ESSD-2023-53: Final response to Reviewers and Commentors

Dear Reviewers and Commentors, we would like to thank you for your time and effort to evaluate our manuscript and dataset HISDAC-ES. We have addressed your comments and suggestions, reflected in the revised version of the manuscript. Your suggestions regarding corrections and changes in the data are very welcome, and we were able to implement most of them. Here, we provide the revised manuscript and also updated the dataset on Figshare, and the code on GitHub. We thank you for your patience and your constructive feedback.

Please find below the responses to your valuable comments and how we address them in the revised manuscript.

REVIEWER 1

HISDAC-ES is a valuable dataset for the study of a variety of dynamic processes in the development of the built environment in Spain. The authors have transformed information from several cadastral datasets into a comprehensive dataset that is far easier to use and directly represents variables likely to be of interest to researchers. The long temporal extent and complete coverage of all of Spain, including urban and rural areas are particularly valuable. The authors have done a laudable job of validating their data to the extent possible, given the dearth of comparable data sources.

Response: Thank you for the positive assessment and for your valuable comments. We appreciate the time and effort invested in reviewing this manuscript.

I have just a few relatively minor suggestions and questions on the manuscript:

R1-1: For readers unfamiliar with the intricacies of Spanish geography, it would be helpful to include a brief background section describing the unique features. This section should highlight the Basque country and Navarra, noting their locations and why they are unique. It should also mention the islands and exclaves that are part of Spain's territory. This section would serve to orient readers when these areas are mentioned later in the manuscript.

Response: Good suggestion. In the revised version, we inserted an extended paragraph introducing the geography, political, historical, and settlement-related characteristics of Spain (lines 104-120):

> Spain is one of the two countries that make up the Iberian Peninsula. It has an area of 506,000 km2. In addition to the peninsular territory, it has two archipelagos, the Canary Islands in the Atlantic Ocean and the Balearic Islands in the Mediterranean Sea, and two exclaves in North Africa, the autonomous cities of Ceuta and Melilla. It is a decentralized state with autonomous communities, seventeen in total, and the aforementioned autonomous cities. The autonomous communities have a high degree of self-government, and several of them are classified as "historic" due to their differential identity associated with their own language. This is the case with Catalonia, Galicia, Valencia, the Balearic Islands, the Basque Country, and Navarre. The latter two, located in the Northern coast of the Iberian Peninsula, also have their own economic agreement and a different fiscal and tax collection system from the rest of the territories. This is the reason why their cadastral data differs from the rest of the country. The administrative organization has four levels: the national level, the autonomous communities (with powers in territorial planning, education, healthcare, primary sector, industry, commerce, and tourism), the provinces (50 in total with limited competencies, mainly coordination and assistance to small municipalities), and the municipalities (8,125), which have powers in urban planning and local services. Spain has had two distinct settlement systems, increasingly diluted, associated with its historical and climatic evolution. In the northwest and the Cantabrian area (Galicia, Asturias, Cantabria, Basque Country, northern Navarre, and part of Catalonia), there has been a traditional dispersal of the rural population in isolated houses and/or small settlements associated with an Atlantic climate, with intensive agriculture and livestock favoured by the presence of abundant water. In contrast, the rest of the territory, with a Mediterranean climate, has experienced concentrated settlements associated with cereal crops, vineyards, and olive groves, as well as extensive livestock farming.

R1-2: Section 2.1.2 mentions in that a "common building function classification scheme" was applied. More details about this scheme would be helpful. What building function categories were included in the source datasets, and how were they harmonized into the common scheme? This could be addressed by a table in an appendix.

Response: Thank you for pointing this out. We added more details on the classification scheme (lines 159-165), and point the reader to a table in our HISDAC-ES GitHub repository, where this mapping is applied:

> the different data sources. Thus, we harmonized the data by renaming columns, and by applying a common building function classification scheme, including the six building function classes "residential", "commercial", "industrial", "agricultural", "public services", and "office". "Public services" is probably the broadest of these categories, including governmental buildings, but also health-related buildings and cultural institutions (e.g., churches or museums). Specifically, building function ontologies differed slightly for the data from the region of Navarra, the province of Alava, and were consistent across the other regions / provinces. For example, commercial buildings in Navarra are called "trade" instead of "commercial". The applied mapping scheme can be accessed on the HISDAC-ES GitHub repository[8].

R1-3: Section 4.3 on the long-term trajectory evaluation points out that correlations are highest in teh Southern region. Another striking feature of Fig. 14 is that the correlations in Madrid peak earlier in the time series than for other regions. Is this possibly also attributable to survivorship bias and more and earlier redevelopment around Madrid?

Response: Good observation. We agree and added this observation to our interpretation of Fig. 14 (lines 438-441):

> of building stock renewal and thus, a weaker effect of the survivorship bias in the HISDAC-ES data. Interestingly, correlations reach an early peak for the Madrid region (1960's for BUFA, 1930's for DEVA) and then drop. Such a decreasing agreement towards recent epochs could be related to heavy (peri-)urban renewal in the Madrid region, which would be less well captured in HISDAC-ES (cf. Section 4.5). Besides this comparison of quantitative measures per grid cell, we also compared the

R1-4: In Fig. 18, discussed in section 4.6, the completeness of the number of dwellings attribute is notably lower than other attributes nationwide. Should users be concerned about this gap?

Response: Good catch. The "number of dwellings" attribute is, per definition, only available for residential buildings. For the statistics in Fig. 18, the completeness is calculated across all buildings (including those of non-residential use), yielding lower numbers. We re-calculated the completeness for residential buildings only and modified Fig. 18 accordingly:

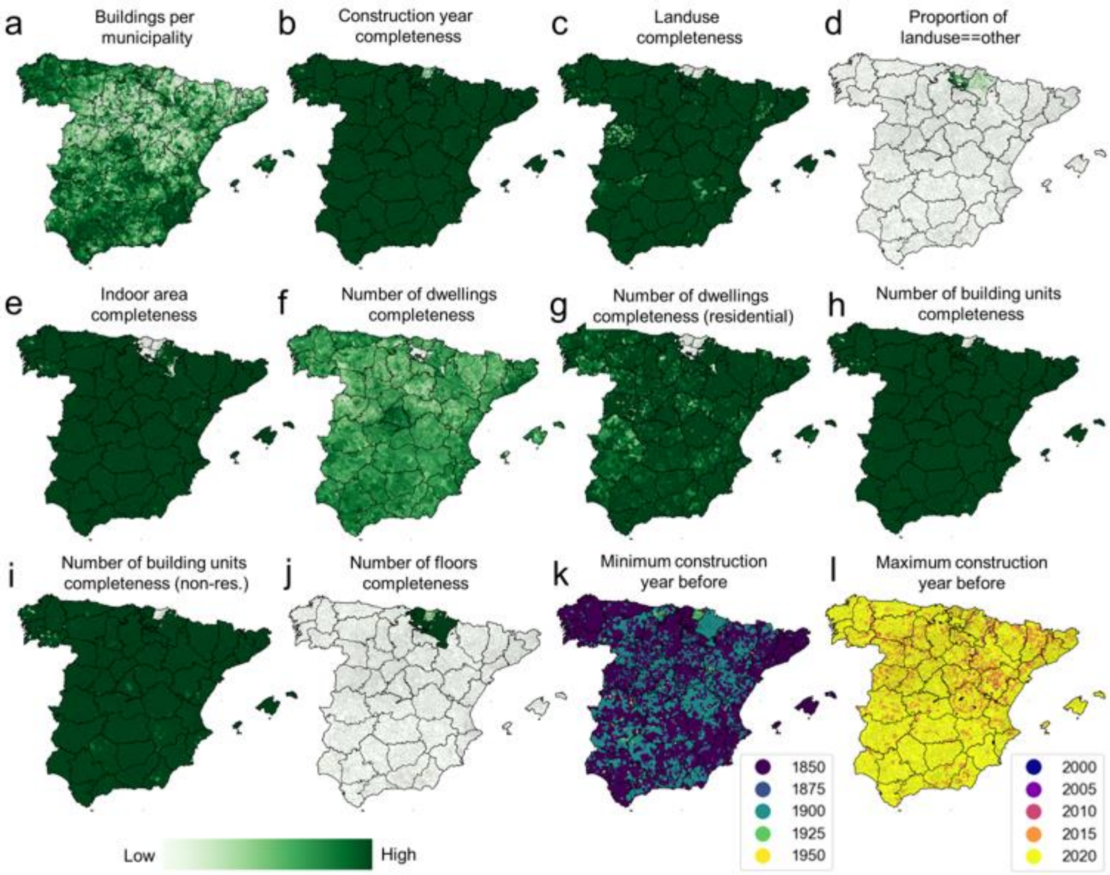


Figure 18: Attribute completeness and construction year coverage at the municipality level. See Appendix Fig. G1 for corresponding maps of the Canary Islands.

R1-5: In the video supplement, animations 1-3 use a different set of municipalities than animations 4-6. Why is this?

Response: Good catch. As can be seen in Fig. 18, land use information is not available in some regions of the Basque Country, including the (major) cities of San Sebastián and Bilbao. For this reason, land use evolution cannot be shown for these two cities, and thus, we replaced these cities in the animations depicting *land use* evolution (animations 4-6) by the cities of Jaén and Cadiz. We did want to include San Sebastián and Bilbao in some of the animations, that is why the set of cities is different. In the revised version, we clarified this in lines 626-631:

> Animations 1-3 are shown from 1920 to 2020. Animations 4-6 are shown from 1950 to 2020. Note that for the land-use related animations (4-6), we binarized the building density layers stratified by land use category (i.e., highlighting grid cells where at least one building of the respective land use class exists). Major cities where no land use information was available (i.e., San Sebastián and Bilbao, shown in animations 1-3) were replaced by the cities of Cadiz and Jaén in animations 4-6. Animations 7 and 8 show the municipality-level aggregates, converted into densities (i.e., built-up area per km², buildings per km²), shown in percentiles based on the data distributions across all years.

R1-6: See also the attached manuscript with minor edits for clarification in the text.

