Water Science & Technology

© 2025 The Authors

Water Science & Technology Vol 91 No 10, 1141 doi: 10.2166/wst.2025.067

Quantifying urban land cover imperviousness as input for flood simulation using machine learning: South African case study

Ione Loots [a,*], Jeffrey Colin Smithers[b] and Thomas Rodding Kjeldsen[b,c]

[a] Department of Civil Engineering, University of Pretoria, Pretoria, South Africa
[b] Centre for Water Resources Research and School of Engineering, University of KwaZulu-Natal, Pietermaritzburg, South Africa
[c] Department of Architecture and Civil Engineering, University of Bath, Bath, UK
*Corresponding author. E-mail: ione.loots@up.ac.za

IL, 0000-0003-0715-6852

ABSTRACT

The imperviousness of urban surfaces is an important parameter in simulating urban hydrological responses, but quantifying imperviousness can be challenging and time-consuming. In response, this study presents a new framework to efficiently estimate the imperviousness of urban surfaces, using satellite images with Red, Green and Blue bands and a land cover dataset with multiple built-up urban classes through remote sensing, machine learning and field verification. The methodology is adaptable to other regions with similar datasets. For a case study in Pretoria, South Africa, major differences in median total impervious area percentages (mTIA%) were identified when compared between land cover groups: residential areas had a lower imperviousness median (mTIA% = 38%) than commercial (mTIA% = 81%) and industrial (mTIA% = 89%) land cover. The mTIA% also varies between 17 and 61% for a range of different formally developed residential classes and between 14 and 43% for a range of different informally developed residential classes. These mTIA% are recommended for any urban area within the South African National Land Cover dataset. These values can be incorporated into hydraulic and hydrological models, which improve the efficiency of parameter estimation for modelling. The methodology successfully quantified temporal imperviousness changes in the study area.

Key words: impervious, informal settlement, land cover, QGIS, remote sensing, urban flood

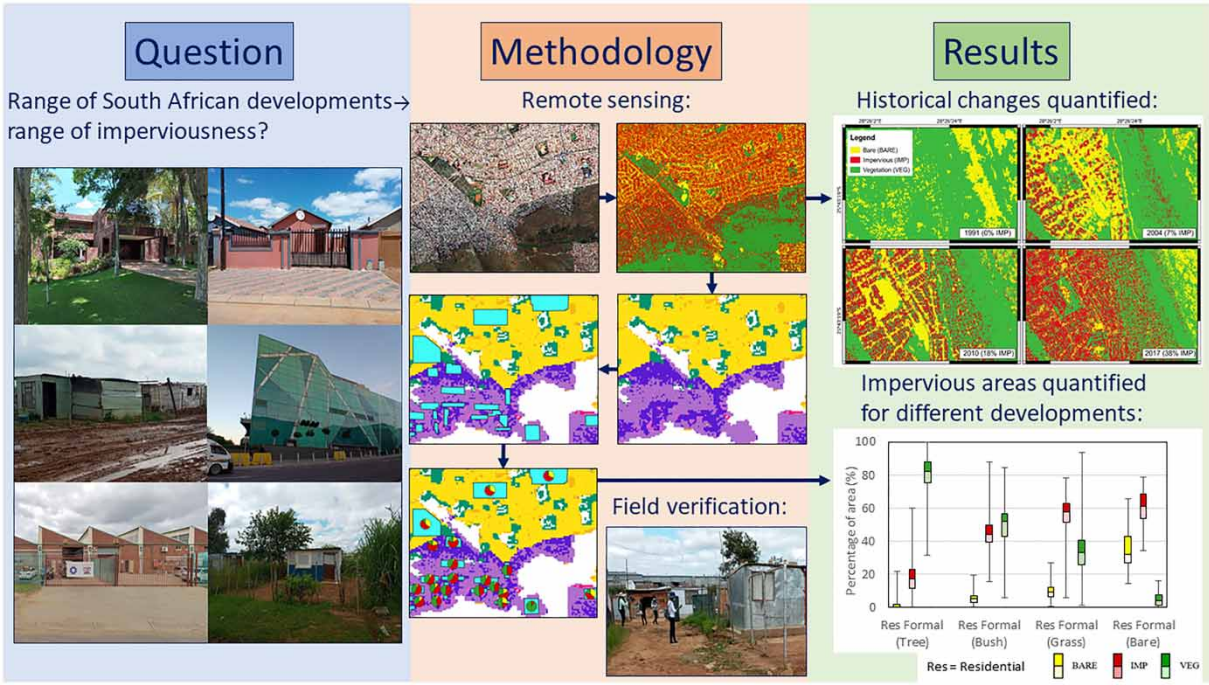
HIGHLIGHTS

- A new framework quantifies urban land cover imperviousness using satellite images with visible bands.
- Impervious areas differ significantly among built-up urban classifications, as well as between formal and informal developments.
- The methodology effectively tracks temporal changes in imperviousness.
- Findings reduce the time and effort for parameter estimation required for hydrological and hydraulic modelling.

This is an Open Access article distributed under the terms of the Creative Commons Attribution Licence (CC BY-NC 4.0), which permits copying, adaptation and redistribution for non-commercial purposes, provided the original work is properly cited (http://creativecommons.org/licenses/by-nc/4.0/).

Downloaded from http://iwaponline.com/wst/article-pdf/91/10/1141/1566930/wst2025067.pdf
by UNIVERSITY OF PRETORIA user

GRAPHICAL ABSTRACT



INTRODUCTION

Hydraulic and hydrological models are often used to simulate, predict and forecast urban flooding (Guo *et al.* 2021; Mignot & Dewals 2022). These models are also used to design urban drainage systems for design events with selected durations and return periods (El Hattab *et al.* 2020; Ambrosio *et al.* 2022; Ferrans *et al.* 2022).

Urbanisation is often linked with more frequent and severe flooding (Jacobson 2011; Braud *et al.* 2013; Sakijege 2013; Ferreira *et al.* 2016; Saurav *et al.* 2021). It is reported that since 2005, more than half of the world's population lives in urban areas, with the proportion of urban dwellers set to rise to about 60% by 2030 (UNDP 2019). The developing continents of Africa and Asia are experiencing the fastest urbanisation rates (Cohen 2006; UNDP 2014; Zhang *et al.* 2015). Urban areas contain many different types of buildings and roads, resulting in large impervious areas. Total impervious area (TIA in this study) affects runoff generation and conveyance (Nguyen *et al.* 2022). Therefore, many widely used models, such as the Hydrologic Engineering Center's River Analysis System (HEC-RAS) and the United States Environmental Protection Agency's Storm Water Management Model (SWMM), require TIA as an input parameter when considering urban areas.

However, it is time-consuming for practitioners to quantify the TIA of large catchment areas. Time and resource constraints are prevalent worldwide, but especially problematic in developing countries like South Africa. Therefore, practitioners tend to estimate TIA visually using satellite photos or 1:50,000 drawn topographical maps (Topo maps) derived from aerial photos. However, this *ad hoc* method can lead to inconsistent results. For example, a study conducted in South Africa by Loots *et al.* (2023) used Topo maps to define the total extent of urban development and found varying runoff responses from catchments with similar urban development levels. The study concluded that the density of impervious areas in different development types could contribute to the inconsistent runoff responses. In many developing countries, the imperviousness and layout characteristics of formally planned, structured and developed suburbs (formal development) differ significantly from those in informal areas, where settlement occurs in an unplanned and disorganised manner (Gogate & Rawal 2015).

Most South African cities have experienced significant population increases in recent years (Mlambo 2018; Mubangizi 2021; StatsSA 2022). People migrate to urban areas hoping for improved livelihoods (Cohen 2006; Geyer *et al.* 2012; Capps *et al.* 2016; UNDP 2019). Geyer *et al.* (2012) found continued growth of informal townships on city outskirts. They also noted a trend in the post-Apartheid era where people migrated away from urban centres towards urban fringes. This combination has led to significant development of both formal and informal settlements on city outskirts. Thompson (2019)

Downloaded from http://iwaponline.com/wst/article-pdf/91/10/1141/1566930/wst2025067.pdf
by UNIVERSITY OF PRETORIA user

reported a 30% increase in the South African urban footprint between 1990 and 2018. In 2022, an estimated 88.8% of the population lived in formal settlements, 8.1% in informal dwellings and 3.1% in traditional dwellings. Gauteng is the most urbanised of South Africa's provinces, where 11% of households were informal dwellings in 2022 (StatsSA 2022). This rapid urban development makes it imperative to quantify the TIA of different development types in South Africa to improve the hydrological modelling of urban areas.

