Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

A Multi-Temporal National Census of Brick Kilns in Bangladesh from 2000 to 2023

Joynal Abedin

j.abedin@lsu.edu

Louisiana State University

Sabiha Sabrina

University of Florida

Jafor Shadik

Leading University

Humaira Akter

University of Dhaka

data-descriptor

Keywords:

Posted Date: November 3rd, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11255983/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Abstract

Brick kilns are a major source of air pollution, land degradation and occupational hazard across South Asia, yet no consistent, long-term record shows where they are and how they change. We present a manually digitized, multi-temporal census of brick kilns for all of Bangladesh at five epochs: 2000, 2010, 2016, 2020 and 2023. Kilns were identified by visual interpretation of very-high-resolution imagery in Google Earth, searched systematically on a 10 × 10 km grid, and recorded as points at the chimney base. After removal of duplicate and out-of-boundary records, the census contains 3,901, 5,512, 7,652, 8,385 and 9,333 kilns, respectively. One-to-one matching between epochs links the 34,783 records to 10,912 persistent kiln sites, each with a presence history, first and last year observed, and administrative codes. Independent digitization of the same kilns agrees to a median of 14 m. The dataset supports research on air quality, exposure, land use, labor and regulation, and provides a baseline for evaluating automated kiln detection.

Background & Summary

Fired clay bricks are the main walling material in Bangladesh, and their production has grown with urbanization and construction. The Department of Environment (DoE) estimated about 7,000 kilns producing nearly 23 billion bricks a year, contributing about 1% of gross domestic product and employing about one million people[1]. Most kilns are fixed-chimney Bull's Trench or derived designs, fired with coal and biomass from November to April[1,2].

Kilns are a major, seasonal source of air pollution. Kiln clusters around Dhaka contribute a large share of dry-season fine particulate matter[3], and South Asian kilns emit particulate matter, black carbon, carbon monoxide and sulfur dioxide at rates that depend on kiln technology and fuel[4,5]. Kilns also consume topsoil from agricultural land[1,6] and employ a large informal workforce whose conditions have drawn attention across South Asia's "brick belt"[7].

Regulating the sector requires knowing where kilns are and how the stock changes. Existing maps are local or single-date. Studies have tracked kiln development in peri-urban Dhaka[8] and the seasonal air-quality impact of Dhaka's kilns[9]. Lee et al.[10] showed that deep learning can locate kilns in satellite imagery across Bangladesh, and Boyd et al.[7] used satellite imagery to estimate kiln numbers across the brick belt. No dataset yet provides a consistent, manually verified national record spanning two decades, with kiln-level identifiers that track openings and closures.

This Data Descriptor presents such a record. We digitized every visible kiln in Bangladesh at five epochs (2000, 2010, 2016, 2020 and 2023), cleaned the layers with documented quality-control rules, and linked kilns across epochs into 10,912 persistent sites (Fig. 1). The dataset can be used to (i) quantify the growth, turnover and spatial diffusion of the sector; (ii) build kiln-proximity exposure metrics for air-quality and health studies; (iii) assess compliance with siting rules; (iv) estimate land and topsoil use; and (v) train and validate automated kiln detectors with multi-year, human-verified labels.

Methods

The census was produced in five steps: imagery selection, digitization, quality control, cross-epoch harmonization and packaging (Fig. 2).

Study area and kiln definition

The census covers the whole land area of Bangladesh (about 147,570 km²), organized into 8 divisions and 64 districts. A kiln was defined as a brick-firing structure visible in imagery: an oval or rectangular firing trench, usually with a central fixed chimney, and adjacent clay-molding and drying yards. These features make kilns distinctive in very-high-resolution imagery[10]. Each kiln was recorded as one point at the base of its chimney, or at the center of the firing trench where no chimney was visible.

