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.

Author-produced version of the article published in Journal of hydroinformatics, 2019, 21 (1): 104-122
The original publication is available at https://iwaponline.com/jh/article/21/1/104/64103/Decomposition-of-2D-polygons-and-its-effect-in
doi : 10.2166/hydro.2018.031

Decomposition of 2D polygons and its effect in hydrological models

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

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

P. Sanzana (corresponding author)

J. Gironás

J. F. Muñoz

S. Vicuña

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

E-mail: ppsanzana@ing.puc.cl

P. Sanzana

J. Gironás

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

J. Gironás

S. Vicuña

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

J. Gironás

S. Vicuña

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

P. Sanzana

I. Braud

F. Branger

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

N. Hitschfeld

J. Romero

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

F. Rodriguez

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

M. Fuamba

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

X. Vargas

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

A. Mejía

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

ABSTRACT

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

Author-produced version of the article published in Journal of hydroinformatics, 2019, 21 (1): 104-122
The original publication is available at https://iwaponline.com/jh/article/21/1/104/64103/Decomposition-of-2D-polygons-and-its-effect-in
doi : 10.2166/hydro.2018.031

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

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

GLOSSARY

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

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

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

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

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

Author-produced version of the article published in Journal of hydroinformatics, 2019, 21 (1): 104-122
The original publication is available at https://iwaponline.com/jh/article/21/1/104/64103/Decomposition-of-2D-polygons-and-its-effect-in
doi : 10.2166/hydro.2018.031

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

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

Initial mesh: output mesh from the initial segmentation step.

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

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

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

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

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

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

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

Shape factor: dimensionless metric used to describe the shape of a 2D polygon, which is independent of its size.

Hydrological response unit (HRU): a simple polygon with hydrological homogeneous properties (i.e., hydrological properties whose variation within an HRU is small compared to that among neighboring HRUs (Flügel 1995)).

Urban hydrological elements (UHE): a simple polygon composed of a cadastral parcel and its corresponding portion of street (Rodriguez *et al*. 2008).

Drainage network: connectivity structure among the hydrological elements within the catchment contributing to the channelized system. It includes overland path flows, streams, ditches, and sewers.

INTRODUCTION

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

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

Author-produced version of the article published in Journal of hydroinformatics, 2019, 21 (1): 104-122
The original publication is available at https://iwaponline.com/jh/article/21/1/104/64103/Decomposition-of-2D-polygons-and-its-effect-in
doi : 10.2166/hydro.2018.031

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

Appropriate meshes for representation of varying size elements

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

Author-produced version of the article published in Journal of hydroinformatics, 2019, 21 (1): 104-122
The original publication is available at https://iwaponline.com/jh/article/21/1/104/64103/Decomposition-of-2D-polygons-and-its-effect-in
doi : 10.2166/hydro.2018.031

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

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

Current technologies in polygonal decomposition

The decomposition of 2D shapes into meaningful parts is a key process in human image understanding, and geometrical computing areas (Lien & Amato 2005; Liu *et al*. 2014). The accepted notion of meaningful parts relies on human perception, and is based on the observation that human vision defines part boundaries (e.g., hands and neck in a human body shape) (Krayevoy & Sheffer 2006). Any non-convex polygon can be decomposed into sub-set parts approximatively convex, which are commonly associated with meaningful parts. Several efficient strategies to decompose polygons into approximatively convex pieces, such as the DuDe (Liu *et al*. 2014) and ACD (Lien & Amato 2005) algorithms have been developed and applied to decompose silhouette-based features. Unfortunately, the convexity criterion may not be enough for terrain representation using irregular meshes for hydrological models. Many models use a terrain representation based on hydrological response units (HRUs) (Flügel 1995) and urban hydrological elements (UHEs) (Rodríguez *et al*. 2008), meaningful physiographic units of the natural and urban landscapes with homogeneous properties, respectively. These units can be highly irregular and bad-shaped, and they eventually need to be decomposed into meaningful small pieces. For example, there can be sliver polygons (i.e., elongated polygons with small area) perfectly convex (i.e., long streets, linear hedges, footpaths, median strips, or narrow green areas) but not suitable for flow routing. A sliver polygon can work as an artificial wall restricting flow routing, because its elevation – which may not be representative of the entire area involved – can be such that flow from lower adjacent polygons is restricted (Jankowfsky 2011). Hence, the criteria of decomposition of HRUs must consider not only convexity but other factors such as form factor or shape factor, circularity, elongation ratios, and compactness coefficient. These criteria can be implemented through tools using a divide-and-conquer approach able to adapt to a variety of geometrical criteria. An example of such a tool is Geo-PUMMA (Sanzana *et al*. 2017), a semi-automatic vectorial toolbox to represent urban and peri-urban small catchments (0.1–10 km^2) and the hydrological connectivity among their components.

Contribution

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

Author-produced version of the article published in Journal of hydroinformatics, 2019, 21 (1): 104-122
The original publication is available at https://iwaponline.com/jh/article/21/1/104/64103/Decomposition-of-2D-polygons-and-its-effect-in
doi : 10.2166/hydro.2018.031

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

BASIC CONCEPTS

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

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

Author-produced version of the article published in Journal of hydroinformatics, 2019, 21 (1): 104-122
The original publication is available at https://iwaponline.com/jh/article/21/1/104/64103/Decomposition-of-2D-polygons-and-its-effect-in
doi : 10.2166/hydro.2018.031

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

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

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

Author-produced version of the article published in Journal of hydroinformatics, 2019, 21 (1): 104-122
The original publication is available at https://iwaponline.com/jh/article/21/1/104/64103/Decomposition-of-2D-polygons-and-its-effect-in
doi : 10.2166/hydro.2018.031

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

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

Shape descriptors

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

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,∞)

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

Table 2 | Representative shape descriptors used in image processing

Author-produced version of the article published in Journal of hydroinformatics, 2019, 21 (1): 104-122
The original publication is available at https://iwaponline.com/jh/article/21/1/104/64103/Decomposition-of-2D-polygons-and-its-effect-in
doi : 10.2166/hydro.2018.031

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	$CI = P_{\mathrm{convex}}/P$	(0,1]
Solidity index	$SI = A/A_{\mathrm{convex}}$	(0,1]
Compactness	$C = ((4/\pi)A)^{0.5}$	(0,1/π]

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

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

The more relevant shape descriptors for the present work are those that allow the identification of non-convex and sliver HRUs. Highly non-convex units characterized by values of $CI = P_{\mathrm{convex}}/P << 1$ typically have their centroid outside of the polygon and non-smoothed contours (P_{convex} is the convex hull perimeter). On the other hand, values of $CI \approx 1$ ensure the centroid is inside the polygons and non-smoothed boundaries. In image processing, the Form factor can be defined in different ways, depending on whether a circle or square is used as the reference shape (FF^* and FF^{**}). Sliver units have values of $FF^{**} << 1$, and $FF^{**} \approx 1$ for well-shaped polygons with a square or pseudo-square shape. From now on, FF^{**} will be simply referred to as *FF* ($FF = 4\pi A / P^2$). To illustrate the use of *CI* and *FF*, Table 3 presents their values for the irregular shape units shown in Figure 1. Note how sliver polygons and non-convex polygons have low values of *FF* and *CI*, respectively.