In order to simplify the estimation of impervious areas for hydraulic and hydrological modelling, the main aim of this study was to develop a new framework to quantify the imperviousness of urban land cover classes from satellite images using the Red, Green and Blue (RGB) bands. The first study objective was to develop a methodology that could be used to derive TIA in any region with a land cover dataset that differentiates between a range of urban land cover classes and with high-resolution satellite imagery containing RGB bands. The second objective was to use this methodology to derive representative estimates of TIA for different South African National Land Cover (SANLC) classes. The third objective was to apply the methodology to quantify temporal changes to TIA in selected study catchments.

METHODS

A combination of remote sensing analysis and ground-truthing was applied to measure the typical TIA of different urban land cover classes and verify the results. The three main components of the methodology that could be used to derive TIA in any region with a land cover dataset that differentiates between a range of urban land cover classes and satellite imagery with RGB bands, as set out in Objective 1 of this study, are introduced below:

(i) The remote sensing measurement of imperviousness is summarised in Figure 1:

   (a) Satellite images with Red (R), Green (G) and Blue (B) bands were used (Figure 1(a)).
   (b) The Environmental Mapping and Analysis Program Box (EnMAP-Box) machine learning plugin available in the Quantum Geographic Information System (QGIS) (ENMAP-Box_Developers 2019) was applied to classify land cover as bare (BARE), impervious (IMP) or vegetation (VEG) (Figure 1(b)).
   (c) A land cover dataset that differentiates between different urban built-up land cover classes was overlaid (Figure 1(c)).
   (d) Sample zones were chosen for each of the studied land cover classes. Sample zones were manually selected from areas with homogeneous land cover classification. Sample zones were selected based on land cover classes only, without considering other data sources, to avoid bias in the selection (Figure 1(d)).
   (e) A zonal histogram classification tool was used to quantify each sample zone's BARE, IMP and VEG area (Figure 1(e)).

(ii) The same methodology was applied using additional images in other areas of the city to verify the results.

(iii) Further verification was undertaken through physical measurement of TIA at independent sample sites in the study area, and the results were compared with the remote sensing results.

This approach is described in the following paragraphs and then applied to case study areas in the city of Pretoria, South Africa. The methodology is first used to derive representative estimates of TIA for different SANLC classes (as set out in Objective 2). The machine learning land cover classification tool is then used to quantify temporal changes to TIA in selected study catchments (as set out in Objective 3). The following sections describe the methodology followed in this study in further detail.

Study area

With over 2.9 million inhabitants (StatsSA 2022) and an area of over 17,000 km^2, the City of Pretoria, located in the Tshwane Metropolitan Municipality, South Africa, was chosen as a case study area. This city contains the full range of South African built-up urban development types. The SANLC classification (Geoterraimage 2019, 2021) was used to identify the different types of urban land cover in this study. The sample areas were chosen to cover all eight built-up urban residential area types as classified by the SANLC, as well as industrial and commercial areas. The distribution of classes was mapped across the study area, as shown in Figure 2, and includes both formal and informal urban areas. It should be noted that the land cover classes for urban recreational areas (including parks and sports fields), landfills, water bodies and wetlands were not included in this study, as these land covers would generally not contribute to the imperviousness of an urban area.

Downloaded from http://iwaponline.com/wst/article-pdf/91/10/1141/1566930/wst2025067.pdf
by UNIVERSITY OF PRETORIA user

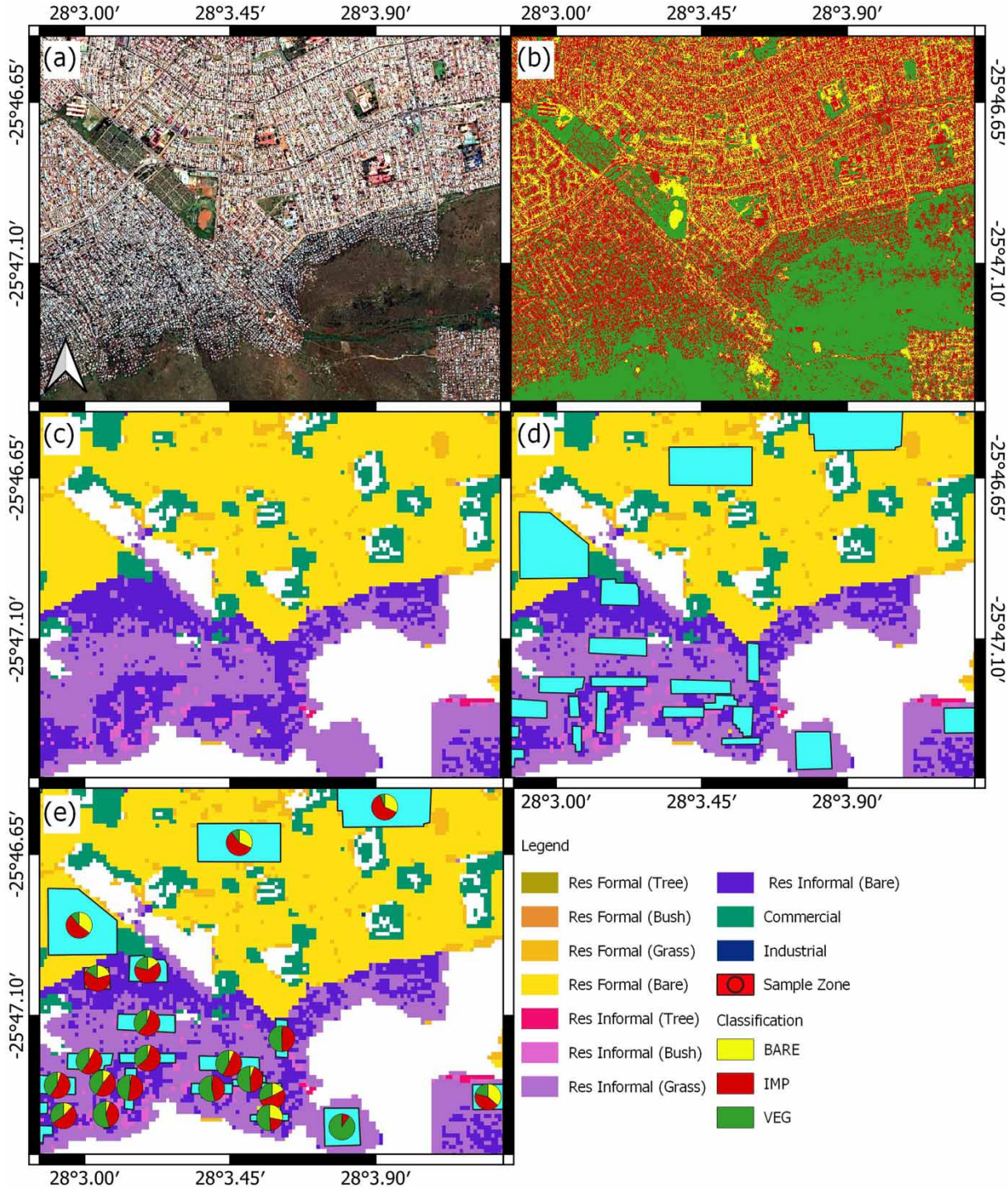


Figure 1 | Imperviousness measurement methodology illustrated using a subset of the sample data: (a) satellite image; (b) machine learning results of land cover classification; (c) land cover classes; (d) sample zones chosen from the land cover classes and (e) pie charts of the imperviousness results obtained using the zonal histogram classification tool.

