Integrada de Desastres Naturales
20 CONICYT/FONDAP/15110017, Avenida Vicuña Mackenna 4860, Santiago,
21 Chile

22 **J. Gironás**

23 **S. Vicuña**

24 Centro Interdisciplinario de Cambio Global, Avenida Vicuña Mackenna 4860,
25 Santiago, Chile

26 **P. Sanzana**

27 **I. Braud**

28 **F. Branger**

29 IRSTEA, UR RiverLy, Centre de Lyon-Villeurbanne, 5 rue de la Doua, CS
30 20244, 69625 Villeurbanne, France

31 **N. Hitschfeld**

32 **J. Romero**

33 Computer Science Department, Facultad de Ciencias Físicas y Matemáticas,
34 Universidad de Chile. Av. Blanco Encalada 2120, Santiago, Chile

35 **F. Rodriguez**

36 LUNAM Université, Institut Français des Sciences et Technologies des
37 Transports de l'Aménagement et des Réseaux (IFSTTAR), Département
38 Géotechnique Eau et Risques et Institut de Recherche en Sciences et
39 Techniques de la Ville (IRSTV), F-44341 Bouguenais, France

40 **M. Fuamba**

41 Department of Civil, Geological, and Mining Engineering, Polytechnique
42 Montréal, Montréal, Canada

43 **X. Vargas**

44 Departamento de Ingeniería Civil, Facultad de Ciencias Físicas y
45 Matemáticas, Universidad de Chile, Blanco Encalada 2002, Santiago, Chile

46 **A. Mejía**

47 Department of Civil and Environmental Engineering, The Pennsylvania State
48 University, University Park, State College, PA 16802, USA

49 **ABSTRACT**

50 2D non-uniform polygonal meshes allow representing the impact of landscape
51 elements and small infrastructures on water flows. The initial vectorial mesh,
52 derived from intersection of several geographical information systems' layers,
53 can have highly non-convex or sliver polygons. These bad-shaped elements
54 compromise accurate numerical flow computation. We propose a flexible

55 divide-and-conquer strategy to decompose polygons into physiographical
56 meaningful parts using shape descriptors to better represent the surface terrain
57 and hydrologic connectivity. We use the convexity index (CI) and the form
58 factor (FF) to consider convex and square like optimum shapes. The strategy
59 was applied to two peri-urban areas whose hydrologic response was simulated
60 using distributed modeling. Good-quality meshes were generated with
61 threshold values of $CI \approx 0.8$ and $FF \approx 0.2$, and $CI \approx 0.95$ and $FF \approx 0.4$ for
62 undeveloped and highly urbanized zones, respectively. We concluded the mesh
63 segmentation facilitates the representation of the spatially distributed
64 processes controlling not only the lumped response of the catchment, but the
65 spatial variability of water quantity and fluxes within it at medium and small
66 scales.

67 **Key words** | peri-urban features, polygonal decomposition, spatial uncertainty
68 in hydrological model, terrain representation

69 **GLOSSARY**

70 **Simple polygon:** polygon delimited by a continue and close polyline in which
71 each pair of edges only intersect at one vertex. The number of vertexes equals
72 the number of edges.

73 **Regular polygon:** an n-sided simple polygon with equilateral edges and
74 equiangular interior angles.

75 **Irregular polygon:** an n-sided simple polygon with non-equilateral edges or
76 non-equiangular interior angles.

77 **Non-uniform mesh:** a mesh composed of irregular polygons of variable size.

78 **Sliver polygon:** an elongated polygon with small area. In this article, a sliver
79 polygon usually represents roads, footpaths, median strips, or narrow green
80 areas.

81 **Well-shaped polygon:** a simple polygon that fulfills geometrical constraints
82 given by threshold values of one or more shape descriptors such as the
83 convexity index or the form factor.

84 **Bad-shaped polygon:** a not well-shaped polygon.

85 **Initial mesh:** output mesh from the initial segmentation step.

86 **Initial segmentation step:** procedure in which main GIS layers are intersected
87 to generate the initial mesh.

88 **Good-quality mesh:** mesh composed just of well-shaped polygons. The
89 concept of good-quality depends on geometrical constraints. In our case, the
90 geometrical constraints are given by numerical restrictions of the distributed
91 hydrological model, and by a good representation of hydrological connectivity
92 (i.e., a representative drainage network).

93 **Pseudo-convex polygon:** a simple polygon with a convexity index value
94 larger than a defined threshold.

95 **Pseudo-square polygon:** a simple polygon with a form factor value larger
96 than a defined threshold.

97 **Pseudo-regular polygon:** a simple polygon with a shape metric value larger
98 than a defined threshold.

99 **Meaningful parts:** the accepted notion of meaningful parts relies on human
100 perception, and is based on the observation that human vision defines part
101 boundaries along negative minima of principal curvatures (Krayevoy &
102 Sheffer 2006). Examples of meaningful parts are hands and the neck in a
103 human body shape.

104 **Meaningful physiographic unit:** any natural or urban feature with
105 homogeneous properties (i.e., same land use and pedo-topo-geological
106 properties).

107 **Shape factor:** dimensionless metric used to describe the shape of a 2D
108 polygon, which is independent of its size.
109 **Hydrological response unit (HRU):** a simple polygon with hydrological
110 homogeneous properties (i.e., hydrological properties whose variation within
111 an HRU is small compared to that among neighboring HRUs (Flügel 1995)).
112 **Urban hydrological elements (UHE):** a simple polygon composed of a
113 cadastral parcel and its corresponding portion of street (Rodriguez *et al.* 2008).
114 **Drainage network:** connectivity structure among the hydrological elements
115 within the catchment contributing to the channelized system. It includes
116 overland path flows, streams, ditches, and sewers.

117 INTRODUCTION

118 The resolution of spatial discretization in hydrological models is directly
119 governed by the objectives of the modeling, the hydrological processes to be
120 represented, and data quality and availability (Dehotin & Braud 2008). For
121 planning, analysis, and design of sustainable urban and peri-urban hydro-
122 landscapes at small scales, high resolutions are highly desirable. Moreover,
123 hydrological models designed to study the hydrology of urban and peri-urban
124 catchments require irregular meshes representing spatial features of highly
125 variable sizes, such as gardens, infiltration strips, and detention ponds
126 (Lagacherie *et al.* 2010; Abily *et al.* 2013; Jankowfsky *et al.* 2014).

127 Landscape features such as agricultural and urban boundaries, streets,
128 buildings, tree lines and hedges, as well as the main channelized
129 infrastructure, can be represented by any 2D vectorial mesh. Normally,
130 hydrological models applied in urban environments use 2D vectorial meshes
131 composed of sub-catchments, such as SWMM (Rossman 2009) and
132 StormCAD (Haestad Methods 1995). Moreover, 2D hydraulics models for
133 computing flooding areas use regular grid and triangular (e.g., MIKE 21 and

134 MIKE 21 FM (DHI 2007a, 2007b)), or non-uniform hex-dominant meshes
135 (3D FVM (Versteeg & Malalasekera 2007)). The intersection of vectorial
136 layers leads to an initial model mesh composed of polygons with
137 homogeneous hydrological properties (i.e., with the same land use and pedo-
138 topo-geological properties) that are representative of the terrain. This mesh
139 preserves all the artificial boundaries at small scales, but contains lots of
140 polygons with irregular shape that must be improved prior to implementing
141 the hydrological model, as they may cause some numerical problems (Moussa
142 *et al.* 2002; Zundel *et al.* 2002; Rodriguez *et al.* 2008; Jankowsky *et al.*
143 2014). Indeed, some equations in hydrological models are based on specific
144 geometrical characteristics that may have erroneous values when computed
145 using bad-shaped polygons (i.e., polygons that do not fulfill certain
146 geometrical constraints). Examples of these equations are the ones used in the
147 PUMMA model (Jankowsky *et al.* 2014) to estimate: (1) lateral fluxes
148 between two polygons using Darcy's law; (2) fluxes between hydrological
149 units (represented by polygons) and rivers (represented by polylines)
150 considering the Dupuit–Forchheimer hypothesis (Miles 1985); and (3) surface
151 fluxes between polygons using the Manning equation, where the flux depends
152 on the distance between the centroids of polygons and/or polylines that may
153 be erroneous for non-convex polygons (see examples in Figure 1). All these
154 drawbacks have motivated the development of new technologies to represent
155 the terrain using good quality meshes. Similar advances have been
156 implemented to discretize complex geometries in other Earth sciences areas,
157 such as earthquakes (Shewchuck 1996) and hydrogeology (Panday *et al.*
158 2013).

159 **Appropriate meshes for representation of varying size elements**

160 Building an appropriate mesh for representing spatial domains and obtaining
161 meaningful results when using numerical models is essential. The main
162 differences among commonly used meshes are related to the type of elements
163 that can be used and the definition of a well- or bad-shaped element. For finite
164 elements' methods, triangulations and quadrilateral meshes are the most
165 common (Kim & Chung 2014; Russo *et al.* 2015). The classical triangulation
166 fulfills the Delaunay condition and has been implemented in widely used
167 algorithms such as Triangle in 2D (Shewchuck 1996) and TetGen in 3D (Si
168 2006). Typical quality criteria used in these algorithms are that the minimum
169 angle of each triangle must be greater than α – or the edge-ratio must be smaller
170 than β – a threshold value. In the case of quadrilateral meshes (Lee *et al.* 2003),
171 common approaches are based on off-setting techniques, while classical
172 quality criteria consider the convexity of the quadrilaterals and a maximum
173 possible angle. In recent years, the polygonal finite element method based on
174 polygonal meshing has become more popular (Panday *et al.* 2013). Polygonal
175 meshing allows the generation of a mesh of variable resolution, in which high
176 and low point densities are combined using graded transitions. The most used
177 polygonal mesh is that composed of polygons defined by the Voronoi regions,
178 which can be generated by building the dual-graph of a Delaunay mesh. The
179 Voronoi regions are convex polygons and commonly have six sides on average
180 (O'Rourke 1998). The constrained Delaunay triangulation (DT) algorithm
181 respects the polygon boundary without inserting new points. The Delaunay
182 condition may not be accomplished by the triangles containing the boundary
183 of the initial polygon. On the other hand, the conforming Delaunay
184 triangulation (DT) algorithm inserts points so that all triangles fulfill the
185 Delaunay condition. This method may be used together with some geometrical
186 criteria involving angles, area, etc.