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

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

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

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

Author-produced version of the article published in Journal of hydroinformatics, 2019, 21 (1): 104-122
The original publication is available at https://iwaponline.com/jh/article/21/1/104/64103/Decomposition-of-2D-polygons-and-its-effect-in
doi : 10.2166/hydro.2018.031

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

DIVIDE-AND-CONQUER STRATEGY TO DECOMPOSE BAD-SHAPED POLYGONS

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

k: index assigned to number of vertexes

Author-produced version of the article published in Journal of hydroinformatics, 2019, 21 (1): 104-122
The original publication is available at https://iwaponline.com/jh/article/21/1/104/64103/Decomposition-of-2D-polygons-and-its-effect-in
doi : 10.2166/hydro.2018.031

l: index assigned to any polygon

m: index assigned to any polygon different than polygon l

i: step index for each iteration

j: index assigned to the final components

n: total number of triangles obtained in the triangulation step

r: index used for the biggest triangle not used in previous steps

s: index used for the biggest neighbor triangle

$\mathbf{T}_r$: bigger triangle not used in the iteration step

$\mathbf{T}_s$: bigger neighbor triangle

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

The following polygons and functions are defined:

- A polygon **C** is a component of **P** if $\mathbf{C} \subset \mathbf{P}$
- $\mathbf{H_P}$ is the convex hull of a polygon **P**, i.e., the convex polygon with the smallest area containing **P**
- **P** is said to be convex if $\mathbf{P} = \mathbf{H_P}$
- **P** is said to be well-shaped pseudo-convex if $CI(\mathbf{P}) > CIT$, i.e., $\text{perimeter}(\mathbf{H_P})/\text{peri}(\mathbf{P}) > CIT$
- $\mathbf{S_P}$ is the equivalent square polygon of **P** if $area(\mathbf{S_P}) = area(\mathbf{P})$
- **P** is said to be square if $\mathbf{P} = \mathbf{S_P}$

Author-produced version of the article published in Journal of hydroinformatics, 2019, 21 (1): 104-122
The original publication is available at https://iwaponline.com/jh/article/21/1/104/64103/Decomposition-of-2D-polygons-and-its-effect-in
doi : 10.2166/hydro.2018.031

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

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

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

A convex decomposition of **P**, **CD**(**P**), is a decomposition of **P** (Lien & Amato 2005) that contains only convex components:

$$\mathbf{CD}(\mathbf{P}) = \left\{\mathbf{C}_l \middle| \mathbf{C}_l \epsilon \mathbf{D}(\mathbf{P}) \text{ and } \mathbf{C}_l = \mathbf{H}_{\mathbf{C}_l}\right\} \quad (2)$$

A well-shaped pseudo-convex decomposition of **P**, **PCD**(**P**), is a decomposition of **P** that contains only convex components or pseudo-convex components:

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

A square decomposition of **P**, **SD**(**P**), is a decomposition of **P** that contains only square components:

$$\mathbf{SD}(\mathbf{P}) = \left\{\mathbf{C}_l \middle| \mathbf{C}_l \epsilon \mathbf{D}(\mathbf{P}) \text{ and } \mathbf{C}_l = \mathbf{S}_{\mathbf{C}_l}\right\} \quad (4)$$

A well-shaped pseudo-square decomposition of **P**, **PSD**(**P**), is a decomposition of **P** that contains only squares or pseudo-squares components:

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

Algorithm for non-convex polygons

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

doi : 10.2166/hydro.2018.031

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

Algorithm for big polygons

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

Author-produced version of the article published in Journal of hydroinformatics, 2019, 21 (1): 104-122
The original publication is available at https://iwaponline.com/jh/article/21/1/104/64103/Decomposition-of-2D-polygons-and-its-effect-in
doi : 10.2166/hydro.2018.031