It was previously established that urban development influences hydrological processes in South African catchments (Loots *et al.* 2023). It is envisaged that the results from the current study could be used in future studies to further analyse the developmental influences in the context of South African development types and corresponding levels of imperviousness. Therefore, urban catchment areas with available flow data were chosen to quantify historical changes in imperviousness. Five

Downloaded from http://iwaponline.com/wst/article-pdf/91/10/1141/1566930/wst2025067.pdf
by UNIVERSITY OF PRETORIA user

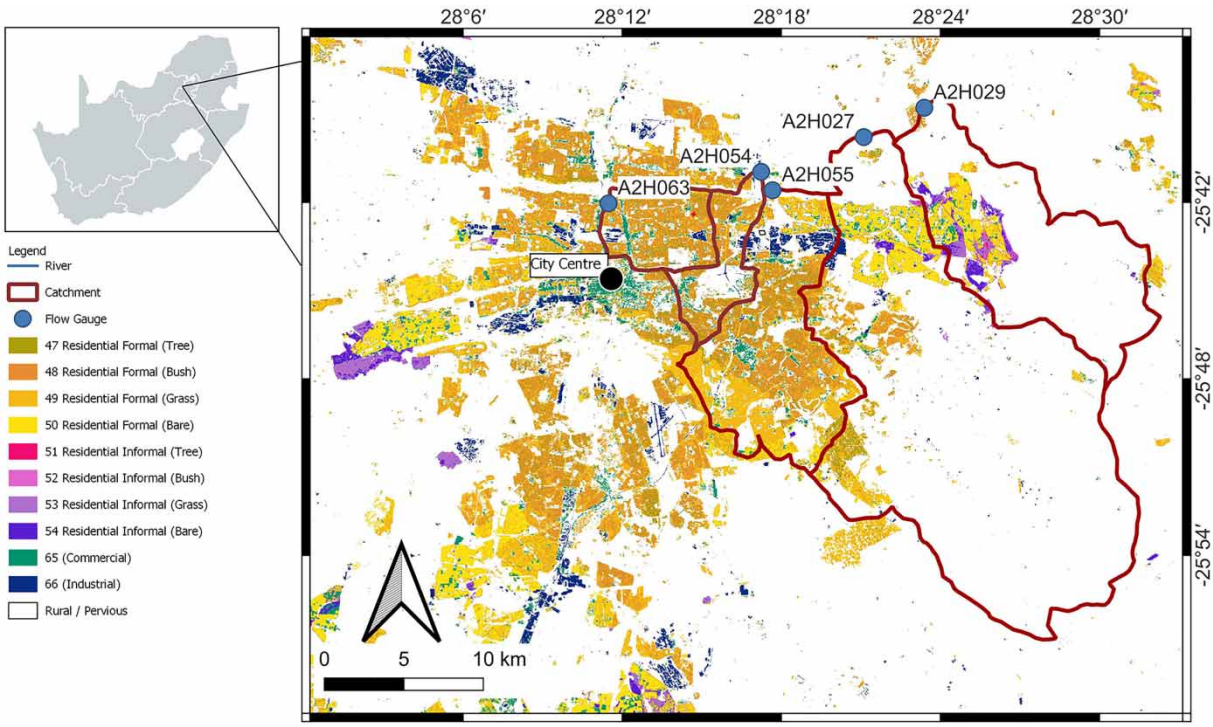


Figure 2 | SANLC classes in the study area, with the five study catchments used in this study.

catchments with at least 14% developed urban area as of 2020 were selected (Figure 2). As shown in Figure 2, the catchments include both formal and informal urban areas. The catchment sizes range from 30 (A2H063) to 357 km^2 (A2H027).

Figure 2 shows that Catchment A2H063 is the furthest west of the five catchments. It is also the closest to the historical urban centre of Pretoria. Moving progressively east are A2H054, A2H055, A2H027 and A2H029. While the three westerly catchments are almost entirely covered by formal development, the two most easterly catchments are less developed overall and contain informal settlements.

Formal residential areas in the study area typically contain houses (Figure 3(a) and 3(c)) or intensive residential developments with high boundary walls (Figure 3(b)). The individual buildings are typically surrounded by a combination of paved and vegetated areas connected to formalised drainage networks. In contrast, informal settlements are typically constructed more haphazardly, with little or no impervious area apart from the roofs of dwellings (Figure 3(d)–3(f)). These types of developments do not typically have formal drainage systems. Commercial (Figure 3(g)) and industrial (Figure 3(h)) areas mostly have high percentages of TIA and are connected to formalised drainage.

Imperviousness quantification using machine learning

The EnMAP-Box machine learning plugin available in the QGIS (ENMAP-Box_Developers 2019) was applied to classify the land cover images in this study. The impervious areas of the high-resolution (1.5 m resolution pixels) Spot 6-7 South African National Space Agency (SANSA) satellite images with RGB bands, dated 2017, were used for this analysis. Land cover sample points were manually selected from visual inspection of the satellite images to achieve an acceptable distribution among the six initially classified general land cover types, being: (1) buildings, (2) trees, (3) grass, (4) bare, (5) roads, and (6) water. These six land cover types were chosen with the primary aim of distinguishing impervious areas from pervious areas. Training and independent validation points were selected to cover different chromatic shades representing these classes in the study area and processed using the EnMAP-Box machine learning tool. Between 50 and 200 points for each general land cover type were selected from each of the four SANSA satellite images used in the study. The number of points depended on the availability of land cover types in the available images. To simplify output, land cover types were grouped into three categories: impervious area (IMP) from buildings and roads, bare area (BARE) from bare and water

Downloaded from http://iwaponline.com/wst/article-pdf/91/10/1141/1566930/wst2025067.pdf
by UNIVERSITY OF PRETORIA user

Figure 3 | Examples of urban land cover classes in South Africa.

Downloaded from http://iwaponline.com/wst/article-pdf/91/10/1141/1566930/wst2025067.pdf
by UNIVERSITY OF PRETORIA user

and vegetated area (VEG) from trees and grass. Figure 1(b) shows an example of the IMP, BARE and VEG categories identified for a subset of the data from the image shown in Figure 1(a) using the machine learning tool.

Since the primary purpose in this case was to quantify impervious area, water was not retained in a separate category for this study. In the SANLC dataset, water is classified in its own class. Hence, the water bodies incorporated in the urban built-up land cover classes investigated in this study are negligible. When this method is applied in regions with many water bodies in the urban environment, it is recommended that the water land cover type be kept in its own classification and quantified separately from the other classification groups.

After the images were classified, the TIA of different SANLC classes was quantified. The 2018 SANLC database, with 20 m resolution pixels, was used, as it is the land cover dataset with the best resolution for South Africa. The 2018 database was used as it was developed within a year of the satellite photos used for this study, as described in the previous paragraph. Sample zones were selected from each of the 10 SANLC urban built-up land use classes considered in this study (examples are shown in Figure 1(c) and 1(d) for a subset of the data). The 2018 SANLC data was used to select sample zones. Sample zones were made as large as possible, with some SANLC land use classes having larger contiguous areas than others. Sample zones were selected based on SANLC classes only, without considering other data sources, to avoid selection bias. The locations of the sample zones are shown in Figure 4. The total area sampled for each SANLC land use class, as well as the average area per sample zone, are summarised in Table 1.

The grouped classification derived using the machine learning tool was overlaid with the sample zones from the SANLC data in order to calculate the IMP, BARE and VEG area of each sample zone (Figure 1(e)), using the Zonal Histogram tool in QGIS. The median, 25th and 75th percentile of impervious area and maximum and minimum percentages of each of the three grouped classifications were calculated for each SANLC class included in this study.

Verification of the remote sensing imperviousness quantification

Verification zones were selected from each of the 10 SANLC urban built-up land use classes. Verification zones were selected randomly from areas in Pretoria, but in suburbs different from those where the sample zones were selected for imperviousness quantification. Verification zones were also selected considering only the land cover class. This was done to avoid bias in the results. The locations of the verification zones are shown in Figure 4.