Response: Thank you very much for these detailed comments. We corrected all minor comments, and also changed the naming of the “temporal” characteristics to “age-related characteristics” throughout. We uploaded the correspondingly renamed TIF files to the data repository.

In spot checking the many available data layers, I discovered a few issues:

R1-7: In the HISDAC-ES_All_LAEA subset, the phys_dwel_sum_v1_100 and phys_dwel_mean layers.

Response: Indeed, we noticed an issue with the physical characteristics layers, in both the LAEA and UTM version of the data. It is related to NaNs in the underlying variables. We reprocessed the data and updated the data in the repository.

R1-8: Also in the HISDAC-ES_All_LAEA subset, the phys_bufa_mean_v1_100 appears to have many more 0 cells than expected and to be generally inconsistent with the evol_bufa_v1_100_2020 layer. (See overlay of these two layers at the end of the attached PDF file.)

Response: This issue is related to what we describe in R1-7 – in the new version of the data this is corrected.

REVIEWER 2

General Comments

This is a very well written paper outlining a very interesting high resolution data set on built-up areas in Spain going backwards in time to 1990 that is also rich in detail, i.e., several variables that correspond to four components related to the state and evolution of the built environment. The introduction is well written and make a clear case for the need for such a data set. The authors have undertaken a considerable evaluation process of the data set using many different sources and acknowledge the limitations, in particular, the survivorship bias. The data are readily available with a doi and are well documented, so it was easy to download and view them. The animated gifs are a nice addition. Overall, this is a really valuable data set with many different potential applications, some of the which the authors refer to in the paper. Having such a data set for all of Europe would be amazing.

Response: Thank you for your time and your valuable comments, and for this detailed assessment.

Specific Comments

R2-1: Line 136, which attributes were retained with the centroids?

Response: As a first step, we calculated the building footprint area (based on polygonal geometries reprojected into LAEA - EPSG:3035) and then converted into centroids, retaining all relevant attributes for subsequent calculations. In the revised version, we clarified this in lines 155-159:

> After downloading and gathering the building data for 8,131 municipalities, covering all regions of Spain, we first calculated and attached the building footprint area (obtained after reprojecting to EPSG:3035) and converted the polygonal building footprint data to centroids, retaining all relevant attributes, to reduce the computational effort for the subsequent data processing. Despite the common INSPIRE framework, attribute names and building function classes differed slightly between

R2-2: Lines 138/139, what was the common building function classification scheme used? Or is this what you refer to later, i.e., residential, commercial, industrial, agricultural, public services, offices?) Where would a building like a church or museum fall?

Response: Good point. We did some manual checks in the data, and, as we expected, museums and churches are labelled as “public services”. We added some more detail on this in lines 160-165:

> classification scheme, including the six building function classes "residential", "commercial", "industrial", "agricultural", "public services", and "office". "Public services" is probably the broadest of these categories, including governmental buildings, but also health-related buildings and cultural institutions (e.g., churches or museums). Specifically, building function ontologies differed slightly for the data from the region of Navarra, the province of Alava, and were consistent across the other regions / provinces. For example, commercial buildings in Navarra are called "trade" instead of "commercial". The applied mapping scheme can be accessed on the HISDAC-ES GitHub repository[8].

R2-3: Line 159, spatial aggregation into 100m grid – does this match the CORINE 100m grid?

Response: Indeed, our 100m LAEA grid aligns with the EEA EUROSTAT grid, which also aligns with the CLC grid. In the revised version, we added this information in line 170:

> ~~We decided to provide~~The gridded surfaces of HISDAC-ES are provided in three different spatial reference systems: (a) ETRS89 UTM Zone 30N (EPSG:25830) for the Spanish mainland, Balearic Islands, as well as the exclaves Ceuta and Melilla located in Northern Africa; (b) REGCAN-95 (EPSG:4083) for the Canary Islands; and (c) ~~for all Spanish territory~~ in the reference grid of the European Environmental Agency (EEA), which is based on the ETRS89 Lambert Azimuthal Equal Area projection (LAEA; EPSG:3035) and is commonly used for pan-European statistical mapping (e.g., Corine Land Cover data), for all Spanish territory. This way, users can refer to the data in UTM / REGCAN for mapping purposes (North-oriented,

R2-4: Line 190, Why did you not compare with the Copernicus Urban Atlas product? It would also have been interesting to use the Copernicus soil sealing product as an additional evaluation even if this is only possible for more recent years.

Response: Our main reason for not using the Copernicus urban atlas (CUA) product is that it only covers around 700 "functional urban areas" (in the 2018 version) and only around 300 urban areas in the 2006 version. Thus, rural areas are not contained in that data product. For the evaluation of HISDAC-ES, it was important to understand the quality of the data across the rural-urban gradient. This is particularly crucial since we believe that HISDAC-ES will be used for many applications in the rural domain (i.e., in tows, villages, sparse rural settlements) not only for urban areas.
Moreover, several LULC classes contained in CUA seem to be contained in the CLC classification scheme as well, and CLC also covers the rural domain. For these reasons and not to further increase the complexity and length of our validation analysis, we chose CLC over CUA. Furthermore, we used GHSL rather than CUA because of its more extended temporal coverage (1975-2014, as opposed to 2006-2018). However, we will certainly use CUA in future work, as it provides a valuable, urban-focused data product.

R2-5: Line 215, you mention that you compared GHS-BUILT with the World Settlement Footprint and there was good agreement but a full evaluation with the latter product would have been useful because it performs better in rural areas than GHS-BUILT, which is what you highlight in your results section. Hence an agreement in urban probably doesn't reflect this better performance in rural areas.

Response: We agree that the WSF data products are very valuable. However, as HISDAC-ES is derived from recent cadastral data, we expect very high accuracy for the contemporary (2020) epoch. Thus, a comparison with WSF v2019 does not seem necessary given the amount of evaluation analysis we already conducted. The main challenge is to evaluate the accuracy of HISDAC-ES over time, as the implicit assumptions and strategies (i.e., measuring historical development through the lens of contemporary building stock age information) cause a potential survivorship bias that needs to be quantified.
Thus, assessing the accuracy of the data over time was our priority. While WSF Evolution is a highly relevant data product, in particular because of the annual temporal resolution and its

robustness due to the use of the full Landsat archive, we chose the GHSL over WSF Evolution, for the following reasons:

- GHSL has longer temporal extent than WSF evolution – the farther we can "look back" in time, the better.
- The GHS-SMOD settlement model allows for a seamless rural-urban stratification of our agreement assessments in a consistent manner.
- At the time of writing, we were aware that GUF outperforms GHSL 2014 (https://www.mdpi.com/2072-4292/10/6/895), and that WSF v2015 outperforms the GHSL v2018 (https://www.nature.com/articles/s41597-020-00580-5#Fig4). However, we were unable to find a multi-temporal accuracy assessment of the WSF evolution data product, whereas the GHSL v2018 has been evaluated multi-temporally (e.g., https://www.sciencedirect.com/science/article/pii/S0034425722002310).
- While it is possible that the WSF evolution outperforms the GHSL in historical epochs (due to the temporally dense underlying Landsat samples, whereas the data underlying the GHSL v2018 is temporally sparse), we decided to use the GHSL, as knowing its drawbacks (i.e., higher omission errors in historical epochs and rural domain) would help us to interpret the performance of the HISDAC-ES over time) whereas the historical performance of the WSF evolution remains unknown to us, impeding a meaningful interpretation of the level of agreement with HISDAC-ES observed in early epochs.
- We agree it would be interesting to conduct an "inverted" analysis and assess the performance of remote-sensing derived multi-temporal settlement products (e.g., WSF evolution, GHSL, …) using the Spanish cadastral data (or the EUBUCCO dataset) as reference data and will consider such extended comparative assessments in future work.

R2-6: Line 222, you refer to Corine being at an original resolution of 30 m but this should be 100m or is this a higher resolution Spanish product that was then provided to the EEA to be harmonized into the 100 m Corine product? There is also a 30m time series product recently produced for CORINE, but you should then reference this.

Response: Apologies for this mistake. We actually did use the 100m CLC data product. As the grid underlying the CLC is the EEA / Eurostat grid, no resampling was necessary. In the revised version, we state this in line 266:

> For most years in which CLC data is available, its estimated accuracy exceeds 85% (Büttner et al., 2021), while in the case of Spain, the accuracy of the CLC versions 2000 and 2006 have an estimated overall accuracy of > 93% (Diaz-Pacheco & Gutiérrez, 2013) in the Madrid region. ~~Specifically~~Herein, we obtained CLC data, available at a spatial resolution of 100m × 100m for the earliest (1990) and most recent (2018) available epoch (Fig. 3e,f, also Fig. A1)~~., and resampled it from the original resolution of 30m to the HISDAC-ES grid of 100m spatial resolution, using a majority resampling rule.~~ As the grid

Technical Corrections

R2-7: Line 59, change 'allow to mitigate' to 'all these two shortcomings to be mitigated'

Response: Done.

R2-8: Line 62, 'for example, (Uhl and Leyk, 2022a)' should be 'for example, Uhl and Leyk (2022a)'

Response: Done.

R2-9: Line 74, change ‘on over’ to ‘of over’

Response: Done.

R2-10: Line 85, change ‘European Union’ to ‘EU’

Response: Done.

R2-11: Lines 113 to 116, numbering of sections described in these lines doesn’t match numbering of the actual sections, e.g., outlook is section 8

Response: Thanks for catching this, we corrected the numbering throughout.

R2-12: Line 124, there is no section 2.3

Response: This should be Section 4 – we corrected this.

R2-13: Line 130, change ‘allow accessing’ to ‘allow the building data to be accessed’

Response: Done.

R2-14: Line 133, add ‘a’ before Web Feature Service

Response: Done.

R2-15: Line 405, remove space before full stop

Response: Done.

R2-16: Line 465, moves from section 4 to section 6 so no section 5

Response: Thanks, we corrected the section numbering throughout.

