

RECORD 2020/10

# 1:500 000 STATE REGOLITH GEOLOGY OF WESTERN AUSTRALIA – COMPILATION METHODOLOGIES

by  
S Jakica, N de Souza Kovacs, J Hogen-Esch and IMT Granado



Government of Western Australia  
Department of Mines, Industry Regulation  
and Safety

EXPLORATION  
INCENTIVE  
SCHEME

Geological Survey of  
Western Australia





Government of **Western Australia**  
Department of **Mines, Industry Regulation  
and Safety**

RECORD 2020/10

# 1:500 000 STATE REGOLITH GEOLOGY OF WESTERN AUSTRALIA – COMPILATION METHODOLOGIES

by

S Jakica, N de Souza Kovacs, J Hogen-Esch and IMT Granada

PERTH 2020



**Geological Survey of  
Western Australia**

**MINISTER FOR MINES AND PETROLEUM**  
**Hon Bill Johnston MLA**

**DIRECTOR GENERAL, DEPARTMENT OF MINES, INDUSTRY REGULATION AND SAFETY**  
**David Smith**

**EXECUTIVE DIRECTOR, GEOLOGICAL SURVEY AND RESOURCE STRATEGY**  
**Jeff Haworth**

**REFERENCE**

**The recommended reference for this publication is:**

Jakica, S, de Souza Kovacs, N, Hogen-Esch, J and Granado, IMT 2020, 1:500 000 State regolith geology of Western Australia – compilation methodologies: Geological Survey of Western Australia, Record 2020/10, 22p.

**ISBN** 978-1-74168-898-6

**ISSN** 2204-4345

Locations in this publication refer to the Geocentric Datum of Australia 1994 (GDA94).



**Disclaimer**

This product uses information from various sources. The Department of Mines, Industry Regulation and Safety (DMIRS) and the State cannot guarantee the accuracy, currency or completeness of the information. Neither the department nor the State of Western Australia nor any employee or agent of the department shall be responsible or liable for any loss, damage or injury arising from the use of or reliance on any information, data or advice (including incomplete, out of date, incorrect, inaccurate or misleading information, data or advice) expressed or implied in, or coming from, this publication or incorporated into it by reference, by any person whatsoever.

**Published 2020 by the Geological Survey of Western Australia**

This Record is published in digital format (PDF) and is available online at <[www.dmirs.wa.gov.au/GSWApublications](http://www.dmirs.wa.gov.au/GSWApublications)>.



© State of Western Australia (Department of Mines, Industry Regulation and Safety) 2020

With the exception of the Western Australian Coat of Arms and other logos, and where otherwise noted, these data are provided under a Creative Commons Attribution 4.0 International Licence. (<http://creativecommons.org/licenses/by/4.0/legalcode>)

**Further details of geoscience products are available from:**

Information Centre  
Department of Mines, Industry Regulation and Safety  
100 Plain Street  
EAST PERTH WESTERN AUSTRALIA 6004  
Telephone: +61 8 9222 3459 Email: [publications@dmirs.wa.gov.au](mailto:publications@dmirs.wa.gov.au)  
[www.dmirs.wa.gov.au/GSWApublications](http://www.dmirs.wa.gov.au/GSWApublications)

**Cover image:** Packing up the campsite in a claypan about 5 km south of Minilya in the southern Pilbara (photo by Olga Blay, DMIRS)

## Contents

Abstract .....	1
Introduction .....	1
Methodology .....	3
Data sources and reliability .....	3
Topology .....	5
Smoothing .....	5
Generalization .....	6
GSWA regolith classification scheme — from old to new .....	6
Downscaling process .....	9
Aggregation .....	9
GeoScaler generalization .....	10
Post-GeoScaler editing .....	15
Coastline adjustments .....	15
Edge matching and topology cleaning .....	17
Physiographic province discrepancy .....	18
Performance and final generalization .....	19
Summary and future development .....	19
Acknowledgements .....	20
Appendix .....	20
References .....	22

## Figures

1. 1:500 000 State regolith geology of Western Australia .....	2
2. The physiographic provinces across Western Australia .....	3
3. Sources used to compile the 1:500 000-scale State regolith map .....	4
4. Example of vertex reduction using generalization .....	6
5. Effect of physiographic province prefixes on regolith codes .....	8
6. Illustration of the aggregation step to preserve information in smaller outcrops .....	9
7. Crosscutting priorities during aggregation of residual or relict (_R) units .....	11
8. GeoScaler Tools menu showing the scripts used in the generalization process .....	13
9. Cellular automata (CA) process explaining the role of Moore's Neighborhood value .....	14
10. Converting the raster dataset to the vector layer .....	15
11. Map downscaling using the process of GeoScaler generalization .....	16
12. Coastline adjustment for 1:500 000-scale State regolith map .....	17
13. Manual edge matching after GeoScaler generalization .....	18
14. Example of regolith polygons containing physiographic province discrepancies .....	18
15. Example illustrating long straight lines in the map due to large polygon splitting .....	19
16. Flow chart summarizing the processes used to create the new State digital regolith map .....	20
17. Example of attributes captured in the State digital regolith map .....	21

## Tables

1. Example of code translation for the 1:500 000 State regolith geology of Western Australia .....	3
2. Eleven landforms the primary regolith codes are based on .....	7
3. Physiographic provinces across Western Australia .....	8
4. List of alluvial units preserved during the aggregation process .....	10
5. Order of importance of residual units .....	12
6. <i>Surficial maps</i> and <i>Bedrock maps</i> modules in the GeoScaler tool .....	13
7. GeoScaler generalization model parameters and their respective values .....	14

## Appendix

*Available with the PDF online as an accompanying digital resource*

List of State regolith code translations from the old system to the current GSWA coding system



# 1:500 000 State regolith geology of Western Australia — compilation methodologies

by

S Jakica, N de Souza Kovacs, J Hogen-Esch and IMT Granada

## Abstract

The digital '1:500 000 State regolith geology of Western Australia' is a synthesis of existing Geological Survey of Western Australia regolith and surface geology maps. This product supersedes the 2003 edition 1:500 000 State regolith map. In comparison to the previous edition, this new map incorporates all regolith coverage available at 1:100 000 and 1:250 000 scales, is based on a revised regolith classification scheme, and was compiled using an automated algorithm for polygon generalization.

Regolith geology from maps at 1:100 000 and 1:250 000 scales was combined to produce a seamless digital regolith coverage suited for use at a nominal 1:500 000 scale. As the synthesis involved downscaling from the original source maps, polygon geometry was modified to suit. The modification of polygon geometries included aggregating small clusters of polygons with the same code into larger shapes, elimination of microfeatures, and simplification of polygon contours using the cellular automata model of the GeoScaler Python script toolbox for ArcGIS. Subsequently, some manual editing was required during edge matching and topology cleaning to improve the polygon geometries consistent with Geological Survey of Western Australia cartographic scale standards.

Coding of regolith map units in the State regolith layer follows the Geological Survey of Western Australia regolith classification scheme with the addition of a suffix representing physiographic province subdivisions across the State. Earlier maps that did not conform to the current scheme were recoded accordingly. These State regolith map codes consist of three parts: i) a primary code specifying the landform and landform qualifier; ii) a secondary code assigning the main compositional information of the regolith material; iii) a physiographic province suffix. For the scale of this product, tertiary codes indicating parent rock or cement were deemed to be too detailed and were therefore incorporated into higher-level codes.

**KEYWORDS:** aggregation, ArcGIS, GeoScaler, map simplification, physiographic provinces, regolith, regolith codes

## Introduction

The blanket of weathered rock, soils and sediments that make up the regolith in Western Australia is not only of scientific and economic importance, it is a fundamental aspect that controls life in Western Australia. According to Eggleton (2001), 'Regolith is the entire unconsolidated or secondarily recemented cover that overlies more coherent bedrock... everything from fresh rock to fresh air.' Regolith amounts to nearly 80% of the surface of Western Australia, comprising depositional and in situ material (Marnham and Morris, 2003). Regolith varies in thickness from <1 m to over 400 m thick. Mineralization hosted in regolith, including resources for gold, iron ore, nickel, bauxite and coal, represented three-quarters of the State's mineral production by value, approximately \$88 billion, in 2019 (Department of Mines, Industry Regulation and Safety [DMIRS], 2019). In addition, all agricultural land and its subsequent production are directly derived from regolith. Finally, water resources, including surface watercourses and groundwater, are also contained in the regolith layer.

In deeply weathered landscapes — such as most of Western Australia — the regolith is extensive and thick, and often acts as a blanket over fresh rock concealing mineralization. However, it also functions as a natural processing plant, breaking down minerals and releasing chemical elements that are either concentrated or dispersed

in the regolith. This process makes the regolith a valuable source of economic resources. This dispersion of chemical elements means it is also an excellent sampling medium where these elements are distributed away from main ore bodies (Butt et al., 2005). Thus, there is a significant need for understanding the regolith as a heterogeneous layer containing in situ and transported weathered materials. These components differ in terms of character, composition, thickness, geological history, position within the landscape, parent rock, and paleo-weathering environment that develops over time.

Attempts have been made to understand how the regolith evolves, with models of regolith development and landscape evolution describing the genesis of regolith types and correlating various regolith units across the landscape (Pain and Ollier, 1992; Pain et al., 2007). By mapping the regolith geology of Western Australia across different geomorphological provinces, 3D models of regolith and landscape evolution can be derived and presented. The compilation of a seamless statewide regolith layer is the foundation for developing those models.

This Record explains in detail the steps taken to compile the new '1:500 000 State regolith geology of Western Australia' layer (Fig. 1; GSWA, 2020). The layer is a seamless, statewide digital map compiled by assembling and downscaling various data sources of published Geological Survey of Western Australia (GSWA) maps.

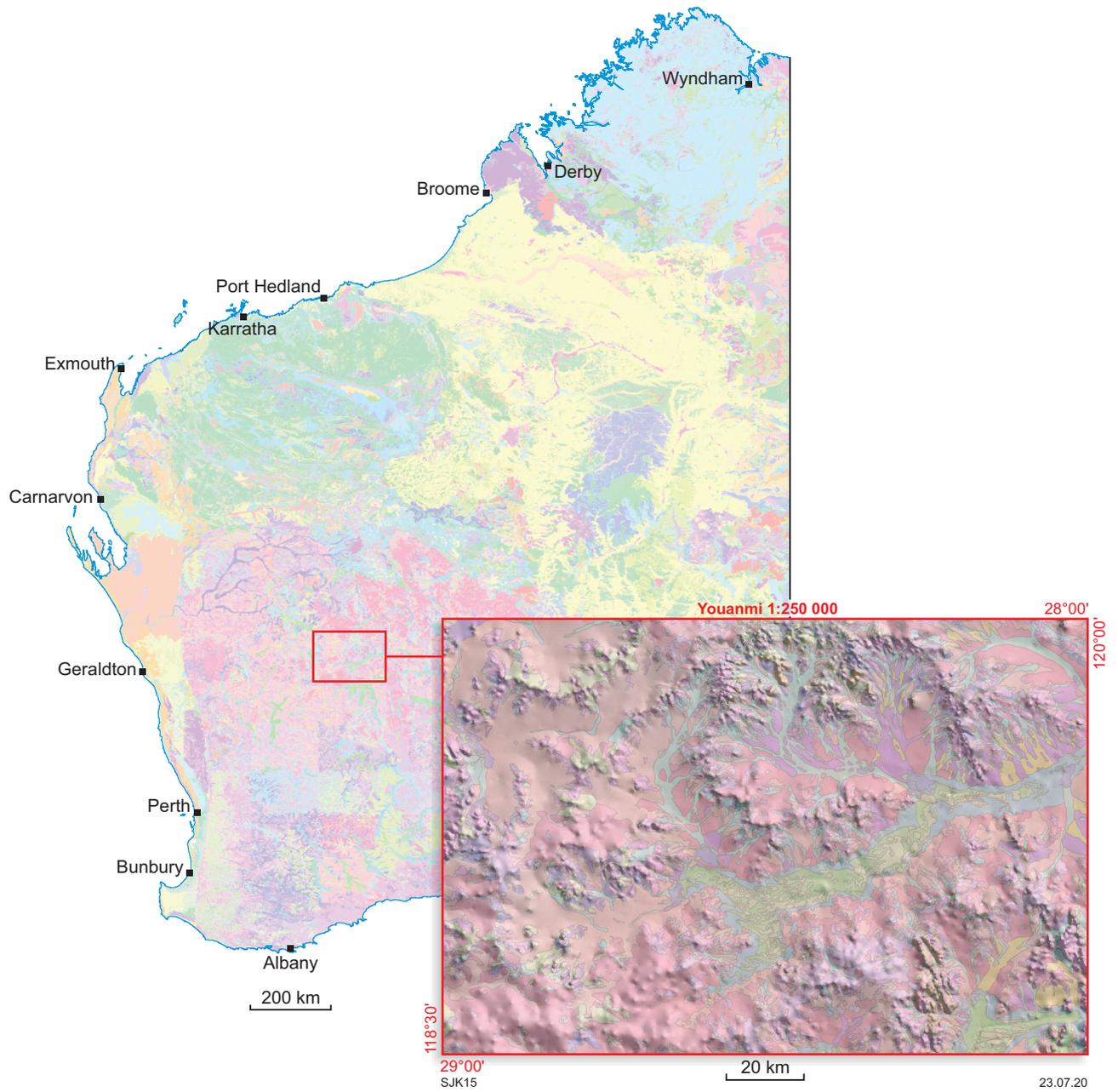


Figure 1. The new 1:500 000 State regolith geology of Western Australia (GSWA, 2020). The map insert shows the 1:250 000-scale YOUNMI map sheet illustrating the level of detail captured. View online via [GeoVIEW.WA](https://www.dmi.wa.gov.au/geoview) and download the map data from [DMIRS Data and Software Centre](https://www.dmi.wa.gov.au/dmirs)

The minimum area (300 000 m<sup>2</sup>), length (1000 m) and width (500 m) of geological polygons conform to the GSWA standard for 1:500 000-scale maps. The resulting map and mapping units depict the regolith according to the GSWA regolith classification scheme (GSWA, 2013). It extends the residual or relict, exposed and depositional (RED) classification scheme of Anand et al. (1993) by expanding the depositional regime into its constituent categories, such as lacustrine or sandplain, specifying composition and provenance of the materials, and adding a geomorphological physiographic province suffix to the code (Fig. 2; Table 1).