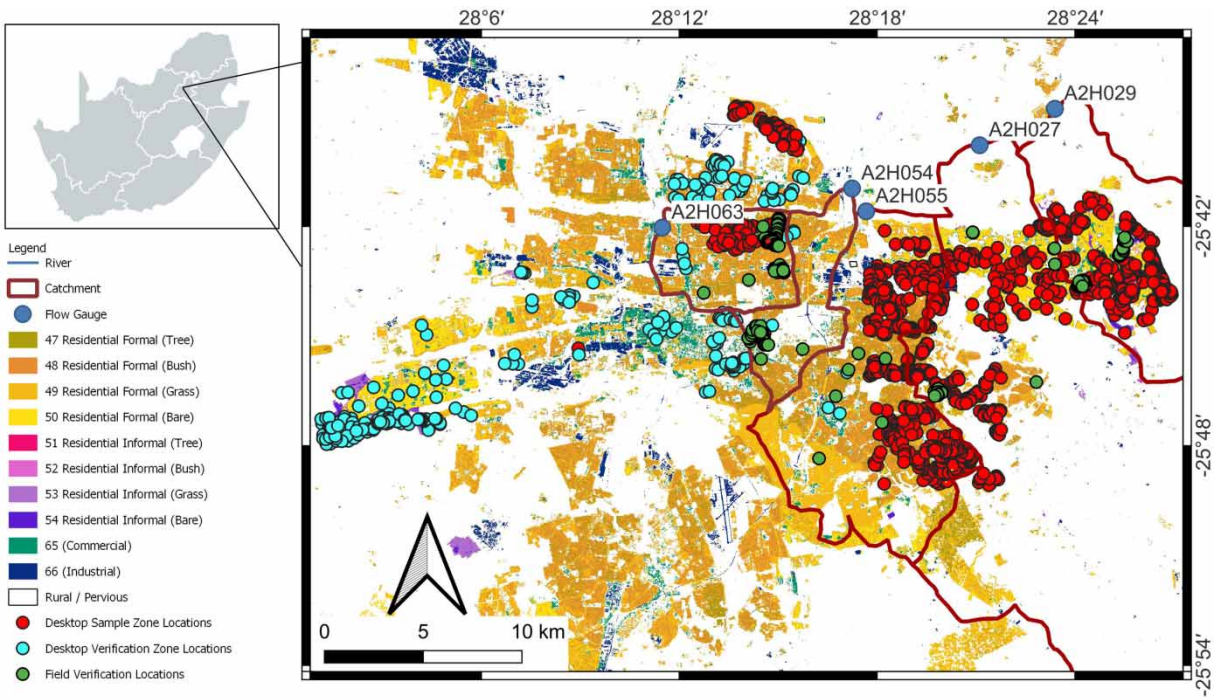


Figure 4 | The locations of sample zones and verification zones used for the remote sensing analysis, as well as locations of field verification sites used in this study.

Downloaded from http://iwaponline.com/wst/article-pdf/91/10/1141/1566930/wst2025067.pdf
by UNIVERSITY OF PRETORIA user

Table 1 | SANLC areas used for sampling and verification of TIA

SANLC class No.	SANLC class description	Remote sensing sample zones			Remote sensing verification zones			Field verification zones
		No. of zones	Total area (m^2)	Avg. area per zone (m^2)	No. of zones	Total area (m^2)	Avg. area per zone (m^2)	No. of parcels per class
47	Residential Formal (Tree)	415	1,151,204	2,774	45	190,248	4,228	87
48	Residential Formal (Bush)	125	560,219	4,482	20	124,673	6,234	84
49	Residential Formal (Grass)	116	790,991	6,819	41	309,280	7,543	67
50	Residential Formal (Bare)	35	1,818,962	51,970	12	1,036,182	86,349	59
51	Residential Informal (Tree)	131	314,109	2,398	6	9,786	1,631	54
52	Residential Informal (Bush)	178	486,299	2,732	51	48,167	945	30
53	Residential Informal (Grass)	40	534,094	13,352	19	853,911	44,943	9
54	Residential Informal (Bare)	37	831,868	30,810	29	401,939	13,860	38
65	Commercial	347	1,922,515	5,540	15	800,948	53,397	32
66	Industrial	80	1,954,666	24,433	24	832,905	34,704	17

The same methodology, as for sample zones, was applied for the imperviousness quantification of verification zones, as summarised in Figure 1(a)–1(e). The median values for the impervious area of the verification zones of each SANLC class were compared with the impervious area median values in the sample zones, and the differences were noted. The total verification areas for the different urban built-up land cover classes are summarised in Table 1.

Further verification through fieldwork

Further verification was achieved through field inspections. Research team members identified and visited independent sample sites for all SANLC classes under investigation. TIA was physically measured at these sites. Sample sites from different neighbourhoods with similar SANLC classes were included in the sampling. Field verification sample sites were chosen independently from remote sensing sample sites or verification sites. The locations of the field verification sites are shown in Figure 4.

The total area with impervious land cover for each sample site was physically measured on-site on properties where access was granted. All paved and roofed areas were included in the TIA calculation. Bare areas were not included in the TIA measurement or calculation. Schematic maps were drawn of the impervious areas at sample sites, and the length and width measurements were taken using a 30-m measuring tape. On properties where access was denied, notes regarding the location of vegetation covering impervious areas were taken from the street, and the TIA was manually delineated using high-resolution (1.8 m) Google Earth images. The total sampled area for the field verification study was 0.45 km^2 from 292 individual sample sites (or stands). A number of the sampled sites cover multiple SANLC classes. These were divided into parcels based on SANLC class coverage. In total, 477 sample parcels were included, with distribution per SANLC class, as shown in Table 1.

The TIA was quantified for each parcel and reported as a percentage of the total parcel area. The TIA percentages were grouped per SANLC class, and the median, 25th percentile and 75th percentile values were calculated. The results were compared with the remote sensing imperviousness median values.

Measurement of historical changes in imperviousness

In order to quantify changes in imperviousness of the study catchments, historical and recent estimates of the TIA were extracted using the EnMap-Box machine learning plugin in QGIS. Prior to 2002, the only available images in the study areas are greyscale aerial photos taken by the South African National Geo-spatial Information (NGI). Photos of the study area were taken in 1984, 1991 and 2001. They were digitised to a resolution of 2.5 m by 2.5 m pixels from the analogue data. Impervious built-up areas were identified from the digitised images through a combination of remote sensing and manual correction of misclassified areas.

Downloaded from http://iwaponline.com/wst/article-pdf/91/10/1141/1566930/wst2025067.pdf
by UNIVERSITY OF PRETORIA user

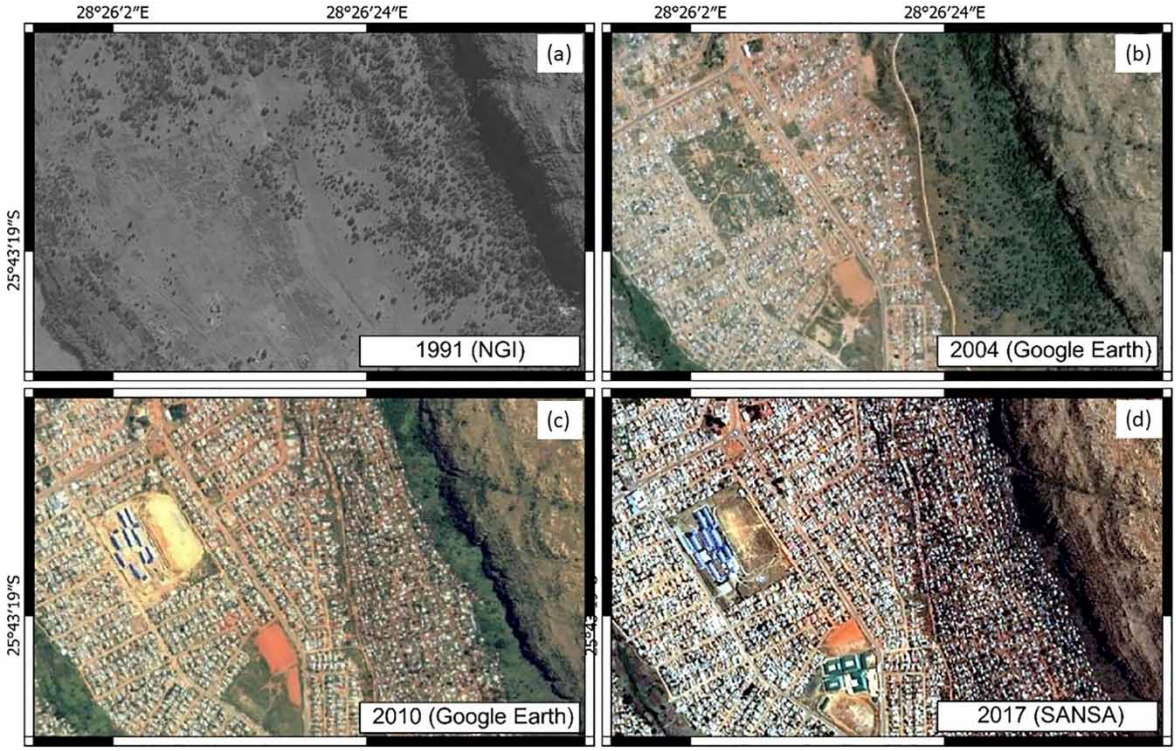


Figure 5 | Examples of imagery used for historical TIA determination. (a) An NGI orthophoto taken in 1991, digitised to a resolution of 2.5 m pixels; (b) a Google Earth satellite image taken in 2004, with 1.8 m pixels; (c) a Google Earth satellite image taken in 2010, with 1.8 m pixels and (d) a SANSA SPOT6-7 satellite image taken in 2017, with 1.5 m pixels.