REVIEWER 3

Review „HISDAC-ES: Historical Settlement Data Compilation for Spain (1900 - 2020)“

The paper is very interesting, well written and results are clearly presented and evaluated. The dataset presented in this paper, the HISDAC-ES, is a valuable contribution to several fields, from demographic studies to urban planning. I do have some general comments/questions and minor comments that I would like the authors to address.

Response: Thank you for your valuable comments and for the positive, detailed assessment.

Comments for the authors:

R3-1: What do the authors mean with “built-up intensity” (lines 25, 78, 281)? is it the same as built-up density? I would suggest to briefly define the concept the first time it is mentioned, so there is a common understanding of the concept.

Response: We agree. In the revised version, we added a clarifying sentence lines 204-206:

industrial, agricultural, public services, and offices) as a proxy measure for ***built-up land use evolution*** from 1900 to 2020. Table 2 provides an overview of the gridded surfaces and spatial variables generated by these data processing steps. These surfaces quantify for example the ***building density*** (i.e., number of buildings per grid cell, BUDENS), the ***built-up surface density*** (i.e., building footprint area per grid cell, BUFA), or the ***built-up intensity*** (i.e., the total building indoor floor area per grid cell, BIA).

R3-2: Line 48: I think it is important to include in the introduction a recent published paper on the effort to homogenize European cadaster data (Milojevic-Dupont, N., Wagner, F., Nachtigall, F. et al. EUBUCCO v0.1: European building stock characteristics in a common and open database for 200+ million individual buildings. Sci Data 10, 147 (2023). https://doi.org/10.1038/s41597-023-02040-2)

Response: Indeed a very relevant dataset, that was not published at the time of writing of the original manuscript. In the revised version, we added a reference in the introduction section (line 49):

2019), mass and material of the building stock (Haberl et al., 2021). Moreover, building-level data is available from industry-generated data sources, such as Google (Sirko et al., 2021), Microsoft[1], from VGI (OpenCityModel[2], Atwal et al., 2022), as well as increasingly from cadastral data sources, for parts of the U.S. (Uhl and Leyk, 2022a) or, recently, for large parts of Europe (EUBUCCO, Milojevic-Dupont et al., 2023). In addition, (commercial) property / real estate data can be obtained through large-scale, data harmonization and dissemination efforts (e.g., ZTRAX[3], Regrid[4], ParcelAtlas[5], EuroGeographics[6]).

R3-3: Line 96: please, include the name of the countries that were compared with available open cadaster data (see Milojevic-Dupont et al paper).

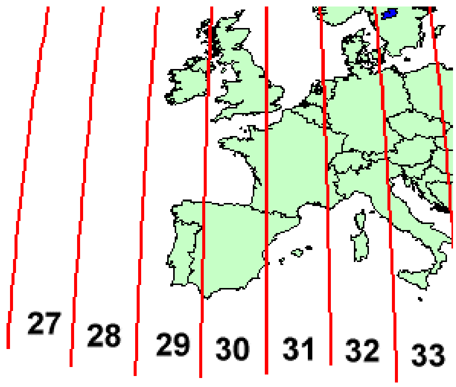
Response: In the revised version, we added some more detail on this process, which was carried out in a non-systematic manner. The work from Milojevic-Dupont is much more detailed (lines 97-99)

Many member countries of the ~~European Union~~EU have made such data publicly accessible (as regulated for instance in the EU PSI directive; ~~(~~European Union 2019)~~,~~ typically derived from cadastral data records, at varying levels of geometric detail and attribute completeness. We roughly compared the building footprint data available for some European countries, in a non-systematic manner (including France, Spain, the Netherlands, and Germany) and found that data from Spain has high levels of data coverage and attribute completeness (for an in-depth study on building data availability across Europe, see Milojevic-Dupont et al., 2023). However, these building data are maintained by different institutions within Spain, i.e., the chartered

R3-4: Line 130 & 140: There are four different UTM Zones in Spain (28-31). Why did not the authors work with geodesic coordinates, WGS84, for the spatial intersection of building centroids and the grid? Besides, I wonder if the grids are created in 25830 or do the authors use an existing grid, such as the one from EEA mentioned? If the grid was created in 25830, how does this affect the size of the grids for UTM zones 28, 29 and 31?

Response: Thank you for pointing us to this issue. We did not work with geodesic coordinates, because a grid in WGS84 is neither equal-area (which is required for statistical purposes), nor does it typically represent shapes in a realistic way (which is preferred for visualization purposes). As we describe in line 170-173, we offer the data in three grids: **1)** EEA LAEA grid (equal area projection for statistical analyses), **2)** REGCAN95 (which uses UTM zone 28N) for visualization purposes in the Canary Islands, and **3)** UTM 30N for visualization purposes in the Iberic peninsula, including the Balearic Islands and the exclaves Ceuta and Melilla.

We understand your concern that using UTM Zone 30N for the entire Iberic peninsula causes slight distortions, that affect meridian convergence and scale in regions outside of UTM Zone 30 (i.e., Galicia and Catalonia, see Fig. below). However, we believe that the deviations from true north and the scaling effect caused by this, is acceptable for visualization purposes. HISDAC-ES users who will use the data for visualization purposes in those regions believe these issues are problematic, we recommend to either reproject the data to the corresponding UTM zones or use an on-the-fly reprojection when visualizing the data within a GIS environment. Creating additional layers for these parts of the Iberic peninsula for each UTM zone would increase the data volume of HISDAC-ES even more. We hope that the Reviewer agrees with this justification and assessment.



Source: https://www.xmswiki.com/images/5/5e/Europe.png

Moreover, we would like to emphasize that the building footprint areas were calculated in LAEA projection and thus, are independent from the distortions discussed above. See line 156 , and 171-173:

After downloading and gathering the building data for 8,131 municipalities, covering all regions of Spain, we first calculated and attached the building footprint area (obtained after reprojecting to EPSG:3035) and converted the polygonal building footprint data to centroids, retaining all relevant attributes, to reduce the computational effort for the subsequent data processing. Despite the common INSPIRE framework, attribute names and building function classes differed slightly between

~~We decided to provide~~The gridded surfaces of HISDAC-ES are provided in three different spatial reference systems: (a) ETRS89 UTM Zone 30N (EPSG:25830) for the Spanish mainland, Balearic Islands, as well as the exclaves Ceuta and Melilla located in Northern Africa; (b) REGCAN-95 (EPSG:4083) for the Canary Islands; and (c) ~~for all Spanish territory~~ in the reference grid of the European Environmental Agency (EEA), which is based on the ETRS89 Lambert Azimuthal Equal Area projection (LAEA; EPSG:3035) and is commonly used for pan-European statistical mapping (e.g., Corine Land Cover data), for all Spanish territory. This way, users can refer to the data in UTM / REGCAN for mapping purposes (mostly North-oriented, angle-preserving); while the datasets in area-preserving LAEA projection can be used for statistical modelling and integration with other datasets (e.g., gridded statistical data from Eurostat[9] or gridded Spanish census data (INE grid[10]). Thus,

R3-5: Line 162: "we calculated the sum and the mean of the building units (BUNITS) per dwelling (DWEL) over all buildings within a given grid cell", I am not sure if I understand well what is it calculated here. Building units are not-residential units, while dwellings are residential units, isn't? since line 152: "the number of dwellings describes the number of housing units in residential buildings, whereas the number of building units counts the number of units within non-residential buildings".

Response: Apologies for the confusion. Indeed, the term building units is intended for units in buildings of non-residential use, while dwellings are meant to be residential units (in residential buildings). Thus, we measure the sum and the mean of the building units per building (BUNITS), and the sum and the mean of the dwellings (DWEL) per building found in each grid cell. In the revised version, we edited accordingly (lines 189-191). This difference is now also clarified using the new maps showing the completeness of DWEL and thus for units within residential buildings only (Fig. 18)

> stratification applied to the different thematic attributes yields a range of different sets of gridded surfaces. For example, we calculated the sum and the mean of the building units (BUNITS) ~~per~~ and dwellings (DWEL) over all buildings within a given grid cell, as well as both the sum and mean building indoor area (BIA) and building footprint area (BUFA)~~, respectively~~, based on the building centroids located within a grid cell. The resulting gridded surfaces represent ***physical features*** of the built

See also related to this our response to comment R4-11.

R3-6: Figure 2: what statistics are calculated at the municipality level? The same ones as for the grid?

Response: Figure 2 aims to give a high-level overview of the data processing. As we describe in Section 3.5: "These datasets contain the zonal sums of grid-cell level variables (i.e., building counts, as well as BUFA, BIA, DWEL, BUNITS) as well as corresponding densities (per municipality area)". We did not provide all statistics at the municipality level, however, in the revised version, we also provide the harmonized building centroid dataset (including municipality identifier), so that users can create their own municipality-level statistics (as per comment R4-16).

R3-7: Lines 209-211: Why is the evaluation performed using the municipality boundary and not using the grid? That would show better the urban-rural gradient.

Response: Apologies for not being very clear here. Unfortunately, we are not entirely sure what the Reviewer means by "not using the grid". The thematic assessment of HISDAC-ES and GHSL was performed at the grid cell level. The simplest case of such a cell-level thematic accuracy/ agreement assessment would be to overlay both surfaces, and calculate the total number of true positives, false positives, etc., and then obtaining an ***overall*** agreement metric (e.g., precision, recall, f-score). In our case, we "group" the grid cells based on the municipality they belong to, and calculate the agreement metrics based on zonal sums of true positives, false positives, etc., per municipality. This is a spatially explicit, ***zonal*** agreement assessment providing an interesting view on the spatial variation of the agreement statistics. In the revised version, we clarified this in lines 249-252:

> layers using Precision, Recall, and F1 score for each epoch and for each municipality. Specifically, we overlaid the binary raster surfaces of DEVA and GHSL and calculated the number of true positive (TP), false positive (FP), and false negative (FN) grid cells within each municipality polygon. These zonal statistics of binary agreement categories were then used to calculate municipality-level Precision, Recall, and F1 score. Moreover, we expected the agreement to vary across the rural-

R3-8: Since the authors evaluate their dataset against other datasets, I think it would be important to mention the accuracy of those datasets.