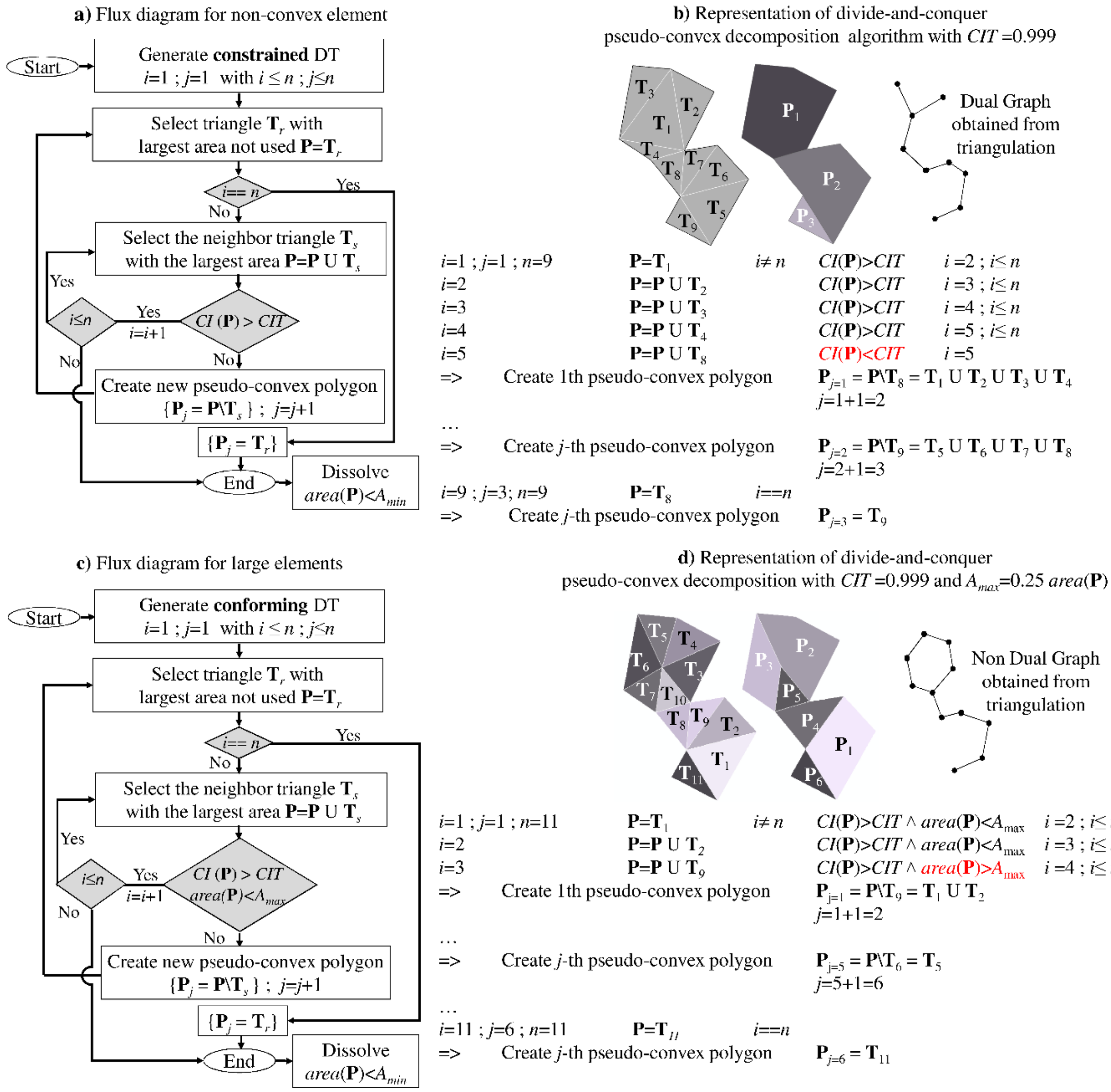


Figure 2 | Flow chart and schematic strategy for non-convex polygons (a) and (b), and big polygons (c) and (d).

Algorithm for sliver polygons

The pseudo-square decomposition algorithm applied to a polygon or HRUs is illustrated in Figures 3(a) and 3(b). In a first step, the initial boundaries are segmented into new ones with maximum separation length of d_{max} (Figure 3(b)). For peri-urban meshes, where sliver polygons are generally associated with streets, paths, river banks, an empirical value of d_{max} = 5 m is recommended to add additional vertexes along the perimeter. In a second step,

Author-produced version of the article published in Journal of hydroinformatics, 2019, 21 (1): 104-122
The original publication is available at https://iwaponline.com/jh/article/21/1/104/64103/Decomposition-of-2D-polygons-and-its-effect-in
doi : 10.2166/hydro.2018.031

the sliver polygon is segmented into triangles ($\mathbf{T}_i$) using a conforming DT without inserting new vertexes inside the polygon (Figure 3(a)). In an iterative manner, the dissolving procedure starts from the biggest triangle associated with the neighboring triangles (Figures 3(a) and 3(b)). The final set **PSD**(**P**) contains only components for which $FF(\mathbf{C}_i) < FFT$.

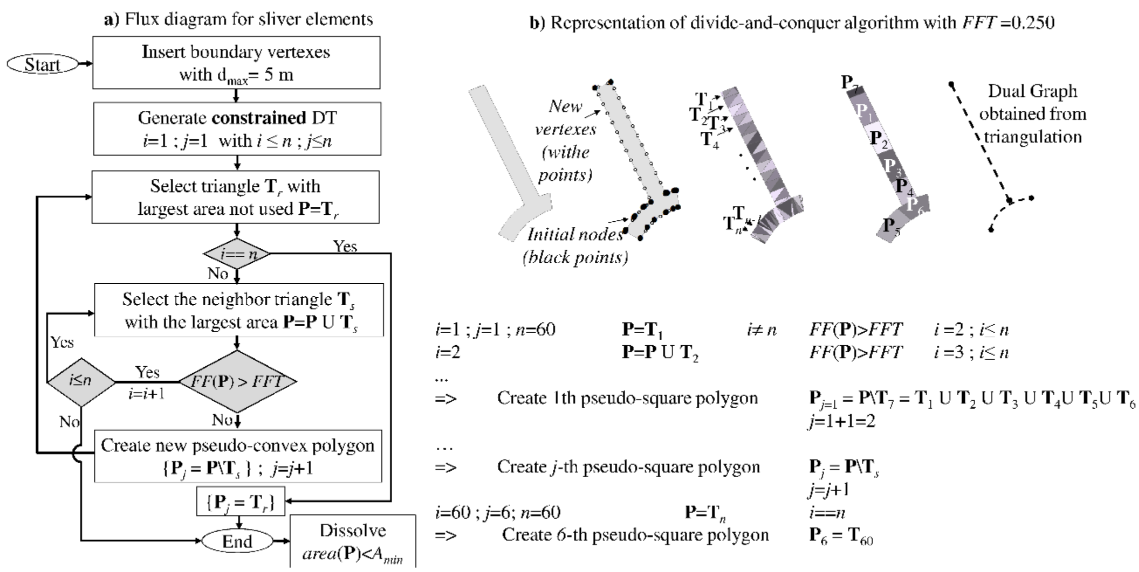


Figure 3 | Flow chart and schematic strategy for sliver polygons (a) and (b).

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

Any bad-shaped polygon can be segmented into triangles ($\mathbf{T}_i$) using a conforming or constrained DT. Therefore, any SHApe DEscriptor (*SHADE*), such as those presented by Karamouz *et al*. (2012) and Russ (2002), can then be used to create a new decomposition of an initial polygon, **D**(**P**). For each *SHADE*, a shape descriptor threshold *SHADET* can be proposed, considering any regular polygon selected by the modeler. Algorithms previously described consider the convexity, area, and form factor as shape constraints, but any other shape descriptor can be selected. Some new vertexes can eventually be added on the boundaries of the polygon (as in the case of pseudo-square decomposition for sliver polygons), or inside the polygon (as in the case of maximum area restriction for bigger polygons). Figure 4 shows the steps to create a pseudo-regular decomposition, **PRD**(**P**), of **P**. This new sub-set contains only regular or pseudo-regular components:

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

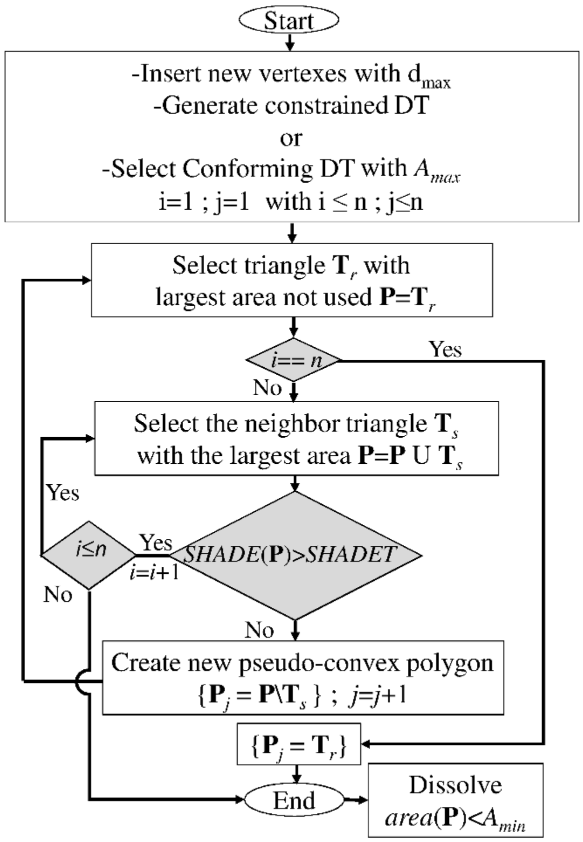


Figure 4 | Flow chart of generalization of divide-and-conquer approach to any shape descriptor.

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

RESULTS

Segmentation of bad-shaped HRUs

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

Author-produced version of the article published in Journal of hydroinformatics, 2019, 21 (1): 104-122
The original publication is available at https://iwaponline.com/jh/article/21/1/104/64103/Decomposition-of-2D-polygons-and-its-effect-in
doi : 10.2166/hydro.2018.031

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

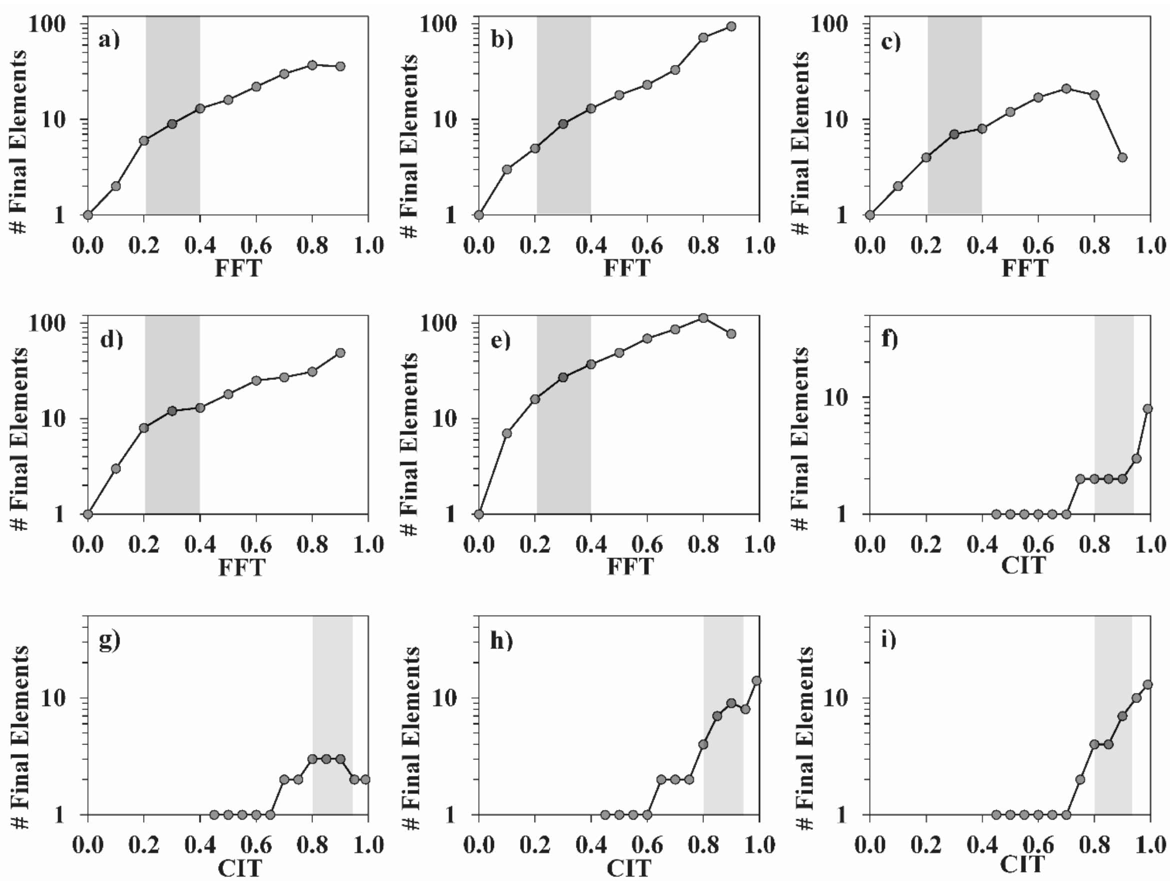


Figure 5 | Number of resulting HRUs after the pseudo-regular decomposition of the polygons from Figure 1, using different values of *FFT* (a)–(e) and *CIT* (f)–(i). For (a), (b), (c), (d) and (e) the band shows the range of *FFT* between 0.2 and 0.4, and for (f), (g), (h) and (i) the band shows the range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

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

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

Author-produced version of the article published in Journal of hydroinformatics, 2019, 21 (1): 104-122
The original publication is available at https://iwaponline.com/jh/article/21/1/104/64103/Decomposition-of-2D-polygons-and-its-effect-in
doi : 10.2166/hydro.2018.031