Post-2002, satellite images from Google Earth were used to detect impervious areas. High-resolution images (1.8 m resolution pixels) were extracted from Google Earth from 2002 to 2021 at approximately 5-year intervals. Images taken on cloud-free days were used. Where possible, images from winter months, when less vegetation covers impervious areas, were used to get as good a representation of impervious area as possible. In some cases, images from different years were used for different catchments, as image selection was based on image quality. The 2017 SANSA SPOT6-7 satellite images were also used. Images for an area with both formal and informal development at the end of the study period are shown in Figure 5 as examples of the images used.

For this analysis, the EnMAP-Box machine learning plugin was applied in QGIS to identify TIA from the historical images. The TIA was extracted for each study catchment from the classified data, for multiple years. The TIA results are described as a percentage of each catchment's area ($TIA_{\%}$).

A comparison was undertaken between the 2001 NGI images and the 2002 Google Earth dataset. An average difference of 2.1% was found between the TIA measured using the two datasets for all study areas. The discrepancy was deemed sufficiently small to combine the results from the different datasets for this study.

RESULTS AND DISCUSSION

This section summarises results obtained in the analysis and verification of TIA in the 10 studied SANLC classes. The derivation of generalised TIA percentages ($TIA_{\%}$) for different urban SANLC classes is also described.

Remote sensing: sample zones

As described above, the TIA was extracted for different 2018 SANLC classes from the 2017 SANSA images. Median values, 25th and 75th percentiles and minimum and maximum values of IMP, VEG and BARE areas are shown using

Downloaded from http://iwaponline.com/wst/article-pdf/91/10/1141/1566930/wst2025067.pdf
by UNIVERSITY OF PRETORIA user

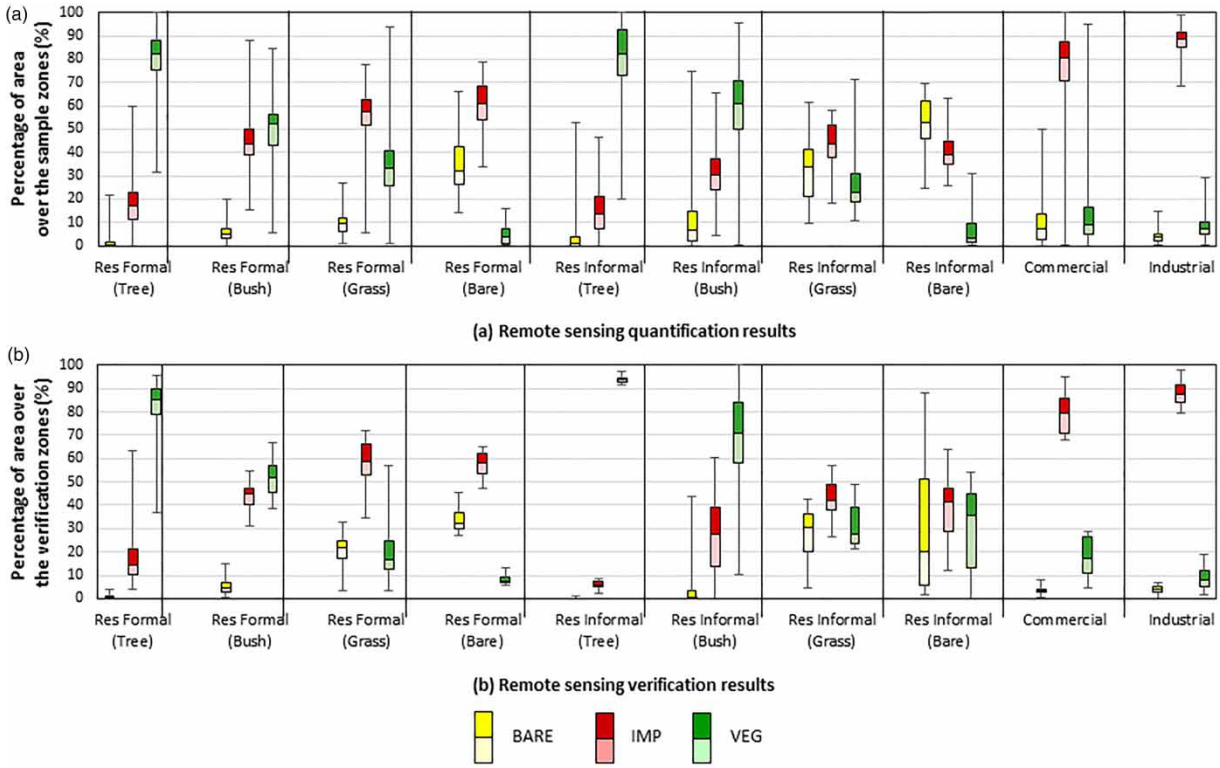


Figure 6 | Results of the (a) remote sensing quantification and (b) remote sensing verification of $TIA_{\%}$.

box-and-whiskers plots in Figure 6(a) for Formal Residential (SANLC 47, 48, 49, 50), Informal Residential areas (SANLC 51, 52, 53, 54), Commercial (SANLC 65) and Industrial (SANLC 66) areas.

It is evident from Figure 6(a) that areas classified as Residential Formal (Trees) have higher percentages of vegetation than other classes. Vegetation percentages decrease progressively from Residential Formal (Tree) to Residential Formal (Bush) to Residential Formal (Grass) to Residential Formal (Bare). The same trend is evident for informal residential classes. This result was expected and confirmed that the methodology works as expected. The impervious area median for Residential Informal (Bare) was lower than expected when compared with other informal residential classes. Closer inspection of the satellite images showed that informal areas classified as Residential Informal (Bare) had high proportions of bare surfaces between dwellings, with small, impervious dwellings. As expected, commercial and industrial areas have high $TIA_{\%}$.

Remote sensing: verification zones

The $TIA_{\%}$ results from the verification zones (Figure 6(b)) compare well with the $TIA_{\%}$ as calculated in the sample zones, with differences between the median IMP values of the sample zones and verification zones of between 0.9 and 3.0% for the four Residential Formal areas (denoted by the red IMP plots in Figure 6), 1.1% for Commercial areas and 1.5% for Industrial areas. The sampled area for Residential Informal (Tree) was the smallest of all 10 areas due to limited land cover with this classification in the study area. The area that was available for verification of TIA in Residential Informal (Tree) was significantly smaller than the sampled area. Less than 1 ha of this land cover class is located in the rest of Pretoria. Therefore, verification was not done for this land cover class. The differences between the median values of the sample zones and verification zones of the verification of impervious areas in the other three Residential Informal SANLC classes were between 1.5 and 3.0%.

It should be noted that seasonal variations between VEG and BARE areas are possible. However, the focus of this research was on quantifying IMP areas. Where possible, images from winter months, less vegetation covers impervious areas, were used to get as good a representation of impervious area as possible. If this methodology is applied in other studies where water bodies are also identified, seasonal variation of water body extent should also be considered.

