



A COMPARISON OF THE ATLAS OF HUMAN SETTLEMENTS AGAINST THE GLOBAL HUMAN SETTLEMENT LAYER

Georgios Ouzounis and Andrew Vekinis

EasyChair preprints are intended for rapid dissemination of research results and are integrated with the rest of EasyChair.

August 28, 2025

A COMPARISON OF THE ATLAS OF HUMAN SETTLEMENTS AGAINST THE GLOBAL HUMAN SETTLEMENT LAYER

Georgios K. Ouzounis, Andrew A. Vekinis

Atlas AI
AI Research
Palo Alto, California, USA

ABSTRACT

The article presents the findings of a quality evaluation exercise between the most recent releases of two planetary-wide built-up basemaps; the Atlas of Human Settlements or AHS of Atlas AI and the Global Human Settlement Layer or GHSL of the Joint Research Center of the European Commission. The exercise was split in two parts; a qualitative and a quantitative analysis, both powered by data collected over 8 regions across the world. The comparison yields an overwhelming lead of the AHS over the GHSL.

Index Terms— ahs, ghsl, built-up, regression, evaluation

1. INTRODUCTION

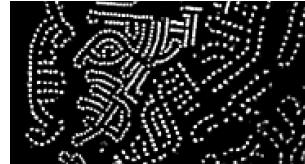
The Atlas of Human Settlements or AHS [10] is a built-up basemap of global extent, delivered at a nominal spatial resolution of 10m and updated annually, with historical records dating back to 2016. The primary data layer, referred to as the Built-up Index or **BuI**, reports the percentage of built surface within a spatial unit of 100 square meters in size. The equivalent term in GHSL [3, 8] is called the Built-up Fraction or **BUFRAC** - Fig. 1. The BuI layer is generated using a state-of-the-art deep-learning model implemented on a U-Net-like, multi-scale convolutional attention encoder to transformer-decoder architecture using multi-scale attention. Further to the BuI, the AHS delivers the Built-up Confidence or **BuC**; a raster image that reports the confidence of the regressor in producing the BuI value for each input pixel.

The AHS is generated from two different models, one that reports built-up (BuI) in the developed world, and a second one tailored for built-up in the developing countries, capturing residential buildings and structures in places where housing and living conditions are poor. It uses a set of model weights that shift the focus on smaller, more dense and radiometrically more diverse patterns of built-up. This analysis is focused on the former model due to limited reference data availability.

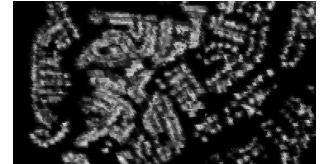
Reporting the built-up systematically, and in a globally consistent manner allows for accurate change detection, free of parallax-related artifacts, which in turn delivers actionable



(a) Google Maps basemap view of a residential neighborhood



(b) The AHS-BuI layer



(c) The GHSL-BUFRAC

Fig. 1. AHS and GHSL views of the S-E Corridor, GA, USA.

data on the growth, regression, or stagnation patterns witnessed in various human settlements across the world. This wide area monitoring (WAM) service [1] is geography, scale, climate, living standards and prior data agnostic and powers a wide range of applications in key industry verticals. Fig. 2 shows an example of the state of built-up in the South-East Corridor, Georgia, USA in early 2024 and built-up change detection between the years 2021 and 2023.

The only known alternative to the AHS, i.e. being of global coverage and delivering the same semantics, is the Global Human Settlement Layer offered by the JRC of the EC. GHSL-BUFRAC is available at 10m resolution for a single year, 2018. The GHSL underwent several model improvements to enhance BUFRAC quality, each one referred to as a *Release* for the 2018 epoch. All references to the GHSL in this article point to the R2023 release/ 2018 epoch.

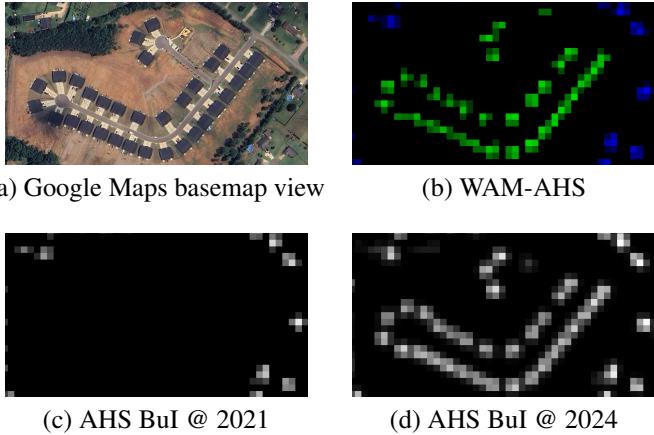


Fig. 2. Change detection using WAM-AHS. The building color coding for (b) is green: new, blue: unchanged.