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

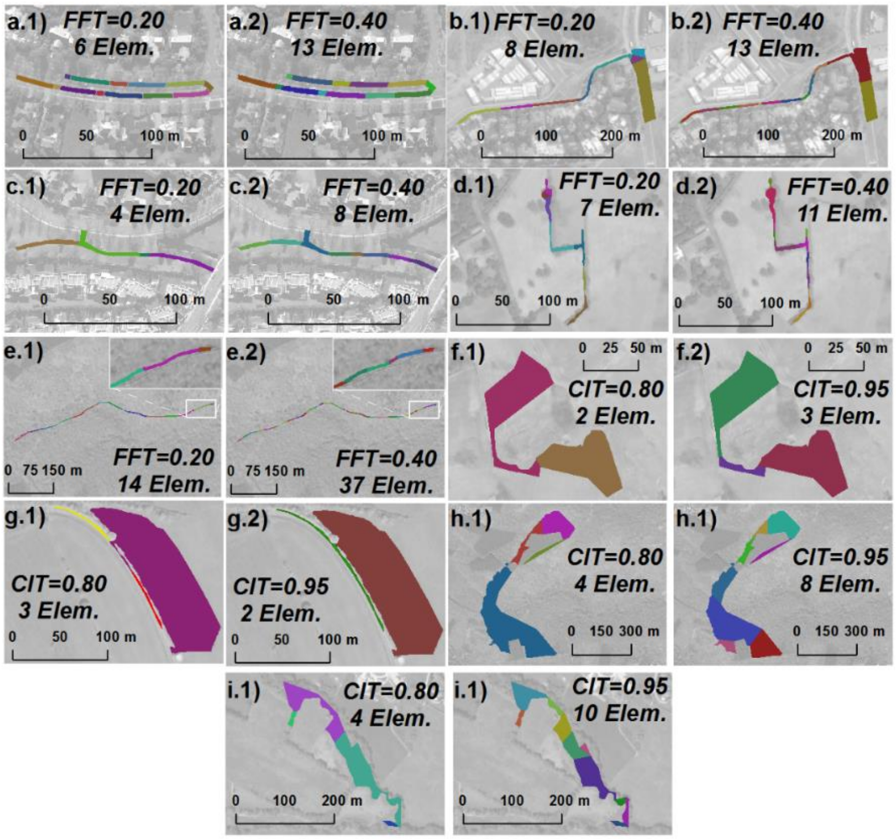
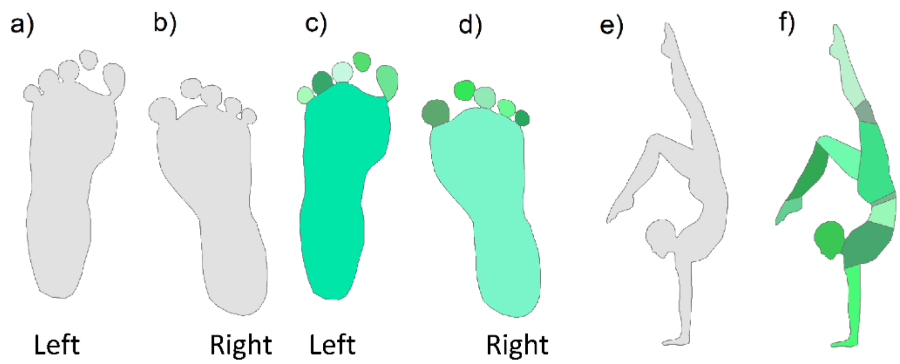


Figure 6 | HRUs segmented with minimum recommended values of *FFT* = 0.20 (panels (a.1)–(e.1)) and *CIT* = 0.80 (panels (f.1)–(i.1)), and maximum vales of *FFT* = 0.40 (panels (a.2)–(e.2)) and *CIT* = 0.95 (panels (f.2)–(i.2)).

Application to images in computer graphics

In disciplines such as human vision or computer graphics, the decomposition of non-convex elements is an active research field. We now evaluate the ability of our approach to decompose non-convex graphical computer images typical of these disciplines, in order to demonstrate its applicability to other fields beyond hydrology. Several pseudo-convex decompositions of these computer images were generated using the *p.convexity.py* script and different convexity index thresholds as it is shown in Figure 7. First, a baby's footprints with initial values of $CI_{initial}$ = 0.83 and 0.86 (Figures 7(a) and 7(b) ,respectively) were decomposed with threshold values of CIT = 0.982 (Figure 7(c)) and CIT = 0.985 (Figure 7(d)). These values allow separating the toes from the sole of each foot. Then, a yoga position silhouette with an initial value of $CI_{initial}$ = 0.66 (Figure 7(e)) was decomposed into a pseudo-regular subset using a value of CIT = 0.985 (Figure 7(f)). In this case, the chosen CIT value allows identifying body parts (i.e., head, arm, torso, waist, etc.). Thus, high values of CIT may be needed to ensure the automatic decomposition of complex non-convex images. The metrics %pieces|$CI_i > CIT$ (i.e., percentage of pieces for which the CI value exceeds CIT) is not 100% because the dissolution of too small elements deforms the shape of well-shaped elements in some cases (i.e., %pieces|$CI_i > CIT$ = 80%, Figure 7(c); %pieces|$CI_i > CIT$ = 50%, Figure 7(d); %pieces|$CI_i > CIT$ = 82%, Figure 7(f)).



$CI_{initial}$	=	0.83	0.86	0.963
CIT	=	0.982	0.985	0.985
#pieces	=	6	6	11
CI_{min}	=	0.979	0.955	0.938
CI_{max}	=	1.00	1.00	1.00
CI_{mean}	=	0.999	0.983	0.987
%pieces|CI_i>CIT	=	80%	50%	82%

Figure 7 | Decomposition of non-convex images into meaningful parts. Images include the left (a) and (c) and right (b) and (d) footprints, and a yoga silhouette (e) and (f).

HYDROLOGICAL ROUTING ASSESSMENT

Qualitative hydrological routing assessment for mesh quality improvement

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

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

Author-produced version of the article published in Journal of hydroinformatics, 2019, 21 (1): 104-122
The original publication is available at https://iwaponline.com/jh/article/21/1/104/64103/Decomposition-of-2D-polygons-and-its-effect-in
doi : 10.2166/hydro.2018.031