Response: We fully agree with this – knowledge of the accuracy of reference data is crucial. However, in this specific case the evaluation of very early epochs (e.g. year 1900), knowledge on the accuracy of reference data is scarce. For the more recent datasets, (GHSL, CLC) we added a few sentences on the estimated accuracy of these datasets from the literature (lines 240-243 and 264-266), and added a reference to studies that quantify the HYDE hindcasting accuracy (lines 278-281), as well as some reflections on the accuracy of the historical map sources (lines 293-295).

R2018A across the rural-urban continuum, and over time (e.g., Liu et al., 2020; Uhl and Leyk 2022b,c), whereas little information is available on the accuracy of newer, multitemporal GHS-BUILT datasets. For example, it has been reported that the GHS-BUILT R2018A yields an average Intersection-over-Union of around 0.35 in 1975, in rural areas, to around 0.65 in 2018, in urban areas, respectively, for selected study areas in the U.S. (Uhl and Leyk 2022b), and correlation coefficients of built-up surface fraction > 0.7, compared to reference data for selected cities in China (Liu et al., 2020).

used Corine Land Cover data (CLC, Büttner 2014) and compared it to the land use / building function layers in HISDAC-ES. For most years in which CLC data is available, its estimated accuracy exceeds 85% (Büttner et al., 2021), while in the case of Spain, the accuracy of the CLC versions 2000 and 2006 have an estimated overall accuracy of > 93% (Diaz-Pacheco & Gutiérrez, 2013) in the Madrid region. ~~Specifically~~Herein, we obtained CLC data, available at a spatial resolution of 100m ×

and available at a spatial resolution of 5' × 5' (approx. 6km × 9km in Spain). While the accuracy of built-up area estimates in the HYDE data is difficult to quantify due to a lack of historical reference data (Klein-Goldewijk & Verburg, 2013), Uhl et al. (2021) find relatively high agreement of urban growth trends extracted from HYDE and from the integrated processing of remote-sensing data and historical maps. Specifically, we used the layer "urban area fraction" from HYDE for each decade

development phases (Fig. 4c). While no quantitative information on the accuracy of these data sources is available, the underlying maps are handcrafted and based on manual interpretation of orthophotos, topographic measurements, or local domain knowledge, and can be deemed to be relatively accurate. Specifically, we manually digitized the areas developed in

R3-9: Section 3.5: how are the statistics derived? using the grids whose centroids are within the municipality? why is it not done by using the centroids of the buildings similarly to the grid approach?

Response: Thanks for this comment. Indeed, we were not very clear in describing the process for creating municipality level statistics. In the revised version, we added some descriptions in the methods section (lines 209-211). In fact, the municipality-level aggregation followed the same strategy as the grid-cell aggregation: We did a spatial join between the harmonized building centroids and the municipality polygons (spatial join based on point-in-polygon query) and transferred the municipality identifier and area to each building record, for subsequent statistical aggregation. Moreover, we added the Python script to create municipality statistics to the HISDAC-ES GitHub repository (https://github.com/johannesuhl/hisdac-es).

provide a time series of zonal statistics, aggregated to the municipality boundaries[11] (see Section 3.5). We created these zonal statistics based on (a) spatially joining the municipality identifier and area to each building centroid of our harmonized building dataset (point-in-polygon query), and (b) deriving statistics (sums, densities) for each municipality. All data processing, as

R3-10: It is unclear why the evaluation with different datasets is done by different spatial units, NUTS, municipalities, etc. and not using, for instance, the level of the datasets that is being compared to the HISDAC-ES or the grid itself.

Response: Or strategy for the evaluation can be outlined as follows:

- Generally, when comparing HISDAC-ES with other datasets, we down-sampled (i.e., aggregated) the finer-resolution dataset to the grid of the lower-resolution dataset:
    - GHSL (30m) was down-sampled to match the HISDAC-ES 100m grid, using an intermediate resolution of 10m.
    - CLC could be used directly, as HISDAC-ES LAEA grid aligns with EEA/CLC grid.
    - HISDAC-ES (100m) was up-sampled to the HYDE 5'x5' grid.
    - Vector data digitized from historical maps were rasterized to match the HISDAC-ES 100m grid.

- This way, we keep modifications to the finer-resolution dataset to a minimum and thus, ***only a minimum of additional uncertainty due to the resampling is induced*** into the agreement assessment.
- ***All agreement metrics are based on grid-cell level comparison*** (thematic or quantity agreement). However, instead of reporting overall agreement metrics across the whole territory under study, ***we spatially constrained our agreement metrics to local or regional strata which we deem meaningful,*** in order to assess regional variations of the measured agreement. These strata are chosen in a way that the sample size (i.e., the number of grid cells) per aggregation unit is sufficiently large. For example, as the HYDE cells have an extent of 6x9km, only one or few cells would cover each municipality and thus, HYDE agreement metrics at the municipality-level would not be statistically robust. For this reason, we used the NUTS-1 regions that would ensure a statistically robust sample (i.e., number of grid cells) per unit.
- Other assessments, such as the attribute completeness, were carried out at the municipality level, as we assume attribute completeness to somewhat follow administrative boundaries (e.g., due to differences in the organization of cadasters, cadastral updating cycles, etc.)

In the revised version of the manuscript, we added a short paragraph explaining this strategy in lines 224-230:

> History Database of the Global Environment), and (c) historical cartographic data (i.e., historical maps and urban atlases) and orthoimagery. For most of these experiments, we implemented the following strategy: ***(1)*** When comparing HISDAC-ES to other gridded datasets, we downsampled the dataset of higher resolution to the dataset of lower resolution. This way, additional uncertainty introduced by resampling is kept to a minimum. ***(2)*** We conducted agreement experiments at the grid cell level, i.e., based on cell-by-cell map comparison or correlation analysis. ***(3)*** From these cell-by-cell level comparisons, calculated agreement metrics within local or regional strata, defined by administrative boundaries or other classificiations, ***granular enough*** that we could assess regional variations of agreement, and ***large enough*** to ensure statistical robustness within each local stratum.

R3-11: The created dataset is "evaluated" against the RS-derived and modeled datasets, but "compared" to the historical maps and orthophotos. Do the authors refer to two different evaluations? Is yes, please, clarify.

Response: Good observation. We generally ***evaluate*** our data (or specific components) by ***comparing*** them to other datasets. The comparison can be quantitative or qualitative. We chose this terminology quite intuitively. In the revised version, we added a clarifying sentence in lines 373-379:

> We compared the layers from HISDAC-ES to a variety of related but independent datasets to evaluate spatial, temporal, and thematic components of our data (Sections 4.1-4.5). The various comparisons carried out are of either quantitative or qualitative nature, and aim to evaluate the quality of the information contained in HISDAC-ES. The chosen evaluation datasets cover a range of data products of different sources (e.g., remote sensing, model-based hindcasting, historical cartographic sources) and different thematic domains (e.g., land use / land cover, built-up areas, urban areas) While none of the evaluation datasets are free from uncertainty, in particular for early points in time, we believe that demonstrating the coherence between the phenomena measured in HISDAC-ES and the respective evaluation datasets will shed light on the quality of HISDAC-ES from various perspectives. These evaluation efforts are summarized in Table 3. Moreover, we assessed the attribute completeness of the building data underlying HISDAC-ES (Section 4.6).

R3-12: Lines 442: what does "the building density in the small, rural communities around Hornillos del Camino is similar to the densities in the center parts of the large cities" mean? That the density is high in the rural areas? Or low in the cities?

Response: In the revised version, we removed this sentence, as it was a side note, and not relevant for the manuscript.

R3-13: Line 460: I am aware that the number of floors is in fact available in the cadaster from Spain. One can check this in the official web map https://www1.sedecatastro.gob.es/cartografia/mapa.aspx. The information can be obtained from the CAT files (See file: Tipo 14: Registro de Construcción – Planta, see: https://www.catastro.minhap.es/documentos/formatos_intercambio/catastro_fin_cat_2006.pdf. Regarding the ATOM files, based on the following document is also available: https://www.catastro.minhap.es/webinspire/documentos/Conjuntos%20de%20datos.pdf. The field “bu-ext2d:numberOfFloorsAboveGround” from “BuildingPart”.

Response: Thank you pointing us to these additional data sources. We are aware of their existence, but we decided not to include them in this version of HISDAC-ES, as harvesting the BuildingPart data requires additional queries that would have made the data acquisition more costly, computationally. However, as we state in the conclusion section, we are planning to incorporate such data in a future version of HISDAC-ES. We expanded that sentence a little and included the reference to the very useful document that the Reviewer has provided in their comment (lines 608-610):

> data in a dasymetric modeling framework could be useful to create fine-grained, historical population estimates (cf. Burghard et al., 2023). Moreover, other components of Spanish ~~INSPIRE-conforming~~cadastral building data could be used, such as sub-building level information (e.g., building parts), to create fine-grained data on building function at the sub-building level, as
>
> 20
>
> well as information on building heights, which are available in a separate data pool[16]. Lastly, with the prospect of increasing
>
> [16] https://www.catastro.minhap.es/webinspire/documentos/Conjuntos%20de%20datos.pdf

Minor comments:

R3-14: Line 37, 43: citing style. The commas are missing.

Response: Done.

R3-15: Line 53-54: revise the commas, is an “and” missing in “, or semantic inconsistencies, incompatibilities”?

Response: We revised this sentence.

R3-16: Line 60: I suggest to use a more recent reference, for example, Milojevic-Dupont et al., (2023).

