Evaluating the quality of the Global Human Settlement Layer R2023A

Johannes H. Uhl

European Commission
Joint Research Centre (JRC)
Unit E.1 - GHSL

The 13th ISDE International Lectures

11/11/2024

European Commission

Global Human Settlement Layer (GHSL) data ecosystem

Dataset name	Variable	Spatial resolution			Year	Nature of the data
		10 m	100 m	1000 m		
GHS-BUILT-S	Total built-up surface [sqm]	●			2018	Measured
GHS-BUILT-S	Total built-up surface [sqm]		●	●	1975-2030	Measured, modelled
GHS-BUILT-H	Average building height [m]		●		2018	Measured
GHS-BUILT-V	Total building volume [m3]		●	●	1975-2030	Derived
GHS-BUILT-S-NRES	Non-residential built-up surface [sqm]	●	●	●	1975-2030, 2018	Measured,modelled
GHS-BUILT-V-NRES	Non-residential building volume [m3]		●	●	1975-2030	Derived
GHS-POP	Resident population		●	●	1975-2030	Measured, modelled
GHS-LAND	Land surface [sqm]	●	●	●	2018	Measured
GHS-SMOD	Degree of Urbanisation settlement model			●	1975-2030	Derived
GHS-MSZ	Morphological Settlement Zone	●			2018	Derived
GHS-BUTYPE	Built-up typology		●		2018	Derived

Continuous estimates
Categorical data
Multitemporal 1975-2030

Global Human Settlement Layer (GHSL) data ecosystem

Dataset name	Variable	Spatial resolution			Year	Nature of the data	Validated
		10 m	100 m	1000 m			
GHS-BUILT-S	Total built-up surface [sqm]	●			2018	Measured	✓
GHS-BUILT-S	Total built-up surface [sqm]		●	●	1975-2030	Measured, modelled	✓
GHS-BUILT-H	Average building height [m]		●		2018	Measured	✓
GHS-BUILT-V	Total building volume [m3]		●	●	1975-2030	Derived	
GHS-BUILT-S-NRES	Non-residential built-up surface [sqm]	●	●	●	1975-2030, 2018	Measured,modelled	✓
GHS-BUILT-V-NRES	Non-residential building volume [m3]		●	●	1975-2030	Derived	
GHS-POP	Resident population		●	●	1975-2030	Measured, modelled	✓
GHS-LAND	Land surface [sqm]	●	●	●	2018	Measured	✓
GHS-SMOD	Degree of Urbanisation settlement model			●	1975-2030	Derived	
GHS-MSZ	Morphological Settlement Zone	●			2018	Derived	
GHS-BUTYPE	Built-up typology		●		2018	Derived	

Continuous estimates
Categorical data
Multitemporal 1975-2030

Global Human Settlement Layer (GHSL) data ecosystem

Dataset name	Variable	Spatial resolution			Year	Nature of the data	Validated
		10 m	100 m	1000 m			
GHS-BUILT-S	Total built-up surface [sqm]	●			2018	Measured	✓
GHS-BUILT-S	Total built-up surface [sqm]		●	●	1975-2030	Measured, modelled	✓
GHS-BUILT-H	Average building height [m]		●		2018	Measured	✓
GHS-BUILT-S-NRES	Non-residential built-up surface [sqm]	●	●	●	1975-2030, 2018	Measured,modelled	✓
GHS-POP	Resident population		●	●	1975-2030	Measured, modelled	✓
GHS-LAND	Land surface [sqm]	●	●	●	2018	Measured	✓

Evaluating the quality of *measured* and *modelled primary* components of the GHSL ecosystem.

Continuous estimates
Categorical data
Multitemporal 1975-2030

Quality evaluation by means of accuracy assessments

	GHSL data	Reference data	Comparative assessment
Built-up surface – Contemporary	GHS-BUILT-S Built-up surface 10m (2018)	VHR/DL/Cadastral vector building footprint data, N=600,000,000; Visual S-2 data inspection, categorical, discrete locations N=1,000,000; Crowd-sourced reference data, continuous, N=50,000	World Settlement Footprint (WSF); ESRI Landcover v2018, v2020; Dynamic World; WorldCover (Contemporary landcover / settlement data)
	GHS-BUILT-S Built-up surface 100m (2020)	VHR/DL/Cadastral vector building footprint data, N=600,000,000	
Historical	GHS-BUILT-S Built-up surface 100m **(1975-2020)**	Building footprint data ***with construction year attribute***, N=170,000,000	GHSL versions 2016-2022; GAUD; WSF Evolution; GAIA; GISD; GISA (Multitemporal settlement (-related) data)
Non-res.	GHS-BUILT-S-NRES Non-residential built-up surface 10m (2018)	Visual S-2 data inspection, categorical, discrecte locations N=1,000,000	
Building height	GHS-BUILT-H Avg. building height 100m (2018)	Cadastral vector building footprint data ***with building height attribute***, N=220,000,000	WSF 3D; ALOS_30m (Global building height data)
Land surface	GHS-LAND Land vs water surface 10m (2018)	Visual S-2 data inspection, categorical, discrecte locations N=1,000,000	ESRI Landcover v2018, v2020; Dynamic World; WorldCover (Contemporary landcover data)
Population	GHS-POP Residential population 100m (1995-2020)	Authoritative gridded population counts (24 countries)	WorldPop (Gridded population data)

Validation | Benchmarking

Reference data

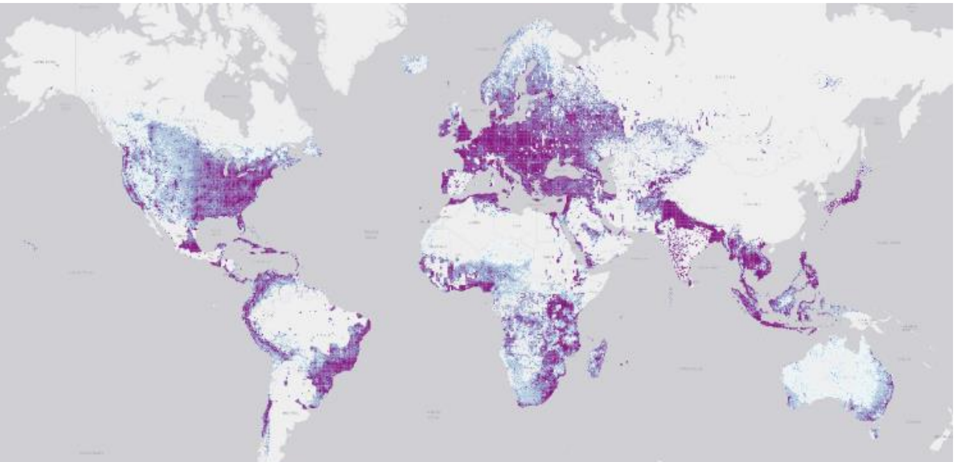
	GHSL data	Reference data			Comparative assessment
Built-up surface: Contem-porary	GHS-BUILT-S Built-up surface 10m (2018)	VHR/DL/Cadastral vector building footprint data, N=600,000,000	Visual S-2 data inspection, categorical, discrete locations N=1,000,000	Crowd-sourced reference data, continuous, N=50,000	World Settlement Footprint (WSF); ESRI Landcover v2018, v2020; Dynamic World; WorldCover Contemporary landcover / settlement data
Built-up surface: Contem-porary	GHS-BUILT-S Built-up surface 100m (2020)	VHR/DL/Cadastral vector building footprint data, N=600,000,000			
Historical	GHS-BUILT-S Built-up surface 100m **(1975-2020)**	Building footprint data ***with construction year attribute***, N=170,000,000			GHSL versions 2016-2022; GAUD; WSF Evolution; GAIA; GISD; GISA Multitemporal settlement (-related) data
Non-res.	GHS-BUILT-S-NRES Non-residential built-up surface 10m (2018)		Visual S-2 data inspection, categorical, discrecte locations N=1,000,000		
Building height	GHS-BUILT-H Avg. building height 100m (2018)	Cadastral vector building footprint data ***with building height attribute***, N=220,000,000			WSF 3D; ALOS_30m Global building height data
Land surface	GHS-LAND Land vs water surface 10m (2018)		Visual S-2 data inspection, categorical, discrecte locations N=1,000,000		ESRI Landcover v2018, v2020; Dynamic World; WorldCover Contemporary landcover data
Population	GHS-POP Residential population 100m (1995-2020)	Authoritative gridded population counts (24 countries)			WorldPop Gridded population data