187 Triangulations (Tucker *et al.* 2001), quadrilateral (grid-cells) meshes
188 (Downer *et al.* 2016), and polygonal meshes based on Voronoi regions (Collon
189 *et al.* 2015) can also be used for terrain representation and hydrological
190 modeling. However, regular polygons such as square grid cells are not
191 adequate to represent the original shape of hydrological elements such as
192 residential lots, agriculture fields, or green areas, nor for representing variable
193 shape boundaries of urban and peri-urban features (i.e., streets, footpaths,
194 hedges). They can only be used with elements bigger than the regular polygon
195 size (i.e., features smaller than the regular raster's cell size cannot be explicitly
196 represented). Other regular polygons (e.g., equiangular and equilateral sides
197 polygons) may represent such complex environments, but the number of final
198 polygons can increase heavily. High resolution 2D runoff modeling based on
199 triangles requires mesh refinement to improve the representation of urban
200 environments' features. This is a demanding task and it is difficult to build
201 models with discretization refinement using standard tools, as time-consuming
202 hand-made operations are still required (Abily *et al.* 2013). In addition, such
203 models can be deployed on small areas of a few hectares, but are
204 computationally too expensive for catchments of a few square kilometers.
205 There is therefore an interest to rely on bad-shaped polygons to better
206 represent the complexity of urban and peri-urban catchments, in order to
207 reduce significantly the number of elements in the final mesh and the
208 computational time of the hydrological models. Nevertheless, following such
209 an approach requires the final mesh to respect criteria ensuring numerical
210 stability and meaningfulness of the geometrical characteristics. Recently, such
211 non-uniform meshes composed of units representing terrain elements have
212 been used in distributed hydrological models such as MHYDAS (Moussa *et*
213 *al.* 2002), URBS (Rodriguez *et al.* 2008) and PUMMA (Jankowsky *et al.*

214 2014), which solve surface, sub-surface, and river flows. Furthermore, GIS-
215 tools such as Geo-MHYDAS (Lagacherie *et al.* 2010), GRIDGEN (Lien *et al.*
216 2015), LumpR (Pilz *et al.* 2017), and Geo-PUMMA (Sanzana *et al.* 2017)
217 have been developed for the implementation of these meshes, which are the
218 focus of the present paper.

219 **Current technologies in polygonal decomposition**

220 The decomposition of 2D shapes into meaningful parts is a key process in
221 human image understanding, and geometrical computing areas (Lien & Amato
222 2005; Liu *et al.* 2014). The accepted notion of meaningful parts relies on
223 human perception, and is based on the observation that human vision defines
224 part boundaries (e.g., hands and neck in a human body shape) (Krayevoy &
225 Sheffer 2006). Any non-convex polygon can be decomposed into sub-set parts
226 approximatively convex, which are commonly associated with meaningful
227 parts. Several efficient strategies to decompose polygons into approximatively
228 convex pieces, such as the DuDe (Liu *et al.* 2014) and ACD (Lien & Amato
229 2005) algorithms have been developed and applied to decompose silhouette-
230 based features. Unfortunately, the convexity criterion may not be enough for
231 terrain representation using irregular meshes for hydrological models. Many
232 models use a terrain representation based on hydrological response units
233 (HRUs) (Flügel 1995) and urban hydrological elements (UHEs) (Rodríguez *et*
234 *al.* 2008), meaningful physiographic units of the natural and urban landscapes
235 with homogeneous properties, respectively. These units can be highly irregular
236 and bad-shaped, and they eventually need to be decomposed into meaningful
237 small pieces. For example, there can be sliver polygons (i.e., elongated
238 polygons with small area) perfectly convex (i.e., long streets, linear hedges,
239 footpaths, median strips, or narrow green areas) but not suitable for flow
240 routing. A sliver polygon can work as an artificial wall restricting flow

241 routing, because its elevation – which may not be representative of the entire
242 area involved – can be such that flow from lower adjacent polygons is
243 restricted (Jankowsky 2011). Hence, the criteria of decomposition of HRUs
244 must consider not only convexity but other factors such as form factor or
245 shape factor, circularity, elongation ratios, and compactness coefficient. These
246 criteria can be implemented through tools using a divide-and-conquer
247 approach able to adapt to a variety of geometrical criteria. An example of such
248 a tool is Geo-PUMMA (Sanzana *et al.* 2017), a semi-automatic vectorial
249 toolbox to represent urban and peri-urban small catchments (0.1–10 km²) and
250 the hydrological connectivity among their components.

251 **Contribution**

252 This paper presents and assesses in detail a flexible divide-and-conquer
253 strategy decomposition for 2D polygons for terrain representation and
254 hydrological routing. The strategy is based on methods partially described
255 previously in the literature. In particular, Sanzana *et al.* (2013) developed tools
256 to improve HRUs considering (1) the high heterogeneity in HRUs' properties
257 derived from a digital elevation model, (2) the existence of concave polygons
258 or polygons with holes inside, (3) the existence of very large polygons, and (4)
259 bad estimations of HRUs' perimeters and distances. Recently, Sanzana *et al.*
260 (2017) developed Geo-PUMMA, a GIS-based tool to generate good quality
261 polygonal meshes composed of HRUs and UHEs for urban and peri-urban
262 catchments. These meshes are composed of the smallest number of properly
263 interconnected polygons (i.e., without spurious pits or unrealistic flow paths),
264 with homogeneous hydrological properties that are representative of the
265 terrain and ensure the efficient application of hydrological models. Based on
266 these previous results, the main contributions of this paper are:

- 267 • The generalization of the work of Sanzana *et al.* (2017) through the
268 proposition and description of a flexible divide-and-conquer strategy for
269 decomposing bad-shaped 2D polygons into smaller meaningful
270 components. This strategy is tested using both urban and peri-urban
271 terrains and silhouette-based features, which exemplifies its potential
272 application beyond the field of hydrology.
- 273 • The assessment of the computational complexity of the algorithms
274 involved in the strategy.
- 275 • An in-depth analysis of the impact of threshold values proposed for the
276 involved shape descriptors, and recommendations of empirical values
277 for these descriptors to be used in the decomposition of bad-shaped
278 polygons.
- 279 • A specific assessment of the effect of 2D polygon decomposition in
280 hydrological models.

281 **BASIC CONCEPTS**

282 The drainage network is a primary input for hydrological models. Generally,
283 this concept refers to the network of pipes and streams conveying flow in a
284 catchment to the outlet. In this paper, this concept refers to the whole
285 connectivity structure among the hydrological elements within the catchment
286 contributing to the channelized system (streams, ditches, and sewers). The
287 drainage network is extracted from a hydrological model mesh representative
288 of the terrain, and the natural and urban features. From now on, we follow the
289 approach by Sanzana *et al.* (2017), which considers a hydrological model
290 mesh to be composed of HRUs and UHEs.

291 The initial mesh is obtained from the direct intersection of GIS layers,
292 such as land use, soil type, geology, sub-catchment, UHEs, and channelized
293 network. The resulting hydrological mesh is formed by bad-shaped polygons,

294 which are more suitable than regular grids for representing human-made
295 features affecting significantly the hydrological processes in a catchment
296 (Lagacherie *et al.* 2010). However, vector-based HRU model meshes come
297 also with specific constraints, depending on the distributed hydrological model
298 to be applied. On the other hand, due to the urban design of the lots, typical
299 UHEs are polygons with regular shape (normally pseudo convex or not sliver).
300 Nevertheless, any UHE with irregular shape can be decomposed with the tools
301 described in this paper.

302 In distributed hydrological meshes, bad-shaped HRUs cause problems
303 when defining the topology of the drainage network (Jankowsky 2011).
304 Typically, the distance between the polygons' centroids and their boundaries is
305 used to calculate flowpath lengths among HRUs, and overland and subsurface
306 flows. This distance is meaningless if the centroid is outside of the polygon,
307 and a modification of the HRUs becomes necessary. Moreover, the HRUs can
308 greatly vary in size, which produces problems in drainage network extraction
309 and the stability of numerical schemes (for instance, by inducing unrealistic
310 water depths on small polygons downstream of larger ones). Thus, the
311 segmentation of very large polygons is recommended to obtain a mesh
312 composed by similar size polygons and avoid flows from big polygons to
313 small ones in hydrologic modeling.

314 Good-quality meshes are composed of well-shaped polygons that are
315 physiographically homogeneous. This representation allows the identification
316 of the hydrologic connectivity among them defined by the terrain. In this
317 paper, a well-shaped HRU is defined as a hydrologically homogeneous, not-
318 sliver and pseudo-convex polygon. Good-quality meshes do not result in the
319 problems previously identified, as they are composed of convex elementary
320 units with the same number of vertexes and boundaries, whose centroids are

321 located inside the polygon. Figure 1 shows examples of bad-shaped HRUs
322 from Estero El Guindo, a peri-urban catchment located in the piedmont of
323 Santiago, Chile (Sanzana *et al.* 2017), and the Mercier and Chaudanne, two
324 peri-urban catchments located in Lyon, France (Sanzana *et al.* 2013). Irregular
325 shapes in this figure include sliver HRUs representing streets (Figure 1(a)),
326 median strips (Figure 1(b)), walking paths (Figures 1(c) and 1(e)), a hedgerow
327 (Figure 1(d)), highly non-convex HRUs corresponding to lots' division
328 (Figures 1(f), 1(h) and 1(i)), and a particular green area (Figure 1(g)). Each
329 feature in Figure 1 plays a relevant role in the water balance and hydrological
330 routing, so preserving their original shape while segmenting them in pseudo-
331 regular pieces becomes relevant.



332

333 **Figure 1** | Example of bad-shaped shape HRUs of urban and peri-urban
334 landscapes.

335 **Shape descriptors**

336 Several basin shape descriptors or metrics have been proposed in the literature
337 to characterize the geomorphology of basins. Either these metrics or the

338 geometric properties used in their calculation have been used to estimate the
 339 hydrological response of a basin from its geomorphology or its time of
 340 concentration, which corresponds to a relevant response parameter used in
 341 hydrologic modeling. Examples of these metrics calculated from the main
 342 stream length (L) and the basin area (A) and perimeter (P) include (Karamouz
 343 *et al.* 2012): form factor (FF_w), basin shape factor (B_s), elongation ratio (R_e),
 344 circularity ratio (R_c) and compactness coefficient (C_c) (Table 1).