Response: Good point. In the revised version, we added some more recent references supporting this claim, including Milojevic-Dupont et al. (lines 60-63):

on building size, material, or function) allow ~~to mitigate~~ these two shortcomings to be mitigated and complement the traditional data sources (e.g., remote sensing data). Cadastral data are increasingly available as open data (von Meyer and Jones, 2013; Haberl et al., 2021; Milojevic-Dupont et al., 2023) and have been used in a variety of geographic, demographic, and economic studies (e.g., Tapp 2010; Leyk et al., 2014; Zoraghein et al., 2016; Nolte 2020; Sapena et al., 2022; Domingo et al., 2023). In

R3-17: Line 61: regarding the demographic applications of cadaster data, a recent study compared the performance of different methods and datasets, RS-derived data versus cadaster information, and the latter produced better results. HISDAC-ES could be used in many applications, for example in the field of population estimations (see: Sapena M, Kühnl M, Wurm M, Patino JE, Duque JC, Taubenböck H (2022) Empiric recommendations for population disaggregation under different data scenarios. PLoS ONE 17(9): e0274504. https://doi.org/10.1371/journal.pone.0274504)

Response: Thank you for pointing us to this very relevant reference. We added the citation at the proposed location (line 63, see R3-16).

R3-18: Line 66: I would avoid the use of "we" when describing previous studies even if they are from the authors. Line 73, for example: "Specifically, in previous work, the Zillow Transaction and Assessment Dataset was employed…". Line 140: also for "we decided".

Response: In the revised version, we edited these sentences accordingly, and went over the whole manuscript to identify and change such phrasing.

R3-19: Line 94: unclosed parenthesis.

Response: Corrected, thanks!

R3-20: I suggest reducing the use of "INSPIRE-conforming" when referring to the cadaster buildings, since once is explain is not necessary information and without it the readability is better.

Response: We fully agree with this observation. In the revised version, we removed most of the occurrences of "INSPIRE-conforming", except in the introduction and in two key locations in the conclusion section.

R3-21: Line 164: are the sum and the mean calculated for both, BIA and BUFA? With "respectively" it seams that the sum is for BIA and mean for BUFA.
Response: Both, sum and mean are calculated for BIA and for BUFA. We reworded this sentence to clarify.

R3-22: Line 236: please, add the level of the NUTS.

Response: We used the NUTS-1 level, and added this information in the revised version.

R3-23: Line 285: I would remove "surfaces" since the authors are referring to the building density, which is not a surface, and BUFA already implies surface in the building footprint.

Response: Good point. When speaking of "surfaces" in the manuscript, we generally refer to "gridded surfaces" (i.e., the raster layers), regardless whether the measured spatial variable represents an area or a count of occurrences / density. We admit that this is confusing, as we also use the term "built-up surface" in the manuscript. In the revised version, we renamed the term "surfaces" to "gridded surfaces" or "layers", in relevant locations, in order to be clear.

R3-24: Lines 362-366: I wonder if INSPIRE land uses or INSPIRE building is the right term to refer to the Spanish cadaster buildings following INSPIRE.

Response: We agree. In the revised version, we reworded accordingly throughout the manuscript.

R3-25: Line 391: typos: “HSDAC” and “sme”

Response: Corrected, thanks.

Figures:

R3-26: Figure 9: I think the maps could be improved by combining the information into one. For example: adding 3 classes, developed land in 1900, in 2020, and not developed for each region.

Response: Thanks for this suggestion. We implemented this suggestion in the revised Fig. 9 (see below). In order to “fill” the freed by merging the two epochs, we added a new panel (c) showing the settlement patterns around the city of León (Fig. 9c).

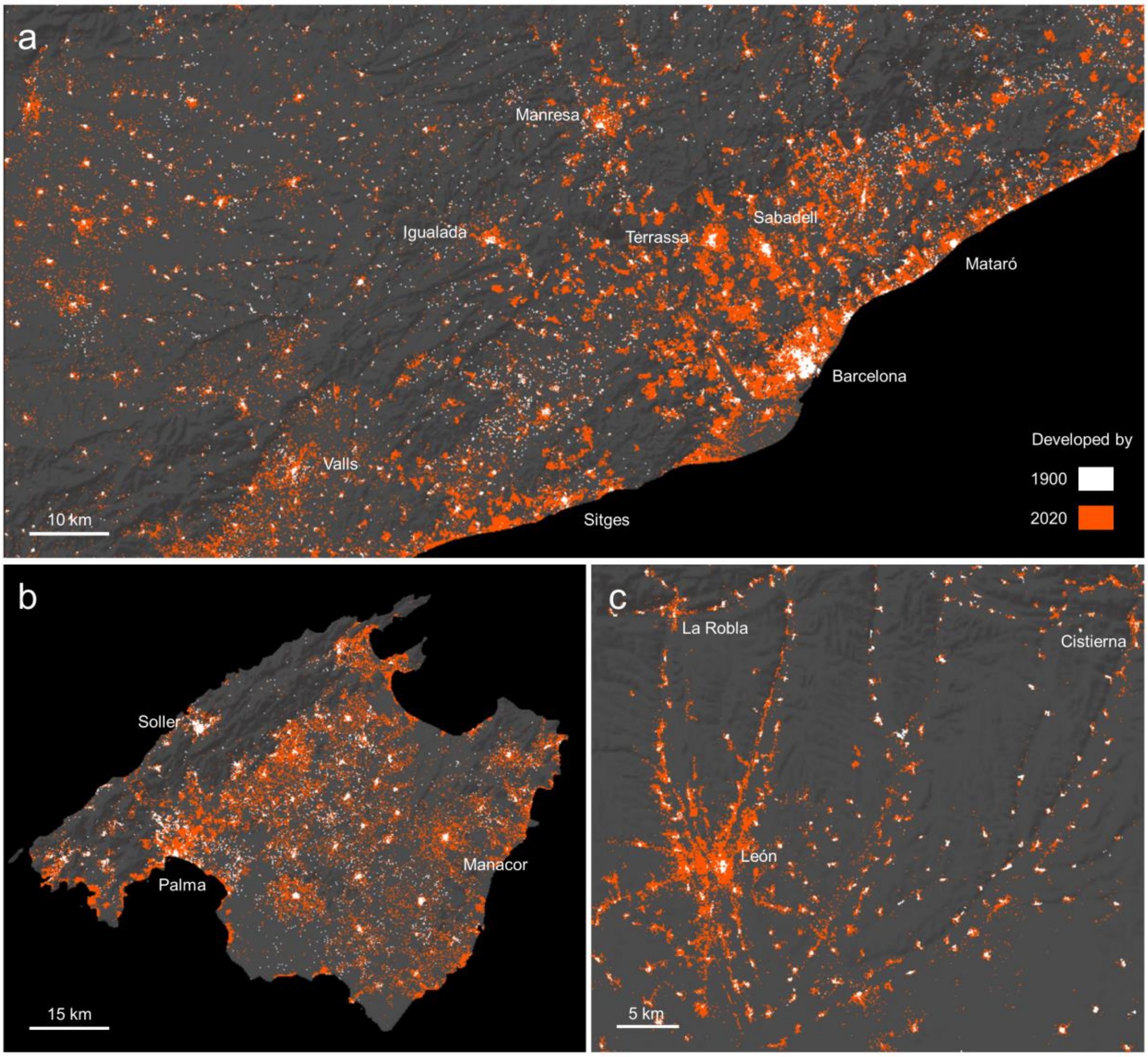


Figure 9: HISDAC-ES multitemporal surfaces of developed areas: DEVA layers for 1900 and 2020, shown ~~for~~(a) for Greater Barcelona, ~~in (a) 1900, and~~ (b) ~~2020, and~~ for the Island of Mallorca, and ~~in~~ (c) ~~1900 and (d) in 2020~~for the Greater León area. Basemap: Esri, USGS, NOAA.

R3-27: "Fig."10 = Figure 10.

Response: Corrected, thanks.

R3-28: Figure 12: As I understand (b) shows the metrics per municipality aggregated by date, what is (c) showing? The global metrics for the entire country?

Response: Fig, 12 b) shows the distribution of ***municipality-level*** agreement metrics by epoch, and Fig. 12 c) show indeed, the ***global*** agreement metrics calculated based on the grid cells within each rural-urban class from the GHS-SMOD classification, aggregated across the whole country, for each epoch. In the revised version, we expanded the caption of Fig. 12 to clarify this:

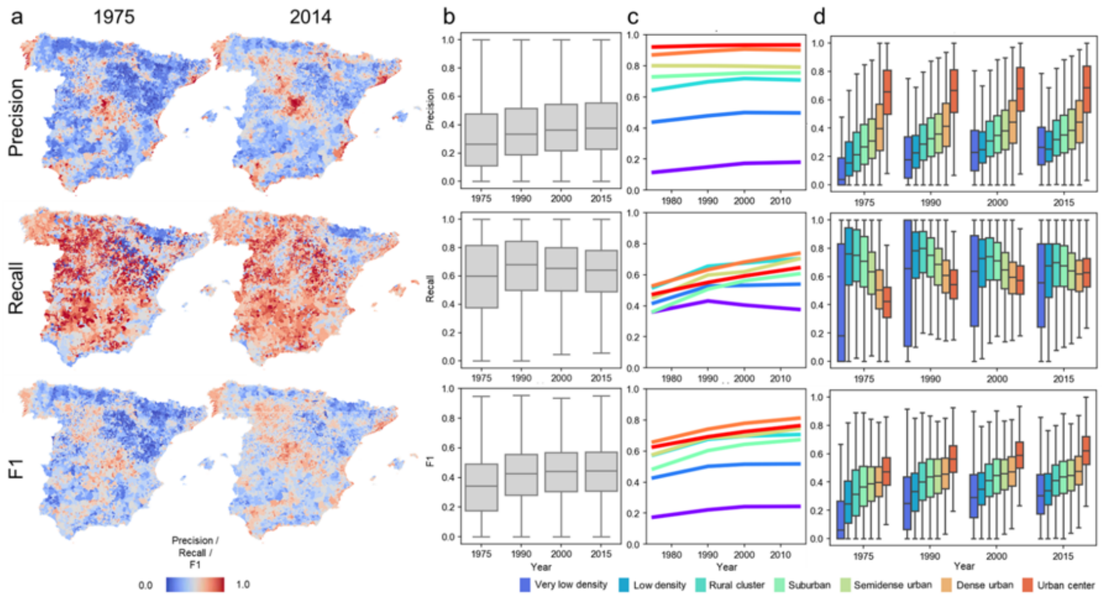


Figure 12: Evaluation of the HISDAC-ES developed area (DEVA) in comparison to built-up areas from the Global Human Settlement Layer per municipality, over time, and across the rural-urban continuum, by means of map comparison. (a) Maps of municipality-level precision, recall, and F1 score in 1975 and 2014, (b) overall temporal trends of municipality-level precision, recall and F1 score, (c) temporal trends of precision, recall and F-1 score calculated globally (i.e., at the country-level) within strata of the seven GHS-SMOD rural-urban classes, and (d) distributions of municipality-level agreement metrics within strata of a GHS-SMOD based rural-urban index calculated per municipality. See Appendix Fig. D1 for corresponding maps of the Canary Islands. All agreement metrics are obtained by map comparison on a cell-by-cell basis at a spatial resolution of 100m × 100m, calculated *locally* within municipality boundaries in (a), (b), and (d), and calculated *globally* (i.e., overall metrics for the whole country) within areas delineated by GHS-SMOD classes in (c).