Downloaded from http://iwaponline.com/wst/article-pdf/91/10/1141/1566930/wst2025067.pdf
by UNIVERSITY OF PRETORIA user

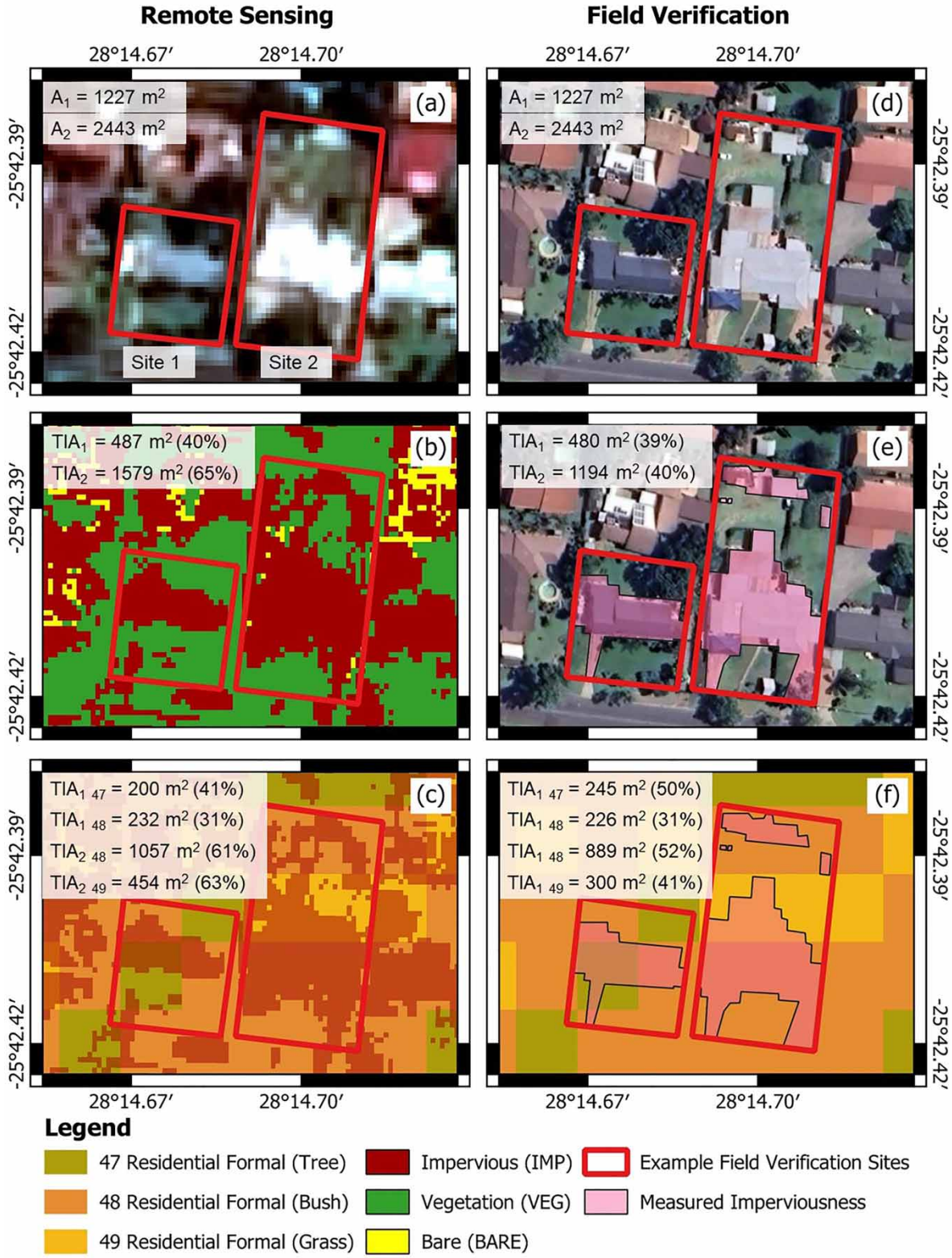


Figure 7 | Comparative example of remote sensing versus field verification results for two sites, with (a) the SPOT6 satellite image used for the remote sensing classification; (b) the remote sensing classification of impervious area (TIA) and (c) showing TIA per site and land cover class for the remote sensing example. The field verification information is shown in (d) the layout of the two example sites; (e) the physically measured TIA and (f) the TIA per site and per land cover class for the field verification example.

Downloaded from http://iwaponline.com/wst/article-pdf/91/10/1141/1566930/wst2025067.pdf
by UNIVERSITY OF PRETORIA user

Field verification

Representative sample sites for all SANLC classes under investigation were identified, and TIA was measured in the field for further verification of the remote sensing results. The field verification measurements and calculations for two example sites are shown in Figure 7. The field verification results of these two sites are compared with the remote sensing of these two sites (although these sites were not specifically included in the remote sensing dataset, the comparison is shown here to illustrate the similarities and differences typically found between remote sensing and field verification results). In this example, the TIA percentages measured remotely and on-site compare well in certain cases (SANLC Class 48 for Site 1), but the TIA percentages for SANLC Class 49 for Site 2 differ by over 50%.

The methodology followed in this study did not specify a direct comparison between remotely sensed and field-verified sample sites. However, based on the results of this example, it is recommended that future studies should consider direct comparison. It is postulated that the discrepancies will average out in larger datasets, but this needs to be confirmed in a future study.

The field verification sites' box-and-whisker plots of $TIA_{\%}$, as measured on the independent sample stands, are shown in Figure 8. The results are compared with the imperviousness results obtained in the sample zones used in the machine learning analysis (as reported in Figure 6(a)).

As shown in Figure 8, the median values of all verification results in residential areas are within 10% of the median of measured results. Therefore, the median results from the field verification compared reasonably well with the $TIA_{\%}$ quantified using the machine learning methodology, with field verification values showing wider ranges of values between the 25th and 75th percentiles of TIA. The broader range of values in the field verification dataset is attributed to the fact that the field verification sites were chosen based on accessibility. This meant that, in contrast to the large homogeneous areas chosen as sample sites for the remote sensing procedure, one site generally contained multiple land cover classes, with some land cover class areas only being a few square metres (an example of this is shown in Figure 7). This heterogeneity, combined with the fact that the land cover layer's pixel resolution is 20 m, resulted in more variability of TIA per land cover class.

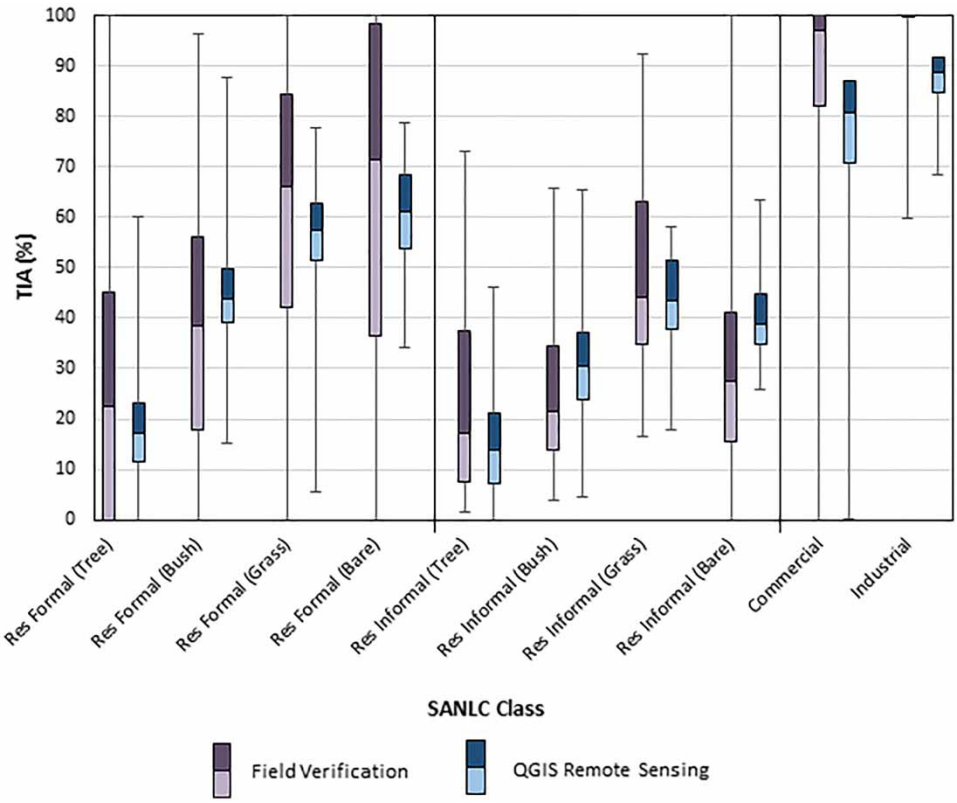


Figure 8 | Field measurement results for TIA in urban built-up SANLC classes.