Imagery and search protocol

Kilns were interpreted visually in Google Earth Pro, using the high-resolution scene closest in date to each target epoch through the historical-imagery tool. To ensure complete coverage, the country was divided into 1,652 grid cells of 10 × 10 km, and each cell was searched systematically for every epoch. Each cell was searched more than once to reduce omissions. For the 2016, 2020 and 2023 epochs, every cell was interpreted from imagery acquired in the target year. For 2010 and 2000, a small number of cells had only cloud-covered or hazy imagery for the target year; these cells were interpreted from the nearest clear scene, acquired in 2011–2012 for the 2010 epoch and in 2001–2002 for the 2000 epoch. Epoch labels therefore denote the target year, with a maximum offset of two years in those cells.

The coordinates show how the layers were built. The 2010 layer was produced by updating the 2000 layer: 3,568 of its points have coordinates identical to 2000 points. The 2016, 2020 and 2023 layers were digitized independently and share no identical coordinates with earlier layers.

Quality control

Two rules were applied to every layer (Table 1).

1. **Exact duplicates.** Points with identical coordinates were reduced to one. The raw 2023 layer contained 691 duplicate pairs, occurring in blocks with constant identifier offsets, which shows that digitization batches had been imported twice.
2. **Boundary check.** Points outside the national boundary (BBS/OCHA ADM1)[11] were removed; this affected six 2023 points along the land border and coast.

Table 1
Records at each processing step, by epoch.

Epoch	Raw records	Exact duplicates removed	Outside boundary	Non-kiln features	Kilns released
2000	3,904	3	0	0	3,901
2010	5,518	6	0	0	5,512
2016	7,657	5	0	0	7,652
2020	8,386	1	0	0	8,385
2023	10,030	691	6	0	9,333

Administrative attribution

Each point was assigned division, district and upazila names and codes by point-in-polygon overlay with the 2020 BBS/OCHA administrative boundaries[11], so that counts can be joined to census and survey data.

Cross-epoch harmonization

Because each kiln was recorded at the same physical feature, a kiln present in two epochs should appear at nearly the same coordinates. Points in consecutive epochs were matched one-to-one within a tolerance of 75 m. The tolerance sits in the trough of the bimodal distribution of distances between epochs (Fig. 4a): records of the same kiln were typically 10–30 m apart, whereas distinct kilns are much farther apart. Within each layer, the median distance between neighboring kilns is about 400 m, and only 0.8–2.3% of kilns have another kiln within 75 m. Where kilns are this close, the matching rule prevents confusion: all candidate pairs were sorted by distance and accepted closest first, and a kiln could be matched only once, so each kiln was paired with its nearest counterpart in the other epoch rather than with a neighboring kiln.

Matched points inherit a persistent site identifier. Sites not re-observed in the next epoch but present again within 75 m in a later epoch were linked across the gap, using the same closest-first, one-to-one rule (1,344 links). Each site therefore has a five-epoch presence history, a first and last year observed, and a flag for intermittent histories. The 34,783 kiln records are resolved to 10,912 sites.

Data Records

The dataset will be deposited in figshare under a CC BY 4.0 license [https://figshare.com/s/128152d58bf3fd37eaba]. All coordinates are in WGS 84 (EPSG:4326) decimal degrees. The files are listed in Table 2.

Table 2
Files in the dataset.

File	Format	Content
bd_brick_kilns_2000_2023.gpkg	GeoPackage	Five point layers (kilns_2000 … kilns_2023) and one site layer (kiln_sites)
kilns_YYYY.csv (5 files)	CSV	One row per kiln per epoch; same fields as the point layers
kiln_sites_2000_2023.csv	CSV	One row per persistent kiln site, with presence history
kiln_counts_district.csv	CSV	Kiln count and density (kilns per 100 km²) by district and epoch
README.txt	Text	Field definitions, processing summary and version history

The point layers share one schema (Table 3). The site layer adds the presence history and observation span of each site.

Table 3
Fields in the point and site layers.