Validation

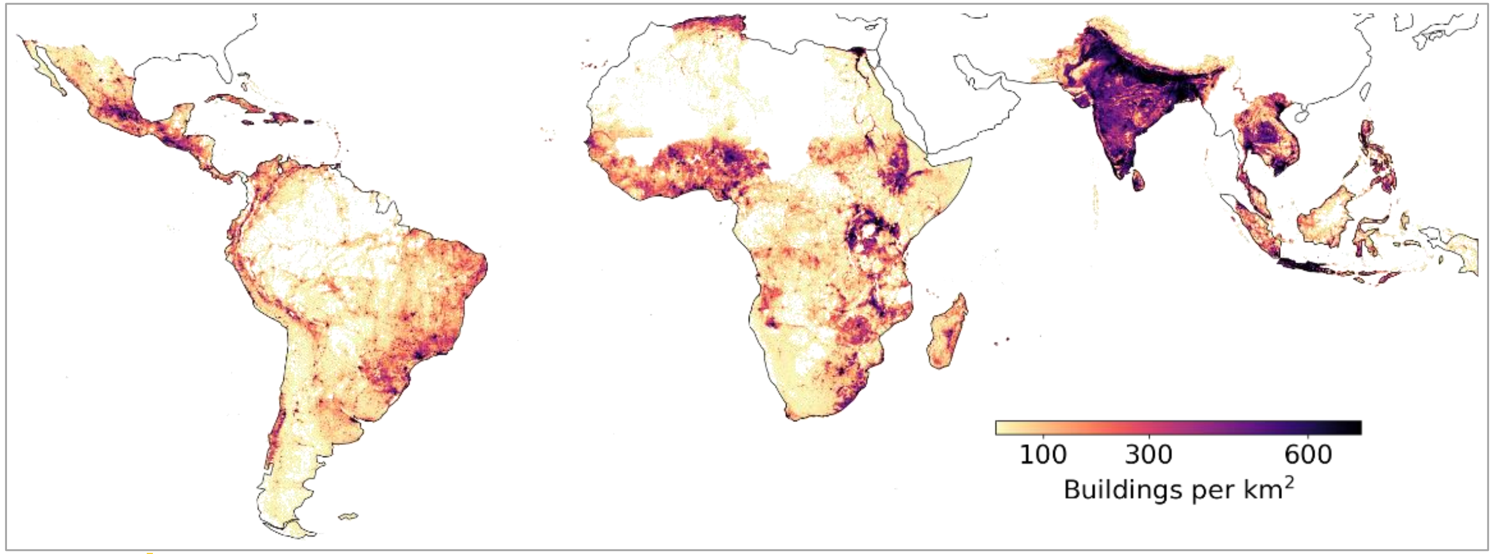
Benchmarking

Reference data I: CNN-based, VHR-derived building footprint data

Microsoft: 1.4B buildings from Bing Maps imagery between 2014 and 2024



Source:https://github.com/microsoft/GlobalMLBuildingFootprints



Source: https://sites.research.google/open-buildings/

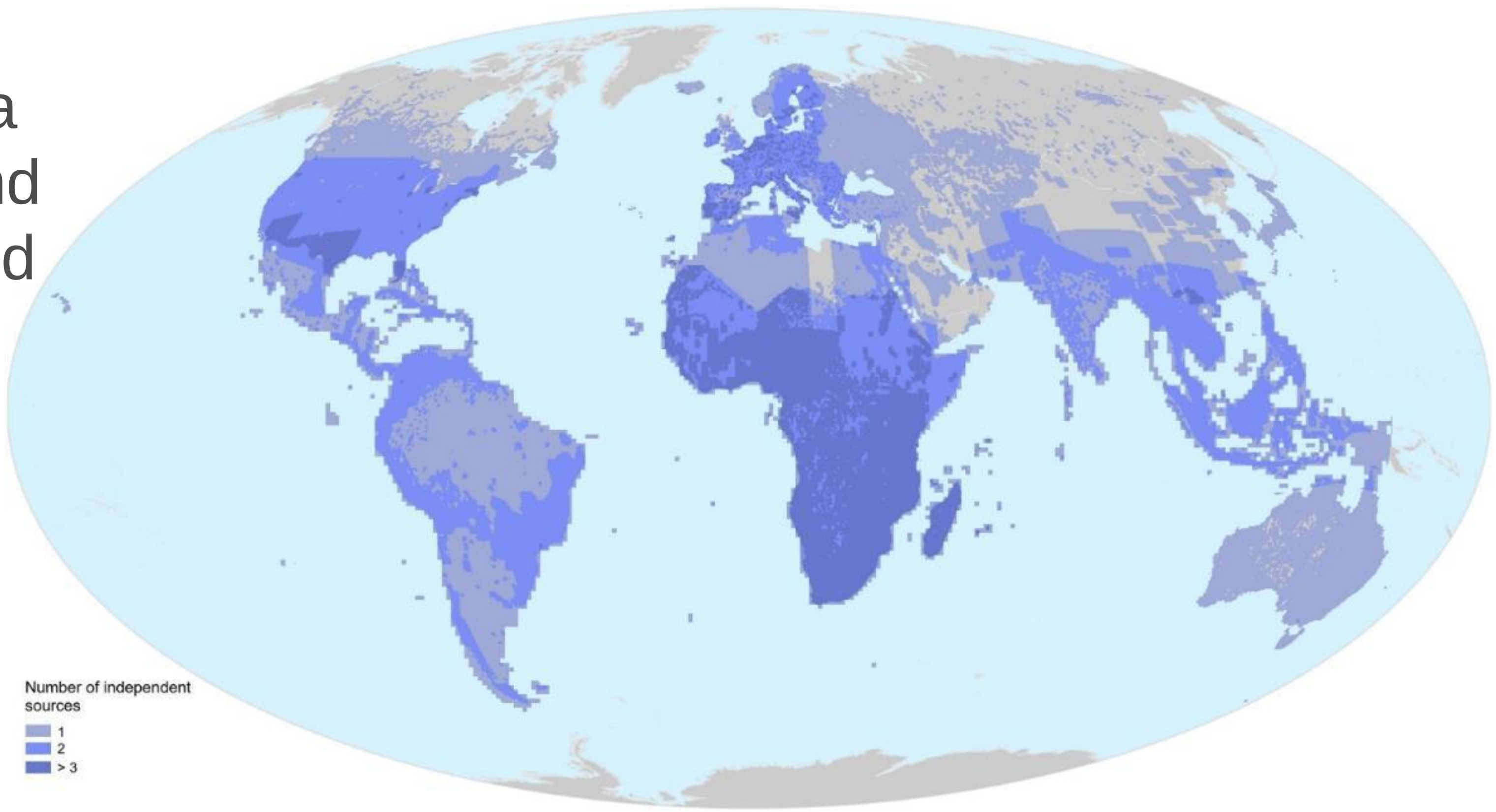
Source:https://github.com/microsoft/GlobalMLBuildingFootprints

Google: The dataset contains 1.8 billion building detections, across an inference area of 58M km^2 within Africa, South Asia, South-East Asia, Latin America and the Caribbean.

European Commission

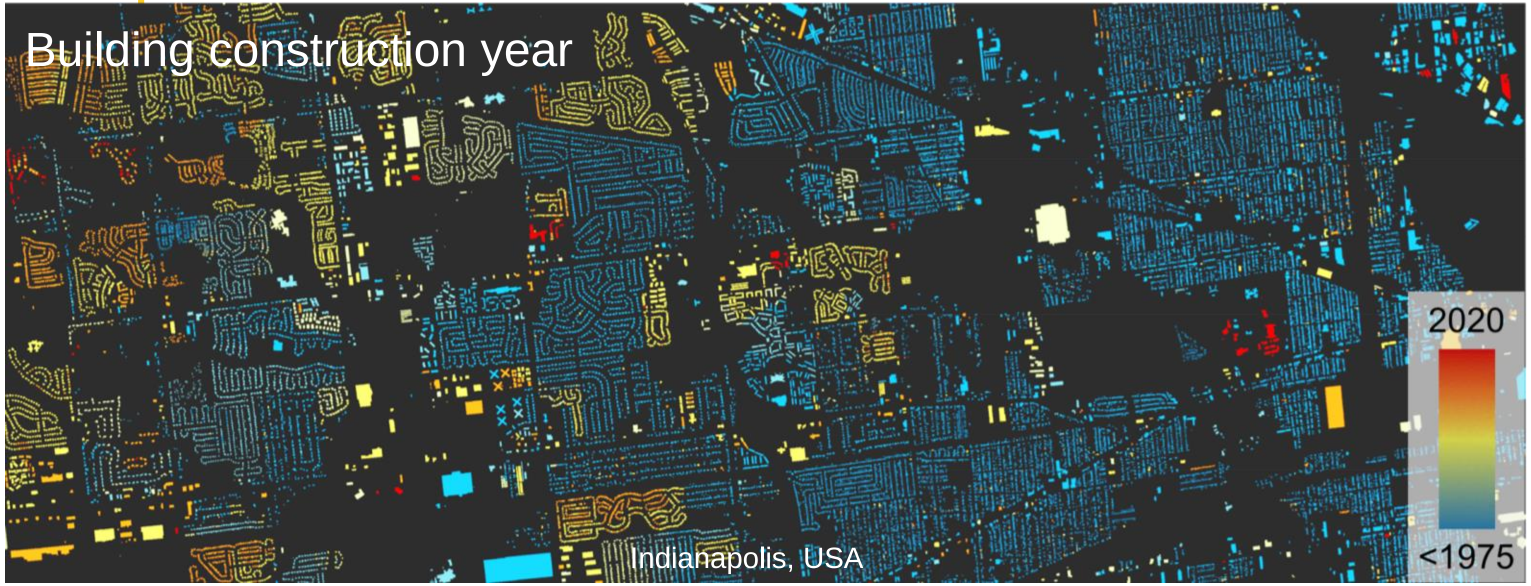
Reference data I: CNN-based, VHR-derived building footprint data

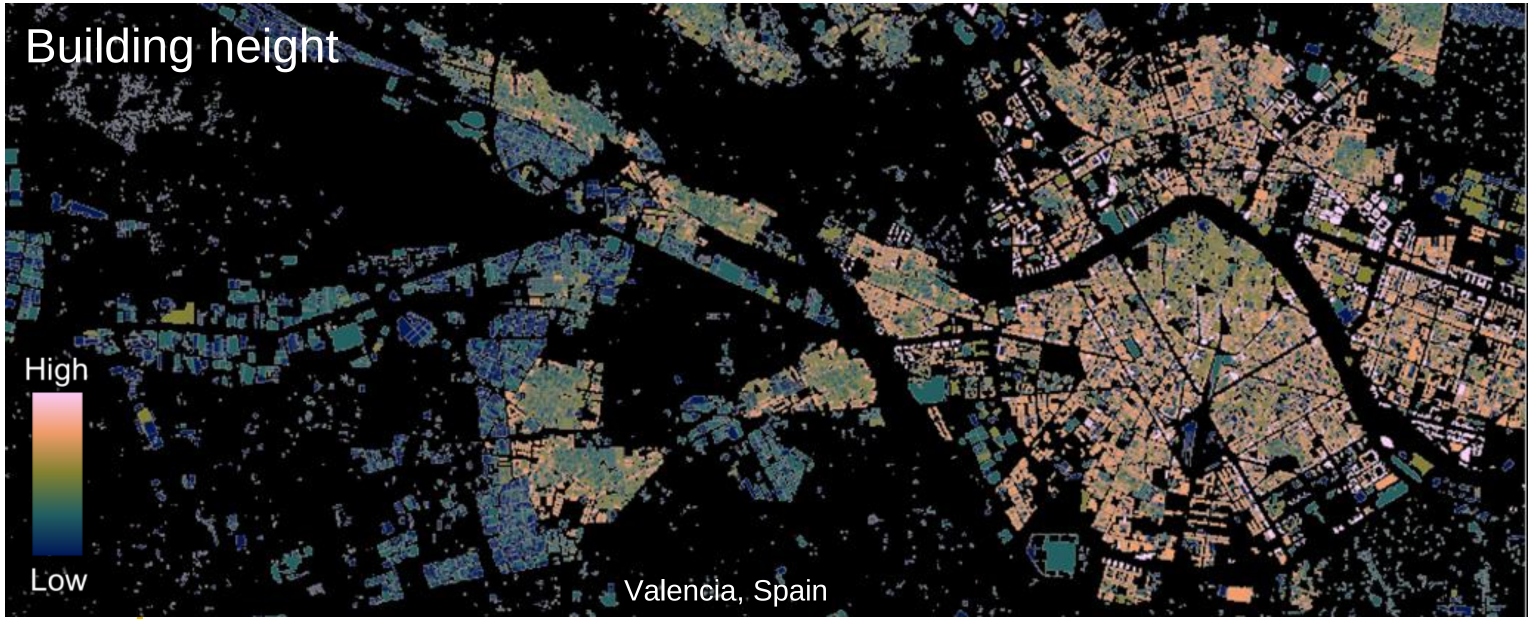
- Sources: Microsoft, Google, Ecopia
- Reference building footprint data were rasterized into 1m, 10m and 100m grids, aligned to GHSL grid
- Analytical unit: blocks of 25x25km, globally distributed
- Agreement calculated at block-level
- Agreement metric: R-accuracy, MAE

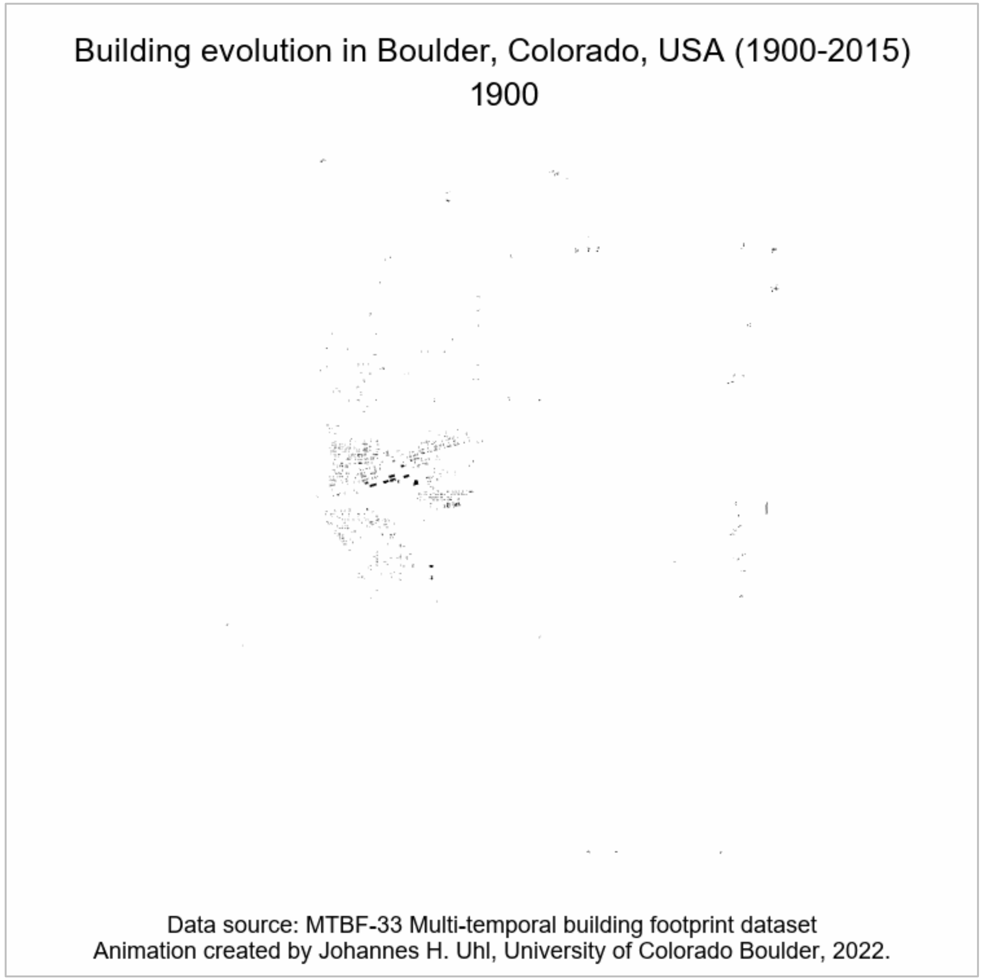


Source: Pesaresi et al. 2024

Reference data II: Cadastral building footprint data attributed with construction year / height information







Data source: MTBF-33,
Visualisation: https://github.com/johannesuhl/shapefile2gif

European Commission

Image source: Pesaresi et al. 2024

Reference data II: Cadastral building footprint data attributed with construction year information