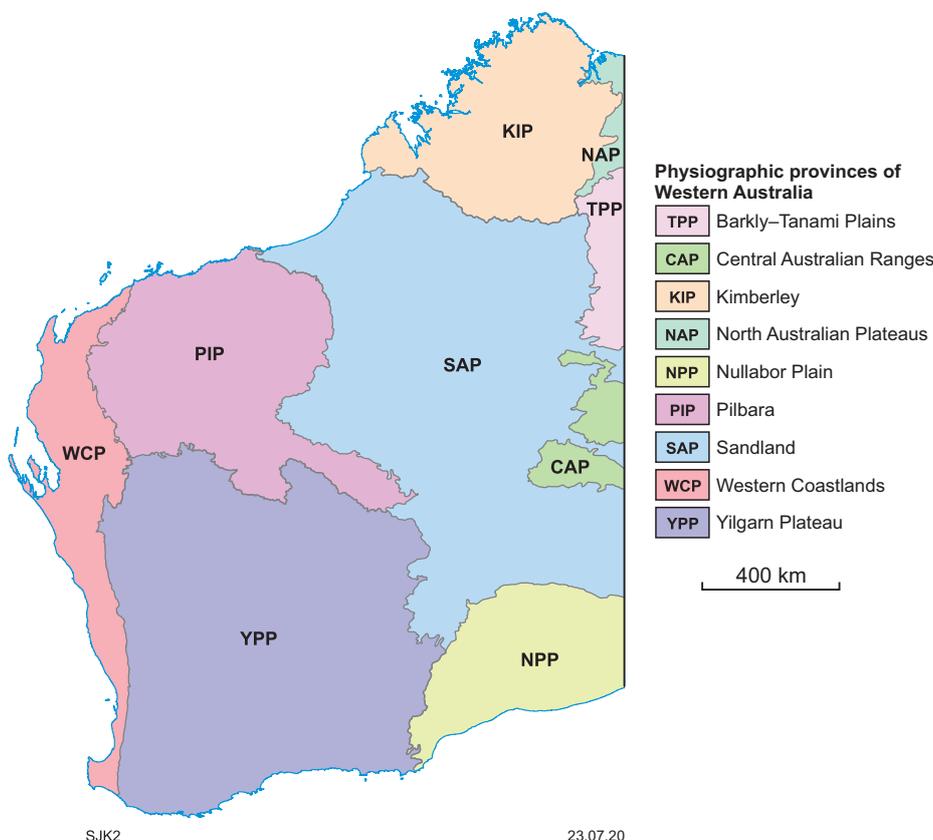
This Record is best read in conjunction with an onscreen view of the 1:500 000 State regolith geology digital layer in [GeoVIEW.WA](http://GeoVIEW.WA).

## Methodology

### Data sources and reliability

To provide full regolith map coverage of the State, several GSWA data source types were used (Fig. 3):

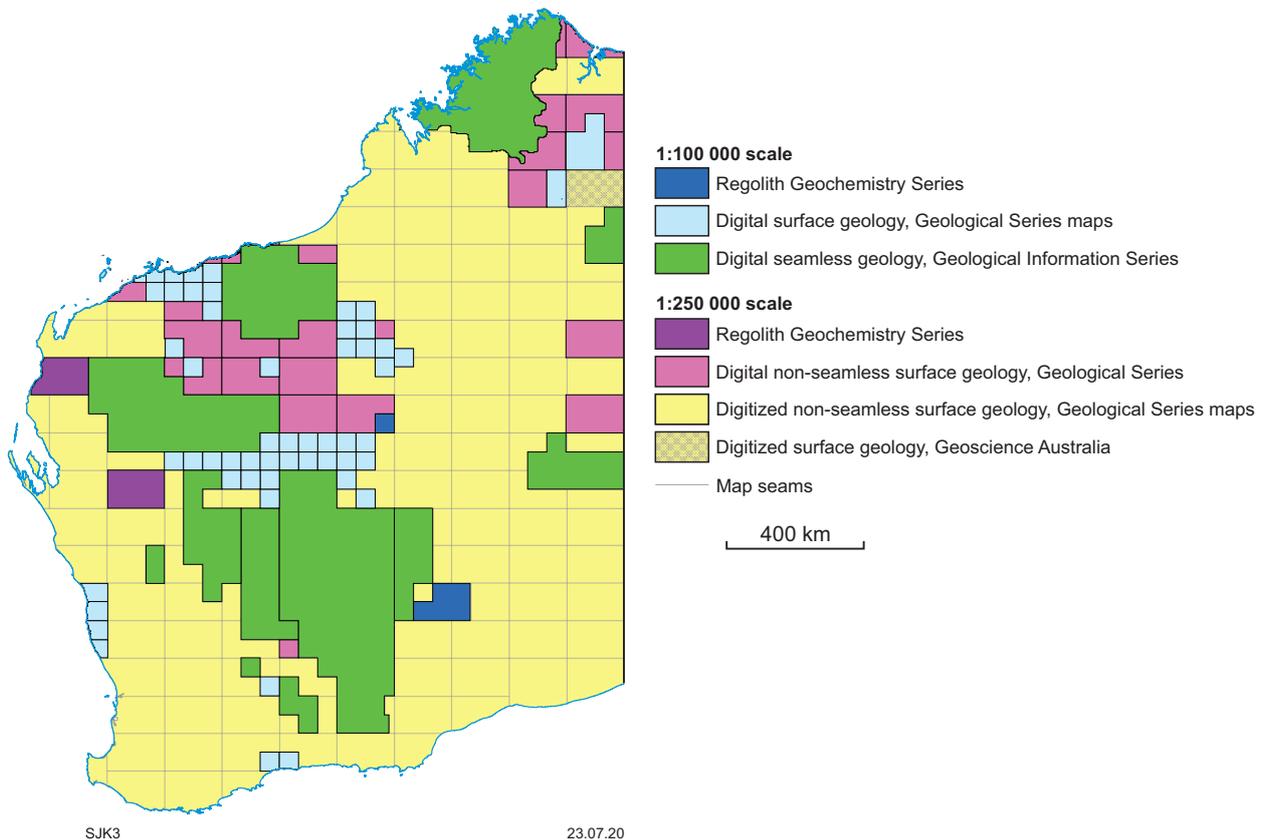
- 1:100 000 Regolith Geochemistry Series; digital seamless regolith geology
- 1:100 000 Geological Series; digital, non-seamless surface geology of individual map sheets
- 1:100 000 Geological Information Series; digital seamless regolith geology



**Figure 2. Physiographic provinces of Western Australia (after Pain et al., 2011) used for the 1:500 000 State regolith map of Western Australia**

**Table 1. An example of code translation for the 1:500 000 State regolith geology of Western Australia layer. The codes follow the GSWA regolith mapping scheme (GSWA, 2013), from which they are simplified to first and second order qualifiers only. The physiographic province suffix is added to the code**

Original codes	Current regolith codes GSWA Record 2013/7	Codes simplified to 1st and 2nd order qualifiers	With physiographic province suffix (Pain et al., 2011)
Ama (amphibolite)	_X-m-m	_X-m	_X-m-YPP
Cczk (calcrete on cemented colluvium derived from ultramafic rock)	_R-k-u	_R-k	_R-k-PIP
CzCb (colluvium; dissected gravel deposits derived from Fortescue Group)	_C-l-m	_C-l	_C-l-PIP



**Figure 3. Map and data sources used to compile the 1:500 000-scale State regolith map**

- 1:250 000 Regolith Geochemistry Series; digital non-seamless regolith geology
- 1:250 000 Geological Series; digital non-seamless surface geology
- 1:250 000 Geological Series; digitized non-seamless surface geology from individual analogue map sheets.

Figure 3 indicates the regions covered by the various sources. Where available, 1:100 000 map data were used. In regions where 1:100 000 mapping was unavailable, 1:250 000 mapping was used. Digital vector data obtained from Geological Information Series (GIS) compilations consisted mainly of 1:100 000-scale regolith geology, which is considered to have the highest level of reliability because it has been mapped in the greatest detail, and because it is the most recent seamless mapping across multiple map sheets. Most of the regolith geology acquired from GIS packages adheres to GSWA Record 2013/7 classification standards (GSWA, 2013), depicting landform and composition, which enabled relatively minimal conversion of regolith classification. These data provided 24% coverage of the State.

Another source that was regarded as having a high level of reliability is the 1:100 000 Regolith Geochemistry Series, but the map data used in the project accounted for less than 1% coverage of the State.

A third source of detailed mapping was the 1:100 000 Geological Series maps. Published as hardcopy maps, the

data consist of individual, digital, non-seamless surface geology that generally were required to be edge matched and undergo comprehensive classification updates. This provided coverage of 6% of the State.

To sufficiently cover the remaining area of the State, 1:250 000-scale mapping was used. The 1:250 000 Regolith Geochemistry Series accounted for only 1% coverage, since overlapping 1:100 000 surface geology mapping was deemed to have a higher level of reliability. The remainder of the mapping coverage used 1:250 000 Geological Series maps, both digital and analogue. Digital surface geology accounted for 10% coverage, whereby vector data existed and met GSWA quality assurance guidelines. Most of these data still required classification conversions. By contrast, the majority (58%) of State 1:250 000 surface geology coverage consisted of historical analogue map sheets that had to be digitized. Individual map sheets had to undergo extensive quality assurance checks. The process used scanned raster images (available via the **Data and Software Centre**) of the original hardcopies, which had to be converted to unpublished vector data.

The quality assurance guidelines for digitizing historical maps ensured:

- correct raster positioning of the map frame; historical projections such as Clarke 1854 and AGD84/AMG84 were re-projected to GDA94 and checked for spatial accuracy

- polygon features were not grouped
- polygons sharing a common boundary were not geologically identical
- clean topology, to a tolerance of 0.0000001 decimal degrees
- no undersize or sliver shapes below 75 000 m<sup>2</sup>
- geology code attribution and polygon shape matched the original publication
- the legend was correctly transferred to a lookup table
- accurate representation of historical overprints and stack codes
- polygon vertex count was not excessively high.

The historical GSWA 1:250 000 Geological Series of maps is considered the source with the lowest level of reliability. The regolith classification systems of this series do not comply with the classification standards of the **Explanatory Notes System (ENS)**. Hence, these digitized source vector data remain unpublished as of February 2020.

ENS is a digital repository of detailed unit descriptions that integrates stratigraphic relationships with links to all tectonic units and events recognized in Western Australia. It replaces the Explanatory Notes previously published to accompany individual 1:100 000 and 1:250 000 Geological Series maps. The system allows incremental updates for new data and refined interpretations, providing a seamless, up-to-date summary of the geology of Western Australia. Explanatory Notes are written for Lithostratigraphic units, Tectonic units, Events and Regolith units.

All map data used were sourced from the GSWA database, with the exception of the GORDON DOWNS, WA Sheet SE52-10 1:250 000 Geological Series map (Smith and Gemuts, 1967) published by Geoscience Australia's predecessor, the Bureau of Mineral Resources.

## Topology

The following was actioned while checking the spatial integrity of all polygon geometries and enforced prior to the use of all source data:

- source data coordinate system – GDA94
- X–Y tolerance of 0.0000001 decimal degrees
- polygons must not overlap
- polygons must not have gaps or holes.

## Smoothing

Automated smoothing removed jagged digitization artefacts from the surface 1:250 000 Geological Series maps. The result more closely resembles the smoothness of the original hand-drawn maps.

The applied tools included a Python 2.7 script developed in-house that uses ET GeoWizards 11.4 toolbox via ArcGIS 10.5. The script enabled batch processing of datasets with consistent results. The method consisted of

disassembling the polygons to individual lines, applying the smoothing algorithm to the linework, followed by rebuilding the polygons. This method produced the intended results, ensuring that node and topological integrity were maintained, and polygon boundary intersections were not manipulated.

The operations undertaken by the script involved:

- ensuring all units were individual units and not grouped, so no multipart features existed (using the ArcGIS tool 'Multipart to single part')
- identifying the geometrical deepest points within polygons, and retaining them for the rebuilding process
- converting the polygon feature boundaries to line features
- removing inconsistencies and duplication of the line features ('Clean'), with a tolerance of 0.0000001 decimal degrees
- applying the smoothing algorithm with the following parameters:
  - algorithm = B-spline
  - smoothness = 7
  - freedom = 5
  - polyline type = closed
- removing excess vertices from the smoothed result, with a generalization tolerance of 0.00004 decimal degrees
- rebuilding the polygon geometry using the previously identified deepest points and the smoothed unit boundary lines
- performing validation check:
  - generating points from the smoothed output
  - creating a comparison file to identify if any rebuilt polygons were assigned the wrong geology or regolith code.