Field	Layer	Type	Description
kiln_id	Points	Text	Unique record identifier, e.g. K2016_00123
site_id	Points, sites	Text	Persistent site identifier linking records across epochs, e.g. BDK00001
epoch	Points	Integer	2000, 2010, 2016, 2020 or 2023
lon, lat	Points, sites	Float	Kiln location (WGS 84); sites use the most recent observation
div_code, div_name	Points, sites	Text	Division (BBS/OCHA ADM1)
dist_code, dist_name	Points, sites	Text	District (ADM2)
upz_name	Points, sites	Text	Upazila (ADM3)
p2000 … p2023	Sites	Integer (0/1)	Presence of the site in each epoch
first_obs, last_obs	Sites	Integer	First and last epoch in which the site was observed
n_obs	Sites	Integer	Number of epochs in which the site was observed (1–5)
intermittent	Sites	Integer (0/1)	1 if the site is missing from an epoch between its first and last observation

The kiln stock grew in every interval, from 3,901 kilns in 2000 to 9,333 in 2023 (Fig. 3a). Of the 10,912 sites, 2,787 were observed in all five epochs. Openings outnumbered sites not re-observed in every interval (Fig. 3b), and growth occurred in all eight divisions, with the largest absolute increases in Dhaka, Chittagong and Khulna divisions (Fig. 3c).

Technical Validation

We assessed positional precision, internal consistency across epochs, and agreement with an official estimate. The tests show that locations are precise to about 10 m and that the 2016 and 2020 layers have modest omissions, which the site panel makes visible.

Positional precision

The 2016, 2020 and 2023 layers were digitized independently, so the distance between matched records of the same kiln measures combined digitizing error. Distances from each kiln to the nearest kiln in the

next layer were strongly bimodal (Fig. 4a). Matched pairs had a median separation of 14 m and a 90th percentile of 29–32 m in all three intervals. Assuming equal, independent errors in two layers, the positional precision of a single record is about 10 m (14 m/√2), well within the footprint of a kiln and its yard.

Duplicates and near-duplicates

After quality control no layer contains points with identical coordinates. The raw 2023 layer stood out with 691 duplicate pairs (Fig. 4b), whereas the other raw layers contained at most six. Pairs of distinct points closer than 10 m are rare (fewer than 20 in any cleaned layer) and are retained, because neighboring kilns sometimes share a yard.

Temporal consistency

A kiln structure that is visible in one epoch and again in a later epoch is unlikely to have disappeared in between. Sites missing from an intermediate epoch therefore give a lower bound on omission in that epoch, although some may reflect demolition and rebuilding on the same site. Such sites numbered 38 in 2010 (0.7% of the expected stock), 413 in 2016 (5.1%) and 943 in 2020 (10.1%). The 2020 layer is therefore the least complete; users estimating change should use the site panel, which records these sites as intermittent, rather than comparing raw counts.

The number of sites classified as persistent depends little on the matching tolerance. Between 2016 and 2023, tolerances of 50 m and 150 m changed the numbers of persistent and new kilns by less than 2% relative to 75 m.

Agreement with official estimates

The DoE estimated about 7,000 kilns in its 2017 national strategy[1]. Our 2016 census contains 7,652 kilns, 9% more; a plausible explanation is that imagery also captures kilns absent from official registers.

Limitations

No independent re-interpretation of a random sample of grid cells has yet been made, so commission and omission rates are not estimated directly. Within an epoch, scenes come from different dates and providers, and in a small number of 2000 and 2010 cells they postdate the target year by up to two years because of cloud or haze. Cross-epoch consistency checks indicate that the 2020 layer misses at least 10% of existing kilns and the 2016 layer at least 5%. Kiln technology (fixed chimney, zigzag, Hoffmann or tunnel) and seasonal operation are not recorded.

Usage Notes

- **Measuring change.** Use the site layer for openings, closures and persistence. Raw epoch counts understate the stock in 2016 and 2020 by the intermittent sites described above; treating intermittent sites as present between their first and last observation gives an upper-bound series.