MTBF-33

Contents lists available at ScienceDirect

Data in Brief

journal homepage: www.elsevier.com/locate/dib

ELSEVIER

Data Article

MTBF-33: A multi-temporal building footprint dataset for 33 counties in the United States (1900 – 2015)

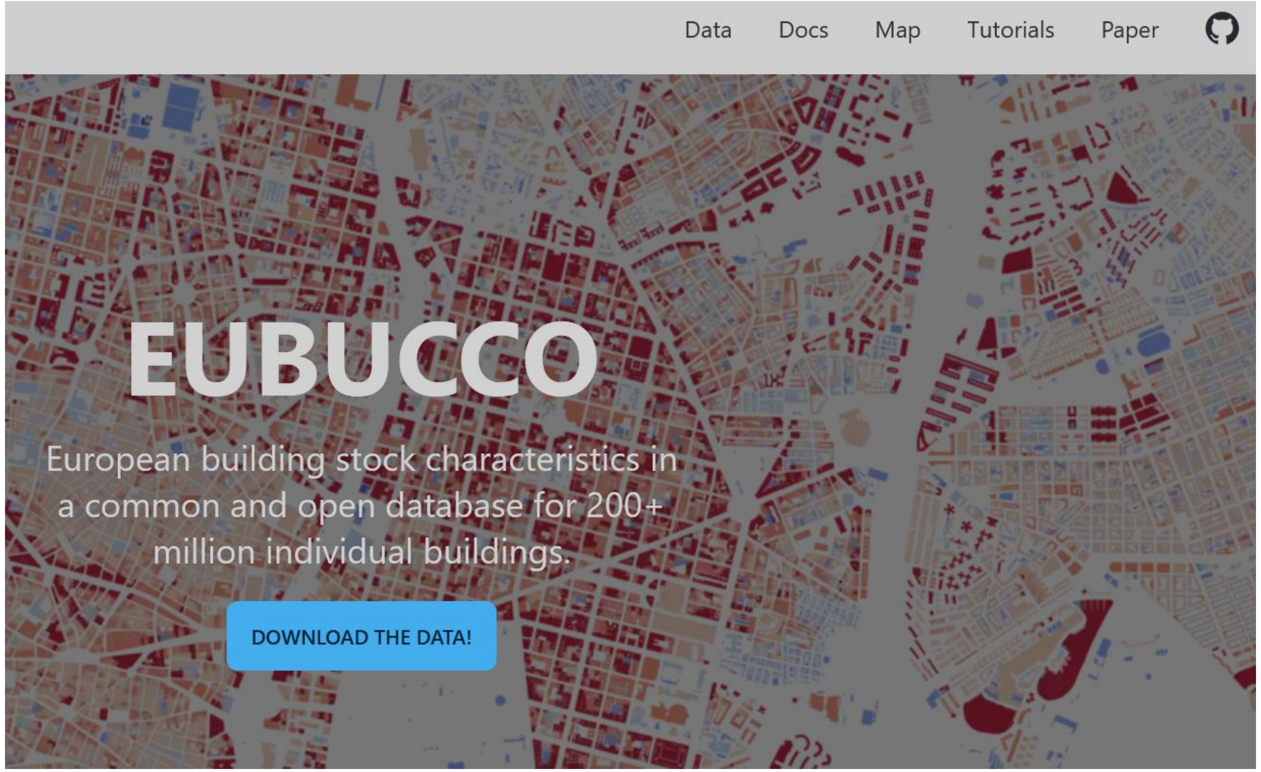
Johannes H. Uhl[a,b,c,*], Stefan Leyk[b,c]

[a] *University of Colorado Boulder, Cooperative Institute for Research in Environmental Sciences (CIRES) 216 UCB, Boulder, CO-80309, USA*
[b] *University of Colorado Boulder, Institute of Behavioral Science, 483 UCB, Boulder, CO-80309, USA*
[c] *University of Colorado Boulder, Department of Geography, 260 UCB, Boulder, CO-80309, USA*

Check for updates

Source: Uhl and Leyk (2022)

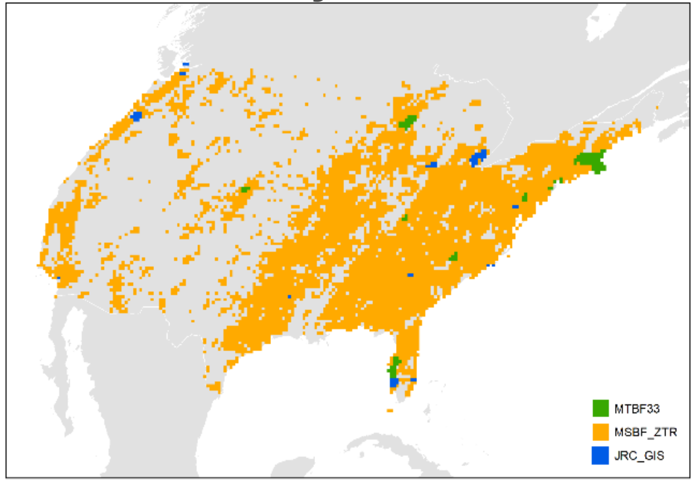
EUBUCCO

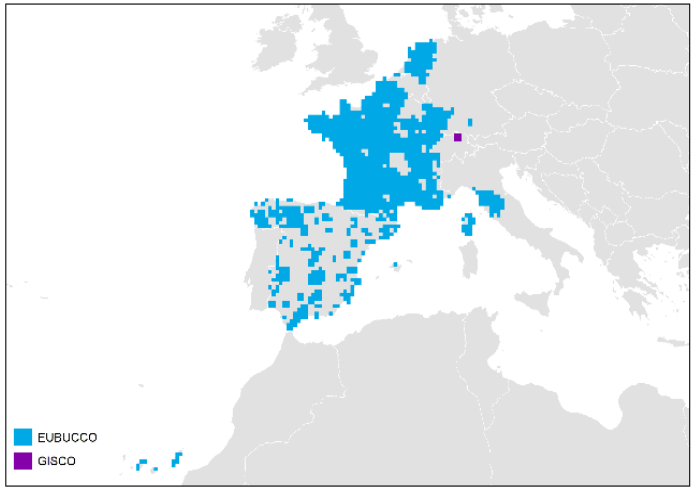


Source: https://eubucco.com/

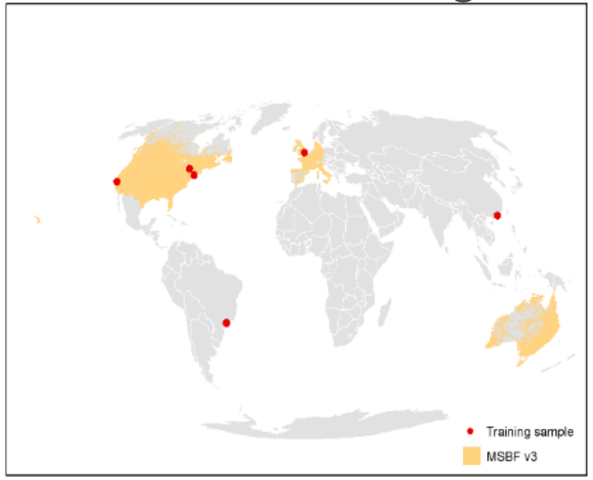
Reference data II: Cadastral building footprint data attributed with construction year / height information

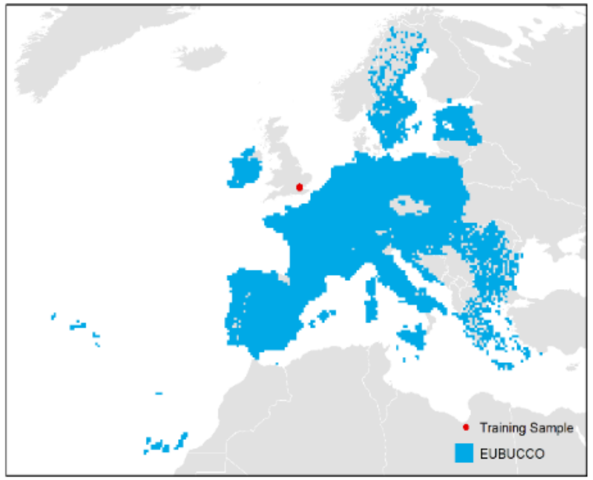
Coverage of BF data with
construction year information

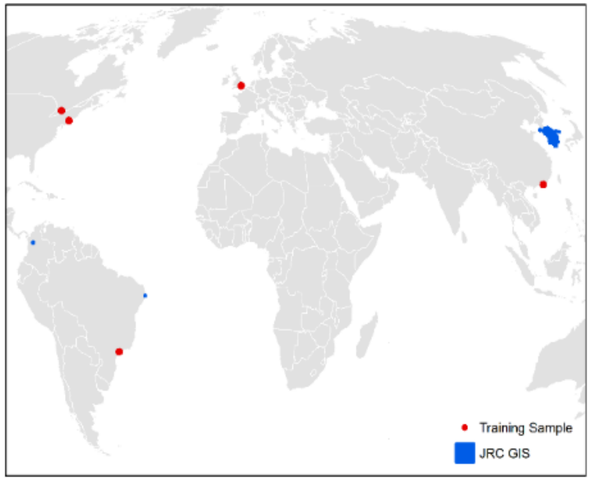




Coverage of BF data with
height information







Source: Pesaresi et al. 2024

Reference data

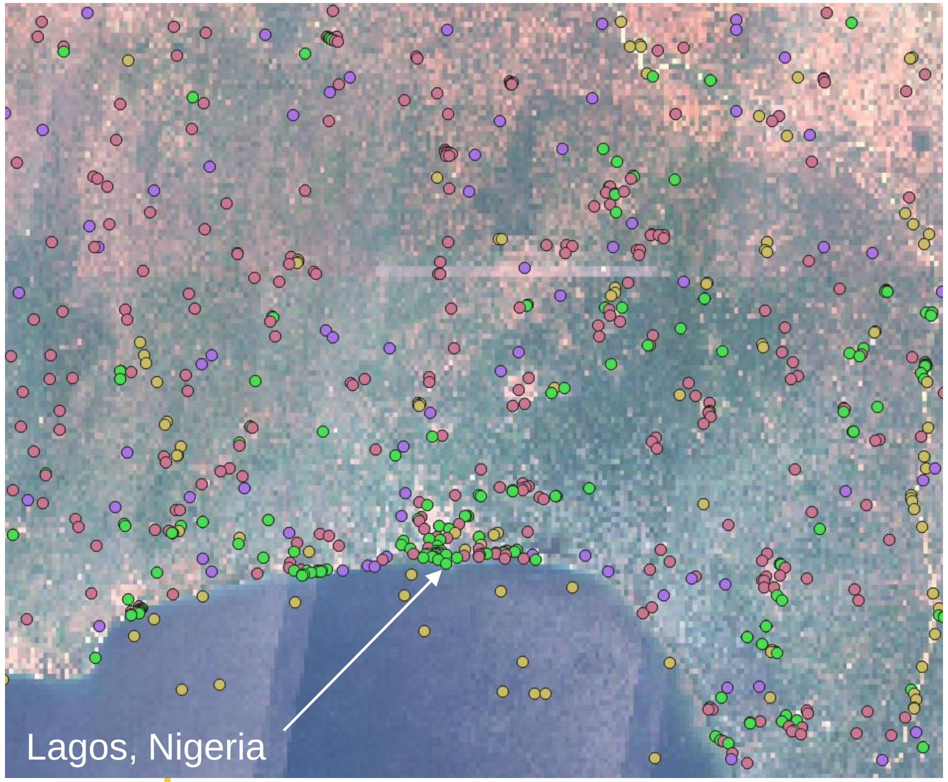
	GHSL data	Reference data	Reference data	Reference data	Comparative assessment
Built-up surface – Contem-porary	GHS-BUILT-S Built-up surface 10m (2018)	VHR/DL/Cadastral vector building footprint data, N=600,000,000	Visual S-2 data inspection, categorical, discrete locations N=1,000,000	Crowd-sourced reference data, continuous, N=50,000	World Settlement Footprint (WSF); ESRI Landcover v2018, v2020; Dynamic World; WorldCover Contemporary landcover / settlement data
Built-up surface – Contem-porary	GHS-BUILT-S Built-up surface 100m (2020)	VHR/DL/Cadastral vector building footprint data, N=600,000,000			
Built-up surface – Historical	GHS-BUILT-S Built-up surface 100m (1975-2020)	Building footprint data ***with construction year attribute***, N=300,000,000			GHSL versions 2016-2022; GAUD; WSF Evolution; GAIA; GISD; GISA Multitemporal settlement (-related) data
Built-up surface – Non-res.	GHS-BUILT-S-NRES Non-residential built-up surface 10m (2018)		Visual S-2 data inspection, categorical, discrecte locations N=1,000,000		
Building height	GHS-BUILT-H Avg. building height 100m (2018)	Cadastral vector building footprint data ***with building height attribute***, N=XXXXX			WSF 3D; ALOS_30m Global building height data
Land surface	GHS-LAND Land vs water surface 10m (2018)		Visual S-2 data inspection, categorical, discrecte locations N=1,000,000		ESRI Landcover v2018, v2020; Dynamic World; WorldCover Contemporary landcover data
Population	GHS-POP Residential population 100m (1995-2020)	Authoritative gridded population counts (24 countries)			WorldPop Gridded population data