Note that B-spline curves do not pass through the original vertices, but instead follow and lie within the convex hull of the original shape. The parameters and algorithm used provided the closest vector representation to the original linework, but manual adjustment was still required to replicate the original mapping. The tolerance value used in this generalization provided an adequate balance between retaining the shape while removing superfluous vertices. This ensured the overall dataset size remained manageable for further processing and amalgamation stages.

Following the automated smoothing process, the digital source maps required manual adjustment. This was performed in ArcGIS 10.5 and included:

- some global, and then local, spatial adjustments to match the vector polygons against the raster map of the scanned original hardcopy. The scope of the regional adjustment depended on the severity of the spatial inconsistency, and generally involved rubber sheet transformation tools and manual adjustment

- reverting smoothed curves back to jagged geometry if the feature was inherently jagged. This involved manual adjustment using map topology toolsets
- final quality assurance checks, including:
  - polygon topology check
  - minimum size checks, especially for sliver polygons
  - confirmation of geology code attribution.

The final quality assurance check validates an accurate replication of the digital vectors that replace the digital raster scans. Note that this check did not validate the geological accuracy of the historically published maps.

## Generalization

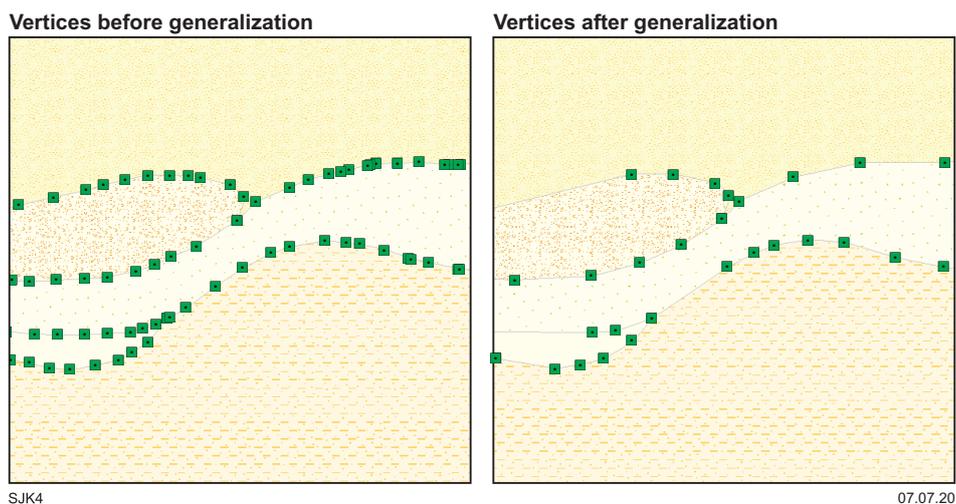
Some unsmoothed vector map sources had too many vertices and required simplification. Likewise, smoothed vector map sources also had many additional vertices added during the smoothing process. These datasets did not necessarily require generalization, since the source mapping had to be simplified in the later downscaling process, but they were potentially more challenging to edit and process, especially with respect to the final nominated 1:500 000 scale of the product. To speed up computer processing and editing time during the quality assurance stage of source digitization, many datasets had the number of vertices reduced using generalization. Figure 4 demonstrates how the geometry of each polygon was simplified, without the shapes being excessively modified.

The tools used for generalization consisted of a Python 2.7 script developed in-house that used ET GeoWizards 11.4 toolbox via ArcGIS 10.5. The method was similar to the smoothing approach, and consisted of disassembling the

polygons to individual lines, applying the generalization to linework, then rebuilding the polygons. This method ensured that topological integrity was maintained and polygon boundary intersections were not manipulated. By default, the geoprocessing tool uses the Douglas–Peucker simplification algorithm to perform the generalization (Douglas and Peucker, 1973). A generalization tolerance of 0.00004 decimal degrees generally suited the data without compromising spatial geometry (with lines modified up to a maximum of 4 m from original positioning). If, in rare cases, spatial geometry would be compromised with that tolerance, the tolerance was reduced to 0.0000001 decimal degrees to avoid this.

## GSWA regolith classification scheme — from old to new

GSWA’s regolith classification scheme and approach to regolith mapping (GSWA, 2013) extends the RED scheme of Anand et al. (1993) by characterizing the regolith as regolith–landform units composed of three parts: i) a primary code describing the landform; ii) where possible, a secondary code describing composition; iii) a tertiary code identifying the parent rock. Due to the scale of the State regolith map, the tertiary code suffixes were too detailed for the intended product and, therefore, were not used for the compilation of the 1:500 000 State regolith map of Western Australia. Instead, a geomorphological physiographic province suffix has been introduced to account for geomorphological differences across the State, which are reflected in the regolith types and their derived landscapes (Pain et al., 2011). For simplicity, this Record only briefly outlines the method used to develop the regolith unit codes for the new regolith map. An additional GSWA Record is in preparation to document the full complexity of building these revised regolith codes.



**Figure 4.** A comparison of geology polygon vertices before automated generalization (left), and after generalization (right). The shapes maintain acceptable spatial integrity, while removing unnecessary vertices

Surface geology maps pre-dating the GSWA 2013 regolith mapping scheme required code translation to the current regolith classification system (Table 1), as per GSWA (2013). The earlier geological maps used a coding system that relied on recording the composition of the regolith, and assigning the age of regolith units as Quaternary, Cenozoic or Tertiary; for example, Cz1 for Cenozoic laterite, Qa for Quaternary sand, Czck for Cenozoic calcrete on cemented colluvium (see Table 1 and Appendix). By compiling all the pre-2013 maps, it was apparent that the same unit code would often correspond to different unit narratives on different maps. This mismatch of unit code and geomorphological setting made it impossible to construct a better geomorphological understanding (see Appendix), especially at statewide scale. Hence, a revised coding scheme was developed that relies on the landform context and geomorphological differences across the State (GSWA, 2013).

Major considerations for development of the new regolith codes were: the landform settings in which regolith is found, the process responsible for its formation, whether the material is in situ or transported, the extent of weathering, the degree of consolidation, and the regolith composition.

Regolith units are assigned to 11 landforms that describe the environment or process responsible for regolith formation. They comprise areas of outcrop, residual or relict units, and nine transported units (Table 2). The landform units can be further subdivided using qualifiers to describe the landform element or patterns, such as colluvial fan (*\_Ca*) or stream channel (*\_Ac*). An underscore is prefixed to the unit code in the GSWA database to distinguish regolith codes from bedrock lithological codes (GSWA, 2013, table 3).

**Table 2. Eleven landforms the primary regolith codes are based on**

<i>Landform unit/code</i>	<i>Landform description</i>
<i>_X</i>	outcrop (including bedrock and weathered rock)
<i>_R</i>	residual or relict material (in situ regolith or remnant deposits from an earlier landscape)
<i>_C</i>	colluvium
<i>_W</i>	sheetwash
<i>_A</i>	alluvial/fluvial
<i>_L</i>	lacustrine
<i>_E</i>	eolian
<i>_S</i>	sandplain
<i>_B</i>	coastal, wave dominated
<i>_T</i>	coastal, tide dominated
<i>_M</i>	marine

To maximize the value of the 1:500 000 State regolith map layer, an optional secondary code was used to describe the regolith material and composition (GSWA, 2013, table 4). For example, the '1' in *\_Ac-1* describes the heterogeneous alluvium in a fluvial channel, and the 'z' in *\_Cl-z* describes silcrete-rich landslide. A suffix representing major physiographic provinces across the State (Pain et al., 2011) was added after the secondary code. The motive behind adding the physiographic description to the regolith unit is to distinguish major physiographic changes across the State by providing additional information on geomorphology and related physical geography. Pain et al. (2011) compiled the map of physiographic provinces using landforms and their processes, as well as potential energy of the landscape. Potential energy of the landscape is influenced by steeper slopes and relatively higher reliefs, representing higher-energy areas with a correspondingly higher rate of erosion and therefore sediment movement. Suffixes used for the physiographic provinces in the 1:500 000 State regolith layer of Western Australia are summarized in Table 3 and illustrated in Figure 2.

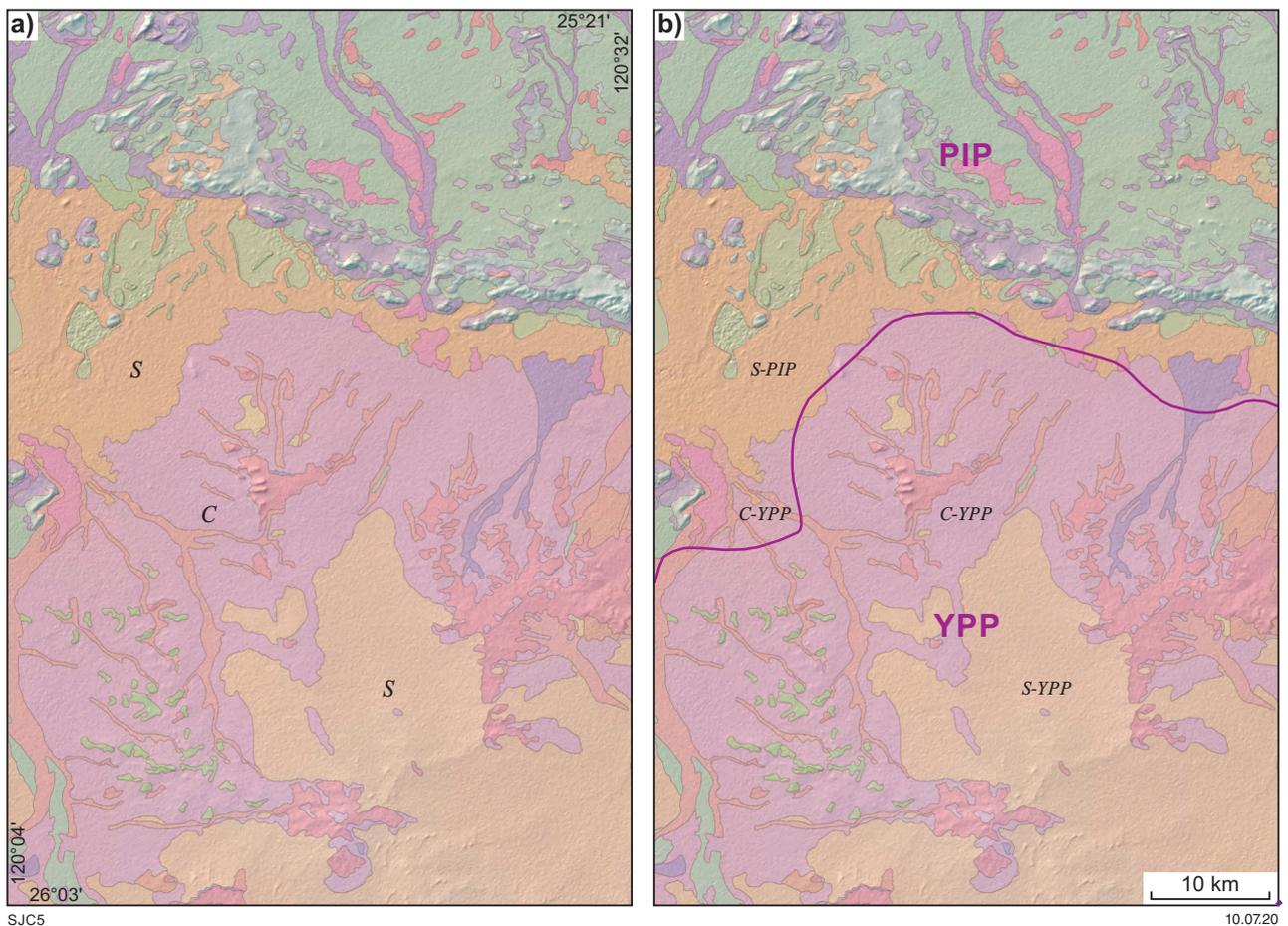
An example of how physiographic provinces are subdivided is shown in Figure 5. In some cases, a polygon will extend beyond the boundary of the designated physiographic region (Fig. 5b). If a polygon extends over two different physiographic regions, as shown in Figure 5, there are two choices: the polygon may require splitting into two polygons representing two neighbouring physiographic regions; or the polygon is preserved and can be maintained with the same name, but a portion of the polygon is permitted to extend beyond the boundary of its physiographic region. Polygon size and geomorphological conditions were considered when deciding the best approach. If a polygon extends only marginally into the neighbouring physiographic region, the polygon and the unit name will stay unchanged. The polygon will keep the name of the dominant unit, in terms of both area and geomorphologic signature.

Regolith units in the legend of the final product (in all available formats) are listed by their physiographic provinces in an order that is guided by an approximate topographical profile. The typical profile progresses from inland to coastal to marine units, and from colluvium through sheetwash, alluvial/fluvial, lacustrine, eolian, sandplain, coastal, marine, residual/relict, to outcrop. Relict or residual units are placed below other regolith units as they are the oldest regolith derived from pre-existing units. Outcrop units are placed below regolith units.

The resulting map layer has regolith units assigned to 11 landforms. They comprise areas of outcrop (*\_X*, including bedrock and weathered rock), residual or relict material (*\_R*, representing in situ regolith or remnant deposits from an earlier or older landscape), and nine transported units: *\_C*, colluvium; *\_W*, sheetwash; *\_A*, alluvial/fluvial; *\_L*, lacustrine; *\_E*, eolian; *\_S*, sandplain; *\_B*, coastal (wave-dominated); *\_T*, coastal (tide-dominated); and *\_M*, marine.