- **Epoch intervals.** Epochs are unevenly spaced (10, 6, 4 and 3 years). Express change per year, or per interval with the interval length stated.
- **Exposure metrics.** Kiln counts within a buffer (for example 1–5 km) of residences, schools or survey clusters can be computed directly from the point layers. Because location precision is about 10 m, buffers of 500 m or more are insensitive to digitizing errors.
- **Area units.** Densities in kiln_counts_district.csv use district areas published by the Bangladesh Bureau of Statistics. Joins to other data should use the codes in the dataset rather than names, because district spellings vary between sources (for example Comilla/Cumilla and Jessore/Jashore).
- **Detector training and validation.** The layers mark kiln presence only. Absence of a point does not certify absence of a kiln, particularly in 2020, so negative training samples should be drawn away from all five layers.
- **Operating status.** Presence in imagery indicates an existing kiln structure, not that the kiln fired in that season. No layer records operating status.

Declarations

Code Availability

The Python code used for quality control, administrative attribution, cross-epoch matching and figure production (brick_kiln_census.py, using geopandas, SciPy and matplotlib) is included in the figshare record with the dataset (https://figshare.com/s/128152d58bf3fd37eaba), together with a README and a list of requirements. Digitization was manual and used no code.

Competing interests

The authors declare no competing interests.

Author Contribution

J.A. conceived and designed the study and directed the project. J.A., S.S., J.S. and H.A. collected the data by digitizing the kiln layers. J.A. performed the quality control, cross-epoch harmonization and validation, created the database, prepared the figures and wrote the first draft of the manuscript. S.S., J.S. and H.A. reviewed and edited the manuscript. All authors read and approved the final manuscript.

Acknowledgements

We thank Google for providing access to historical high-resolution imagery through Google Earth Pro, and the Bangladesh Bureau of Statistics and UN OCHA for making the administrative boundaries publicly

available. This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

Data Availability

The brick kiln census generated in this study, "A Multi-Temporal National Census of Brick Kilns in Bangladesh, 2000–2023", is deposited in figshare under a CC BY 4.0 license and is available at https://doi.org/10.6084/m9.figshare.34063953. The dataset comprises a GeoPackage with five annual point layers and one kiln-site layer, the corresponding CSV files, district-level kiln counts, and a README describing all fields. The administrative boundaries used for attribution are the BBS/OCHA Bangladesh Subnational Administrative Boundaries, publicly available from the Humanitarian Data Exchange (https://data.humdata.org/dataset/cod-ab-bgd). The Google Earth imagery used for digitization is subject to the providers' terms and cannot be redistributed, but it can be viewed through Google Earth Pro's historical-imagery tool.

References

1. Department of Environment (DoE). *National Strategy for Sustainable Brick Production in Bangladesh*. Ministry of Environment and Forests, Government of the People's Republic of Bangladesh, Dhaka (2017).
2. Gomes, E. & Hossain, I. Transition from traditional brick manufacturing to more sustainable practices. *Energy Sustain. Dev.* 7, 66–76 (2003).
3. Guttikunda, S. K., Begum, B. A. & Wadud, Z. Particulate pollution from brick kiln clusters in the Greater Dhaka region, Bangladesh. *Air Qual. Atmos. Health* 6, 357–365 (2013).
4. Weyant, C. et al. Emissions from South Asian brick production. *Environ. Sci. Technol.* 48, 6477–6483 (2014).
5. Haque, M. I., Nahar, K., Kabir, M. H. & Salam, A. Particulate black carbon and gaseous emission from brick kilns in Greater Dhaka region, Bangladesh. *Air Qual. Atmos. Health* 11, 925–935 (2018).
6. Saha, C. K. & Hosain, J. Impact of brick kilning industry in peri-urban Bangladesh. *Int. J. Environ. Stud.* 73, 491–501 (2016).
7. Boyd, D. S. et al. Slavery from space: demonstrating the role for satellite remote sensing to inform evidence-based action related to UN SDG number 8. ISPRS J. Photogramm. Remote Sens. 142, 380–388 (2018).
8. Hassan, M. M., Juhász, L. & Southworth, J. Mapping time-space brickfield development dynamics in peri-urban area of Dhaka, Bangladesh. *ISPRS Int. J. Geo-Inf.* 8, 447 (2019).
9. Mahboob, M., Saha, M. & Hossain, M. I. Spatial assessment of the seasonal impact of brickfields on air pollution in Dhaka metropolitan area using ordination techniques. *Air Qual. Atmos. Health* (2023).