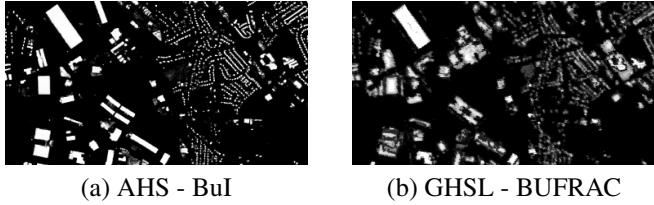


Fig. 3. Example of big building extraction consistency between the two layers

Having only two layers of the kind available in the market today, this article attempts to identify the strengths and weaknesses of each through a qualitative and a quantitative exercise presented in Sections 2 and 3 respectively. A summary of findings and discussion of results is given in Section 4.

2. QUALITATIVE ANALYSIS

To evaluate qualitatively the two layers, we attempted to capture the data scientist user-experience when confronted with both layers as analysis ready data. We consulted four geospatial-data analysts to identify key features that best describe their engagement with the data layers and recorded their experiences when re-visiting each one separately. The findings are discussed in Section 4. The features are:

- **built-up surface completeness:** empirical estimate of completeness of the binary built-up surface,
- **noise in-between built-up:** false BU positives inside the studied settlements,
- **noise outside settlements:** false BU positives outside the studied settlements,
- **visual clarity of built-up:** overall appreciation of the visual scene - Fig. 1 (b, c),
- **big building segmentation:** suitability for segmenting

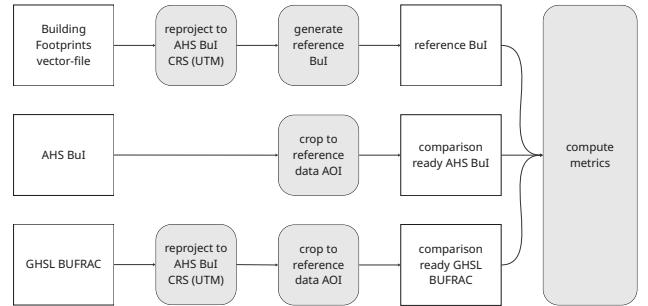


Fig. 4. Evaluation workflow.

clearly distinguishable big buildings- Fig. 3,

- **built-up statistics:** suitability for computing statistics,
- **change detection:** suitability for change detection, Fig. 2.

3. QUANTITATIVE ANALYSIS

The quantitative analysis of both layers was orchestrated as a comparison against reference data. The latter was assembled from 8 areas of interest (AOIs) in 6 different countries: China - Jinxiang, Japan - Hamamatsu, Japan - Kyoto, Japan - Maebashi, Poland - Warsaw, South Africa - Cape Town, UK - Southampton, USA - Worcester (MA). They are referred to as AOI 1 through 8, [5, 2, 4, 6, 9, 7].

Manually delineated building footprints were collected for each AOI and translated into test BuI surfaces, i.e. the same material consumed by both AHS and GHSL models during training. The term test BuI is referred to as reference data instead of *ground truth*, as the quality of the building footprints is subject to human interpretation and skill. Each set of building footprints differs from the others in two ways, the date of production and in the clarity/resolution of the underlying image used to produce them. The reference data used was selected based on the production date being as close to the end of the year 2018 as possible.

To make this a fair comparison we computed the AHS-BuI of each AOI for the year 2018. The quality figures reported do not necessarily reflect the true quality of the AHS (underestimation) but allow for a one-to-one comparison against the GHSL. The results are biased by the fact that we compare findings computed on annual median composites (AHS and GHSL) against the reference data generated for a certain time stamp later in time; $t > 2018$. This propagates the same error to both layers thus does not impact the outcome of the comparison. A further, very minor bias in favor of the AHS comes from the fact that both the reference data and the GHSL layers used, were re-projected to the AHS CRS (UTM) - Fig. 4. The latter was selected over the World-Mollweide of the GHSL for its iso-tropic pixel representation, appreciated in both ML model training and deployment.

To compute the ‘closeness’ of each layer to the reference data we utilized two sets of metrics; one for segmentation and one for regression. The former set was used to evaluate the completeness of the built-up surface, and the latter for the evaluation of the pixel content accuracy. In the following P , N , TP , TN , FP and FN stand for the number of pixels that are *positive* ($BuI>0$), *negative* ($BuI=0$), *true positive*, *true negative*, *false positive*, *false negative* respectively. The segmentation metrics used are as follows:

$$\text{Binary Accuracy} : \frac{TP + TN}{P + N}, \quad (1)$$

$$\text{Precision} : \frac{TP}{TP + FP}, \quad \text{Recall} : \frac{TP}{TP + FN} \quad (2)$$

$$\text{F1 Score} : \frac{2 \times PR \times RE}{RP + RE} = \frac{2 \times TP}{2 \times TP + FP + FN} \quad (3)$$

Matthew's Correlation Coefficient:

$$\frac{TP \times TN - FP \times FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} \quad (4)$$

The regression metrics used are as follows:

$$\text{Round Mean Square Error} : \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{X}_i - X_i)^2} \quad (5)$$

$$\text{Mean Average Error} : \frac{1}{n} \sum_{i=1}^n |\hat{X}_i - X_i| \quad (6)$$

and were computed for each inference image as a whole, and for the $BuI>0$ and $BuI=0$ pixel sets separately.