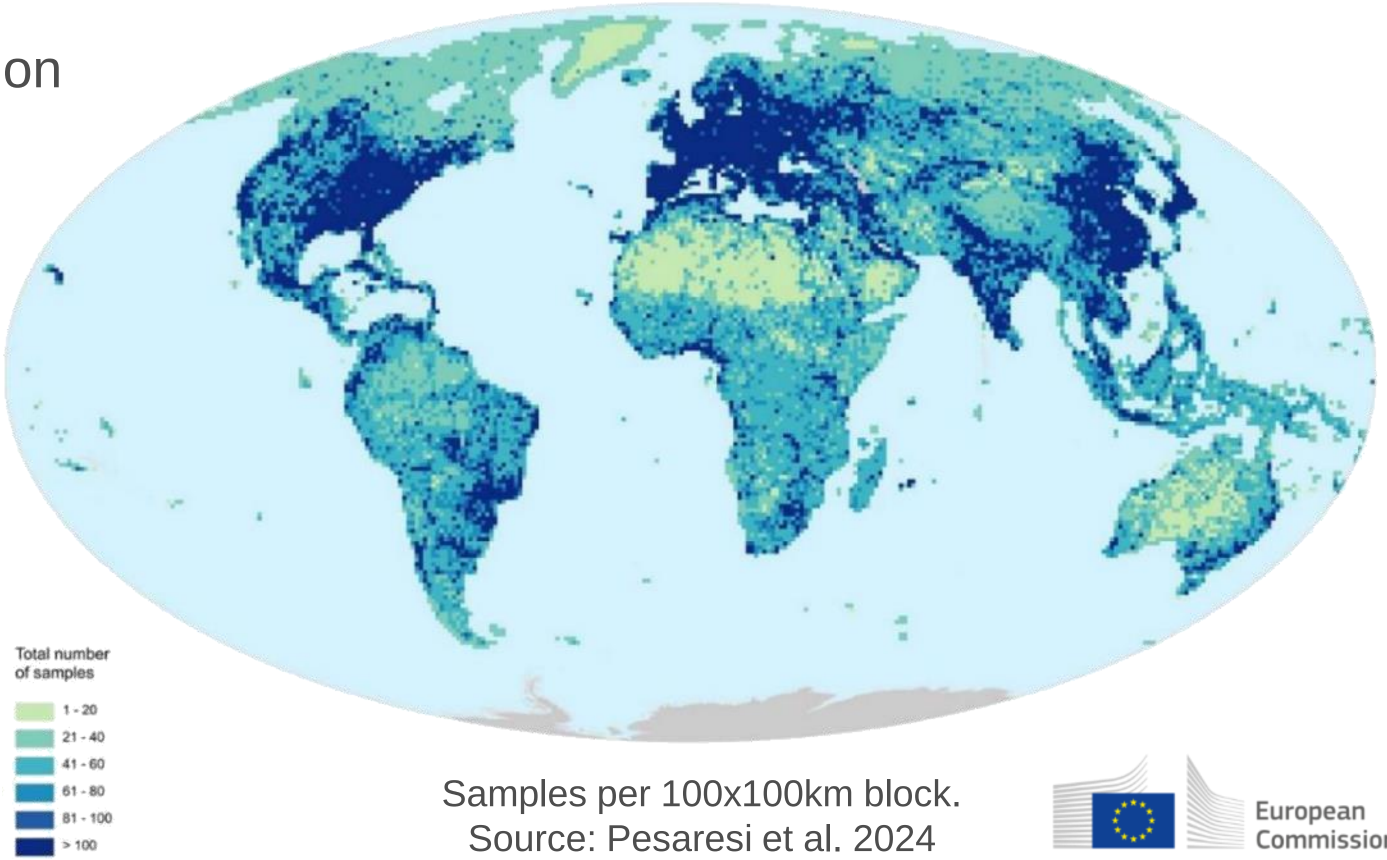
Validation

Benchmarking

Reference data III: Discrete, visual S-2 data inspection, categorical information (N=1,000,000)

- Manually annotated from S-2 data, multiple human inspection, incl. decision confidence information
- Residential, Non-residential built-up, non-built-up,land vs. water
- Agreement metric: Intersection over Union





Samples per 100x100km block.
Source: Pesaresi et al. 2024

European Commission

Reference data

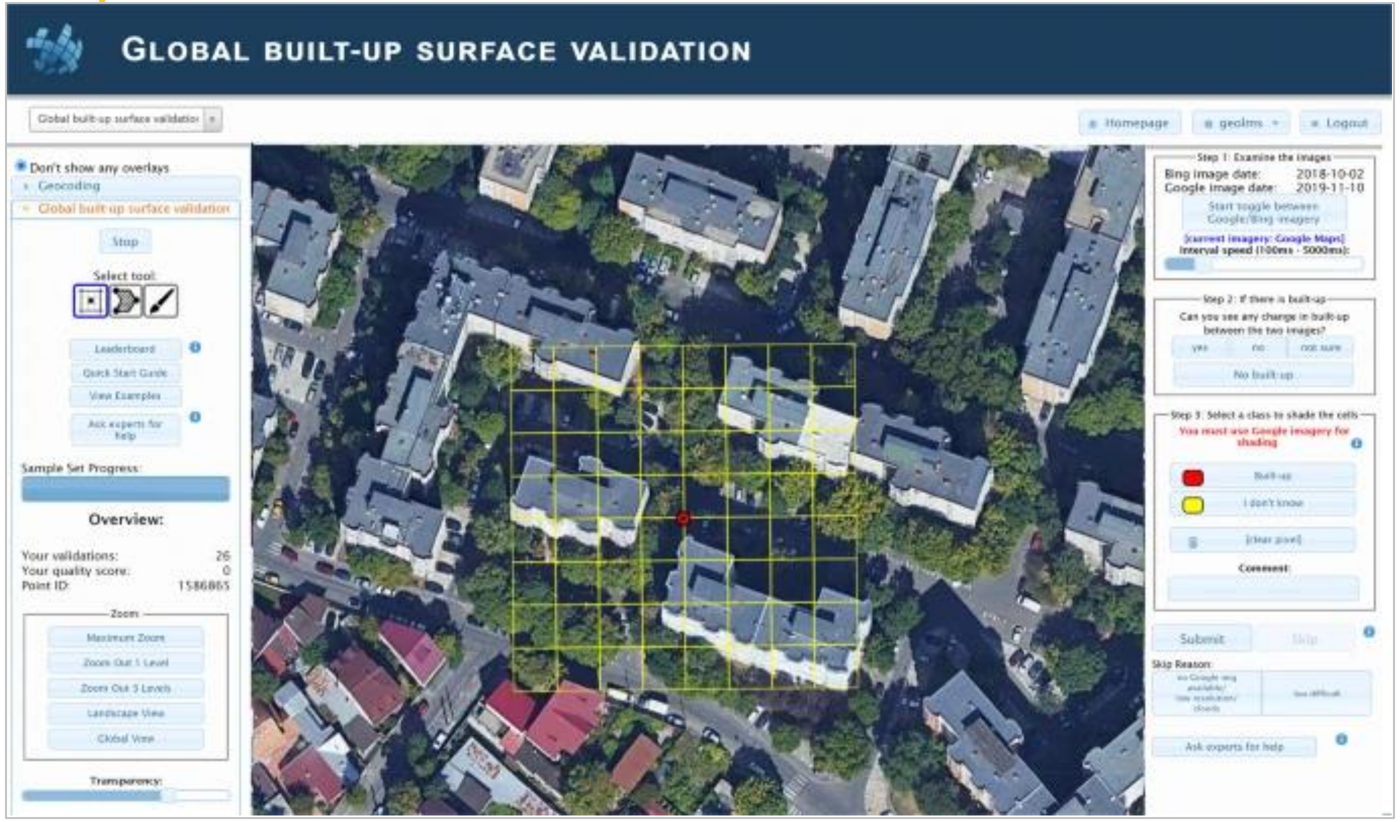
	GHSL data	Reference data			Comparative assessment
Built-up surface – Contem-porary	GHS-BUILT-S Built-up surface 10m (2018)	VHR/DL/Cadastral vector building footprint data, N=600,000,000	Visual S-2 data inspection, categorical, discrete locations N=1,000,000	Crowd-sourced reference data, continuous, N=50,000	World Settlement Footprint (WSF); ESRI Landcover v2018, v2020; Dynamic World; WorldCover — Contemporary landcover / settlement data
Built-up surface – Contem-porary	GHS-BUILT-S Built-up surface 100m (2020)	VHR/DL/Cadastral vector building footprint data, N=600,000,000			
Historical	GHS-BUILT-S Built-up surface 100m (1975-2020)	Building footprint data ***with construction year attribute***, N=300,000,000			GHSL versions 2016-2022; GAUD; WSF Evolution; GAIA; GISD; GISA — Multitemporal settlement (-related) data
Non-res.	GHS-BUILT-S-NRES Non-residential built-up surface 10m (2018)		Visual S-2 data inspection, categorical, discrecte locations N=1,000,000		
Building height	GHS-BUILT-H Avg. building height 100m (2018)	Cadastral vector building footprint data ***with building height attribute***, N=XXXXX			WSF 3D; ALOS_30m — Global building height data
Land surface	GHS-LAND Land vs water surface 10m (2018)		Visual S-2 data inspection, categorical, discrecte locations N=1,000,000		ESRI Landcover v2018, v2020; Dynamic World; WorldCover — Contemporary landcover data
Population	GHS-POP Residential population 100m (1995-2020)	Authoritative gridded population counts (24 countries)			WorldPop — Gridded population data

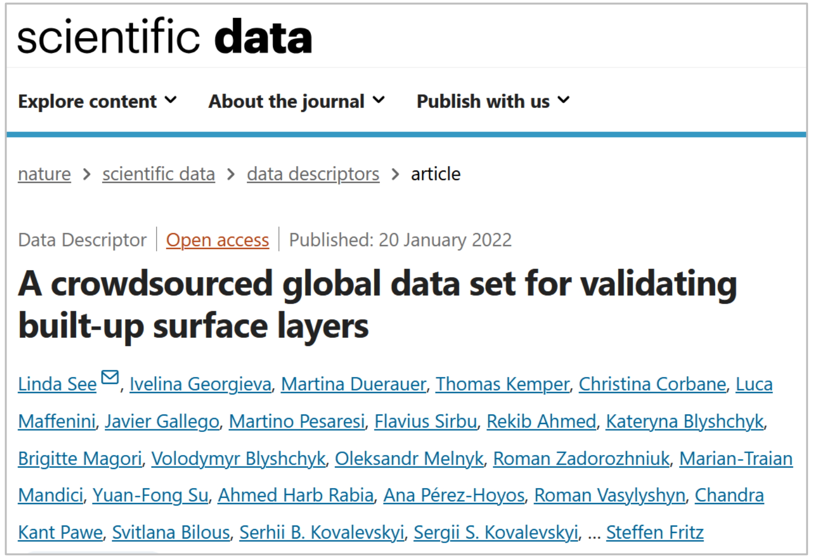
Validation

Benchmarking

European Commission

Reference data IV:
Crowd-sourced built-up surface data




scientific data

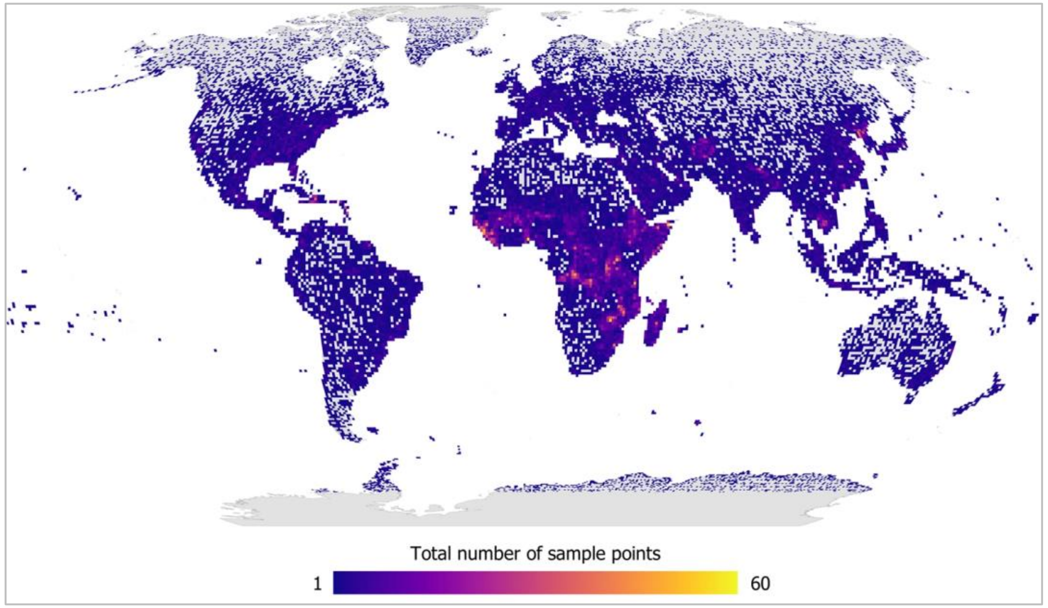
Explore content ⌄ About the journal ⌄ Publish with us ⌄