Quantitative hydrological assessment for mesh quality improvement

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

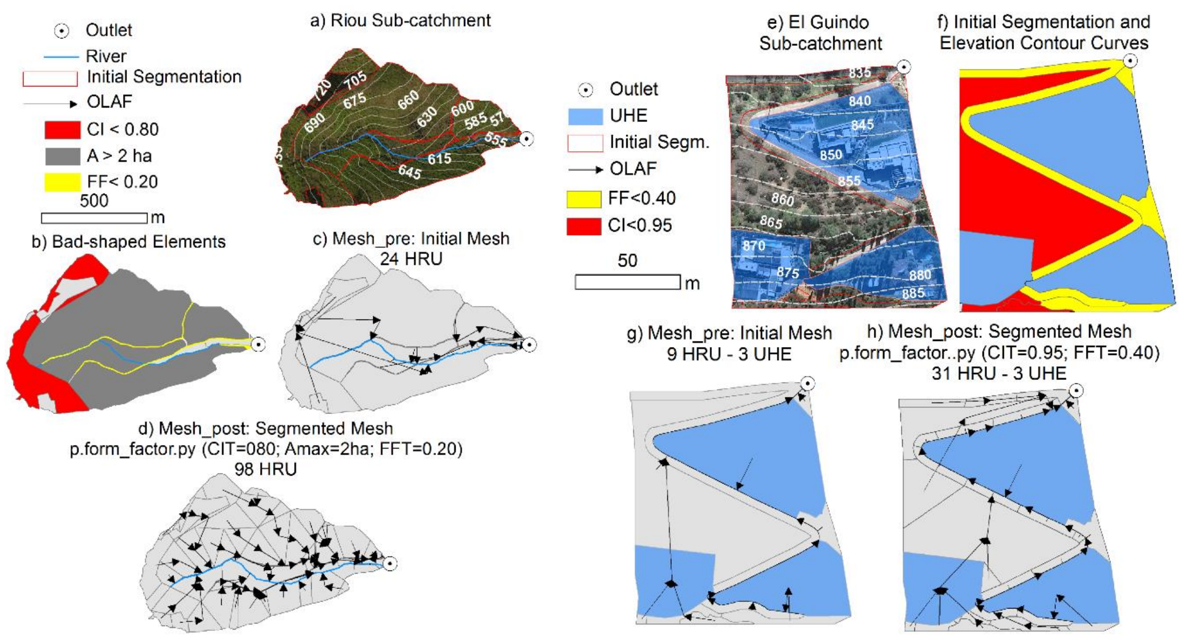


Figure 8 | Riou sub-catchment (a): bad-shaped elements (b), initial mesh (c), segmented mesh (d). El Guindo sub-catchment (a): bad-shaped elements, (b) initial mesh (c), segmented mesh (d).

We used SWMM (Storm Water Management Model) (Rossman 2009; Gironás *et al*. 2010) to simulate the hydrologic response of the Riou and El Guindo catchments. To compare the effect on the hydrographs' outputs, we simulated synthetic rain events of 5-, 10-, and 25-year return period (T) using the alternating block method. The effect of the segmentation is more notorious when using short events with low return periods. This is shown in Figure 9, which compares for both catchments the resulting hydrographs for a 6 h, 2-year storm. Table 4 compares the resulting peak flows for all the synthetic events.

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

Author-produced version of the article published in Journal of hydroinformatics, 2019, 21 (1): 104-122
The original publication is available at https://iwaponline.com/jh/article/21/1/104/64103/Decomposition-of-2D-polygons-and-its-effect-in
doi : 10.2166/hydro.2018.031

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

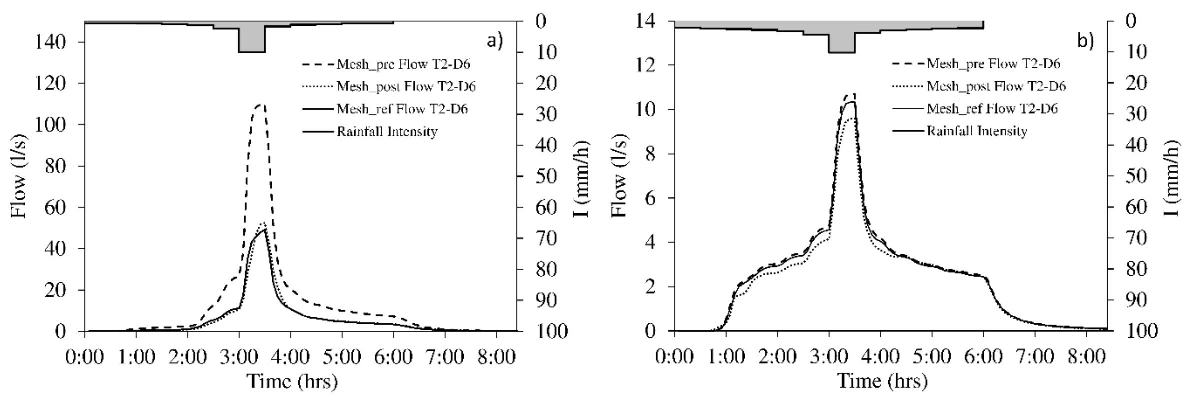


Figure 9 | Simulated hydrograph for Mesh_pre (continuous line) and Mesh_post (dotted line) for Riou (a) and El Guindo (b) sub-catchments using a 6 h, 2-year storm event. The continuos line corresponds to the simulated hydrograph for the reference mesh (Mesh_ref).

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

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

Table 5 | Number of HRUs, river segments, water table interfaces (WTI), and 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%

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

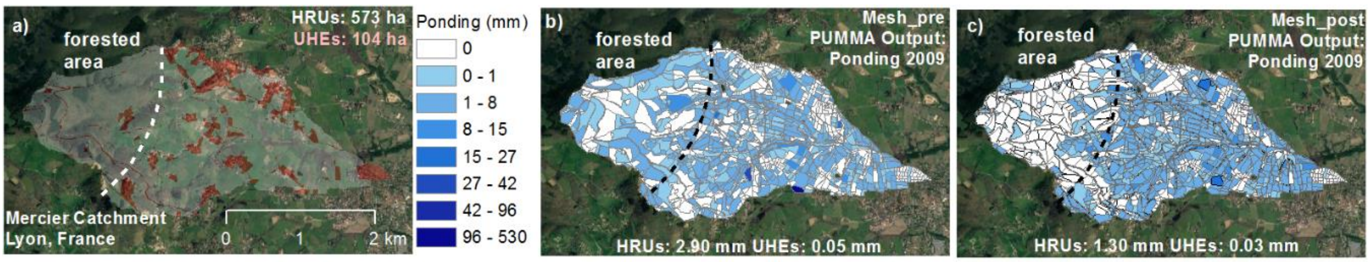


Figure 10 | Application to the Mercier catchment (Lyon, France) for year 2009. (a) The catchment, initial mesh, and location of forested area, (b) average ponding depth obtained from PUMMA model with the initial mesh, and (c) same result obtained using the post-segmented mesh.

CONCLUSION AND FUTURE WORK

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

Author-produced version of the article published in Journal of hydroinformatics, 2019, 21 (1): 104-122
The original publication is available at https://iwaponline.com/jh/article/21/1/104/64103/Decomposition-of-2D-polygons-and-its-effect-in
doi : 10.2166/hydro.2018.031

- The number of final elements of the segmented mesh relates directly to the threshold values for the shape factors. This number increases exponentially with the threshold value of *FF* (*FFT*), and in an irregular manner (i.e., as a step function) with the threshold value of *CI* (*CIT*). Values of *FFT* of 0.2 and 0.4 are recommended for mesh segmentation in undeveloped and highly developed areas, respectively. Conversely, values of *CIT* of 0.8 and 0.95 are recommended for locations with low and high density of green areas, respectively. Very large values of *FFT* or *CIT* can increase significantly the final number of elements in the model mesh, affecting the performance of numerical models using this mesh as a spatial domain.
- High convexity threshold values ($CIT \sim 0.995$) allow obtaining meaningful parts when decomposing silhouette-based features. Because it incorporates the form factor, the proposed strategy is more flexible than existing algorithms, which mostly use the convexity to decompose polygons into pseudo-convex components.
- Overall, mesh segmentation is crucial to avoid bad-shaped elements that affect the realistic representation of hydrological connectivity in distributed hydrological modeling. Moreover, mesh segmentation facilitates the representation of the spatially distributed processes controlling not only the lumped response of the catchment, but the spatial variability of water quantity and fluxes within it.