345 **Table 1** | Basin shape descriptors in hydrology and geomorphology

Watershed shape descriptor	Expression	Range
Form factor	$FF_w = A/L^2$	(0,1)
Basin shape factor	$B_s = L^2/A$	(1, ∞)
Elongation ratio	$R_e = 1.128A^{0.5}/L$	(0,1]
Circularity ratio	$R_c = 12.57A/P^2$	(0,1]
Compactness coefficient	$C_c = 0.2821P/A^{0.5}$	[1, ∞)

346

347 These descriptors are related to the hydrologic response. Thus, larger
 348 peak flows are related to large values of FF_w , R_e , and R_c and small values of B_s
 349 and C_c (Karamouz *et al.* 2012). All these shape descriptors have been used in
 350 basins with areas larger than 1 km²; however, their direct use at urban lot
 351 scales (<1 ha) is not possible, as the main channel is typically not well-defined
 352 to identify the basin length. Conversely, there are shape descriptors used in
 353 image processing which can be used at small and medium scales (Russ 2002):
 354 form factor based on circle (FF^*), form factor based on square (FF^{**}),
 355 roundness (R), aspect ratio (A_r), elongation (E), convexity index (CI), solidity
 356 index (SI), compactness (C) (Table 2).

357 **Table 2** | Representative shape descriptors used in image processing

Image shape descriptor	Expression	Range
Form factor based on circle	$FF^* = 4\pi A/P^2$	(0,1]
Form factor based on square	$FF^{**} = 16A/P^2$	(0,1]
Roundness	$R = 4A/(\pi D_{\max})$	(0,1]
Aspect ratio	$A_r = D_{\max}/D_{\min}$	[1, +∞]
Elongation	$E = L_{\max}/L_{\min}$	[1, +∞]
Convexity index	$CI = P_{\text{convex}}/P$	(0,1]
Solidity index	$SI = A/A_{\text{convex}}$	(0,1]
Compactness	$C = ((4/\pi)A)^{0.5}$	(0,1/π]

358 Note: D_{\max} : maximum diameter, D_{\min} : minimum diameter, L_{\max} , length of the
 359 major axis, L_{\min} : length of the minor axis; L : main stream length; P_{convex} :
 360 convex hull perimeter, A_{convex} : convex hull area (modified from Russ 2002 and
 361 Jiao *et al.* 2012).

362 At scales less than 1 ha, these descriptors have been used to describe the
 363 main features in land use, such as roads, cultivated lands, settlements, rivers,
 364 ponds, and forests and grass lands (Jiao *et al.* 2012).

365 The more relevant shape descriptors for the present work are those that
 366 allow the identification of non-convex and sliver HRUs. Highly non-convex
 367 units characterized by values of $CI = P_{\text{convex}}/P \ll 1$ typically have their
 368 centroid outside of the polygon and non-smoothed contours (P_{convex} is the
 369 convex hull perimeter). On the other hand, values of $CI \approx 1$ ensure the
 370 centroid is inside the polygons and non-smoothed boundaries. In image
 371 processing, the Form factor can be defined in different ways, depending on
 372 whether a circle or square is used as the reference shape (FF^* and FF^{**}).
 373 Sliver units have values of $FF^{**} \ll 1$, and $FF^{**} \approx 1$ for well-shaped
 374 polygons with a square or pseudo-square shape. From now on, FF^{**} will be

375 simply referred to as FF ($FF = 4\pi A / P^2$). To illustrate the use of CI and FF ,
 376 Table 3 presents their values for the irregular shape units shown in Figure 1.
 377 Note how sliver polygons and non-convex polygons have low values of FF
 378 and CI , respectively.

379 **Table 3** | Shape descriptors of HRUs polygons shown in Figure 1

	Unit Physiographic meaning	Area (m²)	<i>CI</i>	<i>FF</i>
a	Segmented street	1,085	<i>0.580</i>	<i>0.057</i>
b	Segmented street	3,439	0.830	<i>0.072</i>
c	Footpath	497	0.940	<i>0.073</i>
d	Green area	680	<i>0.750</i>	<i>0.053</i>
e	Segmented street	2,664	0.950	<i>0.017</i>
f	Lot partition	4,921	<i>0.700</i>	0.241
g	Green area	8,594	<i>0.620</i>	0.204
h	Lot partition	8,9351	<i>0.590</i>	<i>0.130</i>
i	Green area	9,894	<i>0.730</i>	<i>0.126</i>

380 Note: Critical shape descriptor of $CI < 0.80$ and $FF < 0.20$ are in italics and
 381 bold font.

382

383 Jiau & Liu (2012) computed statistics for the main shape metrics of
 384 several land-use classes. Forest and grass lands were the most non-convex
 385 land use classes, while the most convex class is ponds. This is expected as
 386 green areas grow irregularly, whereas pond units such as retention basins are
 387 associated with regular shape units. Furthermore, despite their high CI values,
 388 roads are the most sliver polygons. Thus, convexity cannot be the only
 389 criterion to improve polygonal meshes, and must be complemented by the FF

390 criterion, which allows the decomposition of important units into smaller
391 elements (i.e., small streets, footpaths, backyards, etc.).

392 **DIVIDE-AND-CONQUER STRATEGY TO DECOMPOSE BAD-** 393 **SHAPED POLYGONS**

394 To improve bad-shaped units, we use a divide-and-conquer strategy. The first
395 step is to select bad-shaped elements from the initial mesh using the shape
396 descriptors. The polygons are then segmented into smaller triangles with either
397 a constrained or conforming Delaunay triangulation (DT) algorithm. The
398 resulting triangles are convex elements, but large bulks of relatively small
399 triangles are usually created. To dissolve small triangles, we propose the
400 following algorithms that respect shape descriptors and the overall
401 physiographical-oriented character of the mesh units: (1) an algorithm for non-
402 convex polygons, (2) an algorithm for large polygons, and (3) an algorithm for
403 sliver polygons. These algorithms are implemented in Geo-PUMMA (Sanzana
404 *et al.* 2017) with Python using GRASS functions. The scripts
405 *p.convexity_segmentation.py* and *p.form_factor.py* were modified to apply
406 them not only to GIS features, but also computer graphics' vectorial images.
407 The updated Geo-PUMMA scripts and the dataset used in this paper are
408 available at <https://forge.irstea.fr/projects/geopumma>. Note these algorithms
409 can be easily modified to apply other constraints based on the watershed shape
410 descriptors such as the ones previously identified. What follows is a definition
411 of the different decompositions considered in the strategy (i.e., convex,
412 pseudo-convex, and pseudo-square decomposition), and an explanation of the
413 general notation. In this explanation, polygons are marked with bold font and
414 scalar values with italic font. The following auxiliary indexes and polygons'
415 notation are defined:

416 k : index assigned to number of vertexes

- 417 l : index assigned to any polygon
418 m : index assigned to any polygon different than polygon l
419 i : step index for each iteration
420 j : index assigned to the final components
421 n : total number of triangles obtained in the triangulation step
422 r : index used for the biggest triangle not used in previous steps
423 s : index used for the biggest neighbor triangle
424 T_r : bigger triangle not used in the iteration step
425 T_s : bigger neighbor triangle

426 A simple polygon \mathbf{P} without holes is composed of a set of segments $\delta\mathbf{P}$
427 $= \{\delta\mathbf{P}_0, \delta\mathbf{P}_1, \dots, \delta\mathbf{P}_{n-1}\}$, where $\delta\mathbf{P}_k$ is the segment between two adjacent
428 vertexes, V_k and V_{k+1} , with $k = 0, \dots, n-1$. The area and perimeter of the
429 polygon \mathbf{P} are defined as $area(\mathbf{P})$ and $peri(\mathbf{P})$, respectively, whereas the CI and
430 FF of the polygon \mathbf{P} are defined as $CI(\mathbf{P})$ and $FF(\mathbf{P})$. Pseudo-convex polygons
431 are classified according to their value of CI , and pseudo-square polygons are
432 classified using the FF criterion. The threshold values to create well-shaped
433 polygons for CI and FF are referred as to CIT and FFT , respectively.

434 The following polygons and functions are defined:

- 435 • A polygon \mathbf{C} is a component of \mathbf{P} if $\mathbf{C} \subset \mathbf{P}$
- 436 • $\mathbf{H}_\mathbf{P}$ is the convex hull of a polygon \mathbf{P} , i.e., the convex polygon with
437 the smallest area containing \mathbf{P}
- 438 • \mathbf{P} is said to be convex if $\mathbf{P} = \mathbf{H}_\mathbf{P}$
- 439 • \mathbf{P} is said to be well-shaped pseudo-convex if $CI(\mathbf{P}) > CIT$, i.e.,
440 $perimeter(\mathbf{H}_\mathbf{P})/peri(\mathbf{P}) > CIT$
- 441 • $\mathbf{S}_\mathbf{P}$ is the equivalent square polygon of \mathbf{P} if $area(\mathbf{S}_\mathbf{P}) = area(\mathbf{P})$
- 442 • \mathbf{P} is said to be square if $\mathbf{P} = \mathbf{S}_\mathbf{P}$

- 443 • \mathbf{P} is said to be well-shaped pseudo-square if $FF(\mathbf{P}) > FFT$, i.e.,
444 $4 \cdot \pi \cdot \text{area}(\mathbf{P}) / \text{peri}(\mathbf{P})^2 > FFT$
445 • U_l means union of polygons with sub-index l
446 • \cap means intersection between two polygons
447 • $/$ means subtraction of two polygons
448 • \emptyset means empty set.