nature > scientific data > data descriptors > article

Data Descriptor | Open access | Published: 20 January 2022

A crowdsourced global data set for validating built-up surface layers

Linda See, Ivelina Georgieva, Martina Duerauer, Thomas Kemper, Christina Corbane, Luca Maffenini, Javier Gallego, Martino Pesaresi, Flavius Sirbu, Rekib Ahmed, Kateryna Blyshchyk, Brigitte Magori, Volodymyr Blyshchyk, Oleksandr Melnyk, Roman Zadorozhniuk, Marian-Traian Mandici, Yuan-Fong Su, Ahmed Harb Rabia, Ana Pérez-Hoyos, Roman Vasylyshyn, Chandra Kant Pawe, Svitlana Bilous, Serhii B. Kovalevskyi, Sergii S. Kovalevskyi, ... Steffen Fritz



- 50,000 sample locations
- Sample = 80x80m block with 64 cells
- Data allows to measure continuous built-up surface at each sample area

Source: See et al. 2020

European Commission

Reference data

Built-up surface	GHSL data	Reference data			Comparative assessment
Contem-porary	GHS-BUILT-S Built-up surface 10m (2018)	VHR/DL/Cadastral vector building footprint data, N=600,000,000	Visual S-2 data inspection, categorical, discrete locations N=1,000,000	Crowd-sourced reference data, continuous, N=50,000	World Settlement Footprint (WSF); ESRI Landcover v2018, v2020; Dynamic World; WorldCover — Contemporary landcover / settlement data
Contem-porary	GHS-BUILT-S Built-up surface 100m (2020)	VHR/DL/Cadastral vector building footprint data, N=600,000,000			
Historical	GHS-BUILT-S Built-up surface 100m (1975-2020)	Building footprint data ***with construction year attribute***, N=300,000,000			GHSL versions 2016-2022; GAUD; WSF Evolution; GAIA; GISD; GISA — Multitemporal settlement (-related) data
Non-res.	GHS-BUILT-S-NRES Non-residential built-up surface 10m (2018)		Visual S-2 data inspection, categorical, discrecte locations N=1,000,000		
Building height	GHS-BUILT-H Avg. building height 100m (2018)	Cadastral vector building footprint data ***with building height attribute***, N=XXXXX			WSF 3D; ALOS_30m — Global building height data
Land surface	GHS-LAND Land vs water surface 10m (2018)		Visual S-2 data inspection, categorical, discrecte locations N=1,000,000		ESRI Landcover v2018, v2020; Dynamic World; WorldCover — Contemporary landcover data
Population	GHS-POP Residential population 100m (1995-2020)	Authoritative gridded population counts (24 countries)			WorldPop — Gridded population data

Validation

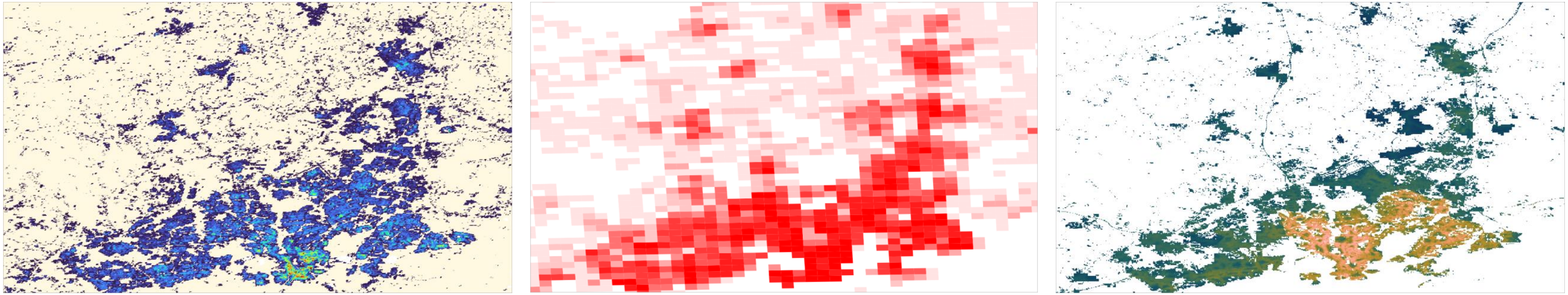
Benchmarking

Reference data V: Authoritative gridded population counts

- First systematic cross-comparison against authoritative gridded population data
- 24 countries, 18 EU countries, plus Brazil, Ecuador, Finland, Japan, Mexico, Republic of Korea
- Agreement metric: R-accuracy, TAA (Total Allocation Accuracy percentage, 1-MAE/2; Batista e Silva et al. 2020)

Accuracy assessment | Benchmarking

GHS-POP 2020 | Official population grid | Worldpop 2020, constrained



Helsinki, Finland

European Commission

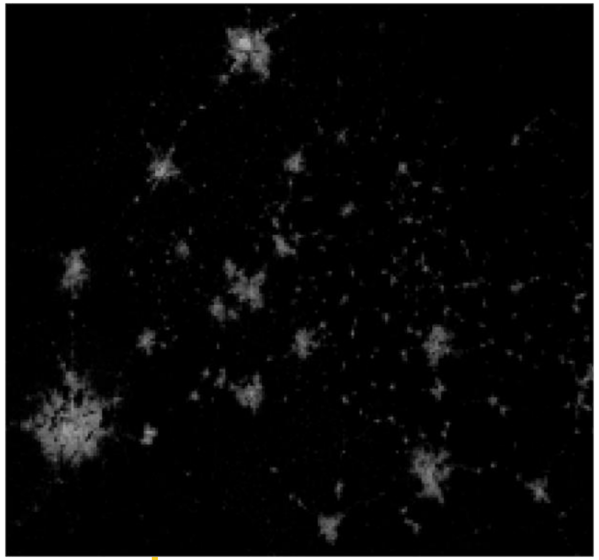
Accuracy assessment of GHS-BUILT-S / -H

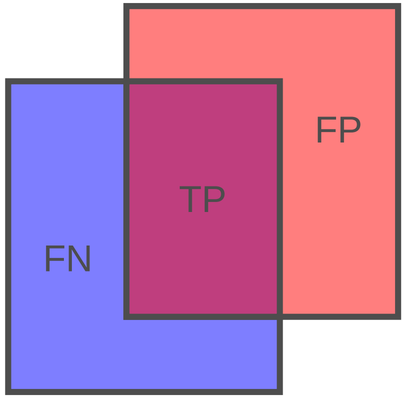
- Comparison of continuous measurements (i.e., subpixel built-up surface in sqm, average building height)

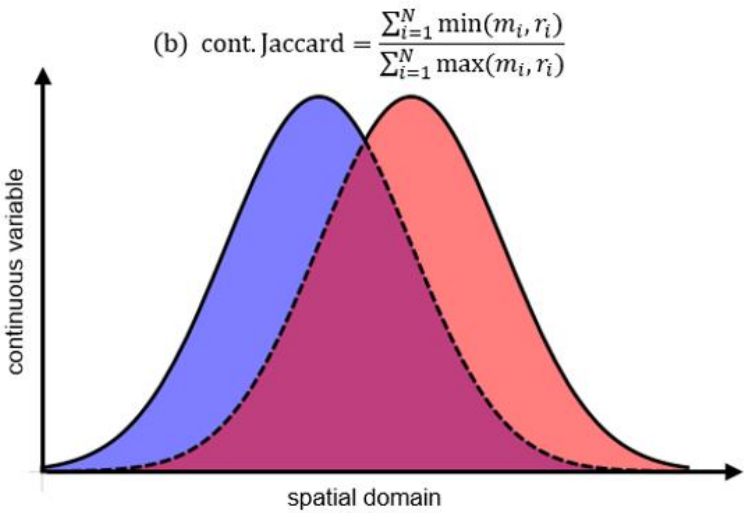
- Need for a metric to compare sparse ratio-scale measurements, without taking into account agreement in the 0-domain (not built-up)

- Mean Absolute Error (MAE) - Drawback: Scale-dependent, does not allow for disentangling commission and omission errors.

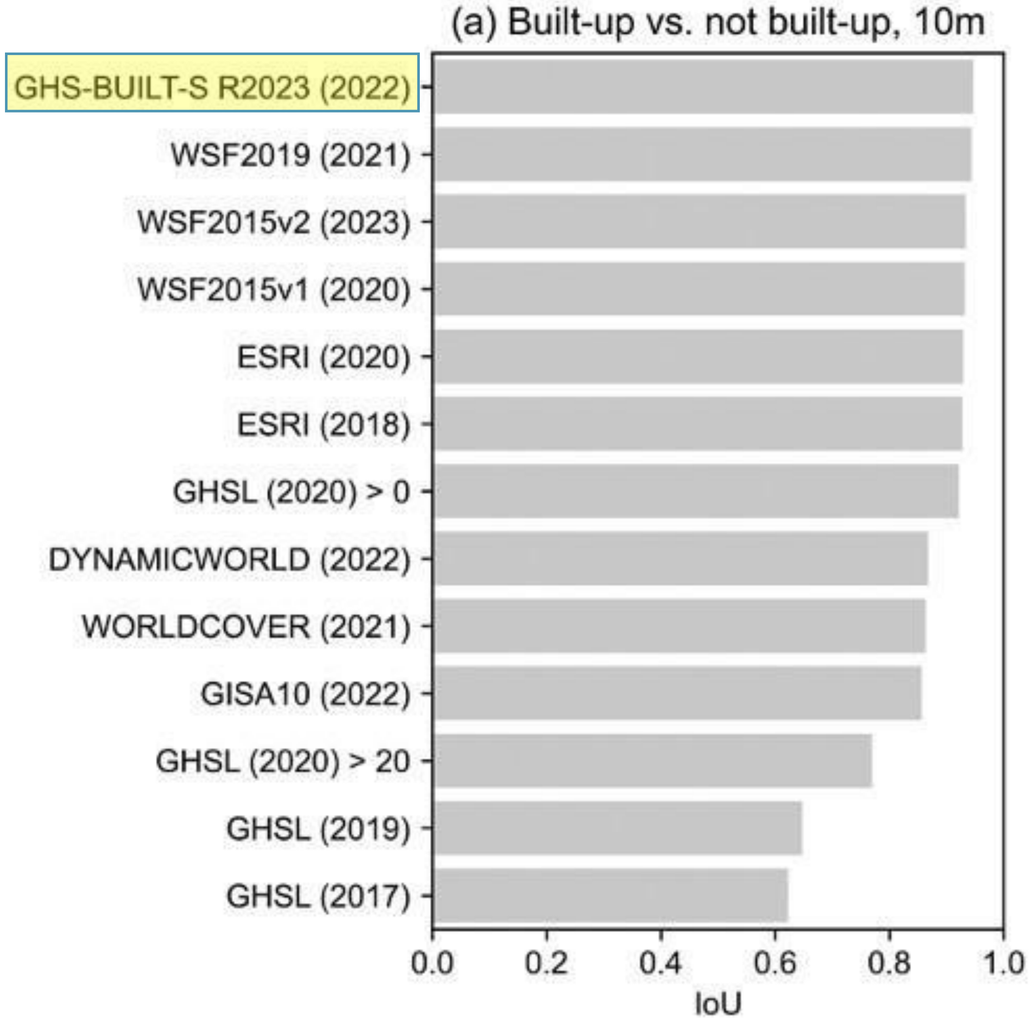
- Using the continuous Jaccard index (R-accuracy; Ružička 1958, Costa 2022, Krasnodębska et al. 2024), a generalization of the Jaccard index to the continuous domain, for grid-based evaluations.