**Table 3. Physiographic provinces across Western Australia**

Suffix	Name	Description
TPP	Barkly–Tanami Plains Physiographic Province	low ranges, plateaus, tablelands and plains
CAP	Central Australian Ranges Physiographic Province	low ranges and tablelands with lowlands and some plateaus
KIP	Kimberley Physiographic Province	plateaus, hills and ridges on resistant rocks, with some valley floors
NAP	North Australian Plateaus Physiographic Province	dissected plateaus and tablelands, some alluvial plains
NPP	Nullarbor Plain Physiographic Province	sandstone and limestone plains, coastal plains and dunes
PIP	Pilbara Physiographic Province	hills and plateaus on resistant rock, some alluvial plains
SAP	Sandland Physiographic Province	low plateaus, sandplains and dunefields
WCP	Western Coastlands Physiographic Province	coastal and alluvial plains, plateaus, scarps and low hills
YPP	Yilgarn Plateau Physiographic Province	erosional plains with sandplains, lakes and alluvial valley floors; breakaways



**Figure 5. Choice of physiographic province suffixes appended to regolith codes in an area divided by a province boundary: a) without the physiographic province boundary, regolith was coded solely on the basis of regolith type; S for sandplain, C for colluvium; b) when added, the physiographic province crosses the colluvial unit; however, because this unit is contiguous and mostly within the Yilgarn Physiographic Province (YPP), the smaller part of the unit north of the province boundary keeps its YPP suffix. In contrast, sandplain units north and south of the boundary take a suffix for the physiographic province in which they are distributed. The regolith map overlies a DEM SRTM 9-second image**

## Downscaling process

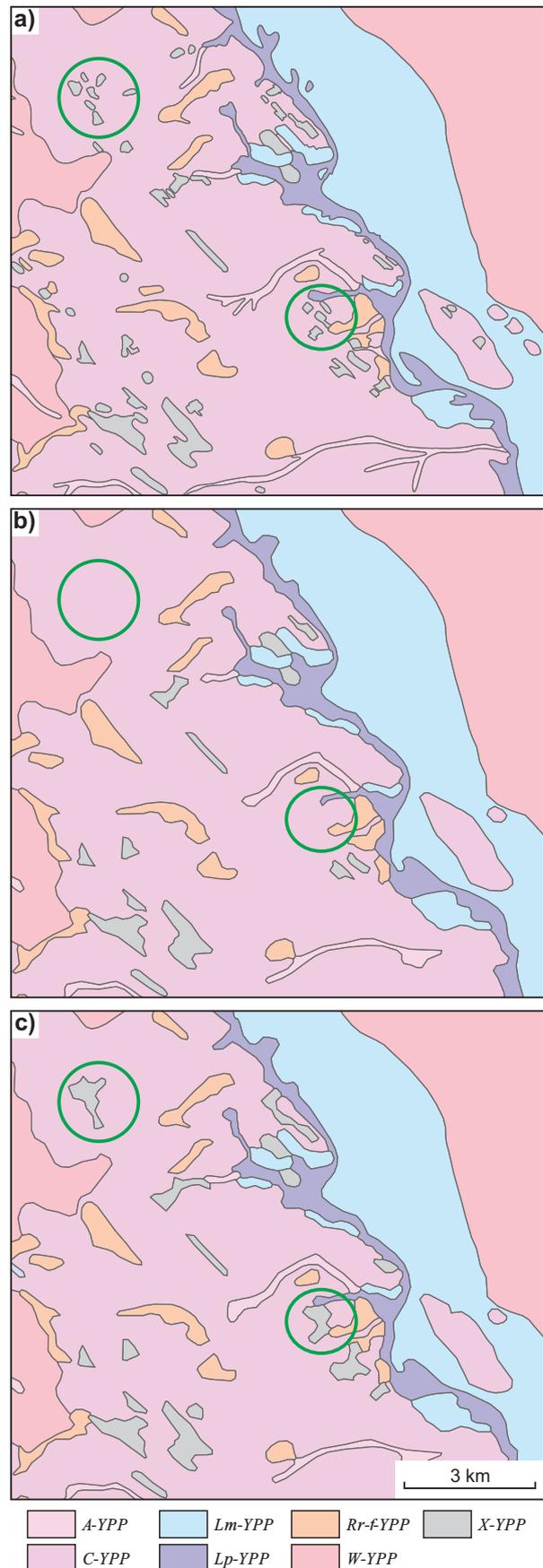
The term ‘downscaling’ is used here to imply map simplification of more detailed mapping (such as 1:100 000 scale) to less detailed mapping (such as 1:500 000 scale) (cf. Smirnov et al., 2012). In some contexts, downscaling is synonymous with the terms ‘simplification’ or ‘generalization’, but these terms can have different meanings when applied by various software or used for a different purpose.

Several aggregation and generalization trials using ArcGIS tools were compared to the cellular automata (CA) model from GeoScaler case studies (Smirnov et al., 2008; Krapf et al., 2012; Smirnov et al., 2012). Standard ArcGIS tools did not provide the intended results in the downscaling process, but the GeoScaler scripts did provide satisfactory test results. An ESRI Spatial Analyst licence was acquired to use specific GeoScaler script tools. GeoScaler script tools that require ESRI Advanced (ArcInfo) or ESRI Standard (ArcEditor) licences were not used in the downscaling process.

To generalize the map from 1:100 000 and 1:250 000 scales to the final scale of 1:500 000, multiple tests were performed to determine the best parameters. Some polygons of scientific or economic significance were lost during the simplification testing process. For example, clusters of small outcrops 200 m apart were completely removed by the GeoScaler process, as they were smaller than the minimum tolerable area set for the particular scale (Fig. 6). This created a map result that was potentially misleading to the user. To preserve these clusters, a process of aggregation was introduced before applying GeoScaler processing.

## Aggregation

To merge a cluster of small polygons of the same regolith unit or bedrock outcrop, several aggregation tools were tested. The aggregation tool from ArcMap toolbox was successful but required an Advanced (ArcInfo) licence. The aggregation tool within ET GeoWizards also produced acceptable results and was the preferred option because it is a more widely available licensing option. An in-house Python script for ArcGIS was developed to automate the selection process of target units, followed by applying the ET GeoWizards geoprocessor to aggregate the units, using a specific order of geological unit priority. This workflow enabled a consistent methodology for the numerous statewide datasets and reduced the complexity or potential errors of manual operation. The script also enabled more efficient processing time than the manual operation.



**Figure 6.** Illustration of the importance of the aggregation step to preserve information in smaller outcrops: a) original layer at 1:100 000 scale showing small clusters of outcropping bedrock units (X-YPP) circled in green; b) GeoScaler generalization to 1:250 000 scale without aggregation step indicating that information on the outcrop is lost (green circle); c) GeoScaler generalization to 1:250 000 scale after aggregation step has been applied; information on the outcrop cluster has been preserved. Regolith codes: \_A-YPP, alluvium; \_C-YPP, colluvium; \_Lm-YPP, Playa and dune; Lp-YPP, playa; \_Rr-f-YPP, ferruginous duricrust; \_W-YPP, sheetwash unit; X-YPP, outcrop

SJK6

23.07.20

The in-house script is specific to GSWA draft datasets and offers no benefit to other external applications, so it is not provided here.

The input file was converted to GDA94 Albers projection prior to running the aggregation script, in order to ensure equal area calculations statewide.

The first step in the aggregation process was to isolate clusters of regolith unit or bedrock outcrop polygons to be merged into a single polygon, which would satisfy the size appropriate for the final map scale. The regolith mapping units of geological and geomorphological significance were all residual materials, weathered rock and outcrops.

Rivers often cut across the outcrop, residual and relict units. To preserve the integrity of the landscape during the aggregation process, the alluvial channel units listed in Table 4 were retained; that is, not deleted where merger of small polygons would have otherwise removed the intervening alluvial channel units. This rule was applied statewide across all physiographic provinces. For simplicity, regolith units in Table 4 do not show physiographic province suffixes.

The second step was to determine suitable circumstances and parameters for polygons to be aggregated or merged, without misinterpreting the regolith geology. After multiple tests, the following script operations and parameters were determined:

- specific geological units were selected by size, such that:
  - for 100 000 to 250 000 aggregation, small polygons selected were up to 140 000 m<sup>2</sup>
  - for 250 000 to 500 000 aggregation, small polygons selected were up to 450 000 m<sup>2</sup>
  - a list of all residual and outcrop units intended for aggregation were reselected
  - a second list of alluvial units was also selected
- selected unit types were separated into individual feature classes for processing
- aggregation of the same type of regolith or outcrop was applied by proximity, such that:
  - for 100 000 to 250 000 aggregation, polygons were within 200 m of each other
  - for 250 000 to 500 000 aggregation, polygons were within 400 m of each other

- regolith unit codes were lost in the aggregation process, so the script reassigned the regolith unit codes to the resulting polygons
- individual aggregated results of residual units and outcrops were combined, and then added back to the original data. Alluvial units were then re-inserted, to ensure they were not compromised.

If the polygon did not fall into the above categories, no changes were applied.

If there were multiple clusters of residual units close to each other, aggregating the clusters into larger polygons might mean these polygons would then crosscut each other (Fig. 7). The type and position of the residual units in the landscape indicate which unit would take priority, and which residual or relict unit overlies or cuts other units (Table 5). The logic used for the order of importance of the residual units was that all relict and transported residual units (\_Rt) would take precedence over other residual units. This rationale was used because relict units are generally found topographically higher in the landscape compared to neighbouring units due to relief inversion. Residual duricrust (\_Rr) and undivided residual or relict units (\_R) generally overlie in situ saprolite, saprolithic rocks, and weathered in situ rocks, hence they were placed as the second most important residual units in this project. Residual or relict sands (\_Rs), which are the product of weathered \_Rr and \_R, can overlie saprolite and weathered rock (\_Ri).

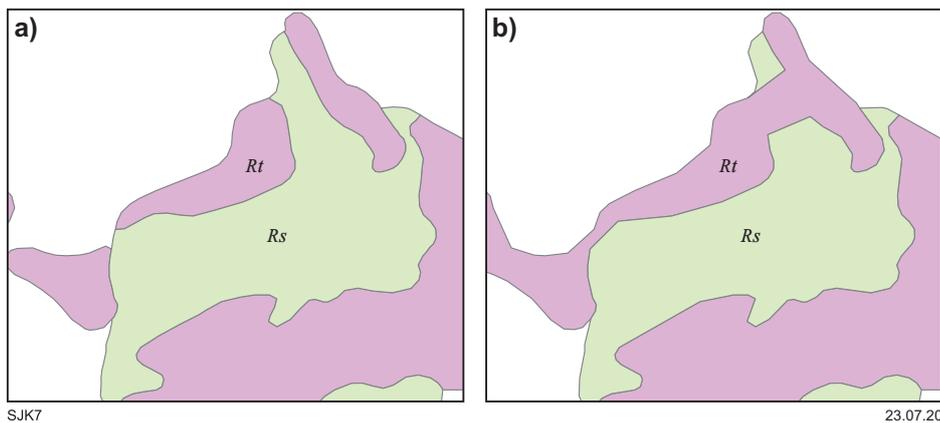
For a practical example of how this ordering affects the construction of the polygons, consider clusters of two neighbouring residual units \_Rt and \_Rs in Figure 7. After aggregation, these clusters will be reduced to two polygons (still \_Rt and \_Rs) that may overlap. Based on the priorities discussed above, polygon \_Rt will overlap or crosscut polygon \_Rs (Fig. 7b). The geomorphological physiographic provinces do not affect the order of importance during the aggregation stage, so the suffix has been disregarded in the codes listed in Table 5.

## GeoScaler generalization

Following the aggregation process, the map was simplified by combining the automated generalization of the *Surficial maps* and the *Bedrock maps* modules of the GeoScaler software. First, all 1:100 000-scale map sheets were downscaled to 1:250 000-scale layers. Subsequently, 1:250 000 layers were downscaled to a 1:500 000-scale layer. The following section describes the parameters

**Table 4. List of alluvial units preserved during the aggregation process in order to keep the integrity of the landscape**

Code	Description	Rationale
_A	undivided alluvial/fluvial	undivided alluvium commonly used to describe river channels and creeks
_Ab	stream bed	river channel or stream bed that commonly cuts through the residual/relict regolith and outcrop units
_Ac	stream channel	river channel or stream bed that commonly cuts through the residual/relict regolith and outcrop units
_Ac-k	stream channel; carbonate rich	river channel or stream bed that commonly cuts through the residual/relict regolith and outcrop units
_Ac-t	stream channel; lithic	river channel or stream bed that commonly cuts through the residual/relict regolith and outcrop units



**Figure 7. Crosscutting priorities during aggregation of relict or residual units: a) before aggregation; examples of different residual units (*\_R*) regolith polygons prior to aggregation; b) after aggregation; polygons crosscutting each other in the correct order of priority after being aggregated**

and methods used to perform downscaling using the two GeoScaler modules and the reasons we chose to use both modules instead of the *Surficial maps* module only.

### GeoScaler software

GeoScaler Tools 10 is a free downloadable Python script toolset intended for ArcGIS 10.0 or 10.1 <<http://geoscan.nrcan.gc.ca/>> (Fig. 8). The tool was developed at Laboratoire de Cartographie Numérique et de Photogrammétrie (LNC), the Quebec division of the Geological Survey of Canada (Huot-Vézina et al., 2012). It has been developed exclusively to generalize surficial and bedrock geology maps. It provides utilities for polygon generalization and considers most of the specifics relevant to geology. Some of the tools required an ESRI Advanced (ArcInfo) or Spatial Analyst licence. The GeoScaler Tools specifically used in this project did not require the Advanced licence, and version 10.5 of ArcGIS was used.