449 A set of components $\{\mathbf{C}_l\}$ is a decomposition of \mathbf{P} , $\mathbf{D}(\mathbf{P})$, if their union
450 is \mathbf{P} and all \mathbf{C}_l are interior disjoint (Lien & Amato 2005), i.e., $\{\mathbf{C}_l\}$ must
451 satisfy:

$$452 \mathbf{D}(\mathbf{P}) = \{\mathbf{C}_l \mid U_l \mathbf{C}_l = \mathbf{P} \text{ and } \forall_{l \neq m} \mathbf{C}_l \cap \mathbf{C}_m = \emptyset\} \quad (1)$$

453 A convex decomposition of \mathbf{P} , $\mathbf{CD}(\mathbf{P})$, is a decomposition of \mathbf{P} (Lien &
454 Amato 2005) that contains only convex components:

$$455 \mathbf{CD}(\mathbf{P}) = \{\mathbf{C}_l \mid \mathbf{C}_l \in \mathbf{D}(\mathbf{P}) \text{ and } \mathbf{C}_l = \mathbf{H}_{\mathbf{C}_l}\} \quad (2)$$

456 A well-shaped pseudo-convex decomposition of \mathbf{P} , $\mathbf{PCD}(\mathbf{P})$, is a
457 decomposition of \mathbf{P} that contains only convex components or pseudo-convex
458 components:

$$459 \mathbf{PCD}(\mathbf{P}) = \{\mathbf{C}_l \mid \mathbf{C}_l \in \mathbf{CD}(\mathbf{P}) \text{ or } CI(\mathbf{P}) \geq CIT\} \quad (3)$$

460 A square decomposition of \mathbf{P} , $\mathbf{SD}(\mathbf{P})$, is a decomposition of \mathbf{P} that
461 contains only square components:

$$462 \mathbf{SD}(\mathbf{P}) = \{\mathbf{C}_l \mid \mathbf{C}_l \in \mathbf{D}(\mathbf{P}) \text{ and } \mathbf{C}_l = \mathbf{S}_{\mathbf{C}_l}\} \quad (4)$$

463 A well-shaped pseudo-square decomposition of \mathbf{P} , $\mathbf{PSD}(\mathbf{P})$, is a
464 decomposition of \mathbf{P} that contains only squares or pseudo-squares components:

$$465 \mathbf{PSD}(\mathbf{P}) = \{\mathbf{C}_l \mid \mathbf{C}_l \in \mathbf{SD}(\mathbf{P}) \text{ or } FF(\mathbf{P}) \geq FFT\} \quad (5)$$

466 **Algorithm for non-convex polygons**

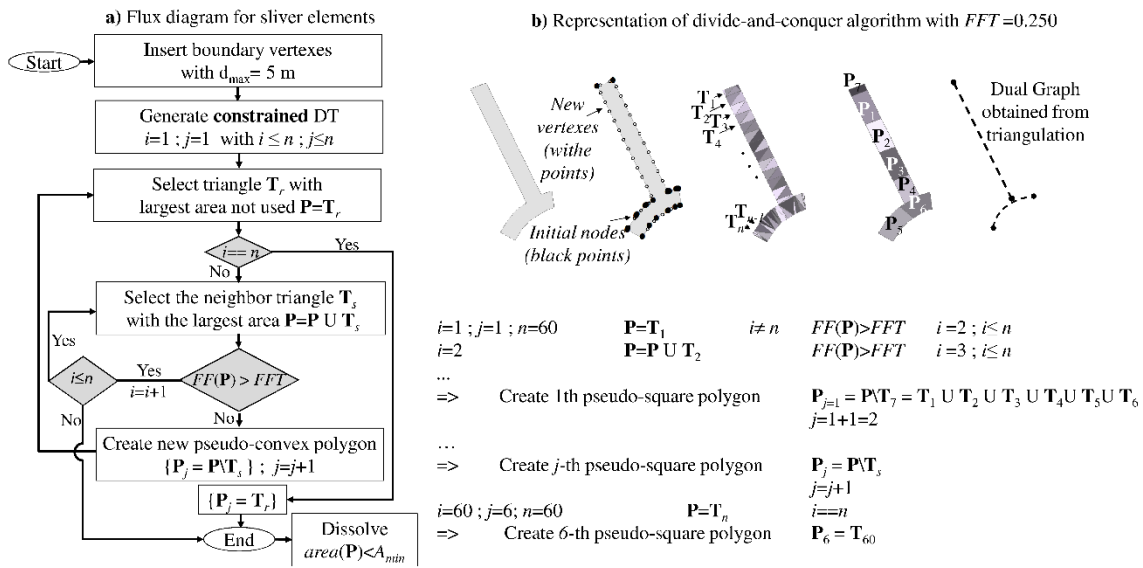
467 The pseudo-convex decomposition algorithm applied to a polygon or HRUs is
468 illustrated in Figures 2(a) and 2(b). In the first step, the bad-shaped polygon is

469 segmented into triangles (\mathbf{T}_i) using a constrained DT implemented within the
470 software Triangle® (Shewchuck 1996). In an iterative manner, the dissolving
471 procedure starts from the largest triangle with the neighboring triangles
472 (Figure 2(a)). The final set $\mathbf{PCD}(\mathbf{P})$ contains only components with a $CI(\mathbf{C}_i) <$
473 CIT . All the final segmentation lines of all polygons, $\delta\mathbf{P}_k$, are connected to the
474 initial vertexes without inserting new vertexes in the initial boundaries (Figure
475 3(b)).

476 **Algorithm for big polygons**

477 The pseudo-convex decomposition algorithm is applied to polygons whose
478 area is larger than a threshold or maximum area A_{max} . Polygons with $area(\mathbf{P}) >$
479 A_{max} can be part of the original mesh, or be obtained after applying the
480 previously described pseudo-convex decomposition algorithm to improve the
481 CI values. Normally, values of $A_{max} = 1$ to 2 ha have been recommended at
482 urban and peri-urban scales (Jankowfsky 2011; Sanzana *et al.* 2013, 2017).
483 The segmentation of big polygons first considers the implementation of a
484 conforming DT using Triangle® (Figure 2(c)). Similarly to the iterative
485 implementation depicted in a previous subsection, the largest triangle and the
486 second largest neighboring triangle are dissolved (Figures 2(c) and 2(d)). The
487 final set $\mathbf{PCD}(\mathbf{P})$ contains only components with a $CI(\mathbf{C}_i) < CIT$ and $area(\mathbf{C}_i)$
488 $< A_{max}$.

499 the sliver polygon is segmented into triangles (T_i) using a conforming DT
 500 without inserting new vertexes inside the polygon (Figure 3(a)). In an iterative
 501 manner, the dissolving procedure starts from the biggest triangle associated
 502 with the neighboring triangles (Figures 3(a) and 3(b)). The final set $PSD(P)$
 503 contains only components for which $FF(C_i) < FFT$.



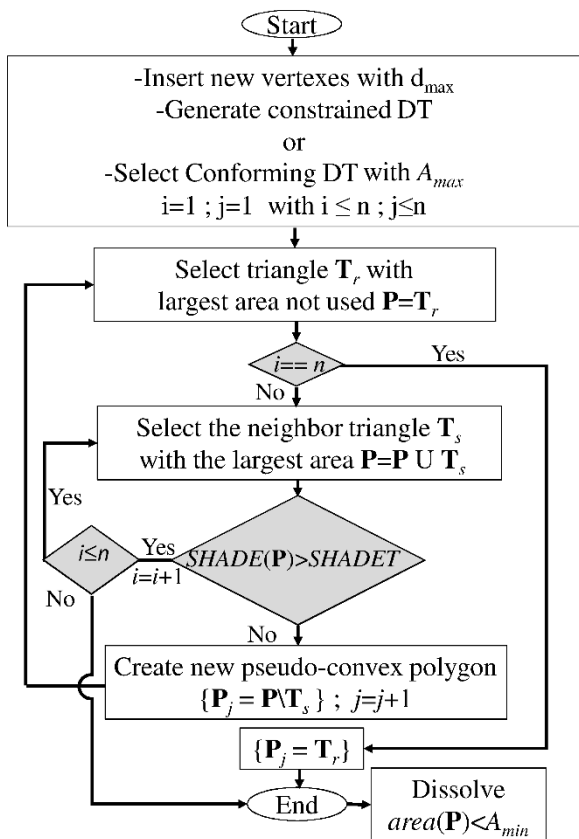
504
 505 **Figure 3 |** Flow chart and schematic strategy for sliver polygons (a) and (b).
 506

507 **Generalization of the divide-and-conquer approach to other shape descriptors**

508 Any bad-shaped polygon can be segmented into triangles (T_i) using a
 509 conforming or constrained DT. Therefore, any SHApe DEscriptor (*SHADE*),
 510 such as those presented by Karamouz *et al.* (2012) and Russ (2002), can then
 511 be used to create a new decomposition of an initial polygon, $D(P)$. For each
 512 *SHADE*, a shape descriptor threshold *SHADET* can be proposed, considering
 513 any regular polygon selected by the modeler. Algorithms previously described
 514 consider the convexity, area, and form factor as shape constraints, but any
 515 other shape descriptor can be selected. Some new vertexes can eventually be
 516 added on the boundaries of the polygon (as in the case of pseudo-square

517 decomposition for sliver polygons), or inside the polygon (as in the case of
 518 maximum area restriction for bigger polygons). Figure 4 shows the steps to
 519 create a pseudo-regular decomposition, $\mathbf{PRD}(\mathbf{P})$, of \mathbf{P} . This new sub-set
 520 contains only regular or pseudo-regular components:

521
$$\mathbf{PRD}(\mathbf{P}) = \{\mathbf{C}_i | \mathbf{C}_i \in \mathbf{RD}(\mathbf{P}) \text{ or } \mathit{SHADE}(\mathbf{C}_i) \geq \mathit{SHADET}\} \quad (6)$$



522
 523 **Figure 4** | Flow chart of generalization of divide-and-conquer approach to any
 524 shape descriptor.