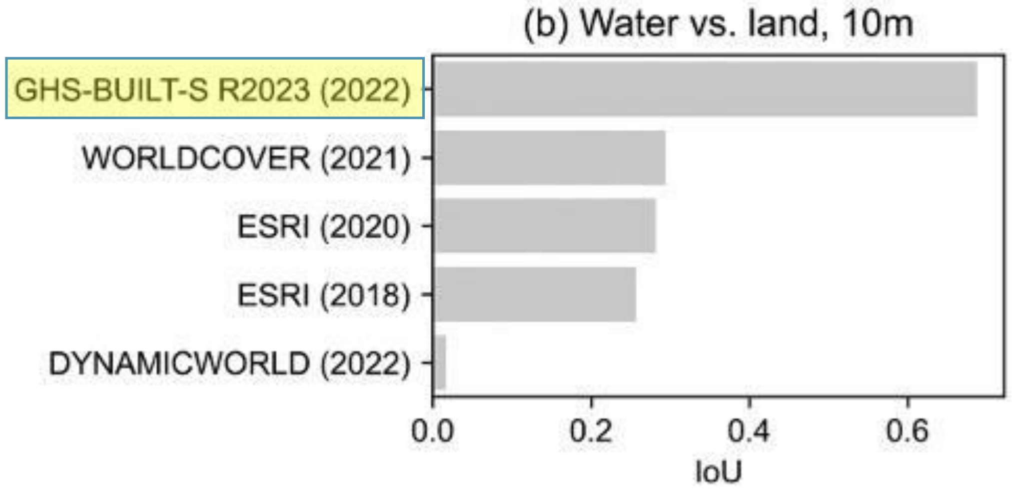




Results I: Categorical agreement at 10m resolution (reference data: visually inspected S-2 data)



Multiclass agreement (Land, Water, Residential, Non-residential):	mIoU = 0.87
built-up vs non-built-up:	IoU = 0.92
Residential vs non-residential:	IoU = 0.80
Water vs Land:	IoU = 0.98



Source: Pesaresi et al. 2024

Results II: Built-up surface accuracy 2020 at 10m resolution, compared with crowd-sourced, continuous reference data

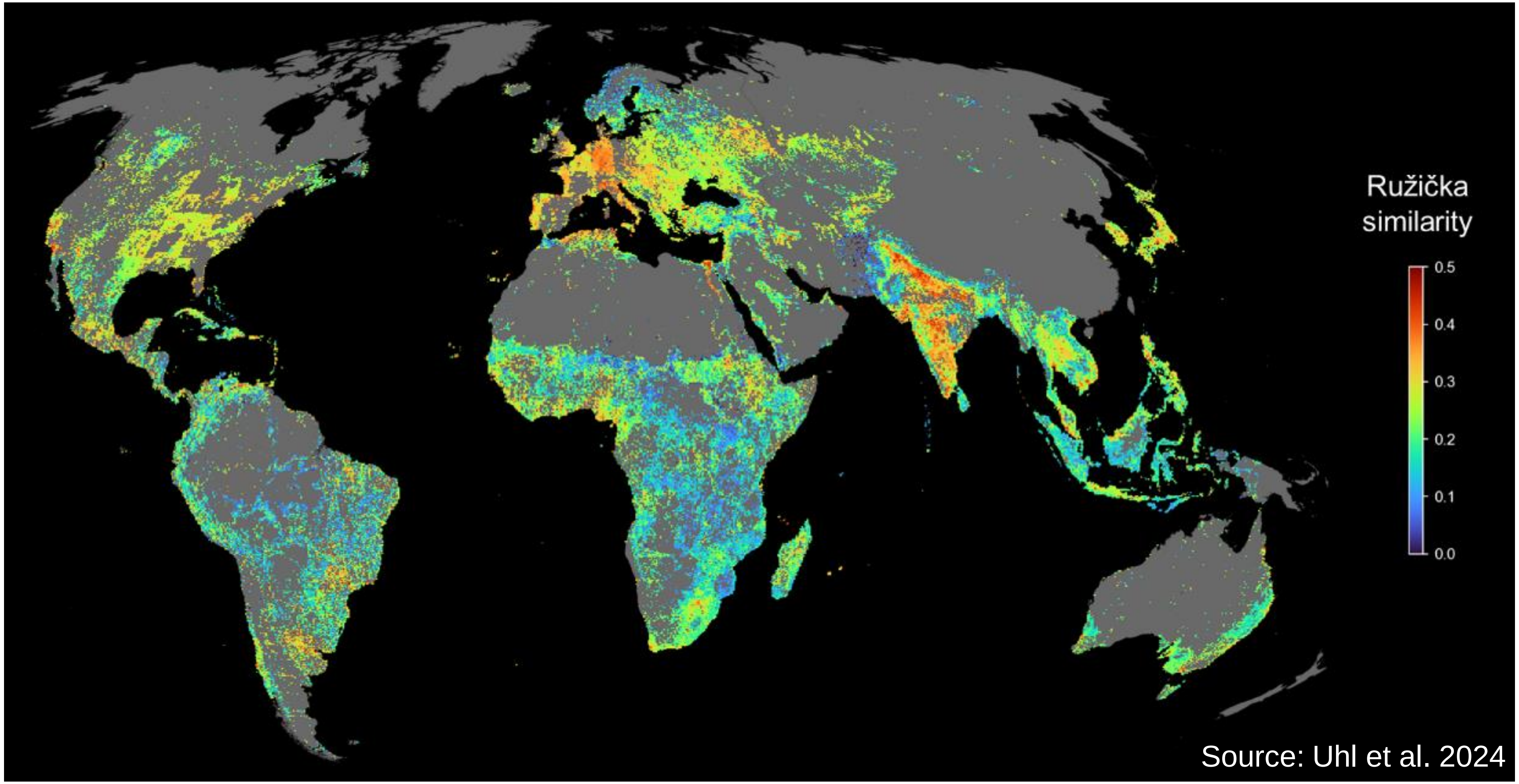
Table 11. Correlation of total built-up surface measured by global models with crowd-sourced reference data (See et al. 2022). The table is sorted by descending Spearman rank correlation coefficient.

No	Model	Spearman	No	Model	Spearman
1	GHS-BUILT-S R2023 E2018	0.8565	14	GHS-BUILT-*P* (2020)	0.7091
2	WSF2019 (2021)	0.8502	15	GHS-BUILT R2019	0.6725
3	WSF2015v2 (2023)	0.8251	16	GAUD (2020)	0.6566
4	DYNAMICWORLD (2022)	0.7989	17	GHS-BUILT (2017)	0.6427
5	ESRI (2020)	0.7811	18	GISAv2 (2022)	0.6316
6	ESRI (2018)	0.7778	19	GISAv1 (2021)	0.6229
7	WSF2015v1 (2020)	0.7574	20	GAIA (2020)	0.6128
8	WSFEVO (2021)	0.7573	21	FROM_GLC10 (2019)	0.5923
9	GHS-BUILT-*P* (2019)	0.7561	22	GHS-BUILT R2016	0.5875
10	WORLDCOVER (2021)	0.7544	23	MODIS (2010)	0.3712
11	CGLC (2020)	0.7177	24	FROM_GLC30 (2013)	0.3474
12	GISD (2022)	0.7123	25	GLOBE_LAND30 (2014)	0.3103
13	GISA10 (2022)	0.7110			

Source: Pesaresi et al. 2024

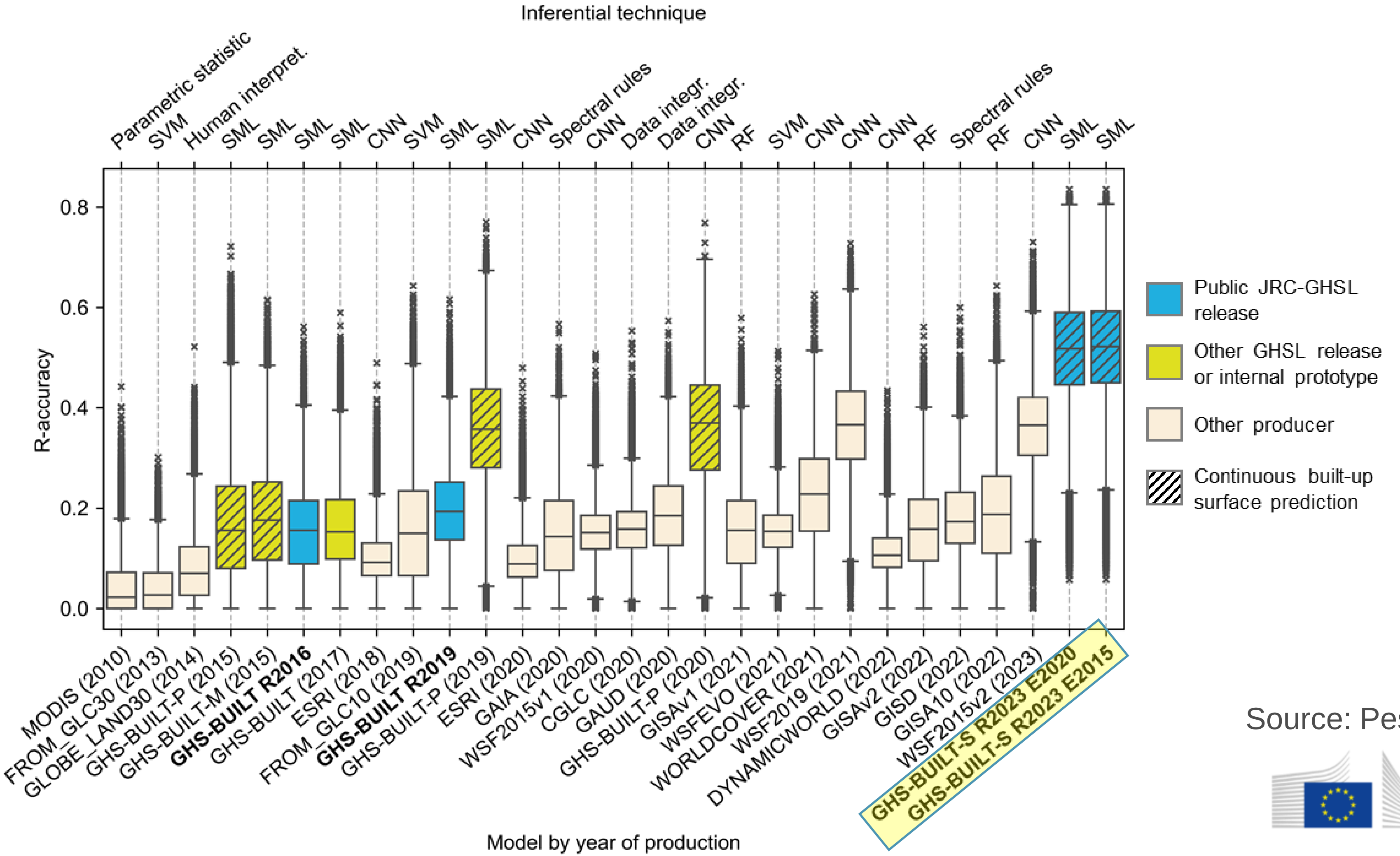
Results III: Built-up surface accuracy 2020 at 100m resolution Reference data: vector building footprint data

- R-accuracy of GHS-BUILT-S at 100m resolution (contemporary case), distributions of 25x25km tile-level estimates



Results III: Built-up surface accuracy 2020 at 100m resolution

- R-accuracy at 100m resolution (contemporary case), distributions of 25x25km tile-level estimates
- Accuracy increase over time across data products



Source: Pesaresi et al. 2024

European Commission

Results III: Built-up surface 2020 accuracy at 100m resolution stratified by region and income class

- GHSL R2023A ranks best in most cases
- Ranking is consistent across most strata (income groups, world regions)
- Income bias: lower income – lower accuracy
- MAE metric shows trends largely consistent with R-accuracy

R-accuracy (by region)