### GeoScaler processing and results

Once aggregation was completed, multiple tests were performed on GeoScaler generalization to see what parameters worked best to downscale the map to the desired 1:500 000 scale. Maps of different scales were used. The first step was to downscale all 1:100 000-scale maps to 1:250 000 scale. From here, the original 1:250 000 maps were merged with downscaled 1:250 000 maps and tested to define the parameters for further downscaling to 1:500 000. The same steps were taken for both map scales during the GeoScaler generalization (Table 6); only the parameters taken for different scales differed (Table 7).

### Modules used

The *GeoScaler environment* module (Fig. 8) was applied before attempts were made at generalization. During this process, the user is required to enter the initial and desired scale, resulting in better data management and overall generalization process. Following this initialization step, we ran the generalization process over the test area using both *Surficial maps* and *Bedrock maps* modules to define which tools from both modules were best suited for the requirements of our product.

### Data preparation

The first step was to choose the value field and specify a cell size for the output raster. If the input file did not have a value field, it was created in ArcMap and populated in the following six steps using the unique geological field:

1. Summarize the codes from the original input file.
2. Add the field to the summarized table naming it, for example, 'Code-ID', and under 'type' choose 'double'.
3. Populate the newly created 'Code-ID' field with the 'Object\_ID' field using the Field Calculator.
4. Go back to the original input file and add a new field called 'Code-ID', choose 'type' to be 'double'.
5. Join the input table with the summarized table by 'Codes' and populate the empty field of 'Code-ID' in the input table with the 'Code-ID' from the summarized table.
6. Remove the summarized table from the original input table. Now the input table is ready to be converted to the raster using the *Data preparation* step.

The desired cell size for the conversion from vector to raster depends on the target scale. It defines the resolution of the output raster. Tests at higher resolution — for example, using a cell size of 15 — resulted in preservation of unnecessary polygons. In contrast, tests at lower resolution — for example, using a cell size of 25 — produced overly simplified maps. The optimal cell size chosen for 1:250 000-scale and 1:500 000-scale target maps was 20 and 25 respectively (Table 7).

### Cellular automata

The second step in GeoScaler generalization processing is *Cellular automata* (Fig. 8). This step generalizes polygons from a raster dataset based on an interactive procedure. It requires several new parameters to be defined:

- input raster
- Moore's Neighborhood Radius — sets the pixel width that cellular automata will be; the value ranges from 1 to 5

**Table 5. Order of importance of residual units based on their position in the landscape**

<i>Code</i>	<i>Description</i>	<i>Rank</i>
_Rt	Relict, transported duricrust	1
_Rt-f	Relict, transported hematitic duricrust	2
_Rt-fl	Relict, transported limonitic duricrust	3
_Rt-q	Relict alluvial deposits; quartz rich	4
_Rt-t	Relict, transported alluvial deposits containing lithic gravel and sand	5
_Rr-xa	Aluminous duricrust	6
_Rr-f	Ferruginous duricrust	7
_Rr-z	Silcrete duricrust (residual or relict)	8
_Rr-zo	Silcrete; includes calcite	9
_Rr-k	Calcrete duricrust (residual or relict)	10
_Rr-kc	Calcrete, nodular to massive; includes calcite and reworked carbonate products	11
_Rr-g	Duricrust, mainly of ferruginous and quartzofeldspathic composition	12
_Rr	Duricrust; undifferentiated	13
_R-f	Residual or relict ferruginous materials	14
_R-k	Residual calcrete and nodular carbonate deposits; includes reworked carbonate products	15
_R-z	Silcrete and brecciated siliceous caprock	16
_R-xm	Residual or relict manganese-rich materials	17
_R-t	Residual or relict lithic materials	18
_R-g	Residual or relict; quartzofeldspathic materials	19
_R-c	Residual or relict; clay-rich materials	20
_R-l	Residual or relict; heterogenous composition	21
_R-d	Residual or relict; lithic materials	22
_Rs-f	Residual, ferruginous sand, with minor clay and silt	23
_Rs-g	Residual, quartzofeldspathic sand, commonly over felsic igneous rocks	24
_Rs-k	Residual sand and calcrete; nodular and pisolitic	25
_Rs-q	Residual, quartz-rich sand, with minor clay and silt	26
_Rs-l	Residual sand of heterogeneous composition	27
_Rs	Residual sand; may contain pisoliths and nodules	28
_Ri-f	In situ weathered rock, saprolite; ferruginous	29
_Ri-fg	Hematite–limonite gossan	30
_Ri-k	In situ weathered rock, saprolite; carbonate rich	31
_Ri-z	Residual siliceous cap rock	32
_Ri-m	Saprolite; ferromagnesian materials derived from weathered mafic rock	33
_Ri-g	Saprolite; quartzofeldspathic sand, silt and clay derived from weathered bedrock	34
_Ri-q	Saprolite; quartz sand, silt and clay derived from weathered bedrock	35
_Ri-ck	Weathered rock dominated by kaolin and quartz; protolith uncertain	36
_Ri-t	In situ weathered rock; rock fragments	37
_Ri-w	Deeply weathered rock; protolith undetermined	38
_Ri-l	In situ weathered rocks; heterogeneous composition	39
_Ri	In situ weathered rocks; undifferentiated	40

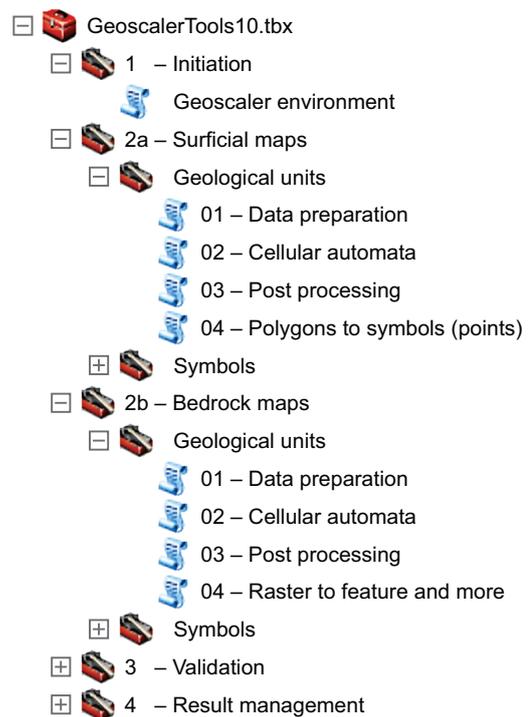
**Table 6. Surficial maps and Bedrock maps modules in the GeoScaler tool used for the State regolith map downscaling. The table indicates the similarities and differences between the four steps of the two modules**

<i>Surficial maps</i>		<i>Bedrock maps</i>	
<i>Geological units</i>			
1	<i>Data preparation</i> Converts any feature into a raster dataset for generalization purposes. In options, it is possible to add hydrographic network	1	<i>Data preparation</i> Converts any feature into a raster dataset for generalization purposes. In options, it is possible to include faults from a given feature class
2	<i>Cellular automata</i> Generalizes polygons from a raster dataset, based on an iterative procedure called Cellular Automata. In options, small polygons can be managed and multiple outputs extracted	2	<i>Cellular automata</i> Generalizes polygons from a raster dataset, based on an iterative procedure called Cellular Automata. In options, faults and geological codes of choice can be preserved
3	<i>Post processing</i> Converts generalized raster dataset into a feature class. In options, it is possible to crop output feature class with border feature layer	3	<i>Post processing</i> Refines results by managing polygon size and preserved codes
4	<i>Polygon to symbol</i> Transfers small polygons, previously eliminated by the generalization procedure, into point symbols	4	<i>Raster to feature and more</i> Converts generalized raster dataset into a feature class. In options, polygon contours can be smoothed and edges cut with any border feature class

- final cellular automata — final iteration value from 1 to 250
- output raster

Advanced options:

- manage small polygons — minimum polygon size in cells based on target scale (*Surficial maps* module only)
- multiple output — the number of intermediate layers that a user would like to keep within the final result (*Surficial maps* module only)
- preserve faults — freezes the boundaries of the polygons along the faults during the generalization process to preserve the break aspect of the fault between the contacts of the units and the faults (*Bedrock maps* module only)
- preserve other codes — freezes the lithological units without delimiting a number of neighbourhood cells (Huot-Vézina et al., 2012). The code of certain units that a user specifies in this step will not be generalized (*Bedrock maps* module only).



SJK16

23.07.20

**Figure 8. GeoScaler Tools menu showing the modules GeoScaler environment, Surficial maps and Bedrock maps with their scripts used in the generalization process of downscaling source data from 1:100 000 and 1:250 000 to the nominated final scale of 1:500 000**

**Table 7. GeoScaler generalization model parameters and their respective values**

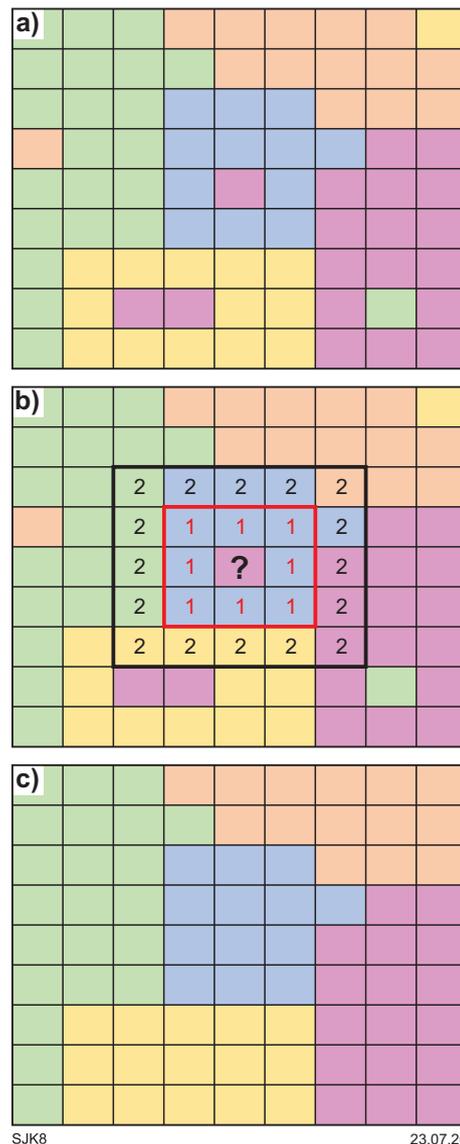
Parameter name	Parameter value	
Initial to target scale	1:100 000 → 1:250 000	1:250 000 → 1:500 000
Cell size (raster resolution)	20	25
Moore's Neighborhood Radius	2	2
Final cellular automata	80	120
Preserve other codes	all _X and _R codes	all _X and _R codes
Manage small polygons	200	475

The value chosen for Moore’s Neighborhood Radius defines the number of neighbour cells on each side of the cell being analysed (Fig. 9; Huot-Vézina et al., 2012). Only the cells within the chosen neighbourhood radius can affect the cell(s) under consideration (Fig. 9b,c). The higher the Moore’s Neighborhood value, the smoother the final result will be (Huot-Vézina et al., 2012). The values used for scale conversions to 1:250 000\* and 1:500 000 layers are listed in Table 7. Final cellular automata defines the number of iterations that the application must perform. The valid range is from 1 to 250 with higher number of iterations yielding greater generalization levels. The value is defined empirically with no set rules to define the optimal number of iterations for a given generalization. To obtain the best results for a given generalization, several tests were completed on the small section of the test area. For the State regolith map of Western Australia with a scaling of 1:250 000 to 1:500 000, we found that 120 iterations were appropriate (Table 7).

The ‘Manage small polygons’ option removes all undesired polygons, and those that are too small at the target scale, from the input dataset. For our new State regolith map, we found that the minimum polygon size value by number of cells is 475. Applying this number, most of the polygons that were smaller than the minimum allowable size for the 1:500 000-scale map (minimum polygon area of 300 000 m<sup>2</sup> and minimum polygon width of 500 m) were removed. However, this removal included outcrop and in situ residual units. To preserve information from these, we decided that those codes must be preserved before removing undesired small polygons. Since the *Surficial maps* module does not have an option to preserve desired codes, we opted to use the *Bedrock maps* process for *Data preparation* and *Cellular automata* steps. To preserve a certain code, a ‘Code-ID’ value for that particular code was entered in the ‘Preserve other codes’ option.

**Post-processing**

Requirements for the *Post processing* step (Fig. 8) differ significantly between *Surficial* and *Bedrock maps* modules (Table 6). In the *Surficial maps* module, this step converts the previously generalized input raster dataset to a feature class. In contrast, in the *Bedrock maps* module the *Post processing* step refines results of the generalized raster dataset by managing polygon size and preserved codes. Since managing small polygons



**Figure 9. Cellular automata (CA) process explaining the role of Moore’s Neighborhood Radius value: a) input raster layer with different colours representing different geological units; b) cell with question mark is a cell under analysis, Moore’s radius of 1 is outlined in red, Moore’s radius of 2 is outlined in thick black line; c) final generation of cells after cellular automata with Moore’s radius 2 applied. The CA will match the state most frequently occurring within its neighbourhood, which includes the two closest cells in every direction**

was missed in the previous cellular automata step of the *Bedrock maps* module, prior to converting the raster file back to vector format, polygons that were too small at the target scale had to be removed.