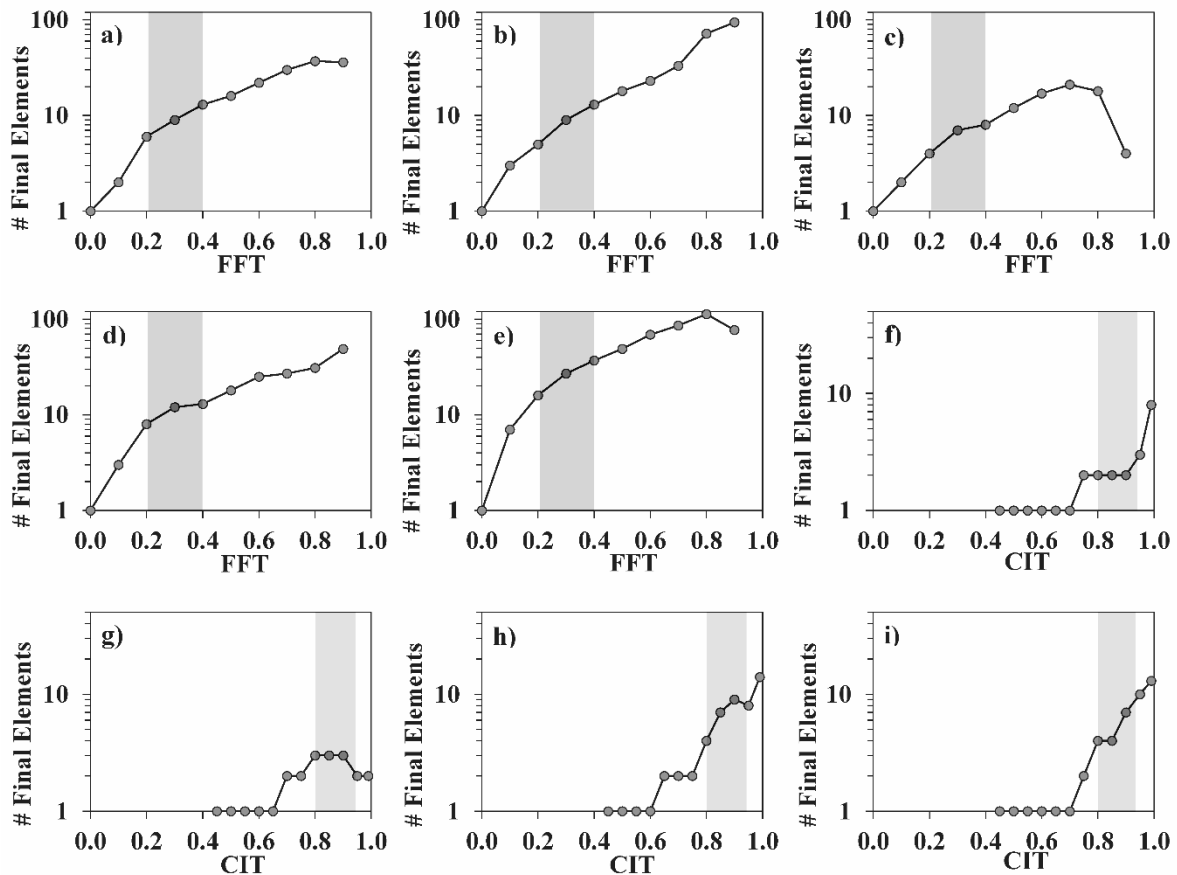
525 In the Appendix, we present the theoretical and empirical order of the
 526 proposed pseudo-convex decomposition strategy.

527 **RESULTS**

528 **Segmentation of bad-shaped HRUs**

529 The bad-shaped HRUs shown in Figure 1 were decomposed into pseudo-
 530 regular polygons using Triangle® and the algorithms *p.convexit.py* and

531 *p.form_factor.py* in Geo-PUMMA. The different panels in Figure 5 illustrate
532 the effect of several *FFT* and *CIT* values on the number of elements in the
533 final segmentation of the corresponding HRUs. Restrictive (i.e., high) values
534 of these indexes increase the final number of elements for each HRU. In fact,
535 the final number of elements in which sliver polygons are decomposed
536 increases exponentially when using the *FF* criterion (Figures 5(a), 5(b), 5(d),
537 and 5(e)), and up to 100 final elements are obtained with values of *FFT* > 0.9.
538 Interestingly, for values of *FFT* ~ 0.8, a maximum number of elements is
539 reached for two curves (i.e., 18 and 113 elements in Figures 5(c) and 5(e),
540 respectively), which shows that the performance of the *p.form_factor.py*
541 algorithm depends on the polygon shape and the presence of narrowing
542 regions, abrupt discontinuities, etc.

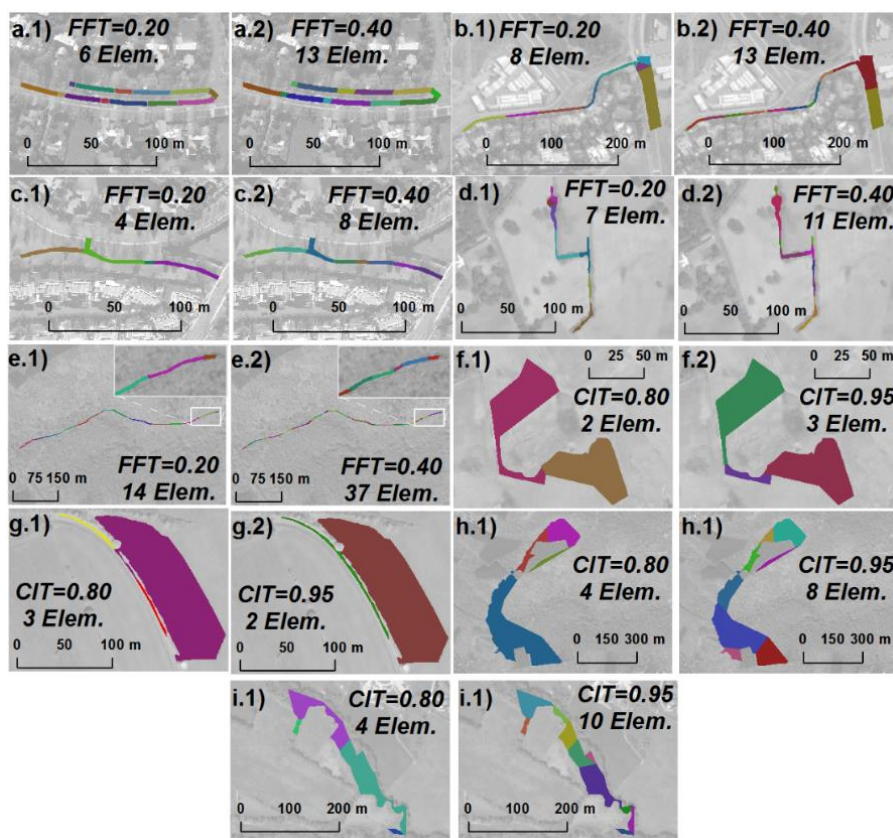


544 **Figure 5** | Number of resulting HRUs after the pseudo-regular decomposition
545 of the polygons from Figure 1, using different values of *FFT* (a)–(e) and *CIT*
546 (f)–(i). For (a), (b), (c), (d) and (e) the **band** shows the range of *FFT* between
547 0.2 and 0.4, and for (f), (g), (h) and (i) the **band** shows the range of *CIT*
548 between 0.80 and 0.95.

549 As expected, the number of elements also increases with the *CIT* value,
550 and up to 14 elements are created for values of $CIT = 0.99$. This increase is
551 less continuous than for the *FFT* case, as the number of elements does not
552 change within certain ranges of *CIT*; nonetheless, the convexity of each final
553 element always improves with larger values of *CIT*. In general, non-strictly
554 increasing curves in Figures 5(f)–5(i) are explained by the existence of
555 irregular narrowing zones in the initial polygons, hence, more than the
556 quantity of vertexes in the initial polygon, their location is what explains the
557 obtained results.

558 Low values of *FF* are typical of land use classes with several small
559 irregular-shaped elements (i.e., roads composed of street, or green areas
560 composed of trees or small shrubs) (Jiau & Liu 2012). Low values of *CI* are
561 associated with green areas such as hedgerows. From an empirical
562 perspective, we propose values of $FFT = 0.20\text{--}0.40$, and $CIT = 0.80\text{--}0.95$ to
563 decompose the bad-shaped HRUs. These values allow the generation of well-
564 shaped HRUs composed of small meaningful parts (e.g., road median,
565 backyard, small set of shrubs or trees) without substantially increasing the
566 final number of elements to be eventually used in a hydrological model. Thus,
567 the minimum of the proposed range of threshold corresponds to a value
568 needed to improve the representation of bad-shaped polygons into a new set of
569 well-shaped HRUs, whereas the maximum of the range corresponds to a limit
570 value that avoids increasing the final number of well-shaped HRUs. Figure 6

571 shows the polygons of Figure 1 decomposed using these minimum and
572 maximum recommended thresholds.



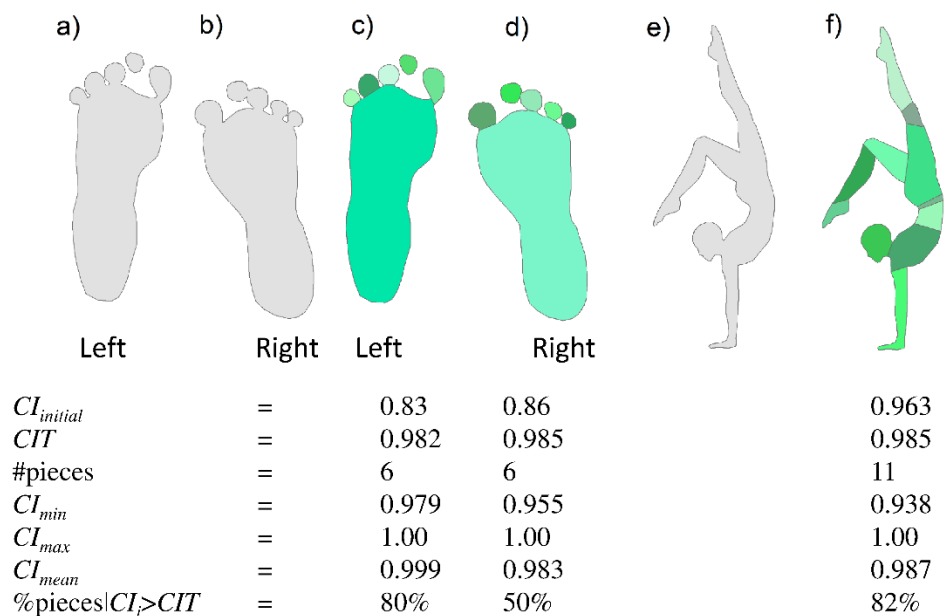
573

574 **Figure 6** | HRUs segmented with minimum recommended values of $FFT =$
575 0.20 (panels (a.1)–(e.1)) and $CIT = 0.80$ (panels (f.1)–(i.1)), and maximum
576 values of $FFT = 0.40$ (panels (a.2)–(e.2)) and $CIT = 0.95$ (panels (f.2)–(i.2)).

577 **Application to images in computer graphics**

578 In disciplines such as human vision or computer graphics, the decomposition
579 of non-convex elements is an active research field. We now evaluate the
580 ability of our approach to decompose non-convex graphical computer images
581 typical of these disciplines, in order to demonstrate its applicability to other
582 fields beyond hydrology. Several pseudo-convex decompositions of these
583 computer images were generated using the *p.convexity.py* script and different
584 convexity index thresholds as it is shown in Figure 7. First, a baby's footprints

585 with initial values of $CI_{initial} = 0.83$ and 0.86 (Figures 7(a) and 7(b)
 586 ,respectively) were decomposed with threshold values of $CIT = 0.982$ (Figure
 587 7(c)) and $CIT = 0.985$ (Figure 7(d)). These values allow separating the **toes**
 588 from the sole of each foot. Then, a yoga position silhouette with an initial
 589 value of $CI_{initial} = 0.66$ (Figure 7(e)) was decomposed into a pseudo-regular
 590 subset using a value of $CIT = 0.985$ (Figure 7(f)). In this case, the chosen CIT
 591 value allows identifying body parts (i.e., head, arm, torso, waist, etc.). Thus,
 592 high values of CIT may be needed to ensure the automatic decomposition of
 593 complex non-convex images. The metrics $\%pieces|CI_i > CIT$ (i.e., percentage
 594 of pieces for which the CI value exceeds CIT) is not 100% because the
 595 dissolution of too small elements deforms the shape of well-shaped elements
 596 in some cases (i.e., $\%pieces|CI_i > CIT = 80\%$, Figure 7(c); $\%pieces|CI_i > CIT =$
 597 50% , Figure 7(d); $\%pieces|CI_i > CIT = 82\%$, Figure 7(f)).