	Sub-Sah. Africa	Austr./NZ	N. America	Lat. America/Carib.	E.+S. Asia	N. Africa/W. Asia	C./S. Asia	Europe	Average
GHS-BUILT-S R2023 E2015	0.31	0.47	0.51	0.48	0.52	0.58	0.56	0.59	0.50
GHS-BUILT-S R2023 E2020	0.31	0.47	0.50	0.47	0.52	0.58	0.56	0.59	0.50
WSF2019 (2021)	0.19	0.35	0.35	0.34	0.38	0.43	0.41	0.43	0.36
GHS-BUILT-P (2020)	0.28	0.25	0.43	0.30	0.33	0.31	0.35	0.39	0.33
WSF2015v2 (2023)	0.22	0.35	0.34	0.35	0.38	0.41	0.41	0.41	0.36
GHS-BUILT-P (2019)	0.27	0.34	0.42	0.32	0.34	0.34	0.39	0.36	0.35
WORLDCOVER (2021)	0.16	0.23	0.19	0.20	0.27	0.19	0.25	0.29	0.23
GHS-BUILT R2019	0.15	0.21	0.14	0.20	0.25	0.25	0.23	0.22	0.21
GISA10 (2022)	0.12	0.14	0.16	0.18	0.21	0.19	0.28	0.24	0.19
GAUD (2020)	0.14	0.18	0.14	0.18	0.21	0.21	0.22	0.22	0.19
GISD (2022)	0.14	0.13	0.15	0.17	0.20	0.22	0.23	0.20	0.18
GHS-BUILT-M (2015)	0.13	0.14	0.19	0.18	0.17	0.18	0.16	0.22	0.17
GHS-BUILT-P (2015)	0.12	0.12	0.20	0.17	0.17	0.16	0.15	0.20	0.16
GHS-BUILT (2017)	0.14	0.15	0.13	0.18	0.20	0.21	0.20	0.16	0.17
FROM_GLC10 (2019)	0.11	0.17	0.15	0.15	0.20	0.11	0.19	0.20	0.16
CGLC (2020)	0.10	0.13	0.16	0.17	0.15	0.20	0.21	0.17	0.16
GISAv2 (2022)	0.12	0.14	0.12	0.16	0.20	0.16	0.20	0.19	0.16
WSFEVO (2021)	0.11	0.13	0.15	0.16	0.14	0.19	0.20	0.17	0.16
GISAv1 (2021)	0.12	0.14	0.12	0.15	0.20	0.15	0.19	0.19	0.16
GHS-BUILT R2016	0.09	0.16	0.13	0.15	0.17	0.17	0.16	0.20	0.16
WSF2015v1 (2020)	0.10	0.14	0.14	0.16	0.15	0.19	0.20	0.17	0.16
GAIA (2020)	0.10	0.16	0.11	0.17	0.17	0.17	0.17	0.18	0.15
DYNAMICWORLD (2022)	0.09	0.10	0.10	0.11	0.11	0.13	0.14	0.14	0.11
ESRI (2018)	0.08	0.08	0.07	0.11	0.11	0.13	0.14	0.12	0.10
ESRI (2020)	0.08	0.08	0.07	0.10	0.10	0.12	0.13	0.12	0.10
GLOBE_LAND30 (2014)	0.06	0.08	0.06	0.10	0.11	0.06	0.08	0.10	0.08
MODIS (2010)	0.03	0.03	0.05	0.05	0.02	0.05	0.05	0.04	0.04
FROM_GLC30 (2013)	0.03	0.05	0.04	0.04	0.03	0.04	0.06	0.05	0.04

MAE [sqm] (by region)

	Sub-Sah. Africa	Austr./NZ	N. America	Lat. America/Carib.	E.+S. Asia	N. Africa/W. Asia	C./S. Asia	Europe	Average
GHS-BUILT-S R2023 E2015	924	712	435	590	658	596	687	541	643
GHS-BUILT-S R2023 E2020	929	718	444	599	662	603	691	551	650
WSF2019 (2021)	1957	1252	895	1114	1179	1212	1380	1199	1274
GHS-BUILT-P (2020)	751	1715	526	1104	1221	1075	1294	787	1059
WSF2015v2 (2023)	1573	1225	762	1009	1135	1092	1298	1164	1157
GHS-BUILT-P (2019)	987	1057	303	1038	1238	948	1052	1062	961
WORLDCOVER (2021)	1294	1108	892	1300	1471	1625	1696	1232	1327
GHS-BUILT R2019	1029	1798	1462	1193	1415	1576	1582	2069	1515
GISA10 (2022)	1006	1447	1152	1152	1367	1517	1189	1453	1285
GAUD (2020)	1116	2100	1318	1225	1268	1656	1335	1925	1493
GISD (2022)	1420	2399	1583	1872	2246	2022	2122	2383	2006
GHS-BUILT-M (2015)	342	664	455	496	508	618	514	448	506
GHS-BUILT-P (2015)	147	506	286	249	303	307	265	271	292
GHS-BUILT (2017)	1134	1832	1289	1102	1254	1705	1543	2464	1540
FROM_GLC10 (2019)	810	1151	1133	1304	1168	1525	1124	1194	1176
CGLC (2020)	1628	1566	1085	1487	1903	1809	1865	1896	1655
GISAv2 (2022)	1075	2166	1508	1434	1427	1900	1608	2125	1655
WSFEVO (2021)	2774	3129	1652	1949	3436	2937	3055	3234	2771
GISAv1 (2021)	1076	2147	1449	1506	1459	1929	1632	2080	1660
GHS-BUILT R2016	1666	2153	1356	1351	1621	1851	1482	2033	1689
WSF2015v1 (2020)	2723	3051	1684	1912	3292	2870	2998	3257	2723
GAIA (2020)	996	1971	1712	1338	1137	1824	1576	1841	1549
DYNAMICWORLD (2022)	3374	3982	3171	3260	4253	3703	3999	4458	3775
ESRI (2018)	3594	4222	3987	3484	4622	4515	4405	4704	4192
ESRI (2020)	3681	4173	4026	3604	4721	4715	4531	4700	4269
GLOBE_LAND30 (2014)	2522	2696	2606	2273	2125	5662	3878	4508	3284
MODIS (2010)	1350	2065	1380	2409	1777	2828	2483	2322	2077
FROM_GLC30 (2013)	1149	1886	1229	1986	1199	1933	2147	2041	1696

R-accuracy (by income class)

	Low Income	Lower Middle	Upper Middle	High Income	Average
GHS-BUILT-S R2023 E2015	0.28	0.50	0.52	0.54	0.46
GHS-BUILT-S R2023 E2020	0.28	0.50	0.51	0.54	0.46
WSF2019 (2021)	0.18	0.36	0.37	0.39	0.32
GHS-BUILT-P (2020)	0.26	0.33	0.33	0.41	0.33
WSF2015v2 (2023)	0.21	0.36	0.37	0.38	0.33
GHS-BUILT-P (2019)	0.26	0.34	0.35	0.39	0.33
WORLDCOVER (2021)	0.15	0.25	0.22	0.24	0.21
GHS-BUILT R2019	0.12	0.22	0.22	0.18	0.19
GISA10 (2022)	0.10	0.21	0.20	0.20	0.18
GAUD (2020)	0.10	0.20	0.20	0.18	0.17
GISD (2022)	0.12	0.21	0.18	0.18	0.17
GHS-BUILT-M (2015)	0.11	0.16	0.19	0.20	0.16
GHS-BUILT-P (2015)	0.09	0.16	0.18	0.20	0.16
GHS-BUILT (2017)	0.11	0.18	0.17	0.16	0.15
FROM_GLC10 (2019)	0.08	0.17	0.16	0.17	0.15
CGLC (2020)	0.09	0.16	0.16	0.17	0.15
GISAv2 (2022)	0.08	0.18	0.17	0.15	0.15
WSFEVO (2021)	0.09	0.16	0.16	0.16	0.14
GISAv1 (2021)	0.08	0.18	0.16	0.15	0.14
GHS-BUILT R2016	0.07	0.15	0.17	0.16	0.14
WSF2015v1 (2020)	0.09	0.16	0.16	0.16	0.14
GAIA (2020)	0.07	0.15	0.17	0.14	0.13
DYNAMICWORLD (2022)	0.08	0.12	0.12	0.12	0.11
ESRI (2018)	0.07	0.11	0.11	0.10	0.10
ESRI (2020)	0.06	0.11	0.10	0.10	0.09
GLOBE_LAND30 (2014)	0.04	0.08	0.10	0.08	0.08
MODIS (2010)	0.02	0.04	0.04	0.05	0.04
FROM_GLC30 (2013)	0.02	0.04	0.04	0.04	0.04

MAE [sqm] (by income class)

	Low Income	Lower Middle	Upper Middle	High Income	Average
GHS-BUILT-S R2023 E2015	811	698	631	491	658
GHS-BUILT-S R2023 E2020	816	702	639	500	664
WSF2019 (2021)	1583	1421	1271	1021	1324
GHS-BUILT-P (2020)	499	1103	1067	707	844
WSF2015v2 (2023)	1216	1296	1169	929	1153
GHS-BUILT-P (2019)	701	1105	1019	702	882
WORLDCOVER (2021)	1033	1509	1380	1088	1253
GHS-BUILT R2019	740	1481	1566	1678	1366
GISA10 (2022)	840	1256	1293	1321	1178
GAUD (2020)	660	1340	1508	1597	1276
GISD (2022)	1071	2120	2178	1864	1808
GHS-BUILT-M (2015)	266	465	469	517	429
GHS-BUILT-P (2015)	119	245	223	341	232
GHS-BUILT (2017)	799	1554	1789	1607	1437
FROM_GLC10 (2019)	682	1096	1274	1202	1063
CGLC (2020)	1388	1845	1834	1363	1608
GISAv2 (2022)	753	1514	1724	1761	1438
WSFEVO (2021)	2203	3248	2810	2230	2623
GISAv1 (2021)	800	1525	1747	1715	1447
GHS-BUILT R2016	1249	1719	1629	1668	1566
WSF2015v1 (2020)	2173	3167	2782	2244	2592
GAIA (2020)	573	1325	1633	1745	1319
DYNAMICWORLD (2022)	2683	4053	4001	3559	3574
ESRI (2018)	2934	4437	4385	4109	3966
ESRI (2020)	3074	4546	4451	4143	4053
GLOBE_LAND30 (2014)	2710	3355	3488	3233	3196
MODIS (2010)	1001	2016	2457	1791	1816
FROM_GLC30 (2013)	825	1629	2036	1519	1502

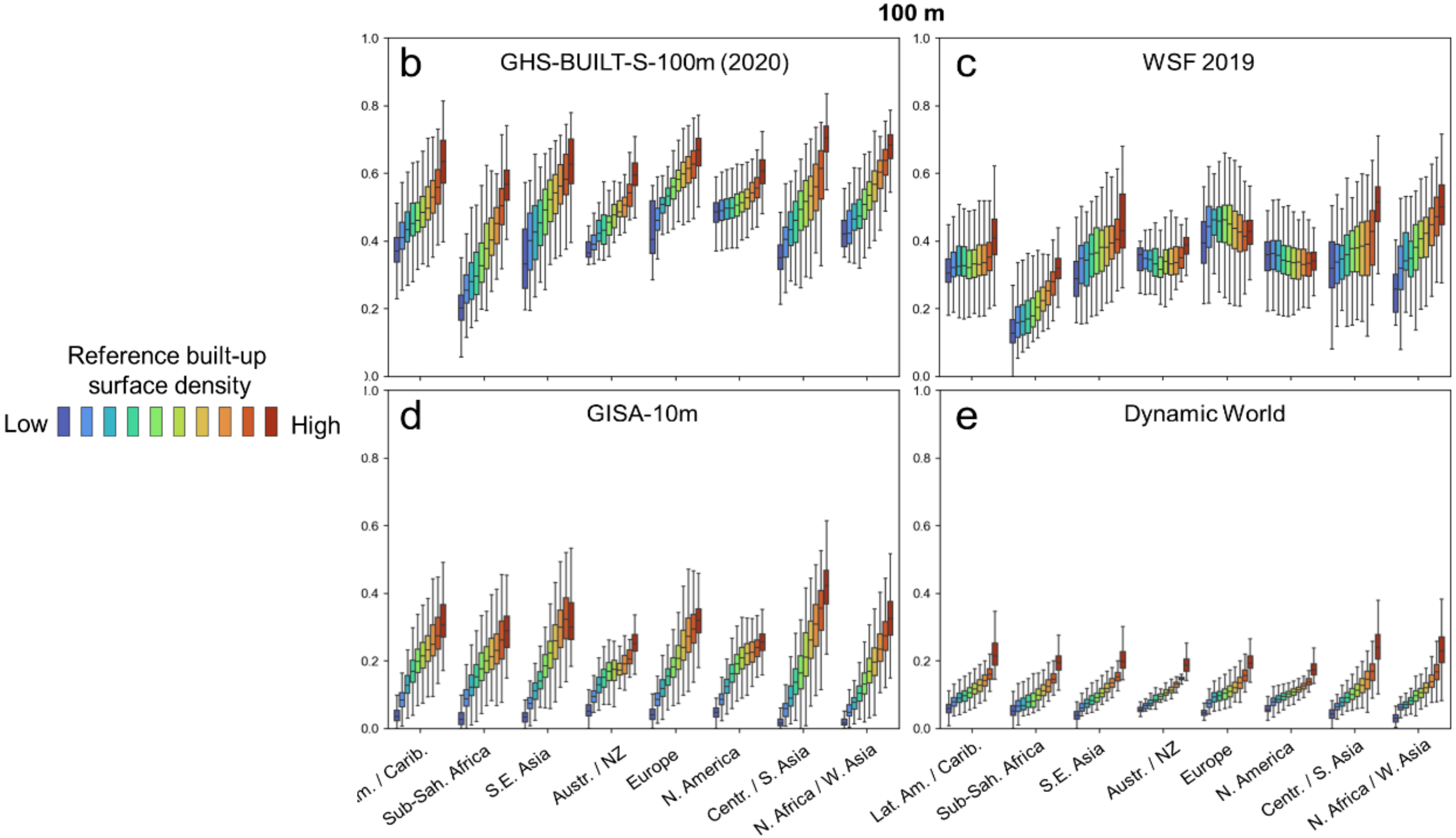
Accuracy

Low High

Source: Pesaresi et al. 2024

Results III – Built-up surface 2020 accuracy at 100m resolution along a rural-urban gradient

- Accuracy increase from rural to urban settings, consistent across world regions, and for most other datasets.