4. DISCUSSION OF RESULTS

Qualitative analysis findings: Responses on the Built-up Surface Completeness suggested that both layers, if treated as binary surfaces, deliver a satisfactory IoU with reference data, i.e. they do not miss and do not over-represent built-up. Some in-between built-up noise appears in both, examples of which are due to highly reflecting road segments, construction materials, aggregation of metal sheets that are not parts of roofs (containers, trucks), parking lots, etc. The GHSL slightly under performs due to increased sensitivity to parking lots/loading bays. Some minor false positives appear outside settlements and are primarily attributed to highly reflecting rocks and in some instances to minor water bodies.

The visual clarity of built-up is the highest discriminator between the two layers. While built-up is mostly captured accurately in both, the GHSL delivers a blurry view making it hard to trace individual buildings. By contrast to the GHSL,

the AHS presents clearly distinguishable buildings, even the smallest ones, in cases they are further than 10m apart, i.e. 1 spatial unit separation. An example is shown in Fig. 1.

Big buildings can be extracted from both layers as stand-alone structures that are useful and in-demand for supply-chain and real-estate management applications. While generally highlighted well in both layers, in the GHSL big buildings appear with textural noise running through their extent that makes it harder for simple/fast computer vision scripts to return a single segment per building, Fig. 3.

Built-up statistics can be computed from both layers, offering insights such as how much of a settlement surface area is built, how dense is the built-up, etc. A limitation that relates to the earlier observation on big buildings is that in the case of the GHSL the built-up cannot be trivially binned to size histograms reporting how many buildings are there for specific size ranges. In case of very dense built-up both layers cannot discriminate between individual buildings and this is primarily due to the spatial resolution of the input data.

Lastly, the GHSL being a single-year release cannot be utilized directly in change detection, where as AHS-WAM detects change robustly and among any two BuI instances from the present date back to 2016. Change detection can be computed by ML models on multispectral image pairs, but at the cost of model retraining for each AOI to prevent drift, and increased sensitivity to local built-up patterns. The unavailability of historical GHSL data weakens its adoption by what appears to be one of the highest priority commercial use-cases.

Quantitative analysis findings: An initial observation from Table 1 is that the GHSL has a major imbalance between precision and recall. Low precision and high recall means that the model is good at finding all the actual positives ($BuI>0$) but among the instances the model predicts as built-up, many are actually false positives. The lack of model sophistication leads to poor generalization erring on the side of predicting positive BuI and pretty much ‘catches everything’ but inaccurately. The AHS maintains a far better balance between the two metrics. Looking at the F1 and MCC scores (holistic view of model performance), the AHS leads by 14% and 13% respectively and this is consistent throughout all 8 AOIs.

Table 2 lists the regression metrics’ scores related to the BuI values. The AHS layers on average and across all pixels show a relative reduction of about 30% in regards to RMSE and about 35% in regards to MAE when compared to the GHSL layer values. These relative reductions persist in empty ($BuI=0$) and non-empty ($BuI>0$) pixel sets too.

5. CONCLUSIONS

In this paper we evaluated the AHS on selected global AOIs, comparing it to the GHSL. Both qualitative and quantitative results show AHS has a competitive edge. As a continually evolving product, further improvements are expected.

Table 1. Segmentation Metrics

metric	layer	AOI1	AOI2	AOI3	AOI4	AOI5	AOI6	AOI7	AOI8	average
Binary Accuracy	GHSL	0.95449	0.95114	0.87118	0.81972	0.78339	0.94021	0.92876	0.85596	0.88810
	AHS	0.97388	0.98036	0.92371	0.93487	0.93890	0.96911	0.96635	0.94827	0.95443
Precision	GHSL	0.46971	0.45751	0.67842	0.43199	0.37874	0.55974	0.45056	0.37651	0.47539
	AHS	0.63336	0.71959	0.81203	0.72438	0.75916	0.74474	0.68024	0.68203	0.71944
Recall	GHSL	0.96417	0.96787	0.98799	0.98565	0.97163	0.98625	0.92083	0.95821	0.96782
	AHS	0.84213	0.86114	0.93088	0.85008	0.79426	0.90091	0.81496	0.77707	0.84642
F1 Score	GHSL	0.63169	0.62132	0.80445	0.60071	0.54503	0.71416	0.60506	0.54060	0.63287
	AHS	0.72297	0.78403	0.86740	0.78221	0.77631	0.81541	0.74153	0.72646	0.77704
MCC Score	GHSL	0.65567	0.64714	0.74130	0.57813	0.52037	0.71788	0.61520	0.54678	0.62780
	AHS	0.71753	0.77725	0.81794	0.74751	0.74121	0.80304	0.72707	0.69984	0.75392