Studies dedicated to building optimal model meshes (Zundel *et al*. 2002) and testing the performance of modeling tools at small scales (Abily *et al*. 2013) emphasize the importance of mesh refinement due to the time-consuming task involved. The approach here presented addresses this issue as they attempt to improve the construction of irregular model meshes for

Author-produced version of the article published in Journal of hydroinformatics, 2019, 21 (1): 104-122
The original publication is available at https://iwaponline.com/jh/article/21/1/104/64103/Decomposition-of-2D-polygons-and-its-effect-in
doi : 10.2166/hydro.2018.031

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

ACKNOWLEDGMENTS

This work was developed within the framework of Project MAPA (IDRC 107081-001) and Project ECOS-CONICYT C14U02. The work presented was part of the Doctoral thesis funded by the CONICYT PCHA/2014-21140685 and “Estadías en el Extranjeros para Tesistas de Doctorado” VRI-UC grants. Funding from Projects FONDECYT Nº1131131 and 1181506, FONDAP 15110020, FONDECYT 1181506 and IRSTEA-Lyon are also acknowledged. Finally, Jorge Gironás also acknowledges FONDAP 15110017. The assistance in the SWMM model construction from graduate students Alexander Hoch and Tomás Bunster is also appreciated. The Mercier catchment is part of OTHU (Observatoire de Terrain en Hydrologie Urbaine). This work was partially developed within the framework of the Panta Rhei Research Initiative of the International Association of Hydrological Sciences.

REFERENCES

Abily, M., Duluc, C. M., Faes, J. B. & Gourbesville, P. 2013 Performance assessment of modelling tools for high resolution runoff simulation over an industrial site. *Journal of Hydroinformatics*, **15**(4), 1296–1311. DOI: 10.2166/hydro.2013.063

Collon, P., Steckiewicz-laurent, W., Pellerin, J., Gautier, L., Caumon, G., Reichart, G. & Vaute, L. 2015 3D geomodelling combining implicit surfaces and Voronoi-based remeshing: A case study in the Lorraine Coal Basin (France). *Computers & Geosciences*, **77**, 29−43. DOI: https://doi.org/10.1016/j.cageo.2015.01.009

Dehotin, J. & Braud, I. 2008 Which spatial discretization for distributed hydrological models? Proposition of a methodology and illustration for medium to large scale catchments. *Hydrology and Earth System Sciences*, **12**(3), 769–796. DOI: https://doi.org/10.5194/hess-12-769-2008

DHI 2007a MIKE 21 FLOW MODEL, Hydrodynamic Module: Scientific Documentation. Danish Hydraulics Institute, Horsholm, p. 58.

DHI 2007b MIKE 21 & MIKE 3 FLOW MODEL FM, Hydrodynamic and Transport Module: Scientific Documentation. Danish Hydraulics Institute, Horsholm, p. 50.

Downer, C., Ogden, F., Neidzialek, J. & Liu, S. 2006 Gridded surface/subsurface hydrologic analysis (GSSHA) model: a model for simulating diverse streamflow-producing process. In: *Watershed Models* (V. P. Singh and D. K. Frevert, eds). Taylor and Francis, Boca Raton, FL, USA, pp. 131–157.

Fernández-Pato, J., Caviedes-Voullième, D. & García-Navarro, P. 2016 Rainfall/runoff simulation with 2D full shallow water equations: Sensitivity analysis and calibration of infiltration parameters. *Journal*

Author-produced version of the article published in Journal of hydroinformatics, 2019, 21 (1): 104-122
The original publication is available at https://iwaponline.com/jh/article/21/1/104/64103/Decomposition-of-2D-polygons-and-its-effect-in
doi : 10.2166/hydro.2018.031

of Hydrology, **536**, 496–513. DOI: https://doi.org/10.1016/j.jhydrol.2016.03.021

Flügel, W. A. 1995 Delineating hydrological response units by geographical information system analyses for regional hydrological modelling using PRMS/MMS in the drainage basin of the River Bröl, Germany. *Hydrological Processes*, **9**, 423–436. DOI: https://doi.org/10.1002/hyp.3360090313

Fuamba, M., Branger, F., Braud, I., Batchabani, E., Sanzana P., Sarrazin, B. & Jankowfsky, S. Value of distributed water level and soil moisture data in the evaluation of a distributed hydrological model. Application to the PUMMA model in the semi-rural Mercier catchment (6.6 km^2) in France. *Journal of Hydrology* (submitted).

Gironás J., Roesner L. A., Rossman L. A. & Davis J. A. 2010 A new applications manual for the storm water management model (SWMM). *Environmental Modelling and Software*, **25**(6), 813–814. DOI: https://doi.org/10.1016/j.envsoft.2009.11.009

Gonzalez-Sosa, E., Braud, I., Dehotin, J., Lassabatère, L., Angulo-Jaramillo, R., Lagouy, M., Branger, F., Jacqueminet, C., Kermadi, S. & Michel, K. 2010 Impact of land use on the hydraulic properties of the topsoil in a small French catchment. *Hydrological Processes*, **24**(17), 2382–2399. DOI: https://doi.org/10.1002/hyp.7640

Haestad Methods 1995 StormCAD for Windows User's Guide. Waterbury, CT, USA.

Jankowfsky, S. 2011 Understanding and modelling of hydrological processes in small peri-urban catchments using an object oriented and modular distributed approach. Application to the Chaudanne and Mercier sub-catchments (Yzeron Catchment, France). École Doctorale Terre, Univers, Environnement. L'Institut National Polytechnique de Grenoble, Grenoble, France.

Jankowfsky, S., Branger, F., Braud, I., Rodriguez, F., Debionne, S. & Viallet, P. 2014 Assessing anthropogenic influence on the hydrology of small peri-urban catchments: Development of the object-oriented PUMMA model by integrating urban and rural hydrological models. *Journal of Hydrology*, **517**, 1056–1071. DOI: https://doi.org/10.1016/j.jhydrol.2014.06.034

Jiao, L., Liu , Y. & Li, H. 2012 Characterizing land-use classes in remote sensing imagery by shape metrics. *ISPRS Journal of Photogrammetry and Remote Sensing*, **72**, 46–55. DOI: https://doi.org/10.1016/j.isprsjprs.2012.05.012

Karamouz, M., Nazif, S. & Falahi, M. 2012 *Hydrology and Hydroclimatology: Principles and Applications*, CRC Press, Boca Raton, FL, USA.