R3-29: Figure 13: since the authors added "columns" I would also add "rows" for the Corine classes in the caption.

Response: It is somewhat unclear what the Reviewer is asking for. The caption of Fig. 13 is the following: *"Figure 13: Comparison of the HISDAC-ES land use data (**columns**) to land cover classes from Corine Land Cover (**rows**) for the years 1990 and 2018."* – containing references to both, to the columns and to the rows.

- Tables:

R3-30: Table 2: "Building indoor" without capital letter. "surface name" since not all parameters are surfaces, I wonder if there is a better way to call this column.

Response: Thank you for spotting this detail. We fixed the capital letter, and renamed "surface name" to "layer name" – "surface" here referred to "gridded surface" (see reviewer comment R3-23) –

regardless whether the variable actually is an areal measure. To avoid confusion, we use the term “layer” here.

R3-31: Table 3: I would include all the dates that are available: 1975, 1990, 2000 and 2014, instead of 1975-2014, otherwise might seem like an annual product.

Response: Good point. We added the individual years, where possible, or provided more information on the temporal sampling.

Table 3: Overview of the comparative evaluation efforts for HISDAC-ES.

HISDAC-ES variable	Evaluation data product (+URL)	Evaluated time period	Evaluated area	Section
DEVA	GHS-BUILT, GHS-SMOD (https://ghsl.jrc.ec.europa.eu/)	1975, 1990, 2000, -2014	Spain	4.1
Land use	Corine Land Cover (https://land.copernicus.eu/pan-european/corine-land-cover)	1990, -2018	Spain	4.2
BUFA, DEVA	HYDE v3.2 (https://doi.org/10.17026/dans-25g-gez)	1900-2010 (decadal) + 2015	Spain	4.3
MINCOY	Historical urban extents (Zornoza-Gallego 2022a, Remírez et al., 1988)	~~1900-2020~~3-5 epochs between 1900 and 2020	Alicante, Madrid, Valencia	4.4
DEVA (qualitative)	Historical maps (Minutas catastrales, 1:50,000) (http://www.ign.es/wms/minutas-cartograficas?request=GetCapabilities&service=WMS)	1910	Several cities / villages	4.4
DEVA (qualitative)	Historical orthophotos 1956-1957 (https://centrodedescargas.cnig.es/CentroDescargas/catalogo.do?Serie=FPNOA)	1955	Alicante, Madrid, Santiago de Compostela	4.5

R3-32: Table 4: avoid two times “digitized”.

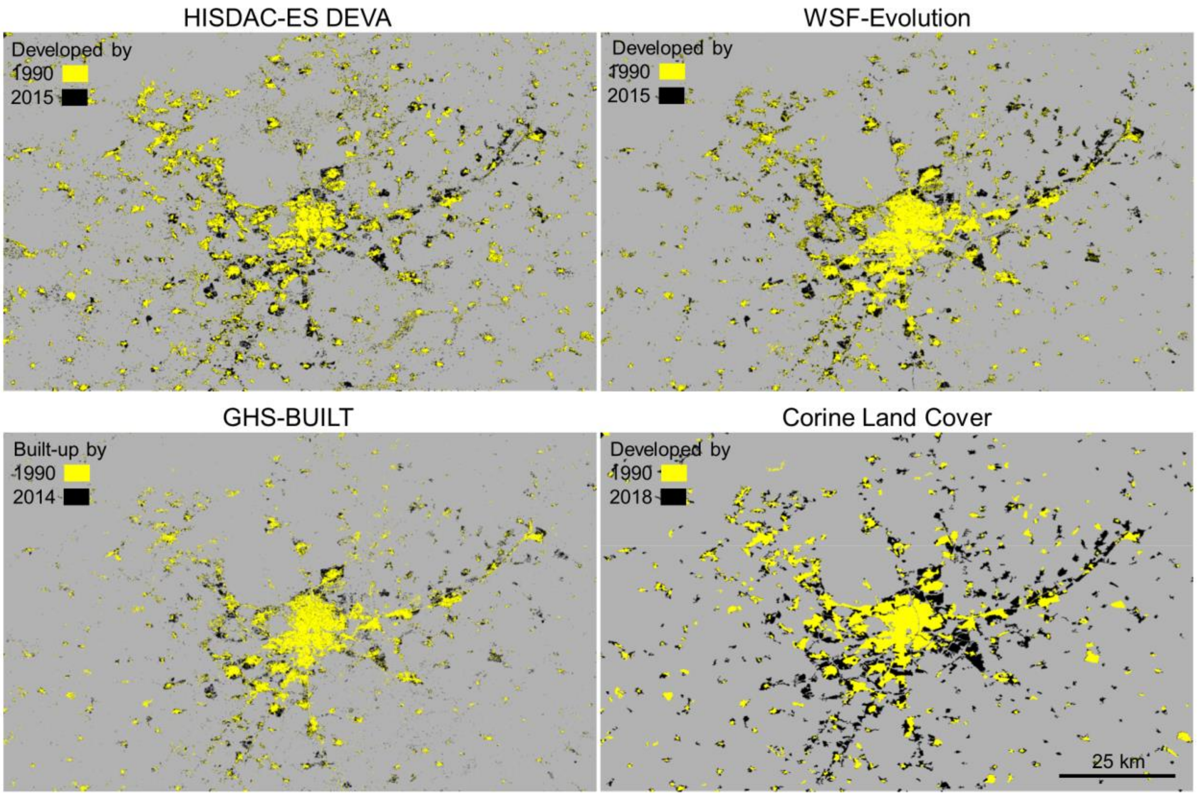
Response: Thanks. Corrected.

R3-33: Appendices: I suggest to give a brief description/title to each appendix A, B, etc.

Response: In the revised version, we added headings to each appendix. Given that the paper is already lengthy, we would like to refrain from adding additional descriptions to each appendix. However, if the Reviewer believes we should still add them, we will be happy to do so.

R3-34: Figure A1: I would combine the 2015 and 1990 map into one, to show better the growth and the differences between these datasets.

Response: In the revised version of the manuscript, we implemented the suggested changes, and we think this is an improvement (see below). We also added a forth panel to this figure, to show the CLC land cover data for the same extent and epochs, also in response to Reviewer comment R3-8 – illustrating the agreement between HISDAC-ES, GHS, CLC and WSF for a qualitative insight on the accuracy / coherence of these reference datasets:



Appendix Figure A1: Visual comparison of HISDAC-ES DEVA, WSF-Evolution, ~~and~~ GHS-BUILT, and Corine Land Cover, in 1990 and approximately 2015. For Corine Land Cover, only the classes "continuous urban fabric", "discontinuous urban fabric", "industrial or commercial units", "sport and leisure facilities", "construction sites", and "port areas" are shown, which are loosely related to developed / built-up areas.

R3-35: Figure B2: Similar to the comment above, I think that combining this 4-time-step maps into one per city will show better the evolution.

Response: Thanks for this comment. We agree that the evolution can be depicted more concisely by combining the four time steps in a single map. This is, however, what is shown in Fig. 8 already. The purpose of Fig. B2 is not primarily to show the evolution of a city, but rather to show that HISDAC-ES is capable to provide insights into the historical spatial configuration of a city. In other words, while Fig. 8 aims to illustrate a *longitudinal* characterization of cities, Fig. B2 illustrates historical, *cross-sectional* information on cities or urban areas. For example, Fig. B2 shows that historical, binary layers from HISDAC-ES could be used to calculate *landscape metrics or urban-morphometric indicators* to describe the historical spatial configuration of a city, while this is not directly evident from Fig. 8. In the revised version, we added some clarifying text (line 350): CITE

including towns, villages, and scattered, unincorporated settlements, as exemplified by the DEVA surfaces (Fig. 9). The DEVA ~~surfaces~~ layers reveal further detail on the spatial configuration of cities in early years, allowing e.g., for the computation of historical, urban morphological indicators (Appendix Fig. B2). We also provide several supplementary animated data

COMMUNITY COMMENT 1

R4-1. HISDAC-ES is a dataset with great potential, both for its coverage and for the period it covers (1900 - 2020). One of the major contributions is the integration of the 5 cadastres of Spain. Four of them cover only one of the 52 provinces and have -each of them- a different data model from the cadastre of the rest of Spain -which covers the remaining 48 provinces-.

Response: Thank you very much for your valuable comments and the thorough assessment of the HISDAC-ES dataset, as well as the provided aggregated statistics.

R4-2. In my opinion many of the details of the data models of the different cadastres should be briefly explained somewhere, since -as seems natural- the criteria guiding the elaboration of the database -functional categories or the distinction between dwellings and building units, for example- are determined by the cadastre with the largest coverage, which results in a lower representativeness of certain variables in the Basque Country and Navarre. In fact, the only totally homogeneous variable is the footprint of buildings (bufa).