Table 2. Regression Metrics

metric	layer	AOI1	AOI2	AOI3	AOI4	AOI5	AOI6	AOI7	AOI8	average
RMSE	GHSL	0.10310	0.09340	0.21453	0.17004	0.18895	0.12891	0.10707	0.11957	0.14069
	AHS	0.09198	0.06411	0.17369	0.11263	0.12570	0.09689	0.07332	0.09188	0.10377
MAE	GHSL	0.02259	0.02074	0.10674	0.07090	0.08300	0.03612	0.02743	0.04053	0.05100
	AHS	0.01846	0.01217	0.08121	0.03771	0.03909	0.02454	0.01638	0.02450	0.03175
RMSE - non empty	GHSL	0.40256	0.33715	0.33785	0.33721	0.35086	0.38365	0.33130	0.30600	0.34832
	AHS	0.37423	0.27614	0.31414	0.28217	0.30835	0.30553	0.26902	0.26250	0.29901
MAE - non empty	GHSL	0.31843	0.26816	0.27299	0.26933	0.27451	0.30701	0.26004	0.23602	0.27581
	AHS	0.29383	0.21599	0.25382	0.21822	0.23220	0.24036	0.20739	0.19934	0.23264
RMSE - empty	GHSL	0.06654	0.06691	0.14866	0.12764	0.15276	0.08033	0.07613	0.08580	0.10059
	AHS	0.05542	0.03452	0.07821	0.05219	0.06771	0.05342	0.03892	0.05622	0.05457
MAE - empty	GHSL	0.01047	0.01073	0.04835	0.04136	0.05590	0.01504	0.01401	0.02345	0.02741
	AHS	0.00718	0.00393	0.02060	0.01080	0.01180	0.00775	0.00538	0.00924	0.00958

REFERENCES

- [1] Atlas AI. Introducing aperture® pulse: A new era of scalable economic change detection. <https://www.atlasai.co/blog/introducing-aperture-pulse>, 2025.
- [2] Source Cooperative. Japanese building footprint data. <https://source.coop/repositories/pacificspatial/flateau/description>, 2024.
- [3] Christina Corbane, Vasileios Syrris, Filip Sabo, Panagiotis Politis, Michele Melchiorri, Martino Pesaresi, Pierre Soille, and Thomas Kemper. Convolutional neural networks for global human settlements mapping from sentinel-2 satellite imagery. *Neural Computing and Applications*, 33(12):6697–6720, 2021. doi: [10.1007/s00521-020-05449-7](https://doi.org/10.1007/s00521-020-05449-7).
- [4] geoportal.gov.pl. Topographic objects database (bdot10k). <https://www.geoportal.gov.pl/en/data/topographic-objects-database-bdot10k/>, 2023.
- [5] Planetek Hellas. Jinxiang building footprints dataset. manual annotation, 2023. URL <https://www.planetek.gr/>.
- [6] City of Cape Town. 2d building footprints. <https://odp-cctegis.opendata.arcgis.com/datasets/>
- cctegis::2d-building-footprints/about, 2023.
- [7] MassGIS (Bureau of Geographic Information). Massgis data: Building structures (2-d). <https://www.mass.gov/info-details/massgis-data-building-structures-2-d#downloads->, 2024.
- [8] Martino Pesaresi, Marcello Schiavina, Panagiotis Politis, Sergio Freire, Katarzyna Krasnodebska, Johannes H. Uhl, Alessandra Carioli, Christina Corbane, Lewis Dijkstra, Pietro Florio, Hannah K. Friedrich, Jing Gao, Stefan Leyk, Linlin Lu, Luca Maffenini, Ines Mari-Rivero, Michele Melchiorri, Vasileios Syrris, Jamon Van Den Hoek, and Thomas Kemper and. Advances on the global human settlement layer by joint assessment of earth observation and population survey data. *International Journal of Digital Earth*, 17(1):2390454, 2024. doi: [10.1080/17538947.2024.2390454](https://doi.org/10.1080/17538947.2024.2390454).
- [9] Ordnance Survey. Great britain's national geographic database. <https://www.ordnancesurvey.co.uk/>, 2024.
- [10] Abe Tarapani. Introducing the fourth release of the atlas of human settlements. <https://www.atlasai.co/blog/introducing-the-fourth-release-of-the-atlas-of-human-settlements>, 2024.