Kim, J. & Chung, J. 2015 Untangling polygonal and polyhedral meshes via mesh optimization. *Engineering with Computers*, **31**(3), 617–629. DOI: https://doi.org/10.1007/s00366-014-0379-5

Krayevoy, V. & Sheffer, A. 2006 Variational, meaningful shape decomposition. In: Proceedings SIGGRAPH '06 ACM SIGGRAPH 2006 Sketches, Article 50, Boston, MA, USA. DOI: https://doi.org/10.1145/1179849.1179912

Lagacherie, P., Rabotin, M., Colin, F., Moussa, R. & Voltz, M. 2010 Geo-MHYDAS: a landscape discretization tool for distributed hydrological modeling of cultivated areas. *Computers & Geosciences*, **36**(8), 1021–1032. DOI: https://doi.org/10.1016/j.cageo.2009.12.005

Lee, K.-Y., Kim, I.-I., Cho, D.-Y. & Wan Kim, T. 2003 An algorithm for automatic 2D quadrilateral mesh generation with line constraints.

Computer-Aided Design, **35**(12), 1055–1068. DOI: https://doi.org/10.1016/S0010-4485(02)00145-8

Lien, J.-M. & Amato, N. M. 2005 Approximate convex decomposition of polygons. *Computational Geometry: Theory & Applications*, **35**(1–2), 100–123. DOI: https://doi.org/10.1016/j.comgeo.2005.10.005

Lien, J.-M., Liu, G. & Langevin, C. D. 2015 GRIDGEN version 1.0 - A computer program for generating unstructured finite-volume grids: U.S. Geological Survey Open-File Report 2014–1109, 26 p.

Liu, G., Xi, Z., & Lien, J. M. 2014 Dual-space decomposition of 2D complex shapes. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 23–28 June, pp. 4154–4161.

Manus, C. 2008 Hydrate project - contribution to WP6 “analyze of major flash flood events”. Technical report, LTHE, IRSTEA (ex-CEMAGREF).

Miles, J. 1985 The representation of flows to partially penetrating river using groundwater flow models. *Journal of Hydrology*, **82**(3–4), 341–355.

Moussa, R., Voltz, M. & Andrieux, P. 2002 Effects of the spatial organization of agricultural management on the hydrological behavior of a farmed catchment during flood events. *Hydrological Processes*, **16**, 393–412. DOI: https://doi.org/10.1002/hyp.333

O’Rourke, J. 1998 *Computational Geometry in C*, 2nd ed. Cambridge University Press, Cambridge, UK, 376 pp.

Panday, S., Langevin, C. D., Niswonger, R. G., Ibaraki, M. & Hughes, J. D. 2013 MODFLOW–USG version 1: An unstructured grid version of MODFLOW for simulating groundwater flow and tightly coupled processes using a control volume finite-difference formulation. *U.S. Geological Survey Techniques and Methods*, book 6, chap. A45, 66 p.

Pilz, T., Francke, T., & Bronstert, A. 2017 lumpR 2.0.0: an R package facilitating landscape discretisation for hillslope-based hydrological models. *Geoscientific Model Development*, **10**, 3001–3023. DOI: https://doi.org/10.5194/gmd-10-3001-2017

Rodriguez, F., Andrieu, H. & Morena, F. 2008 A distributed hydrological model for urbanized areas – model development and application to case studies. *Journal of Hydrology*, **351**, 268–287. DOI: https://doi.org/10.1016/j.jhydrol.2007.12.007

Rossman, L. A. 2009 Storm Water Management Model User's Manual Version 5.0. EPA/600/R-05/040, National Risk Management Research Laboratory, United States Environmental Protection Agency, Cincinnati, OH, USA.

Russ, J. C. 2002 *The Image Processing Handbook*, 4th ed. CRC Press, Boca Raton, FL, USA, 732 pp.

Russo, B., Sunyer, D., Velasco, M. & Djordjević, S. 2015 Analysis of extreme flooding events through a calibrated 1D/2D coupled model: the case of Barcelona (Spain). *Journal of Hydroinformatics*, **17**(3), 473–491. DOI: 10.2166/hydro.2014.063

Sanzana, P., Jankowfsky, S., Branger, F., Braud, I., Vargas, X., Hitschfeld, N. & Gironás, J. 2013 Computer-assisted mesh generation based on hydrological response units for distributed hydrological modeling. *Computers & Geosciences*, **57**, 32–43. DOI: https://doi.org/10.1016/j.cageo.2013.02.006

Sanzana, P., Gironás, J., Braud, I., Branger, F., Rodriguez, F., Vargas, X., Hitschfeld, N., Muñoz J. F., Vicuña, S. & Jankowfsky, F. 2017 A GIS-based urban and peri-urban landscape representation toolbox for hydrological distributed modeling. *Environmental Modelling &*

Author-produced version of the article published in Journal of hydroinformatics, 2019, 21 (1): 104-122
The original publication is available at https://iwaponline.com/jh/article/21/1/104/64103/Decomposition-of-2D-polygons-and-its-effect-in
doi : 10.2166/hydro.2018.031

Software, **91**, 168–185. DOI: https://doi.org/10.1016/j.envsoft.2017.01.022

Shewchuck, J. 1996 Triangle: Engineering a 2D Quality Mesh Generator and Delaunay Triangulator. In: M. C. Lin and D. Manocha (eds.), *Applied Computational Geometry: Towards Geometric Engineering*. Springer, Berlin, Heidelberg, Germany, pp. 203–222.

Shewchuk, J. R. 2002 Delaunay refinement algorithms for triangular mesh generation. *Computational Geometry: Theory and Applications*, **22**(1–3), 21–74. DOI: https://doi.org/10.1016/S0925-7721(01)00047-5

Si, H. 2006 TetGen, A quality tetrahedral mesh generator and three-dimensional delaunay triangulator. Weierstrass Institute for Applied Analysis and Stochastic, Berlin, Germany.

Tucker, G. E., Lancaster, S. T., Gasparini N. M., Bras R. L. & Rybarczyk S. M. 2001 An object-oriented framework for hydrologic and geomorphic modeling using triangulated irregular networks. *Computers & Geosciences*, **27**(8), 959–973. DOI: https://doi.org/10.1016/S0098-3004(00)00134-5

Versteeg, H. K. & Malalasekera, W. 2007 *An Introduction to Computational Fluid Dynamics, The Finite Volume Method*, 2nd ed. Pearson, Harlow, UK.

Zundel, A. K., Nelson, E. J. & Christiansen, J. R. 2002 Stamped features: creation of engineered structures in conceptual models. *Journal of Hydroinformatics*, **4**(1), 63–71. http://jh.iwaponline.com/content/4/1/63.abstract