Response: Good point. In the revised version, we added a paragraph to the conclusions, critically reflecting on such potential inconsistencies between cadastral systems (line 178-180): In lines 590-595, we also elaborate on the distinction between the number of dwellings and building units.

number of building units, building indoor area, and building footprint area (Fig. 1). For clarity, the number of dwellings describes the number of housing units in residential buildings, whereas the number of building units ~~counts~~ also includes the number of units within non-residential buildings (e.g., number of commercial businesses within a building complex, etc.). Furthermore, the building indoor area represents the attribute "official area" and measures the gross indoor area (across all stories) within a building.

number of floors, building volume, average floor height) could also be exploited for such data imputation efforts (cf. Fig. 18). Importantly, as the cadastral data used to create HISDAC-ES originate from different cadastral systems, there may be inconsistencies in the definition and in the measurement of specific attributes. For example, the way how building indoor area (BIA) is defined and measured, could vary across the different cadastral systems, despite conforming to the specifications of the INSPIRE directive aiming to homogenize cadastral data across the EU. Also, the definition and measurement of building units or number of dwellings could be affected by such inconsistencies, where the building footprint area (the input data for the BUFA layers) can be expected to be least affected by differences in cadastral systems. Thus, future work needs to thoroughly assess (and account for) such potential inconsistencies between the different cadstral systems. Similarly, the

R4-3. The validation effort is enormous, although limited by the arguments put forward by the authors. As described in the title, this is more a compilation than a harmonisation. The effort to include the cadastres of the Basque Country and Navarre is important, but there is still an effort to harmonise variables of the type being carried out by databases such as EUBUCCO v0.1 (https://www.nature.com/articles/s41597-023-02040-2) with the development of methodologies to complete variables (https://dx.plos.org/10.1371/journal.pone.0242010) based on urban morphology. Clearly this is outside the scope of the paper, but it represents the next step given the enormous amount of information contained in the cadastres.

Response: Indeed, the imputation of missing data could be a crucial next step. We added this to the outlook on future work (line 585-590):

state of a building (i.e., as of the year 2020), but these properties may have changed since the construction year on record, which may introduce additional uncertainty in the evolutionary layers in HISDAC-ES. Moreover, missing attributes in the cadastral data underlying HISDAC-ES could be estimated using specific data imputation strategies (e.g., Milojevic-Dupont et al., 2020). The complementary nature of certain, related building attributes (such as building indoor area, building height, number of floors, building volume, average floor height) could also be exploited for such data imputation efforts (cf. Fig. 18). Importantly, as the cadastral data used to create HISDAC-ES originate from different cadastral systems, there may be

R4-4. Analysis of a small part of the huge amount of information provided reveals small discrepancies which, while probably not affecting the underlying trends in the data, are difficult to understand from the point of view of the user who wants to make use of the data. The numbers below come from an attempt to generate population grids for census years since 1900 with a methodology similar to that used in the GHSL-POB from the information provided. Additional details are available if required. All calculations mentioned below use the contours provided by the database, which interestingly has a lot of slivers -slivers that are not present in the boundary line database of the National Geographic Institute (Centro Nacional de Información Geográfica)-.

Response: Thank you for performing this detailed assessment. In the revised version of the data, we obtained the latest, official municipality dataset, and recreated all municipality-level statistics, which are updated in the data repository.

Minor inconsistencies in the information

R4-5. It is not true that the zonal statistics provide information on the 8,131 municipalities currently existing in Spain (section 3.5). What the analysis of this information reveals is that there is only information on 8,124 municipalities, those existing on 01/01/2018.
Response: Thank you for pointing this out. We revised the creation of municipality-level statistics (see also comment R4-4) – and corrected accordingly. According to the official municipality shapefile (August 2023), there are 8,217 municipality polygons which we used as input for the data processing, out of which 8,159 intersect with the building data. We reworded accordingly.

R4-6. Furthermore, in the "hisdac_es_municipality_stats_completeness_v1" files, there are 8,169 records, as there are 45 records -only 6 of them with buildings- which correspond to territories not belonging to municipalities - -condominiums or “territories mancomunados”- all of them in Navarre. There are other territories in Spain with these characteristics in other provinces, which, however, they do not appear in the database. Note that the cadastral databases also have information on municipal boundaries, which do not coincide exactly with the boundary lines of the National Geographic Institute (Centro Nacional de Información Geográfica).

Response: As mentioned in response to comment R4-4, we are reprocessing all municipality-level statistics based on official municipality boundaries (https://centrodedescargas.cnig.es/CentroDescargas/index.jsp#; “Límites municipales, provinciales y autonómicos”). We hope that such issues have been resolved with this reprocessing.

R4-7. CatastRo package, mentioned in section 7, only allows the download of the Cadastre of the General Directorate of Cadastre, 48 provinces, but not of the provinces of the Basque Country and Navarre. CatastRoNav package (https://ropenspain.github.io/CatastRoNav/) can be used by R users to download data from the cadastre of Navarra. There is no such facility for the cadastres of the Basque Country.

Response: Thank you for pointing us to the CatastRoNav tool. In the revised version, we added the link to the code availability section.

R4-8. There are some numerical discrepancies between the raster information and that of the descriptive statistics files at the municipality level, at least for the bufa variable. In the descriptive statistics files, we always find more built-up area (bufa_sum) than in the raster files. These discrepancies are about 5% at the beginning of the period, but exceed 11% by the end of the period, which is not negligible, and has no clear explanation.

Response: After communication with the Commenter, we understand this issue and would like to thank for this consistency check, which helped us to identify a bug in the code to create the BUFA layers. In the revised version of the data, this bug has been fixed. Now the total sums of BUFA per year obtained from all grid cells, as compared to the sums of BUFA per year obtained from the vector data, are almost identical. Some remaining differences (of around 0.3% for residential BUFA, and around 0.2% for all BUFA, across all years) can be explained by edge effects, e.g., building centroids outside of municipality areas (e.g., in coastal settlements), or settlements near borders.

Below we report the results of the proposed consistency check, performed on our end after fixing the bug (as documented in the GitHub repository):

year	type	totalsum_muni	totalsum_raster	percent_differnce	year	type	totalsum_muni	totalsum_raster	percent_differnce
1900	resbufa	163,926,000	163,320,154	0.370	1900	allbufa	253,850,085	253,261,729	0.232
1910	resbufa	178,491,375	177,830,130	0.370	1910	allbufa	274,915,794	274,256,269	0.240
1920	resbufa	218,420,927	217,610,353	0.371	1920	allbufa	336,345,434	335,485,979	0.256
1930	resbufa	266,470,377	265,479,278	0.372	1930	allbufa	406,227,763	405,172,154	0.260
1940	resbufa	323,535,355	322,334,247	0.371	1940	allbufa	494,043,058	492,741,525	0.263
1950	resbufa	395,321,270	393,858,652	0.370	1950	allbufa	614,663,761	613,094,927	0.255
1960	resbufa	493,318,597	491,506,680	0.367	1960	allbufa	785,772,626	784,410,182	0.173
1970	resbufa	642,869,669	640,599,783	0.353	1970	allbufa	1,072,527,691	1,070,729,031	0.168
1980	resbufa	860,681,405	857,868,586	0.327	1980	allbufa	1,547,610,313	1,545,104,010	0.162
1990	resbufa	1,058,068,356	1,054,696,777	0.319	1990	allbufa	1,929,692,002	1,926,602,306	0.160
2000	resbufa	1,270,898,148	1,266,968,750	0.309	2000	allbufa	2,376,311,296	2,372,698,286	0.152
2010	resbufa	1,541,144,348	1,536,635,853	0.293	2010	allbufa	2,921,531,812	2,917,319,033	0.144
2020	resbufa	1,581,913,410	1,577,299,798	0.292	2020	allbufa	3,027,806,728	3,023,579,521	0.140

R4-9. The analysis of internal consistency (completeness of attributes in section 4.6) relies on the visual impression of figure 18, but it is likely that tables aggregated to province or regional level would be more illustrative here. These tables reveal clear problems in some variables in the cadastres of the Basque Country and Navarra, with more heterogeneity within these cadastres.

Response: We agree with this comment, and thank you very much for providing these aggregated completeness statistics. As we observe in line 537/538, indeed completeness seems to follow administrative boundaries, in part, due to the different cadastral systems. However, in order to keep the data volume of HISDAC-ES to a reasonable level, we would like to refrain from adding additional datasets. We think that users can aggregate these statistics easily.

R4-10. Also, the number of floors of the building (floors) could have been used to estimate the indoor area (bia), as there is a clear complementarity between these variables in terms of missingness.

Response: A good suggestion. For the scope of the present work, we prefer not to add too many analytical additions, but we added this idea to the outlook (lines 588-590):

al., 2020). The complementary nature of certain, related building attributes (such as building indoor area, building height, number of floors, building volume, average floor height) could also be exploited for such data imputation efforts (cf. Fig. 18). Importantly, as the cadastral data used to create HISDAC-ES originate from different cadastral systems, there may be

R4-11. The number of dwellings is much less representative than the other variables in the database. At the national level the percentage of buildings with no value for this variable (dwellings) is 28%. This fact contrasts with the high completeness for the variable building units (bunits). However, from my point of view, it is not clear from the text (line 151 and 152, page 6, and then line 162) how these two variables are calculated from the original information (which classifies a building according to its use and, given that, the number of dwellings and the number of building units are stated- this for the General Directorate of Cadastre).