598

599 **Figure 7** | Decomposition of non-convex images into meaningful parts.

600 Images include the left (a) and (c) and right (b) and (d) footprints, and a yoga
 601 silhouette (e) and (f).

602

HYDROLOGICAL ROUTING ASSESSMENT

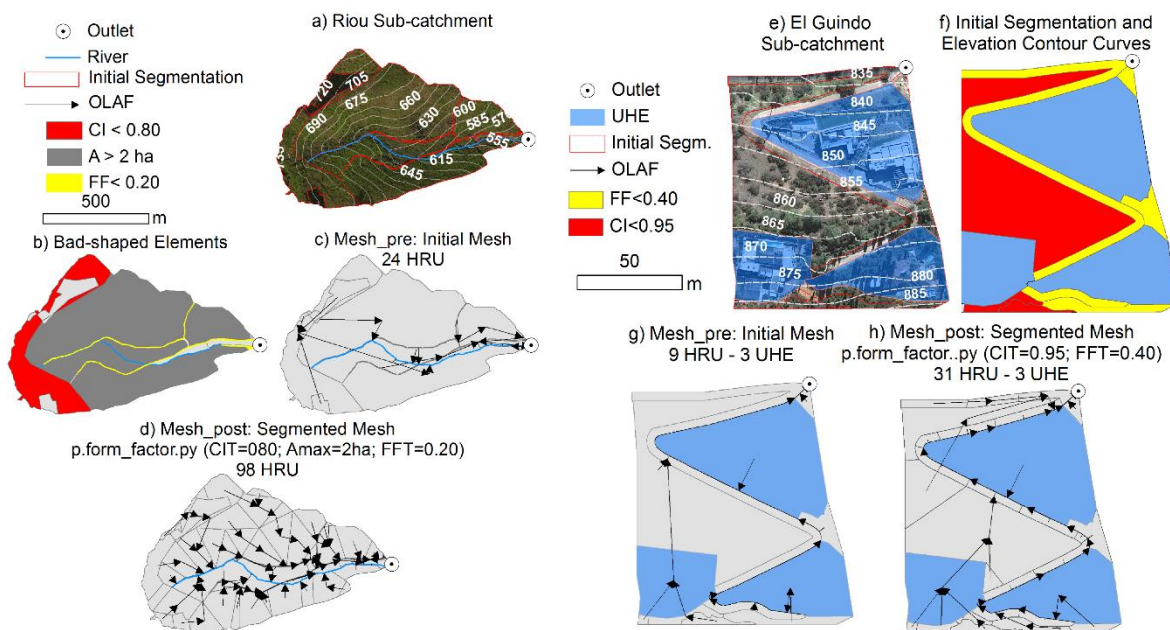
603 **Qualitative hydrological routing assessment for mesh quality** 604 **improvement**

605 In the case of meshes generated to represent hydro-landscapes, the
606 segmentation adopted affects directly the hydrological connectivity among the
607 elements. This connectivity is given by the flow paths obtained using the
608 overland routing routine *p.olaf.py* (Sanzana *et al.* 2017). This script routes the
609 flow between HRUs and the drainage network following topography towards
610 the lowest neighbor HRU or drainage reach. As an example, Figure 8 presents
611 two catchments, the Riou (Lyon, France) and El Guindo (Santiago, Chile),
612 which were adopted as case studies to analyze the effect of the segmentation
613 over the hydrologic response. This figure shows the elevation contours, the
614 initial and segmented model meshes (Mesh_pre and Mesh_post, respectively),
615 the presence and characteristics of bad-shaped polygons, and the
616 corresponding routing structures calculated by the *p.olaf.py* routine. Note that
617 the initial mesh is segmented using the *CI*, *FF* and maximum area criteria.

618 In both cases, a visual inspection shows that the hydrological
619 connectivity tends to be more realistic after segmentation, and flow paths
620 crossing rivers or streets are avoided. In both cases, non-convex and sliver
621 elements were segmented without substantially increasing the final number of
622 elements. For the more rural Riou catchment, values of $CIT = 0.80$ and $FFT =$
623 0.20 were chosen, while for the more urbanized El Guindo catchment, whose
624 streets are better defined, values of $CIT = 0.95$ and $FFT = 0.40$ were adopted.
625 Based on our results, we recommend these pairs of *CIT* and *FFT* values for
626 more natural and urbanized landscapes, respectively. Note these threshold
627 values are higher (i.e., more restrictive) than those presented by Sanzana *et al.*
628 (2013) ($CI < 0.75$ and $CI < 0.88$) and Sanzana *et al.* 2017 ($CI < 0.75$; $FF <$
629 0.20).

630 Quantitative hydrological assessment for mesh quality improvement

631 To further illustrate the effects of the decomposition strategy, we simulated the
632 hydrologic response of two catchments represented using both the initial and
633 segmented meshes. Our objective is not to explore in detail the capabilities of
634 the models, but to assess the effects of the segmentation on their outputs.
635 Because land use and soil parameters for the initial and segmented meshes are
636 the same, the effects of the different meshes and their respective hydrologic
637 connectivity on flow routing are well isolated. Although two particular models
638 are used in this assessment, other semi-distributed and fully distributed
639 hydraulic tools can be used with the modeling meshes prepared with the
640 strategy proposed in this work.



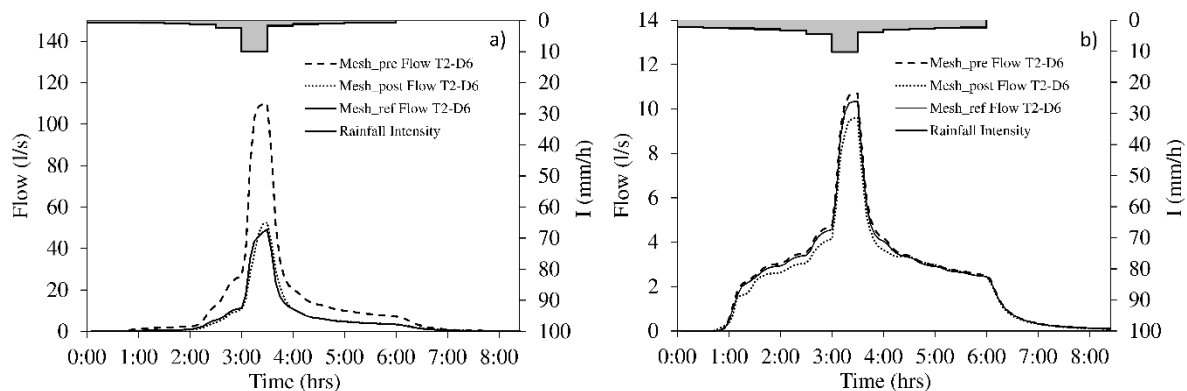
641
642 **Figure 8** | Riou sub-catchment (a): bad-shaped elements (b), initial mesh (c),
643 segmented mesh (d). El Guindo sub-catchment (a): bad-shaped elements, (b)
644 initial mesh (c), segmented mesh (d).

645 We used SWMM (Storm Water Management Model) (Rossman 2009;
646 Gironás *et al.* 2010) to simulate the hydrologic response of the Riou and El
647 Guindo catchments. To compare the effect on the hydrographs' outputs, we

648 simulated synthetic rain events of 5-, 10-, and 25-year return period (T) using
649 the alternating block method. The effect of the segmentation is more notorious
650 when using short events with low return periods. This is shown in Figure 9,
651 which compares for both catchments the resulting hydrographs for a 6 h, 2-
652 year storm. Table 4 compares the resulting peak flows for all the synthetic
653 events.

654 Interestingly, for the Riou catchment, the peak flows from the most
655 segmented model mesh (Mesh_post) are ~50% smaller than those simulated
656 using the initial model mesh (Mesh_pre). For the El Guindo catchment, the
657 peak flow is lower in the segmented mesh just for T = 2 years; for the other
658 return periods, the peak flow actually slightly increases. A possible
659 explanation for such difference in the effect of the mesh segmentation may be
660 the way SWMM infiltrates overland flow in the downstream elements.
661 SWMM is a semi-distributed hydrologic model in which the excess runoff
662 from one element or subcatchment is added homogeneously to the
663 precipitation over the downstream element (Rossman 2009). Hence, the
664 infiltration representation may differ from that of a distributed hydraulic
665 model able to solve the continuity and momentum equations, in which the
666 excess runoff from one element enters the downstream element only through
667 its upstream end. Thus, infiltration is enhanced in SWMM as more
668 subcatchments connected to each other are used to represent the catchment,
669 and is exacerbated when flatter slopes are considered, as is precisely the case
670 of the Riou catchment. Nevertheless, it is important to note that overland
671 routing among many subcatchments is not a common practice in SWMM, as
672 individual subcatchments are typically connected to channelized elements
673 instead. For the sake of comparison, Figure 9 also presents the hydrograph
674 simulated when using the best modeling mesh obtained from the available

675 information (i.e., the reference mesh), which allows the best topographic
 676 fidelity with the largest number of units, while avoiding topological problems.
 677 More details about the generation of this mesh are provided in Sanzana *et al.*
 678 (2017). Interestingly, the hydrographs simulated by this reference mesh and
 679 the segmented mesh are very alike, which seems to demonstrate an ultimate
 680 infiltration capacity as the mesh becomes finer and finer. Curiously, such
 681 behavior is not observed for the Guindo case, for which the reference mesh
 682 and the initial mesh produce more similar results. Overall, these results show
 683 the impact of the modeling mesh when coupling an infiltration model with a
 684 hydrologic overland flow model over the local calculation of the infiltration
 685 rate. This issue is also relevant when a 2D hydraulic routing model is used
 686 instead (Fernández-Pato *et al.* 2016).