Accordingly, the next step in generalizing the State regolith map of Western Australia, was to run the *Post processing* step from the *Bedrock maps* module to remove all small polygons (sizes listed in Table 7). Empirical tests were required to define the optimum size of polygons. It was also possible to preserve certain codes from removal, even if they are smaller than the desired target size, by using the ‘Keep preserved codes’ option. In our case, this step was not used since outcrop and residual codes were already preserved in the previous cellular automata step.

### Convert the raster dataset to feature class

Following the generalization process, the final step was to convert the geological units of the raster dataset back to a vector representation. To do so, one can use either the *Post processing* step in the *Surficial maps* module, or the *Raster to feature and more* step in the *Bedrock maps* module (Fig. 8). After running both steps over our test area, we found that the *Raster to feature and more* step was producing irregular polygon outlines in the resulting vector layer (Fig. 10a). The *Post processing* step in *Surficial maps* module was used instead and proved to give better results with smooth polygon outlines (Fig. 10b).

Figure 11 summarizes the process by illustrating the change in the map undergoing downscaling from the original 1:100 000 scale to the desired 1:500 000 scale. The process illustrates the original 1:100 000-scale map that has first been aggregated and then generalized using GeoScaler Tools to a 1:250 000-scale map. The resultant 1:250 000-scale map has then been aggregated and generalized to the target 1:500 000-scale regolith map. The quality and geological details of the map undergoing downscaling from the original 1:100 000 to target 1:500 000 scale are very well preserved.

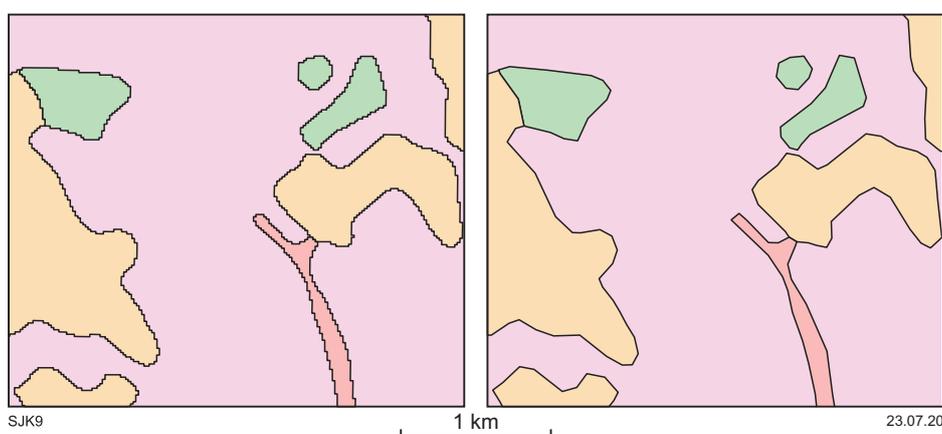
## Post-GeoScaler editing

### Coastline adjustments

The coastline for the 1:500 000-scale State regolith map was adjusted to partially align with the 1:500 000 State interpreted bedrock geology of Western Australia, 2016 (IBG) coastline (GSWA, 2016), where the coastlines were coincident. Due to the inclusion of marine (\_M), tide-dominated coastal (\_T) and wave-dominated coastal (\_B) units, there are some differences between the two coastlines (Fig. 12). The coastline as presented in the State regolith map is not an administrative boundary and should not be used as such or in a legal context.

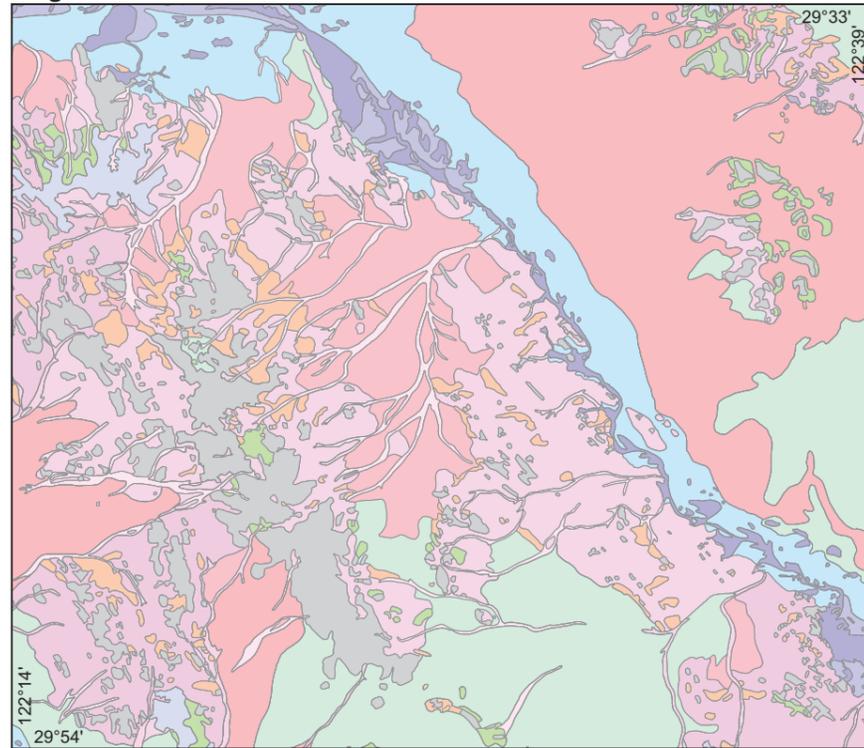
The coastline boundary adjustment was conducted using the following in-house guidelines (also illustrated in Figure 12):

1. ‘update’ areas — regolith coastline was adjusted to the IBG when alignments were within 500 m of each other
2. ‘keep’ areas with the following primary codes: \_M (marine), \_T (tidal), \_B (coastal) — marine units, tide-dominated coastal and wave-dominated coastal units were extended, in an offshore direction, to match the IBG coastline, but not reduced (in an onshore direction) by the IBG. These units tend to extend beyond the IBG coastline
3. transition between ‘update’ and ‘keep’ areas — regions of transition were checked against orthophotos, and Landsat and Google Earth imagery, to establish how to adjust the boundaries
4. islands — offshore islands with an area less than 300 000 m<sup>2</sup> were removed, based on GSWA cartographic guidelines for 1:500 000-scale mapping
5. channels — IBG drainage channels and estuaries were considered too detailed and complex for 1:500 000-scale depiction, especially when attempting to align the regolith geology. Drainage channels and estuaries less than 500 m wide were removed by trimming the estuaries and extending the regolith unit.

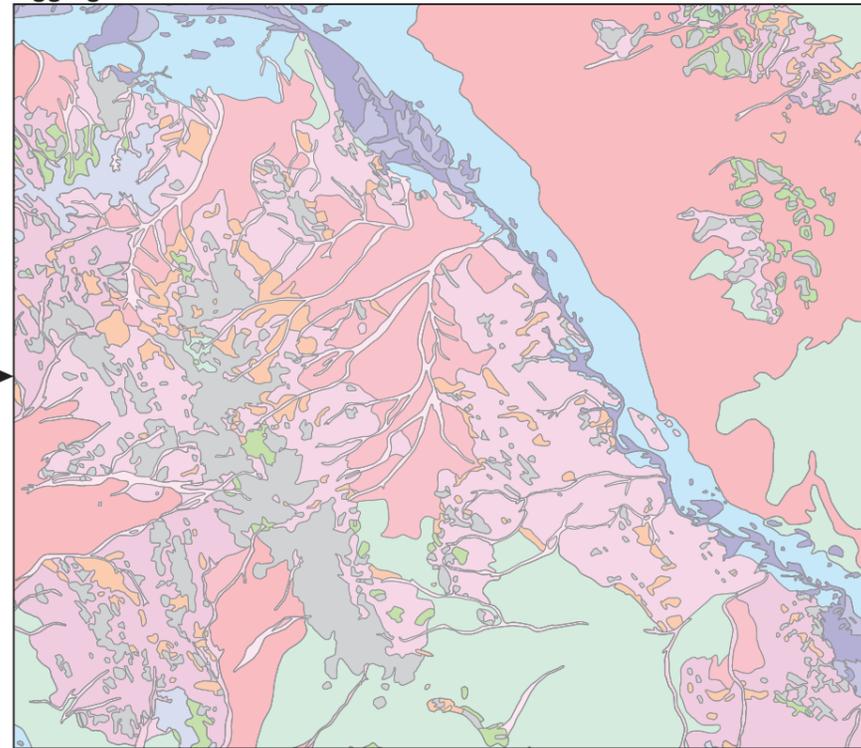


**Figure 10.** Converting the raster dataset to the vector layer: a) the *Raster to feature and more* step in *Bedrock maps* module produced irregular polygon outlines in the resulting vector layer; b) the *Post processing* step in *Surficial maps* module produced smooth polygon outlines in the resulting vector layer

Original 1:100 000



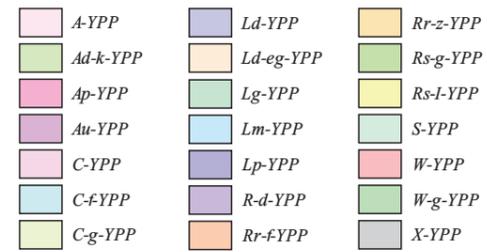
Aggregated 1:100 000



GeoScaler generalization to 1:250 000



10 km



Aggregated 1:250 000



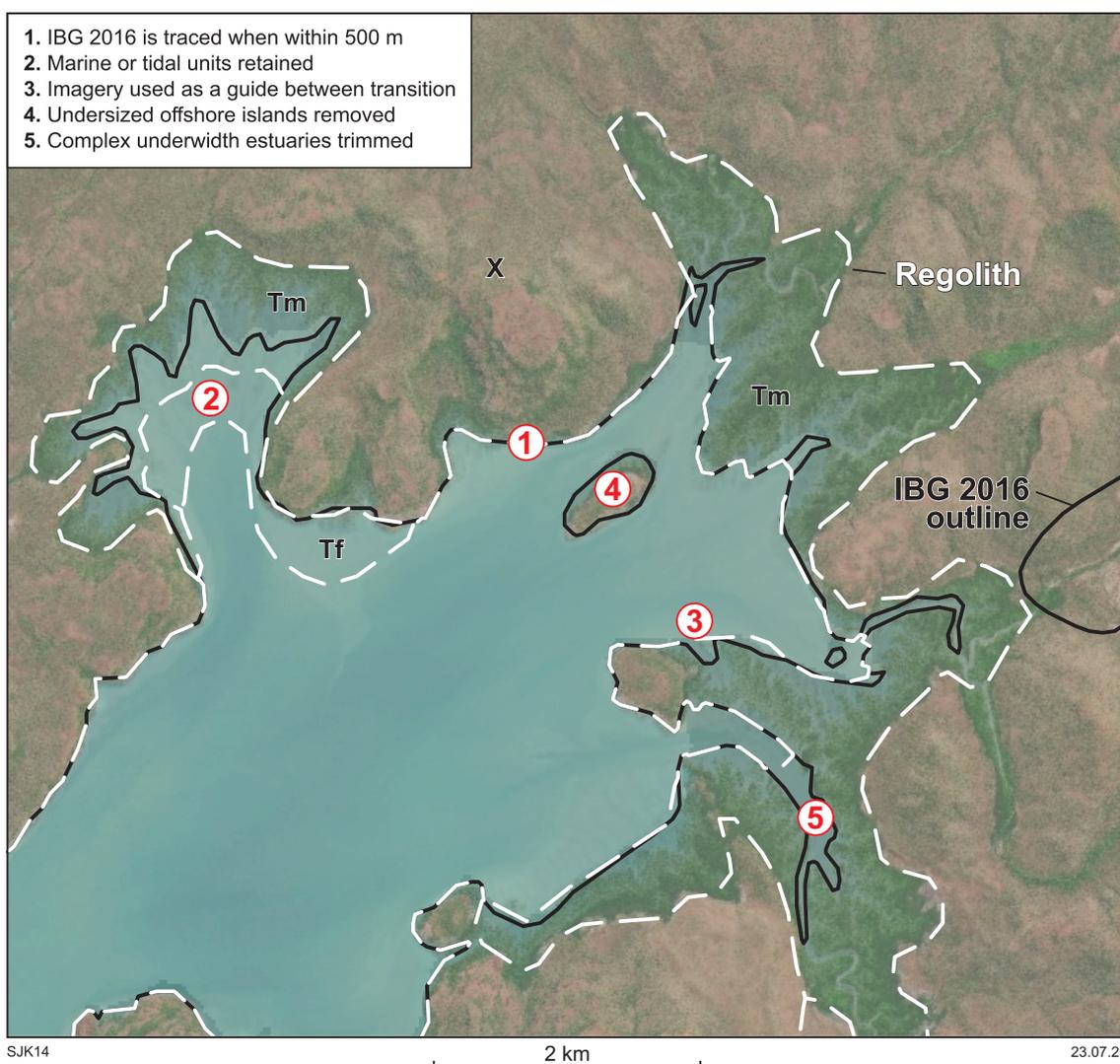
GeoScaler generalization to 1:500 000



Figure 11. Map downscaling from the original 1:100 000 scale to the target 1:500 000 scale. The example illustrates the changes that the map underwent in the aggregation and GeoScaler generalization steps to reach the final product. The geological details stay very well preserved. Information on regolith units in legend is available in [GeoVIEW.WA](#)

SJK10

24.07.20



**Figure 12.** Coastline adjustment for 1:500 000-scale State regolith map. Examples numbered 1–5 illustrate the steps described in the text to adjust the State regolith map coastline so it complies with the coastline of the 1:500 000 State interpreted bedrock geology, 2016 (GSWA, 2016)

### Edge matching and topology cleaning

Following aggregation and GeoScaler generalization of 1:100 000- and 1:250 000-scale maps to 1:500 000-scale State regolith geology layer, manual editing was required during the edge matching and topology cleaning to improve the polygon line work to comply with GSWA cartographic scale standards.