Response: Thank you for this comment. In the revised version of the paper, we calculated the completeness of the NUMDWEL variable with respect to residential buildings only, and the completeness of the BUNITS variable with respect to non-residential buildings only, see the revised Fig. 18:

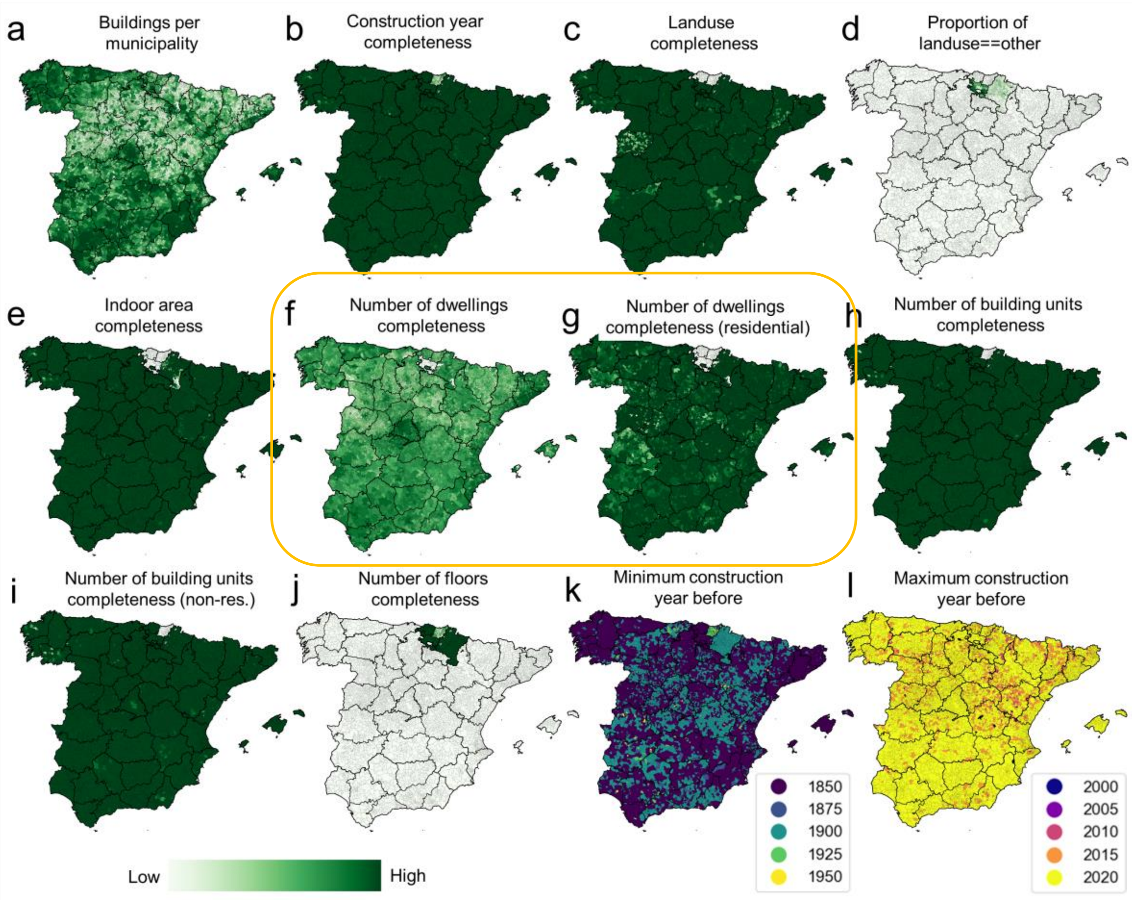


Figure 18: Attribute completeness and construction year coverage at the municipality level. See Appendix Fig. G1 for corresponding maps of the Canary Islands.

Furthermore, we tried to clarify the ambiguous relationship between these two metrics, in particular for buildings of mixed use. We advise data users to be careful with these variables, and possibly using a combination of both for specific modelling purposes, along with suitable methods for uncertainty propagation:

the BUFA layers) can be expected to be least affected by differences in cadastral systems. Thus, future work needs to thoroughly assess (and account for) such potential inconsistencies between the different cadstral systems. Similarly, the variables DWEL (number of dwellings) and BUNITS (building units) need to be treated carefully, due to their potential semantic overlap: for example while DWEL only contains residential units, BUNITS may contain both, residentially and non-residentially used building units, for example in the case of buildings of mixed use. Generally, we advice to be cautious when ingesting the HISDAC-ES data layers into demographic modelling applications, where the progagation of uncertainty from the input data to the outputted products needs to be taken into account (e.g., Goehrlich-Gisbert & Marti, 2017). Finally, the

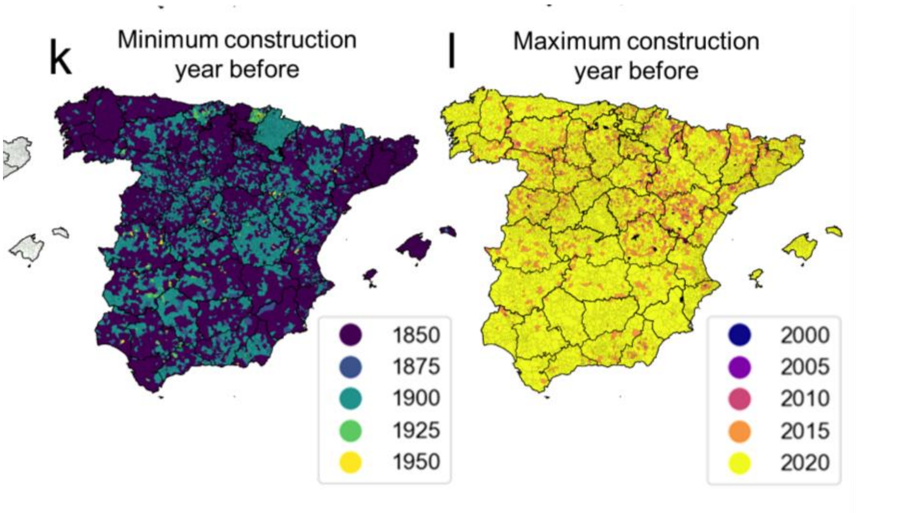
R4-12. Note, in passing, that in the General Directorate of Cadastre, INSPIRE ATOM services, there is information on the number of floors in the "BuildingPart" files. So, this information exists generally, but in another place.

Response: Good point. As we state in line 608-610, we are aware of the BuildingPart files, and plan to include this information in future work.

data in a dasymetric modeling framework could be useful to create fine-grained, historical population estimates (cf. Burghard et al., 2023). Moreover, other components of Spanish INSPIRE-conformingcadastral building data could be used, such as sub-building level information (e.g., building parts), to create fine-grained data on building function at the sub-building level, as well as information on building heights, which are available in a separate data pool[16]. Lastly, with the prospect of increasing availability of INSPIRE-conforming cadastral building data, HISDAC-related efforts will be expanded to other European countries where cadastral building data is of similarly high completeness, quality, and thematic richness.

R4-13. In 1900 there are 183 municipalities without buildings (in the file of zonal statistics, in the rasters it happens only in 172 municipalities). All municipalities have built-up area (bufa) only from 1970 onwards. Numerical analyses of this style may shed more light on the survival bias, mentioned by the authors, and the quality of the data at the beginning of the 20th century. An Excel file with some of this information is attached.

Response: Thank you for this analysis. Indeed, this is interesting information for the data users. In the revised version, we graphically show some of this information in the revised Fig. 18 k and l, and describe this in lines 543-539:



Conversely, information on the number of floors is highly complete in the Navarra region (Fig. 18j), but otherwise not covered in the remaining provinces, and thus, has not been used in this version of HISDAC-ES.

We also assessed the temporal coverage of construction year information at the municipality level, in order to better understand potential survivorship bias in the data. As can be seen in Fig. 18k, most municipalities have the earliest construction year on record <1850, or <1900. Only a few regions have minimum construction years between 1900 and 1925 (e.g., regions around San Sebastián and Bilbao), whereas very few, scattered municipalities have earliest construction years between 1925 and 1950. In these municipalities, data users should be careful when conducting long-term analyses, as survivorship bias may be high. Likewise, we mapped out the the maximum construction year per municipality (Fig. 18l), indicating the recentness of the cadastral building data underlying HISDAC-ES, as in most municipalities, the most recent construction year on record is between 2015 and 2020. Generally, these high levels of attribute completeness and temporal coverage are encouraging and indicate that the layers derived from these attributes are expected to be highly reliable at least for recent points in time. We made these municipality-level attribute completeness statistics available in the data repository.

The mentioned discrepancies between raster- and vector-based municipality-level statistics are likely an artifact due to edge effects caused by grid cells overlapping across municipality borders and resulting misassignment of records to neighboring municipalities.

R4-14. In the statistics by municipality appears the variable municipal area (muni_area_sqm). One would expect this variable to be invariant over time. However, there are some municipalities with value 0 in some years, which coincide exactly with the municipalities and years with no buildings. It is not clear where this variable comes from. In addition, this variable is superfluous, as the municipal area can be calculated from the vector layer.

Response: We agree. In the revised version of the data, we removed this column from the table.

Potentially useful additional information

R4-15. It would be useful to know the date of download of the data. The General Directorate of Cadastre updates the INSPIRE Cadastre data twice a year.

Response: Good point. We acquired the cadastral data in June 2021. We deem this to be of sufficient actuality, as the temporal extent of HISDAC-ES is 1900-2020. In the revised version, we added the data acquisition data to the data availability statement.

R4-16. Since the code generating the information is public (https://github.com/johannesuhl/hisdac-es), it would be useful to make the original data available. Although the summary of the information is adequate, another treatment of the original data might be more suitable for certain purposes. For example, for the generation of historical population grids by dasymetric methods, it would be useful to have the built-up area (bufa) by residential use -currently this variable is only available in density format- or the building height -bia- by years and/or use.

Response: We agree that for some applications, specific stratifications of the data may be useful. In order to facilitate such analyses, we uploaded the harmonized building centroids as an additional vector file (GPKG format) to the HISDAC-ES repository (https://github.com/johannesuhl/hisdac-es). Moreover, in the revised version, we created time series of raster layers measuring BUFA and BIA based on **residential** buildings only (RES_BUFA, RES_BIA layers), in the LAEA grid, and also in the municipality-level statistic files, to facilitate dasymetric / demographic modelling efforts. The script for the rasterization of residential building attributes over time is based on a modified concept (i.e., it reads the country-level harmonized building centroid vector dataset rather than each of the municipality-level shapefiles). This causes a considerable speed-up in data processing time, but is memory-intensive. This script can be used together with the country-level vector dataset and is available at https://github.com/johannesuhl/hisdac-es.

Other changes to the manuscript done by the authors:

We noticed a mistake in Fig. 3, showing the downsampled HISDAC-ES rather than the HYDE data. In the revised version, we corrected this.