687
 688 **Figure 9** | Simulated hydrograph for Mesh_pre (continuous line) and
 689 Mesh_post (dotted line) for Riou (a) and El Guindo (b) sub-catchments using
 690 a 6 h, 2-year storm event. The **continuous line** corresponds to the simulated
 691 hydrograph for the reference mesh (Mesh_ref).

692 **Table 4** | Peak flow (l/s) for return period T for each model mesh

	Riou		El Guindo	
T (year)	Mesh_pre	Mesh_post	Mesh_pre	Mesh_post
2	110	51	11	10

5	144	69	17	19
25	314	154	59	62

693

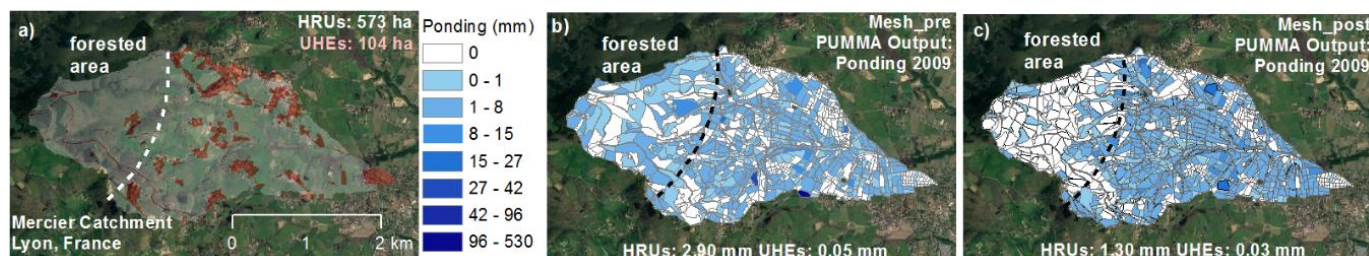
694 We also used the hydrological models PUMMA (Peri-Urban Model for
 695 landscape Management) (Jankowsky *et al.* 2014) to study the response of the
 696 Mercier catchment (Lyon, France) (Figure 10(a)). The model and setup for the
 697 initial mesh is described in Sanzana *et al.* (2017) and Fuamba *et al.* (2017),
 698 while a second model was built for the segmented mesh. The initial mesh is
 699 composed by 1,626 HRUs and 289 UHEs, but only HRUs were segmented.
 700 Both models were run for the year 2009, and the resulting spatial distributions
 701 of ponding depth were compared. The initial mesh produced artificial and
 702 unrealistically large values in some polygons, especially in forested areas in
 703 the west (Figure 10(b)), despite the high infiltration capacity in the area
 704 (Gonzalez-Sosa *et al.* 2010). Ponding reduces 50% when using the segmented
 705 mesh (Figure 10(c)), and polygons with ponding larger than 0.1 m disappear.
 706 These improvements are directly related to changes in the number of elements
 707 in the mesh. The segmentation not only increases the number of HRUs in
 708 20%, but also all the polyline features (i.e., river segments in 459%; water
 709 table interfaces (WTI), in 25%; and water table river interfaces (WTRI), in
 710 47%), as shown in Table 5.

711 **Table 5** | Number of HRUs, river segments, water table interfaces (WTI), and
 712 water table river interfaces (WTRI) in the Mercier catchment

Mesh	River segments	HRUs	WTI	WTRI
Mesh_pre	128	1,626	4,763	580
Mesh_post	716	1,952	5,945	852
Increase (%)	459%	20%	25%	47%

713

714 Note that 12 h of simulation in PUMMA were needed for the initial
715 mesh, while 96 h were used to simulate the post-segmented mesh. Hence,
716 avoiding excessively high threshold values for the shape factors controls the
717 number of segmented HRUs and reduces the computing time used in
718 hydrological modeling.



719

720 **Figure 10** | Application to the Mercier catchment (Lyon, France) for year
721 2009. (a) The catchment, initial mesh, and location of forested area, (b)
722 average ponding depth obtained from PUMMA model with the initial mesh,
723 and (c) same result obtained using the post-segmented mesh.

724 CONCLUSION AND FUTURE WORK

725 This paper presents a flexible divide-and-conquer strategy to create good-
726 quality meshes (composed of simple polygons that fulfill certain geometrical
727 constraints) for distributed hydrological models composed of pseudo-regular
728 elements. Although based on threshold values of the convexity index (CI) and
729 form factor (FF), the proposed strategy can be generalized to other watershed
730 shape descriptors of relevance in hydrology (e.g., basin shape factor,
731 elongation ratio, or compactness coefficient) or any other shape descriptor
732 commonly used in image processing (i.e., roundness, aspect ratio, elongation,
733 or compactness). The algorithms were applied to small catchments in France
734 and Chile, and further tested with silhouette-based features. We conclude the
735 following:

- 736 • The number of final elements of the segmented mesh relates directly to
737 the threshold values for the shape factors. This number increases
738 exponentially with the threshold value of FF (FFT), and in an irregular
739 manner (i.e., as a step function) with the threshold value of CI (CIT).
740 Values of FFT of 0.2 and 0.4 are recommended for mesh segmentation
741 in undeveloped and highly developed areas, respectively. Conversely,
742 values of CIT of 0.8 and 0.95 are recommended for locations with low
743 and high density of green areas, respectively. Very large values of FFT
744 or CIT can increase significantly the final number of elements in the
745 model mesh, affecting the performance of numerical models using this
746 mesh as a spatial domain.
- 747 • High convexity threshold values ($CIT \sim 0.995$) allow obtaining
748 meaningful parts when decomposing silhouette-based features. Because
749 it incorporates the form factor, the proposed strategy is more flexible
750 than existing algorithms, which mostly use the convexity to decompose
751 polygons into pseudo-convex components.
- 752 • Overall, mesh segmentation is crucial to avoid bad-shaped elements that
753 affect the realistic representation of hydrological connectivity in
754 distributed hydrological modeling. Moreover, mesh segmentation
755 facilitates the representation of the spatially distributed processes
756 controlling not only the lumped response of the catchment, but the
757 spatial variability of water quantity and fluxes within it.
- 758 Studies dedicated to building optimal model meshes (Zundel *et al.*
759 2002) and testing the performance of modeling tools at small scales (Abily *et*
760 *al.* 2013) emphasize the importance of mesh refinement due to the time-
761 consuming task involved. The approach here presented addresses this issue as
762 they attempt to improve the construction of irregular model meshes for

763 hydrological modeling. Because the proposed scripts are written in Python
764 using GRASS functions, typical GIS features and formats can be easily
765 utilizable in their application. Unfortunately, GRASS functions may increase
766 the running time required to run the scripts. Although not a big issue when
767 processing small urban and peri-urban catchments, this limitation can be
768 relevant when decomposing many complex non-convex features. As an
769 alternative, geometrical libraries can be used to improve the running time of
770 the proposed scripts. Finally, other examples of decomposition of non-convex
771 polygons into “approximately convex” elements have been implemented by
772 Lien & Amato (2006) (ACD algorithm) and Liu *et al.* (2014) (DuDe
773 algorithm). It would be interesting as future work to compare these approaches
774 in landscapes features’ decomposition, as proposed by Sanzana *et al.* (2017).

775 **ACKNOWLEDGMENTS**

776 This work was developed within the framework of Project MAPA (IDRC
777 107081-001) and Project ECOS-CONICYT C14U02. The work presented was
778 part of the Doctoral thesis funded by the CONICYT PCHA/2014-21140685
779 and “Estadías en el Extranjeros para Tesistas de Doctorado” VRI-UC grants.
780 Funding from Projects FONDECYT N°1131131 and 1181506, FONDAP
781 15110020, FONDECYT 1181506 and IRSTEA-Lyon are also acknowledged.
782 Finally, Jorge Gironás also acknowledges FONDAP 15110017. The assistance
783 in the SWMM model construction from graduate students Alexander Hoch
784 and Tomás Bunster is also appreciated. The Mercier catchment is part of
785 OTHU (Observatoire de Terrain en Hydrologie Urbaine). This work was
786 partially developed within the framework of the Panta Rhei Research Initiative
787 of the International Association of Hydrological Sciences.

788 **REFERENCES**

- 789 Abily, M., Duluc, C. M., Faes, J. B. & Gourbesville, P. 2013 Performance
790 assessment of modelling tools for high resolution runoff simulation
791 over an industrial site. *Journal of Hydroinformatics*, **15**(4), 1296–1311.
792 DOI: 10.2166/hydro.2013.063
- 793 Collon, P., Steckiewicz-laurent, W., Pellerin, J., Gautier, L., Caumon, G.,
794 Reichart, G. & Vaute, L. 2015 3D geomodelling combining implicit
795 surfaces and Voronoi-based remeshing: A case study in the Lorraine
796 Coal Basin (France). *Computers & Geosciences*, **77**, 29–43. DOI:
797 <https://doi.org/10.1016/j.cageo.2015.01.009>
- 798 Dehotin, J. & Braud, I. 2008 Which spatial discretization for distributed
799 hydrological models? Proposition of a methodology and illustration for
800 medium to large scale catchments. *Hydrology and Earth System*
801 *Sciences*, **12**(3), 769–796. DOI: [https://doi.org/10.5194/hess-12-769-](https://doi.org/10.5194/hess-12-769-2008)
802 2008
- 803 DHI 2007a MIKE 21 FLOW MODEL, Hydrodynamic Module: Scientific
804 Documentation. Danish Hydraulics Institute, Horsholm, p. 58.
- 805 DHI 2007b MIKE 21 & MIKE 3 FLOW MODEL FM, Hydrodynamic and
806 Transport Module: Scientific Documentation. Danish Hydraulics
807 Institute, Horsholm, p. 50.
- 808 Downer, C., Ogden, F., Neidzialek, J. & Liu, S. 2006 Gridded
809 surface/subsurface hydrologic analysis (GSSHA) model: a model for
810 simulating diverse streamflow-producing process. In: *Watershed*
811 *Models* (V. P. Singh and D. K. Frevert, eds). Taylor and Francis, Boca
812 Raton, FL, USA, pp. 131–157.
- 813 Fernández-Pato, J., Caviedes-Voullième, D. & García-Navarro, P. 2016
814 Rainfall/runoff simulation with 2D full shallow water equations:
815 Sensitivity analysis and calibration of infiltration parameters. *Journal*