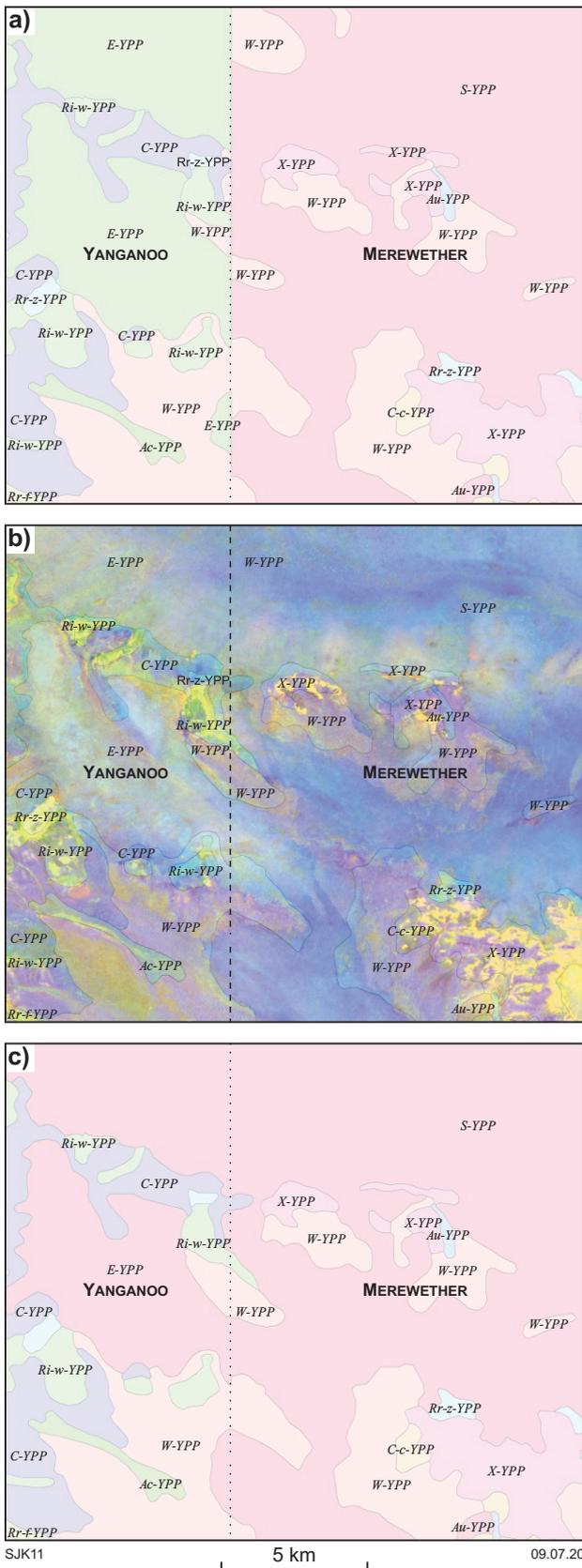
Since the layer was a compilation of maps from different generations and projections, such as Clarke 1854 and AGD84/AMG84, the new layer processed by GeoScaler was reprojected to GDA94 and spatial accuracy was verified.

Most of the polygon editing was along the margins of 1:100 000 and 1:250 000 map sheets. These map seams are indicated in Figure 3. Data were visually compared against published geology maps as well as Landsat and orthophoto images to check the capture and attribution. Typical problems were polygons that terminated at the

100 000 and 250 000 map margins, polygons misaligned at the map sheet margins, and codes mismatched along the margins (Fig. 13a). This was corrected by manually editing the polygon shape or changing the code and merging it with the neighbouring polygon by checking the line work against imagery (Fig. 13b,c).

The most significant topology errors were gaps between polygons or polygons overlapping. This was corrected by running the topology error test with the following rules: ‘must not overlap’ and ‘must not have gaps’. Where such errors occurred, the line work was compared against Landsat images or geology maps, and manually corrected.

Finally, the map was rechecked for small polygons; specifically, those smaller than the minimum tolerable area for a 1:500 000-scale map sheet complying with GSWA cartographic scale standards, and which had not been generalized during the GeoScaler process. There were a number of such polygons at the State map margins and these were mostly dissolved into the background polygon.

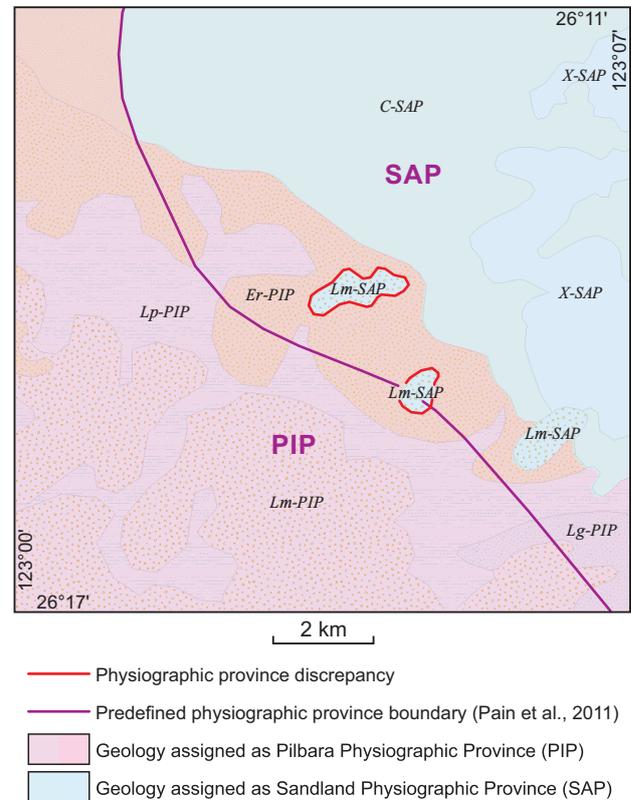


**Figure 13. Manual edge matching after GeoScaler generalization; example along 1:100 000 YANGANOO and MEREWETHER map sheet margins: a) polygon misalignment, polygon terminating along the map sheet margin, code mismatch; b) line work from (a) checked against Landsat imagery and manually corrected; c) resulting map after edge matching**

Other polygon units were revisited and visually compared against the Landsat data, orthophoto images or geology maps, and were manually redrawn into larger polygons.

### Physiographic province discrepancy

The physiographic provinces as defined by Pain et al. (2011) are drawn at the scale of the Australian continent. The introduction of physiographic province boundaries to the State regolith at the more detailed 1:500 000 scale therefore resulted in polygons that straddle province boundaries. These polygons were either split along the lines of the physiographic province boundaries or were simply recoded within the physiographic province that occupies the largest section of the polygon. However, smaller polygons within these larger polygons may have been assigned to a different physiographic province, resulting in an apparently conflicting geological interpretation (Fig. 14). It is estimated that approximately 500 polygons that lie near physiographic province boundaries are assigned to a different province from that of the surrounding polygon. This represents only 0.3% of the 213 000 polygons within the State regolith layer. Although some of these allocations may be legitimate geologically, these instances have not been systematically verified. This does not affect the reliability of the regolith landform or composition identification. These discrepancies will be progressively addressed in future updates.



SJK12 23.07.20

**Figure 14. An example of regolith polygons containing physiographic province discrepancies (outlined in red). The Lm-SAP polygons are surrounded by Er-PIP, but remain unverified**

## Performance and final generalization

Towards project completion, geology polygons were dissolved statewide to remove obsolete boundaries where neighbouring polygons contained the same geology attribute. However, for the purpose of rendering performance, large polygons that exceeded 60 000 vertices were split along the boundaries of 1:250 000 and 1:25 000 tiles, creating long vertical and horizontal boundaries in the map (Fig. 15). This does not affect the interpretation of regolith geology.

To assist with rendering performance, a statewide generalization process was applied, using the same generalization method as the data source preparation stage. A generalization tolerance of 0.00004 decimal degrees was used. This resulted in a 10% reduction of vertices, from a total of 17.5 million vertices down to 15.8 million vertices while remaining suitable for scale.

## Summary and future development

The new 1:500 000 State regolith map of Western Australia has synthesized existing geological information using a revised regolith classification scheme. The map was compiled using GeoScaler, an automated process for polygon generalization. Figure 16 shows a flow chart summarizing these processes. The digital package can be downloaded from the DMIRS [Data and Software Centre](#) in the following formats: ESRI shapefile, MapInfo Tab files and QGIS layer file. It is also available in [GeoVIEW.WA](#). Figure 17 is a snip from the [GeoMap.WA](#) Feature info and More details windows showing an example of the information captured and displayed in this dataset. The nomenclature and hierarchy for the regolith units are based on weekly updates from ENS.

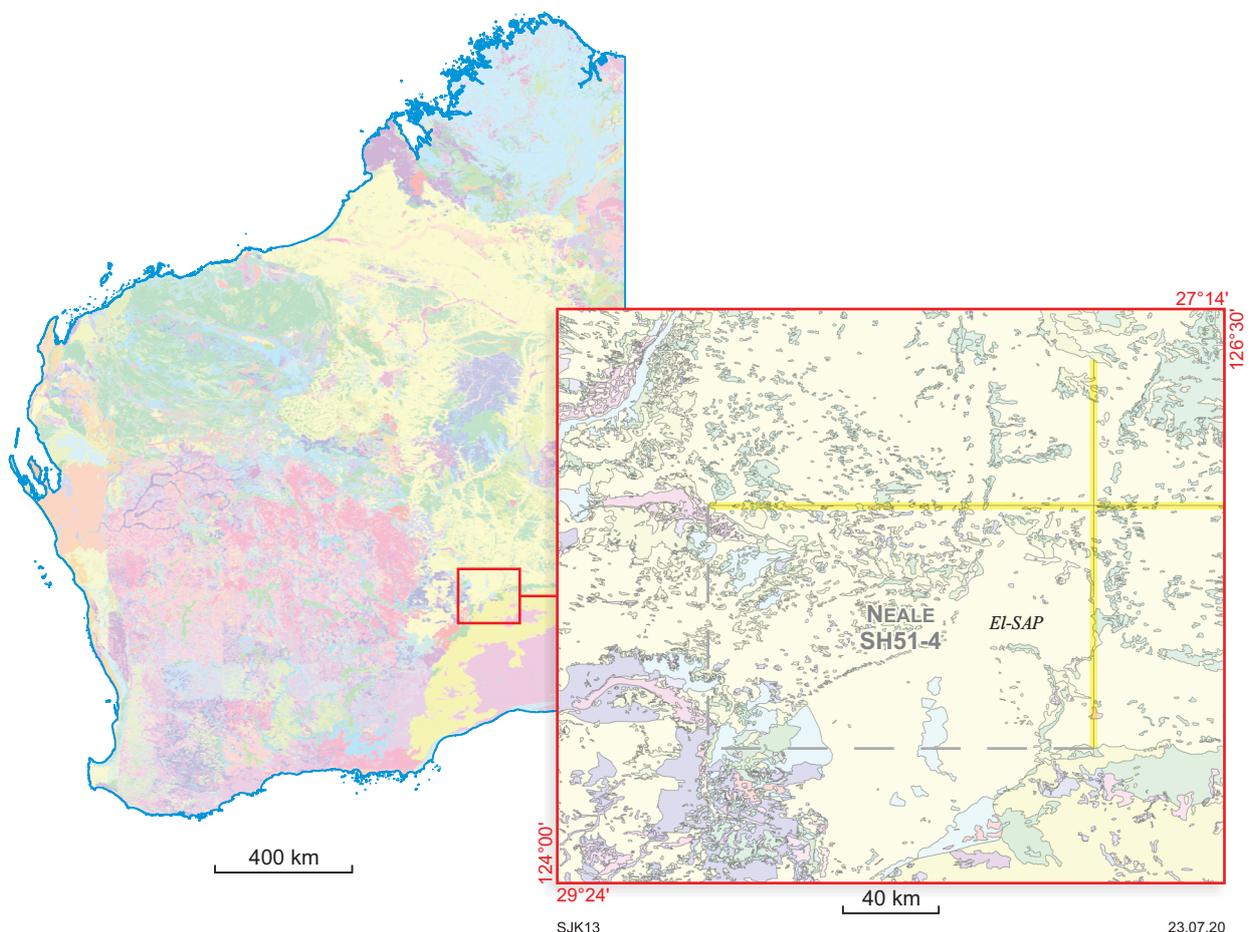


Figure 15. Example showing a large *\_EI-SAP* polygon split along the boundary between 1:250 000 map sheets creating long straight lines. Polygons of such size have been split for performance purposes, so the layer draws better in software platforms such as ArcMap or GeoVIEW.WA. This does not affect interpretation of the regolith geology