Source: Uhl et al. 2024

Results IV: Accuracy of multi-temporal built-up surface estimates (1975-2020)

- For continuous built-up surface estimation at 100m resolution, GHS-BUILT-S ranks highest compared to existing multitemporal, global settlement-related datasets

- Accuracy decreases towards early epochs.

R-accuracy

	1975	1980	1985	1990	1995	2000	2005	2010	2015	2020	Avg.
GHS-BUILT R2023	0.25	0.31	0.31	0.35	0.34	0.40	0.39	0.44	0.43	0.43	0.36
GHS-BUILT R2022	0.21	0.25	0.25	0.27	0.28	0.31	0.35	0.40	0.42	0.39	0.31
GISD (2022)	N/A	N/A	0.14	0.15	0.15	0.16	0.16	0.17	0.17	0.17	0.16
GAUD (2020)	N/A	N/A	0.13	0.14	0.14	0.15	0.16	0.17	0.16	N/A	0.15
GHS-BUILT R2019	0.11	N/A	N/A	0.14	N/A	0.15	N/A	N/A	0.16	N/A	0.14
WSFEVO (2021)	N/A	N/A	0.13	0.14	0.13	0.14	0.13	0.15	0.14	N/A	0.14
GAIA (2020)	N/A	N/A	0.12	0.13	0.12	0.13	0.13	0.14	0.13	0.12	0.13
GHS-BUILT R2016	0.10	N/A	N/A	0.13	N/A	0.14	N/A	N/A	0.15	N/A	0.13
GISAv1 (2021)	0.07	0.09	0.12	0.13	0.12	0.13	0.13	0.14	0.14	0.14	0.12
GISAv2 (2022)	0.07	0.09	0.10	0.13	0.12	0.13	0.13	0.14	0.14	0.14	0.12

MAE [sqm]

	1975	1980	1985	1990	1995	2000	2005	2010	2015	2020	Avg.
GHS-BUILT R2023	64	62	66	64	71	71	76	74	79	81	71
GHS-BUILT R2022	63	61	63	62	65	67	68	69	74	88	68
GISD (2022)	N/A	N/A	224	233	249	266	298	321	334	351	284
GAUD (2020)	N/A	N/A	212	222	243	264	284	297	314	N/A	262
GHS-BUILT R2019	217	N/A	N/A	261	N/A	318	N/A	N/A	359	N/A	289
WSFEVO (2021)	N/A	N/A	215	243	275	302	332	348	373	N/A	298
GAIA (2020)	N/A	N/A	176	218	257	282	308	325	369	392	291
GHS-BUILT R2016	186	N/A	N/A	275	N/A	310	N/A	N/A	352	N/A	281
GISAv1 (2021)	142	163	205	218	242	262	284	294	306	311	243
GISAv2 (2022)	140	160	196	223	246	266	291	304	317	333	248

Accuracy
Low High

N = 7754 tiles of 25 km × 25 km

Source: Pesaresi et al. 2024

Results V: Accuracy of gridded population estimates, 100m resolution, 1995-2020

Table 16. R-accuracy and %TAA (in brackets) for each population grid (WPC: WorldPop UN adjusted constrained; WPU: WorldPop UN adjusted unconstrained) for each reference year available (census grids). Aggregated and RMSE columns refer to data for all available epochs aggregated.

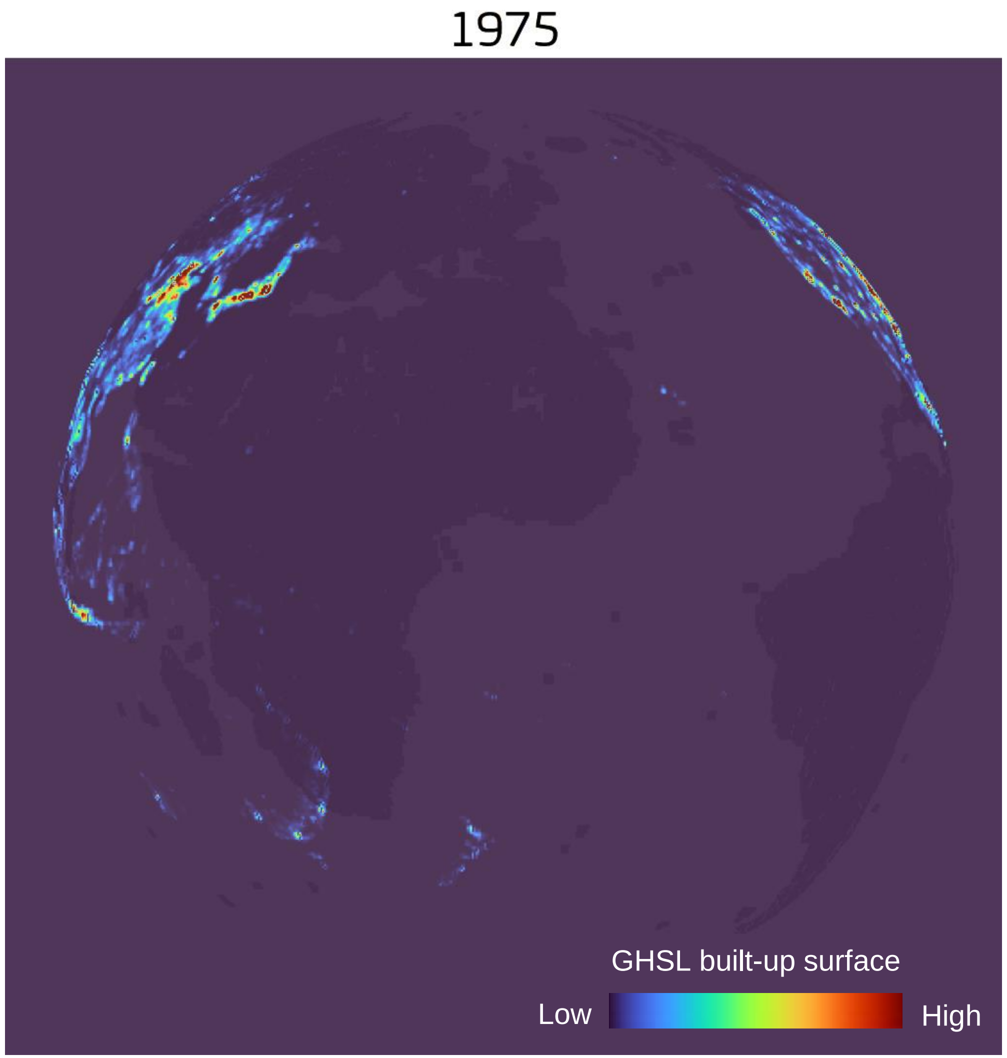
	1995	2000	2005	2010	2015	2020	Aggregated	RMSE
GHS-POP R2023	0.71 (83%)	0.69 (82%)	0.73 (85%)	0.73 (84%)	0.74 (85%)	0.68 (81%)	0.71 (83%)	45.4
GHS-POP R2022	0.69 (82%)	0.65 (79%)	0.70 (83%)	0.69 (82%)	0.69 (82%)	0.61 (76%)	0.66 (80%)	55.4
GHS-POP R2019	-	0.61 (76%)	-	-	0.64 (78%)	-	0.62 (77%)	590.9
WorldPop (constrained)	-	-	-	-	-	0.59 (74%)	0.59 (74%)	205.6
WorldPop (unconstrained)	-	0.53 (70%)	0.58 (73%)	0.57 (72%)	0.61 (76%)	0.55 (71%)	0.56 (72%)	62.7

- GHS-POP achieves higher agreement than WorldPop when compared to authoritative gridded population estimates across 24 countries, in all evaluated epochs (1995-2020).
- Accuracy estimates stable over time.

Source: Pesaresi et al. 2024

Conclusions

- We gathered large amounts of independent reference data that were not available / not used during production of the GHSL R2023A.
- We carried out an accuracy assessment of the measured and modelled data components of the GHSL R2023A data ecosystem, and cross-compared to related, global remote-sensing based datasets.
- Empirical findings suggest that multi-temporal GHSL R2023A built-up surface and resident population estimates are the most accurate global data sources **for continuous variable estimation enumerated at 100m resolution.**
- E.g., accurate SDG indicator 11.3.1 estimation (land use efficiency).
- Our analyses aim to contribute to a more informed, uncertainty-aware usage of global settlement data and beyond.



European Commission

References

- Batista e Silva, Filipe, Sérgio Freire, Marcello Schiavina, Konštantín Rosina, Mario Alberto Marín-Herrera, Lukasz Ziemba, Massimo Craglia, Eric Koomen, and Carlo Lavalle. "Uncovering temporal changes in Europe's population density patterns using a data fusion approach." Nature communications 11, no. 1 (2020): 4631.
- Costa, L. da F., 2022. On similarity. Physica A: Statistical Mechanics and its Applications 599, 127456. https://doi.org/10.1016/j.physa.2022.127456
- EUBUCCO building dataset: https://eubucco.com/
- Google Open Buildings dataset: https://sites.research.google/open-buildings/
- Krasnodębska, Katarzyna, Wojciech Goch, Johannes H. Uhl, Judith A. Verstegen, and Martino Pesaresi. 2024. "Advancing Precision, Recall, F-Score, and Jaccard Index: An Approach for Continuous Gridded Data." Recall, F-Score, and Jaccard Index: An Approach for Continuous Gridded Data, https://doi.org/10.2139/ssrn.4865121
- Microsoft building footprint data: https://github.com/microsoft/GlobalMLBuildingFootprints
- MTBF-33 building dataset: https://data.mendeley.com/datasets/w33vbvjtdy/1, https://doi.org/10.1016/j.dib.2022.108369
- Pesaresi, M., Schiavina, M., Politis, P., Freire, S., Krasnodębska, K., Uhl, J.H., Carioli, A., Corbane, C., Dijkstra, L., Florio, P. and Friedrich, H.K., 2024. Advances on the global human settlement layer by joint assessment of earth observation and population survey data. International Journal of Digital Earth, 17(1), p.2390454.
- Ružička, M. 1958. Anwendung Mathematisch-Statistischer Methoden in Der Geobotanik (Synthetische Bearbeitung von Aufnahmen). Biologia, Bratislava 13:647–661.
- See, Linda, Ivelina Georgieva, Martina Duerauer, Thomas Kemper, Christina Corbane, Luca Maffenini, Javier Gallego et al. "A crowdsourced global data set for validating built-up surface layers." Scientific data 9, no. 1 (2022): 13.
- Uhl, J.H., Pesaresi, M., Politis, P., Goch, K., Melchiorri, M. and Kemper, T., 2024, July. Towards a Quasi-Global Accuracy Assessment of Built-Up Surface Estimates Derived From Sentinel-2 Multispectral Data. In IGARSS 2024-2024 IEEE International Geoscience and Remote Sensing Symposium (pp. 4697-4700). IEEE.
- Uhl, J.H., and S. Leyk. "MTBF-33: A multi-temporal building footprint dataset for 33 counties in the United States (1900–2015)." Data in Brief 43 (2022): 108369.

Thank you for your attention

Global Human Settlement Layer
https://human-settlement.emergency.copernicus.eu/
jrc-ghsl@ec.europa.eu

EU Science Hub
Joint-research-centre.ec.europa.eu

European Commission