- 816 *of Hydrology*, **536**, 496–513. DOI:
817 <https://doi.org/10.1016/j.jhydrol.2016.03.021>
- 818 Flügel, W. A. 1995 Delineating hydrological response units by geographical
819 information system analyses for regional hydrological modelling using
820 PRMS/MMS in the drainage basin of the River Bröl, Germany.
821 *Hydrological Processes*, **9**, 423–436. DOI:
822 <https://doi.org/10.1002/hyp.3360090313>
- 823 Fuamba, M., Branger, F., Braud, I., Batchabani, E., Sanzana P., Sarrazin, B. &
824 Jankowfsky, S. Value of distributed water level and soil moisture data
825 in the evaluation of a distributed hydrological model. Application to
826 the PUMMA model in the semi-rural Mercier catchment (6.6 km²) in
827 France. *Journal of Hydrology* (submitted).
- 828 Gironás J., Roesner L. A., Rossman L. A. & Davis J. A. 2010 A new
829 applications manual for the storm water management model (SWMM).
830 *Environmental Modelling and Software*, **25**(6), 813–814. DOI:
831 <https://doi.org/10.1016/j.envsoft.2009.11.009>
- 832 Gonzalez-Sosa, E., Braud, I., Dehotin, J., Lassabatère, L., Angulo-Jaramillo,
833 R., Lagouy, M., Branger, F., Jacqueminet, C., Kermadi, S. & Michel,
834 K. 2010 Impact of land use on the hydraulic properties of the topsoil in
835 a small French catchment. *Hydrological Processes*, **24**(17), 2382–2399.
836 DOI: <https://doi.org/10.1002/hyp.7640>
- 837 Haestad Methods 1995 StormCAD for Windows User's Guide. Waterbury,
838 CT, USA.
- 839 Jankowfsky, S. 2011 Understanding and modelling of hydrological processes
840 in small peri-urban catchments using an object oriented and modular
841 distributed approach. Application to the Chaudanne and Mercier sub-
842 catchments (Yzeron Catchment, France). École Doctorale Terre,

- 843 Univers, Environnement. L'Institut National Polytechnique de
844 Grenoble, Grenoble, France.
- 845 Jankowfsky, S., Branger, F., Braud, I., Rodriguez, F., Debionne, S. & Viallet,
846 P. 2014 Assessing anthropogenic influence on the hydrology of small
847 peri-urban catchments: Development of the object-oriented PUMMA
848 model by integrating urban and rural hydrological models. *Journal of*
849 *Hydrology*, **517**, 1056–1071. DOI:
850 <https://doi.org/10.1016/j.jhydrol.2014.06.034>
- 851 Jiao, L., Liu , Y. & Li, H. 2012 Characterizing land-use classes in remote
852 sensing imagery by shape metrics. *ISPRS Journal of Photogrammetry*
853 *and Remote Sensing*, **72**, 46–55. DOI:
854 <https://doi.org/10.1016/j.isprsjprs.2012.05.012>
- 855 Karamouz, M., Nazif, S. & Falahi, M. 2012 *Hydrology and Hydroclimatology:*
856 *Principles and Applications*, CRC Press, Boca Raton, FL, USA.
- 857 Kim, J. & Chung, J. 2015 Untangling polygonal and polyhedral meshes via
858 mesh optimization. *Engineering with Computers*, **31**(3), 617–629.
859 DOI: <https://doi.org/10.1007/s00366-014-0379-5>
- 860 Krayevoy, V. & Sheffer, A. 2006 Variational, meaningful shape
861 decomposition. In: Proceedings SIGGRAPH '06 ACM SIGGRAPH
862 2006 Sketches, Article 50, Boston, MA, USA. DOI:
863 <https://doi.org/10.1145/1179849.1179912>
- 864 Lagacherie, P., Rabotin, M., Colin, F., Moussa, R. & Voltz, M. 2010 Geo-
865 MHYDAS: a landscape discretization tool for distributed hydrological
866 modeling of cultivated areas. *Computers & Geosciences*, **36**(8), 1021–
867 1032. DOI: <https://doi.org/10.1016/j.cageo.2009.12.005>
- 868 Lee, K.-Y., Kim, I.-I., Cho, D.-Y. & Wan Kim, T. 2003 An algorithm for
869 automatic 2D quadrilateral mesh generation with line constraints.

- 870 *Computer-Aided Design*, **35**(12), 1055–1068. DOI:
871 [https://doi.org/10.1016/S0010-4485\(02\)00145-8](https://doi.org/10.1016/S0010-4485(02)00145-8)
- 872 Lien, J.-M. & Amato, N. M. 2005 Approximate convex decomposition of
873 polygons. *Computational Geometry: Theory & Applications*, **35**(1–2),
874 100–123. DOI: <https://doi.org/10.1016/j.comgeo.2005.10.005>
- 875 Lien, J.-M., Liu, G. & Langevin, C. D. 2015 GRIDGEN version 1.0 - A
876 computer program for generating unstructured finite-volume grids:
877 U.S. Geological Survey Open-File Report 2014–1109, 26 p.
- 878 Liu, G., Xi, Z., & Lien, J. M. 2014 Dual-space decomposition of 2D complex
879 shapes. In: Proceedings of the IEEE Conference on Computer Vision
880 and Pattern Recognition, 23–28 June, pp. 4154–4161.
- 881 Manus, C. 2008 Hydrate project - contribution to WP6 “analyze of major flash
882 flood events”. Technical report, LTHE, IRSTEA (ex-CEMAGREF).
- 883 Miles, J. 1985 The representation of flows to partially penetrating river using
884 groundwater flow models. *Journal of Hydrology*, **82**(3–4), 341–355.
- 885 Moussa, R., Voltz, M. & Andrieux, P. 2002 Effects of the spatial organization
886 of agricultural management on the hydrological behavior of a farmed
887 catchment during flood events. *Hydrological Processes*, **16**, 393–412.
888 DOI: <https://doi.org/10.1002/hyp.333>
- 889 O’Rourke, J. 1998 *Computational Geometry in C*, 2nd ed. Cambridge
890 University Press, Cambridge, UK, 376 pp.
- 891 Panday, S., Langevin, C. D., Niswonger, R. G., Ibaraki, M. & Hughes, J. D.
892 2013 MODFLOW–USG version 1: An unstructured grid version of
893 MODFLOW for simulating groundwater flow and tightly coupled
894 processes using a control volume finite-difference formulation. *U.S.*
895 *Geological Survey Techniques and Methods*, book 6, chap. A45, 66 p.

- 896 Pilz, T., Francke, T., & Bronstert, A. 2017 lumpR 2.0.0: an R package
897 facilitating landscape discretisation for hillslope-based hydrological
898 models. *Geoscientific Model Development*, **10**, 3001–3023. DOI:
899 <https://doi.org/10.5194/gmd-10-3001-2017>
- 900 Rodriguez, F., Andrieu, H. & Morena, F. 2008 A distributed hydrological
901 model for urbanized areas – model development and application to
902 case studies. *Journal of Hydrology*, **351**, 268–287. DOI:
903 <https://doi.org/10.1016/j.jhydrol.2007.12.007>
- 904 Rossman, L. A. 2009 Storm Water Management Model User’s Manual Version
905 5.0. EPA/600/R-05/040, National Risk Management Research
906 Laboratory, United States Environmental Protection Agency,
907 Cincinnati, OH, USA.
- 908 Russ, J. C. 2002 *The Image Processing Handbook*, 4th ed. CRC Press, Boca
909 Raton, FL, USA, 732 pp.
- 910 Russo, B., Sunyer, D., Velasco, M. & Djordjević, S. 2015 Analysis of extreme
911 flooding events through a calibrated 1D/2D coupled model: the case of
912 Barcelona (Spain). *Journal of Hydroinformatics*, **17**(3), 473–491. DOI:
913 [10.2166/hydro.2014.063](https://doi.org/10.2166/hydro.2014.063)
- 914 Sanzana, P., Jankowsky, S., Branger, F., Braud, I., Vargas, X., Hitschfeld, N.
915 & Gironás, J. 2013 Computer-assisted mesh generation based on
916 hydrological response units for distributed hydrological modeling.
917 *Computers & Geosciences*, **57**, 32–43. DOI:
918 <https://doi.org/10.1016/j.cageo.2013.02.006>
- 919 Sanzana, P., Gironás, J., Braud, I., Branger, F., Rodriguez, F., Vargas, X.,
920 Hitschfeld, N., Muñoz J. F., Vicuña, S. & Jankowsky, F. 2017 A GIS-
921 based urban and peri-urban landscape representation toolbox for
922 hydrological distributed modeling. *Environmental Modelling &*

- 923 *Software*, **91**, 168–185. DOI:
924 <https://doi.org/10.1016/j.envsoft.2017.01.022>
- 925 Shewchuck, J. 1996 Triangle: Engineering a 2D Quality Mesh Generator and
926 Delaunay Triangulator. In: M. C. Lin and D. Manocha (eds.), *Applied*
927 *Computational Geometry: Towards Geometric Engineering*. Springer,
928 Berlin, Heidelberg, Germany, pp. 203–222.
- 929 Shewchuk, J. R. 2002 Delaunay refinement algorithms for triangular mesh
930 generation. *Computational Geometry: Theory and Applications*, **22**(1–
931 3), 21–74. DOI: [https://doi.org/10.1016/S0925-7721\(01\)00047-5](https://doi.org/10.1016/S0925-7721(01)00047-5)
- 932 Si, H. 2006 TetGen, A quality tetrahedral mesh generator and three-
933 dimensional delaunay triangulator. Weierstrass Institute for Applied
934 Analysis and Stochastic, Berlin, Germany.
- 935 Tucker, G. E., Lancaster, S. T., Gasparini N. M., Bras R. L. & Rybarczyk S.
936 M. 2001 An object-oriented framework for hydrologic and geomorphic
937 modeling using triangulated irregular networks. *Computers &*
938 *Geosciences*, **27**(8), 959–973. DOI: [https://doi.org/10.1016/S0098-](https://doi.org/10.1016/S0098-3004(00)00134-5)
939 [3004\(00\)00134-5](https://doi.org/10.1016/S0098-3004(00)00134-5)
- 940 Versteeg, H. K. & Malalasekera, W. 2007 *An Introduction to Computational*
941 *Fluid Dynamics, The Finite Volume Method*, 2nd ed. Pearson, Harlow,
942 UK.
- 943 Zundel, A. K., Nelson, E. J. & Christiansen, J. R. 2002 Stamped features:
944 creation of engineered structures in conceptual models. *Journal of*
945 *Hydroinformatics*, **4**(1), 63–71.
946 <http://jh.iwaponline.com/content/4/1/63.abstract>
947