Downloaded from http://iwaponline.com/wst/article-pdf/91/10/1141/1566930/wst2025067.pdf
by UNIVERSITY OF PRETORIA user

TIA percentages for built-up urban SANLC classes

The results were, therefore, reasonable, and median values were chosen as recommended estimates of $TIA_{\%}$, with the 25th and 75th percentile values of the median (M) suggested as lower (L) and upper (U) confidence intervals. The recommended $TIA_{\%}$ values for urban SANLC classes are summarised in Table 2. It is emphasised that the results from this study are limited to one geographic area due to the limited availability of observed data in urban areas in South Africa. However, it is postulated that the results apply other regions, given the relationships developed with national land cover classes in this study. Thus, until additional data are available and further research is undertaken to improve the results, the $TIA_{\%}$ values in Table 2 are proposed for the hydrological modelling of South African urban areas.

The methodology applied in this study can also be applied in any area with high-resolution (1.5 m resolution or better) satellite imagery with visible RGB bands and a land cover dataset with multiple classes of urban built-up areas. Site-specific values are recommended for hydraulic and hydrological modelling where possible.

Table 2 | Proposed TIA percentages for urban SANLC classes

SANLC class No.	SANLC description	TIA (%) L	M	U
47	Residential Formal (Tree)	11	17	23
48	Residential Formal (Bush)	39	44	50
49	Residential Formal (Grass)	51	58	63
50	Residential Formal (Bare)	54	61	68
51	Residential Informal (Tree)	7	14	21
52	Residential Informal (Bush)	24	30	37
53	Residential Informal (Grass)	38	43	51
54	Residential Informal (Bare)	35	40	45
65	Commercial	71	81	87
66	Industrial	85	89	92

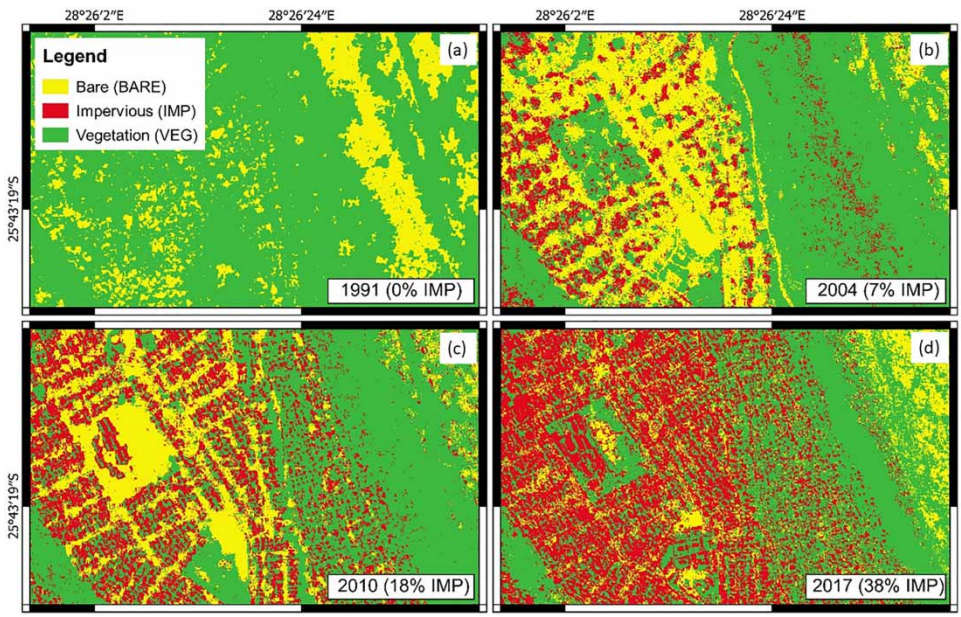


Figure 9 | Classification examples for different years from the images shown in Figure 5, with (a) having 0% imperviousness in 1991, (b) having 7% imperviousness in 2004, (c) having 18% imperviousness in 2010 and (d) having 38% imperviousness in 2017.

Downloaded from http://iwaponline.com/wst/article-pdf/91/10/1141/1566930/wst2025067.pdf
by UNIVERSITY OF PRETORIA user

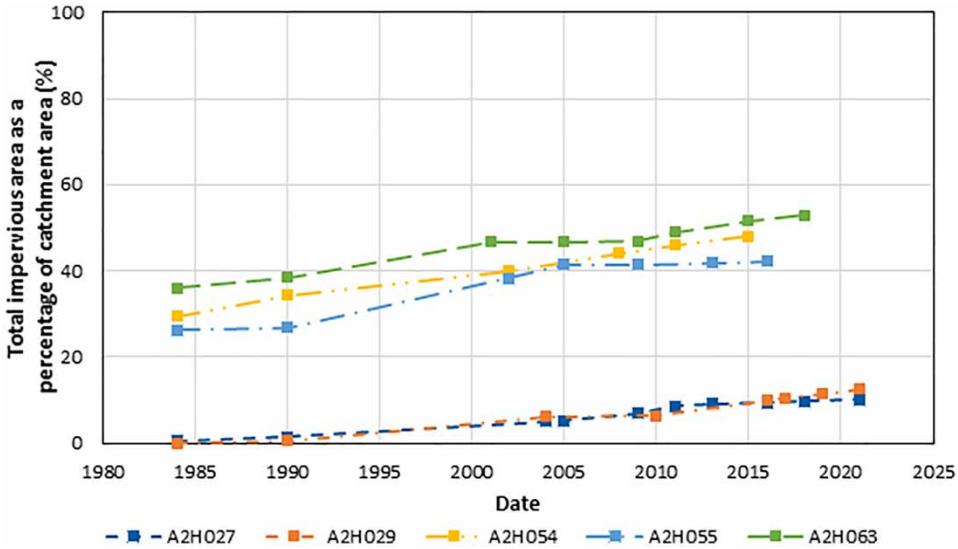


Figure 10 | Total impervious areas as a percentage of catchment area in the five study catchments from 1984 to 2021.

Temporal changes in imperviousness

Figure 9 shows examples of the impervious cover identified using the EnMAP-Box machine learning tool in QGIS on images taken in four different years. The figure shows that the bare and vegetated areas vary. This variation is attributed to temporal vegetative cover changes.

The temporal changes in imperviousness are shown in Figure 10 for all five study catchments. Catchment A2H063 shows the highest TIA% over the study period, increasing from 36.0 to 55.2%. Catchment A2H054 (29.5–50.0%) and Catchment A2H055 (26.2–43.1%) show similar TIA% trends to Catchment A2H063. Catchments A2H027 and A2H029 have the lowest total TIA%, but with significant change after 1994 (2.3–10.1% and 1.9–12.5%, respectively).

These results were expected, based on previous reports that urbanisation is still increasing in developing countries (Cohen 2006; Geyer *et al.* 2012; Capps *et al.* 2016; UNDP 2019), with the continued growth of informal townships on city outskirts (Geyer *et al.* 2012) and a trend in the post-Apartheid era (post-1994) where people migrated away from urban centres, towards urban fringes (Geyer *et al.* 2012).

All five study catchments have observed flow data dating back to 1985 or earlier. Therefore, runoff from these catchments should be compared with impervious areas, to determine if trends exist and quantify the effect of urban built-up areas in developing countries on runoff.

CONCLUSIONS

Previous studies have found that imperviousness significantly influences runoff characteristics. However, determining imperviousness at a catchment scale can be time-consuming and expensive. In order to simplify the estimation of impervious areas for hydraulic and hydrological modelling, this study aimed to develop a methodology that would provide consistent estimates of the imperviousness of built-up urban land cover classes by applying a machine learning tool.

The first objective was achieved by developing the methodology for this study. The methodology applied a combination of remote sensing analysis using the EnMAP-Box QGIS plugin machine learning tool and field verification. This methodology can be applied to derive imperviousness in any region with a land cover dataset that differentiates between a range of urban land cover classes, and with high-resolution satellite images with RGB bands.

The second objective was to use this methodology to derive representative estimates of TIA for different SANLC classes. Major differences were found in median total impervious area percentages ($m_{TIA\%}$) between the different urban built-up land cover classes studied, with urban residential ($m_{TIA\%} = 38\%$) having lower median imperviousness than commercial ($m_{TIA\%} = 81\%$) and industrial ($m_{TIA\%} = 89\%$) classes. Furthermore, $m_{TIA\%}$ varies between 17 and 61% for different formal residential classes and between 14 and 43% for different informal residential classes. $TIA_{\%}$ of formal and informal urban land cover

Downloaded from http://iwaponline.com/wst/article-pdf/91/10/1141/1566930/wst2025067.pdf
by UNIVERSITY OF PRETORIA user

types were generalised based on a South African case study, to derive typical imperviousness values for different SANLC classes. These values are recommended as default values for hydraulic and hydrological modelling of South African formal and informal urban areas, or for use in similar land cover classes in other developing countries, where site-specific information is not readily available. It is recommended that direct comparison of remotely sensed and field-verified data be expanded in future studies.