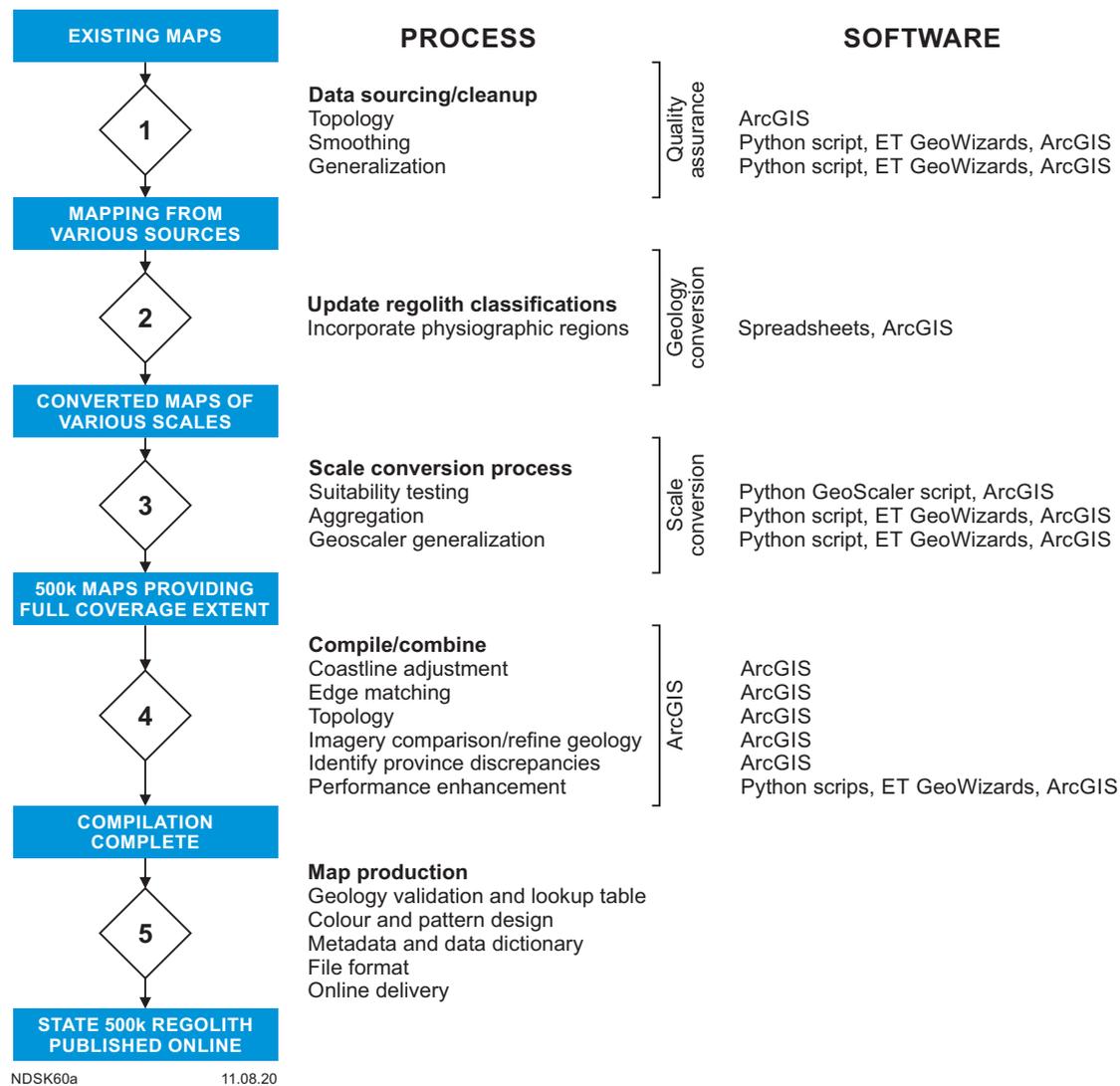


Figure 16. Flow chart summarizing the processes used to create the Western Australia State digital regolith map layer, compilation of which relied on GeoScaler, an automated algorithm for polygon generalization

This is a new map product, not a new edition of the previous 1:500 000 State regolith map (Marnham and Morris, 2003). Due to the complexity of Western Australia’s regolith material and landforms, the 1:500 000-scale State regolith map is designed to provide a broad regional overview and should be used as a guide only when deriving conclusions about conditions at specific locations. Comprehensive regolith mapping requires detailed field and laboratory work, and subsequent data analysis. That detailed work has not been undertaken for compilation of this map.

GSWA intends to continue regolith mapping as an integral part of mapping under cover by understanding the regolith in the remote regions of Western Australia. The 1:500 000-scale State regolith map is an evolving digital layer that will be refined with time as new data and interpretations become available. Compilation of the

1:500 000 State regolith map of Western Australia was funded by the Exploration Incentive Scheme (EIS).

## Acknowledgements

Gabriel Hout-Vézina from Geological Survey of Canada is thanked for the discussions regarding the GeoScaler tools.

## Appendix

A Microsoft Excel table with old codes, narratives and their corresponding maps, and their equivalent new GSWA regolith–landform codes, can be downloaded from DMIRS eBookshop as an accompanying spreadsheet.

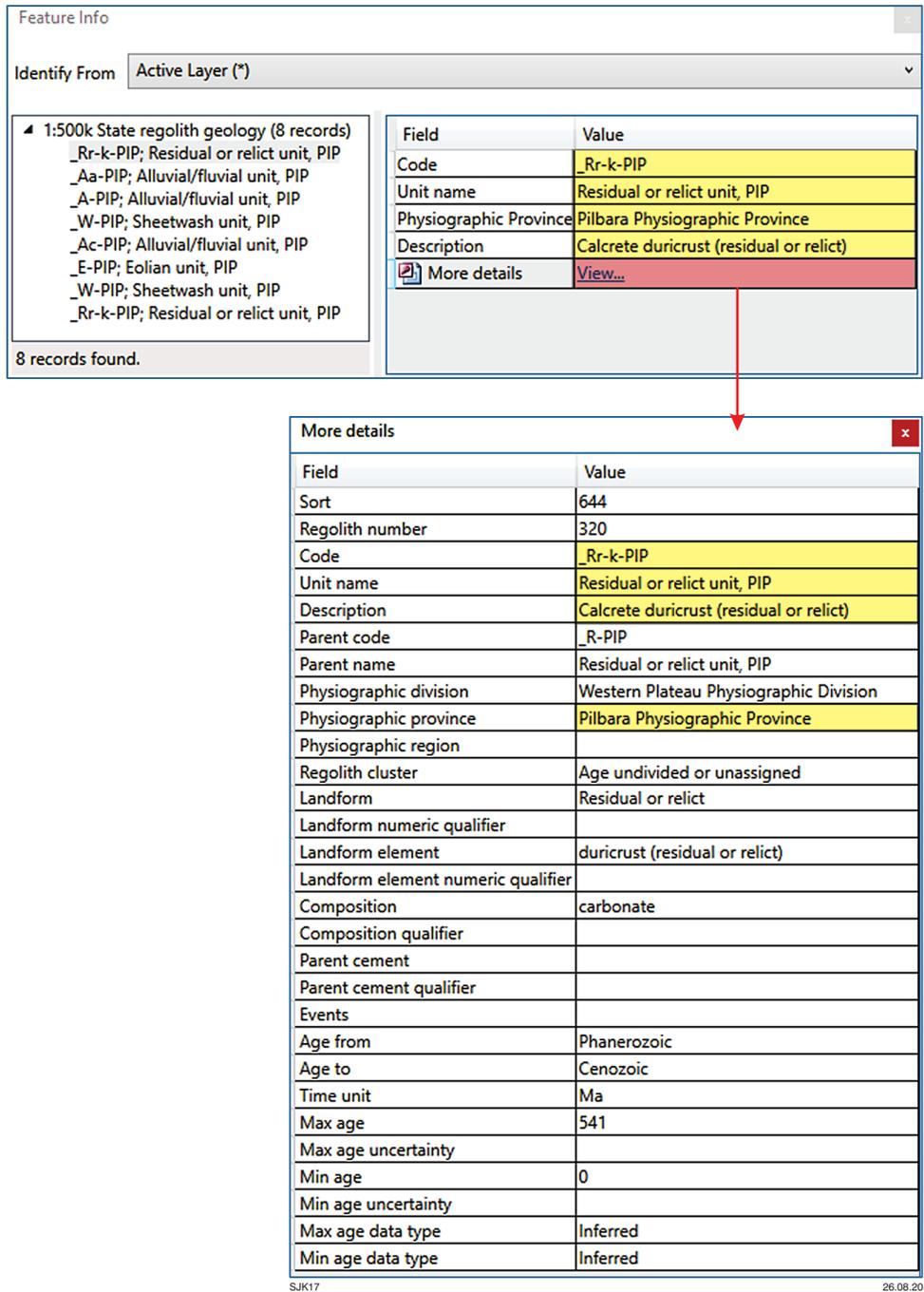


Figure 17. An example of attributes captured for and linked to the Western Australia State regolith digital data layer. The Feature info and More details windows shown here are displayed when the layer is queried in GeoMap.WA. Similar information is returned when the layer is queried in GeoVIEW.WA or ESRI ArcMap

## References

- Anand, RR, Churchward, HM, Smith, RE, Smith, K, Gozzard, JR, Craig, MA and Munday, TJ 1993, Classification and atlas of regolith–landform mapping units: exploration perspectives for the Yilgarn Craton, Australia: CSIRO, Division of Exploration and Mining, Restricted Report 440R, <<https://researchdata.ands.org.au/classification-atlas-regolith-yilgarn-craton/448031>>.
- Butt, CRM, Robertson, IDM, Scott, KM and Cornelius, M (editors) 2005, Regolith expression of Australian ore systems: Cooperative Research Centre for Landscape Environments and Mineral Exploration (CRC LEME), Perth, Western Australia, 431p.
- Department of Mines, Industry Regulation and Safety 2019, Industry activity indicators: Department of Mines, Industry Regulation and Safety, accessed 19 February 2019, <[https://www.dmp.wa.gov.au/Documents/About-Us-Careers/Stats\\_Digest\\_2018-19.pdf](https://www.dmp.wa.gov.au/Documents/About-Us-Careers/Stats_Digest_2018-19.pdf)>.
- Douglas, DH and Peucker, TK 1973, Algorithms for the reduction of the number of points required to represent a line or its caricature: The Canadian Cartographer, v. 10, no. 2, p. 112–122.
- Eggleton, RA (editor) 2001, The regolith glossary: surficial geology, soils and landscapes: Cooperative Research Centre for Landscape Environments and Mineral Exploration (CRC LEME), Canberra, Australian Capital Territory, 144p.
- Geological Survey of Western Australia 2013, Revised classification system for regolith in Western Australia, and the recommended approach to regolith mapping: Geological Survey of Western Australia, Record 2013/7, 26p.
- Geological Survey of Western Australia 2016, 1:500 000 State interpreted bedrock geology of Western Australia, 2016: Geological Survey of Western Australia, digital data layer, <[www.dmirs.wa.gov.au/geoview](http://www.dmirs.wa.gov.au/geoview)>.
- Geological Survey of Western Australia 2020, 1:500 000 State regolith geology of Western Australia: Geological Survey of Western Australia, digital data layer, <[www.dmirs.wa.gov.au/geoview](http://www.dmirs.wa.gov.au/geoview)>.
- Huot-Vézina, G, Boivin, R, Smirnoff, A and Paradis, SJ 2012, GeoScaler: Generalization tool (2nd edition): Geological Survey of Canada, Open File 6231, 82p., doi:10.4095/291993.
- Krapf, CBE, Irvine, JA and Cowley, WC 2012, Compilation of the 1:2 000 000 State Regolith Map of South Australia — a summary: Department for Manufacturing, Innovation, Trade, Resources and Energy, Adelaide, South Australia, Report Book, 17p.
- Marnham, JR and Morris, PA 2003, A seamless digital regolith map of Western Australia: a potential resource for mineral exploration and environmental management, in Geological Survey of Western Australia Annual Review 2002–03, p. 27–33.
- Pain, C, Chan, R, Craig, MA, Gibson, D, Kilgour, P and Wilford, J 2007, RTMAP regolith database field book and user's guide: Cooperative Research Centre for Landscape Environments and Mineral Exploration (CRC LEME), Open File Report 231.
- Pain, C, Gregory, L, Wilson, P and McKenzie, N 2011, The physiographic regions of Australia — Explanatory notes: Australian Collaborative Land Evaluation Program and National Committee on Soil and Terrain, Canberra, Australian Capital Territory, 36p.
- Pain, CF and Ollier, CD 1992, Ferricrete in Cape York Peninsula, north Queensland: BMR Journal of Australian Geology and Geophysics, v. 13, no. 3, p. 207–212.
- Smirnoff, A, Huot-Vézina, G, Paradis, SJ and Boivin, R 2012, Generalizing geological maps with the GeoScaler software: the case study approach: Computers & Geosciences, v. 40, p. 66–86, doi:10.106/j.cageo.2011.07.013.
- Smirnoff, A, Paradis, SJ and Boivin, R 2008, Generalizing surficial geological maps for scale change: ArcGIS tools vs. cellular automata model: Computers & Geosciences, v. 34, p. 1550–1568, doi:10.1016/j.cageo.2007.10.013.
- Smith, JW and Gemuts, I (compilers) 1967, Gordon Downs, WA Sheet SE52-10: Bureau of Mineral Resources (Geoscience Australia) and Geological Survey of Western Australia, 1:250 000 Geological Series.

This Record is published in digital format (PDF) and is available as a free download from the DMIRS website at <[www.dmirs.wa.gov.au/GSWApublications](http://www.dmirs.wa.gov.au/GSWApublications)>.

Further details of geoscience products are available from:

Information Centre  
Department of Mines, Industry Regulation and Safety  
100 Plain Street  
EAST PERTH WESTERN AUSTRALIA 6004  
Phone: +61 8 9222 3459 Email: [publications@dmirs.wa.gov.au](mailto:publications@dmirs.wa.gov.au)  
[www.dmirs.wa.gov.au/GSWApublications](http://www.dmirs.wa.gov.au/GSWApublications)