The third objective was to apply the methodology to quantify temporal changes to TIA in selected study catchments. The results showed that the most western catchment had the highest level of TIA over the study period, with catchments further east having progressively less TIA. This result was expected, with the western catchment being closest to the city centre. As expected from the literature, the results also showed increasing TIA in all the study catchments over time. All five study catchments have observed flow data dating back to 1985 or earlier. Therefore, runoff from these catchments should be compared with impervious areas to determine trends and quantify the effect of urban built-up areas in developing countries on runoff.

ACKNOWLEDGEMENTS

The contributions of G. Mogale, L. Moleka and T. Mposhomali, who assisted with data collection in the field and processing of remotely sensed imperviousness data, are gratefully acknowledged. The contributions of the anonymous reviewers are gratefully acknowledged.

FUNDING

The research presented in this manuscript emanated from a study funded by the Water Research Commission (WRC) (Project K5/2747), whose support is acknowledged with gratitude.

AUTHOR CONTRIBUTIONS

I.L. contributed formulating the research question, designing the methodology, collecting and processing data, analysing data and preparing the manuscript. J.S. contributed to formulating the research question, designing the methodology and preparing the manuscript. T.K. contributed to the design of the methodology and preparing the manuscript.

DATA AVAILABILITY STATEMENT

The SANLC data used for this study is available from https://egis.environment.gov.za/sa_national_land_cover_datasets. The SANSA satellite images can be requested from https://www.sansa.org.za/research/.

CONFLICT OF INTEREST

The authors declare there is no conflict.

REFERENCES

Ambrosio, R. D., Balbo, A., Longobardi, A. & Rizzo, A. (2022) Re-think urban drainage following a SuDS retrofitting approach against urban flooding: a modelling investigation for an Italian case study, *Urban Forestry & Urban Greening*, **70**, 127518. https://doi.org/10.1016/j.ufug.2022.127518.

Braud, I., Breil, P., Thollet, F., Lagouy, M., Branger, F., Jacqueminet, C., Kermadi, S. & Michel, K. (2013) Evidence of the impact of urbanization on the hydrological regime of a medium-sized periurban catchment in France, *Journal of Hydrology*, **485**, 5–23. http://dx.doi.org/10.1016/j.jhydrol.2012.04.049.

Capps, K. A., Bentsen, C. N. & Ramirez, A. (2016) Poverty, urbanization, and environmental degradation: urban streams in the developing world, *Freshwater Science*, **35** (1), 429–435. https://doi.org/10.1086/684945.

Cohen, B. (2006) Urbanization in developing countries: current trends, future projections, and key challenges for sustainability, *Technology in Society*, **28**, 63–80. https://doi.org/10.1016/j.techsoc.2005.10.005.

El Hattab, M. H., Theodoropoulos, G., Rong, X. & Mijic, A. (2020) Applying the systems approach to decompose the SuDS decision-making process for appropriate hydrologic model selection, *Water*, **12** (3), 632. https://doi.org/10.3390/w12030632.

ENMAP-Box_Developers (2019) *EnMAP-Box 3 – A QGIS Plugin to Process and Visualize Hyperspectral Remote Sensing Data*. Humboldt-Universität zu Berlin under contract by the Helmholtz Centre Potsdam GFZ, Berlin, Germany.

Ferrans, P., Torres, M. N., Temprano, J. & Rodriguez Sanchez, J. P. (2022) Sustainable Urban Drainage System (SUDS) modeling supporting decision-making: a systematic quantitative review, *Science of the Total Environment*, **806**, 150447. https://doi.org/10.1016/j.scitotenv.2021.150447.

Downloaded from http://iwaponline.com/wst/article-pdf/91/10/1141/1566930/wst2025067.pdf
by UNIVERSITY OF PRETORIA user

Ferreira, C. S. S., Walsh, R. P. D., Nunes, J. P. C., Steenhuis, T. S., Nunes, M., Pedrosso de Lima, J. L. M., Coelho, C. O. A. & Ferreira, A. J. D. (2016) Impact of urban development on streamflow regime of a Portuguese peri-urban Mediterranean catchment, *Journal of Soils and Sediments*, **16** (11), 2580–2593. https://doi.org/10.1007/s11368-016-1386-5.

Geoterraimage (2019) *2018 South African National Land-Cover Dataset*. Pretoria, South Africa: Department of Environmental Affairs.

Geoterraimage (2021) *2020 South African National Land-Cover Dataset*. Pretoria, South Africa: Department of Environmental Affairs.

Geyer, H. S. J., Geyer, H. S., Du Plessis, D. J. & Van Eeden, A. (2012) Differential urbanisation trends in South Africa – regional and local equivalents, *Environment and Planning A*, **44** (12), 2940–2956. https://doi.org/10.1068/a4528.

Gogate, N. G. & Rawal, P. M. (2015) Identification of potential stormwater recharge zones in dense urban context: a case study from Pune City, *International Journal of Environmental Resolution*, **9** (4), 1259–1268.

Guo, K., Guan, M. & Yu, D. (2021) Urban surface water flood modelling – a comprehensive review of current models and future challenges, *Hydrology and Earth System Sciences*, **25**, 2843–2860. https://doi.org/10.5194/hess-25-2843-2021.

Jacobson, C. R. (2011) Identification and quantification of the hydrological impacts of imperviousness on urban catchments: a review, *Journal of Environmental Management*, **92**, 1438–1448. https://doi.org/10.1016/j.jenvman.2011.01.018.

Loots, I., Smithers, J. C. & Kjeldsen, T. R. (2023) Quantifying the influence of urban development on runoff in South Africa, *Urban Water Journal*, **20** (10), 1541–1554. doi:10.1080/1573062X.2022.2027472.

Mignot, E. & Dewals, B. (2022) Hydraulic modelling of inland urban flooding: recent advances, *Journal of Hydrology*, **609** (2022), 127763. https://doi.org/10.1016/j.jhydrol.2022.127763.

Mlambo, V. (2018) An overview of rural-urban migration in South Africa: its causes and implications, *Archives of Business Research*, **6** (4), 63–70. https://doi.org/10.14738/abr.64.4407.

Mubangizi, B. C. (2021) Rural-urban migration and smart cities: implications for service delivery in South Africa, *African Renaissance*, **18** (1), 181–201. https://doi.org/0.31920/2516-5305/2021/18n1a9.

Nguyen, H. H., Gericke, A. & Venohr, M. (2022) Importance of different imperviousness measures for predicting runoff and nutrient emissions from non-urban and urban land-uses at large spatial coverage, *Journal of Environmental Management*, **315** (2022), 115105. https://doi.org/10.1016/j.jenvman.2022.115105.

Sakijege, T. (2013) Managing flood risks: lessons from Keko Machungwa informal settlement in Dar Es Salaam, Tanzania, *Indonesian Journal of Geography*, **45** (1), 1–14.

Saurav, K. C., Shrestha, S., Ninsawat, S. & Chonwattana, S. (2021) Predicting flood events in Kathmandu Metropolitan City under climate change and urbanisation, *Journal of Environmental Management*, **281**, 111894. https://doi.org/10.1016/j.jenvman.2020.111894.

StatsSA (2022) *Census 2022 – Statistical Release*. Pretoria, South Africa: Department: Statistics South Africa.

Thompson, M. (2019) *2018 South African National Land-Cover Change Assessments*. Pretoria, RSA: Geoterraimage.

UNDP (2019) *World Urbanization Prospects: The 2018 Revision*. New York, USA: United Nations, Department of Economic and Social Affairs, Population Division.

UNDP (2014) *World Urbanization Prospects: The 2014 Revision*. United Nations, New York, USA.

Zhang, H, Chen, Y. & Zhou, J. (2015) Assessing the long-term impact of urbanization on run-off using a remote-sensing-supported hydrological model, *International Journal of Remote Sensing* **36** (21), 5336–5352.

First received 3 December 2024; accepted in revised form 10 April 2025. Available online 16 May 2025

Downloaded from http://iwaponline.com/wst/article-pdf/91/10/1141/1566930/wst2025067.pdf
by UNIVERSITY OF PRETORIA user