10. Lee, J. et al. Scalable deep learning to identify brick kilns and aid regulatory capacity. Proc. Natl Acad. Sci. USA 118, e2018863118 (2021).
11. UN OCHA & Bangladesh Bureau of Statistics. Bangladesh – Subnational Administrative Boundaries (bgd_admbnda_bbs_20201113). *Humanitarian Data Exchange* https://data.humdata.org/dataset/cod-ab-bgd (2020).

Figures

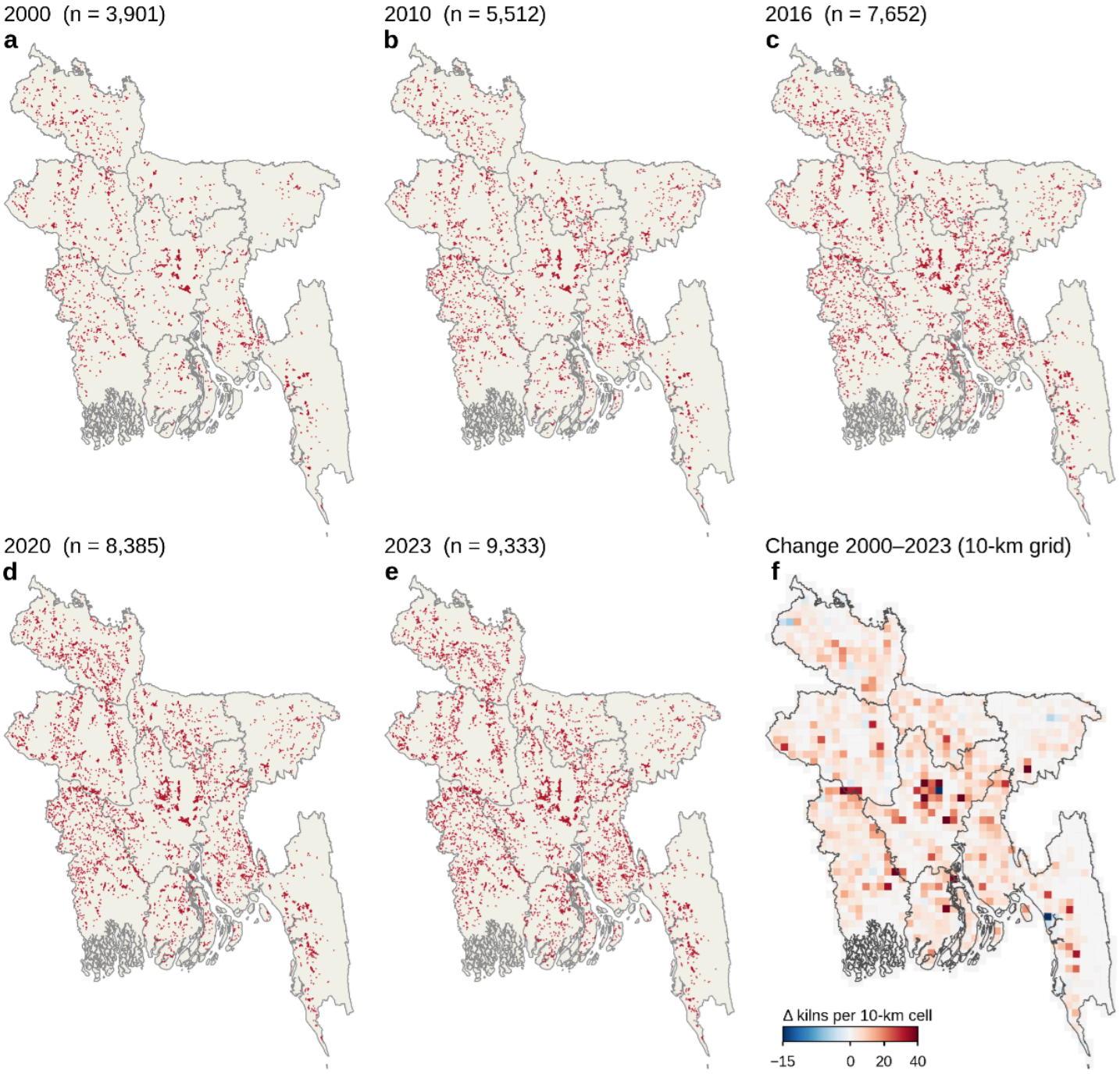


Figure 1

Brick kilns in Bangladesh at five epochs after quality control (a-e) and the net change in kiln numbers per 10 × 10 km grid cell between 2000 and 2023 (f). Gray lines are division boundaries (BBS/OCHA ADM1)[11].

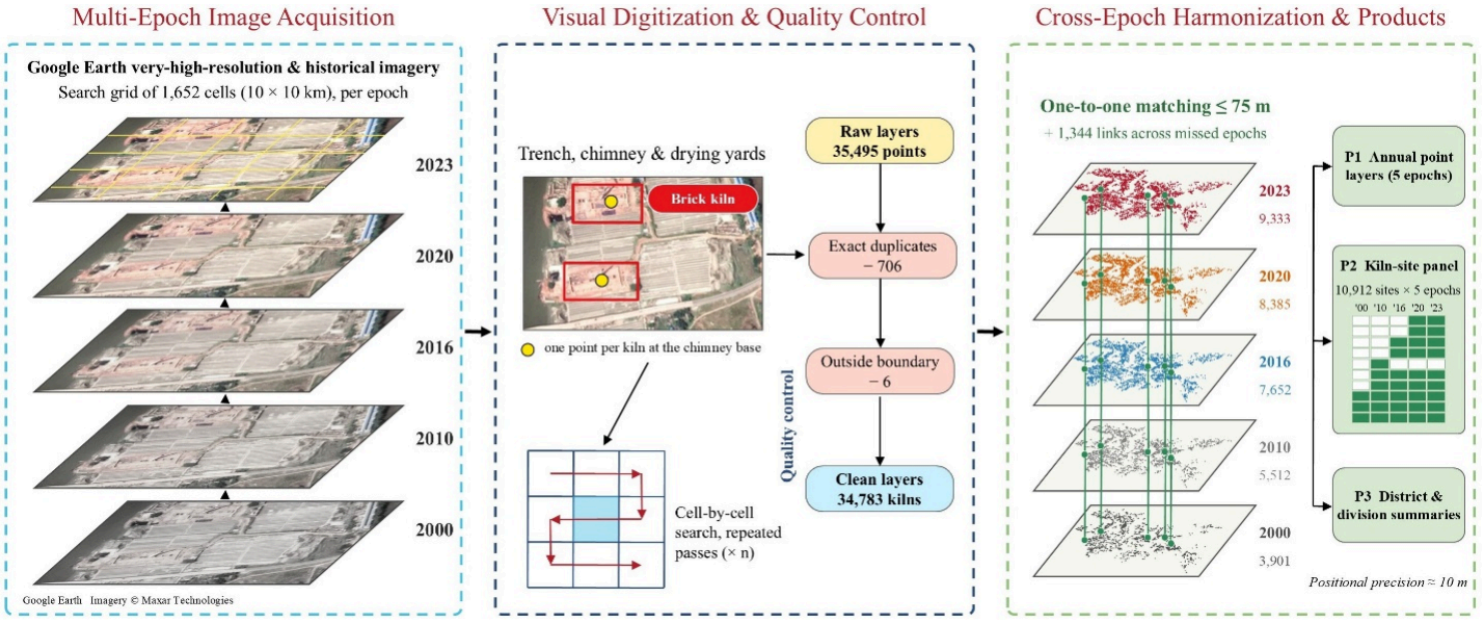


Figure 2

Workflow used to produce the census.

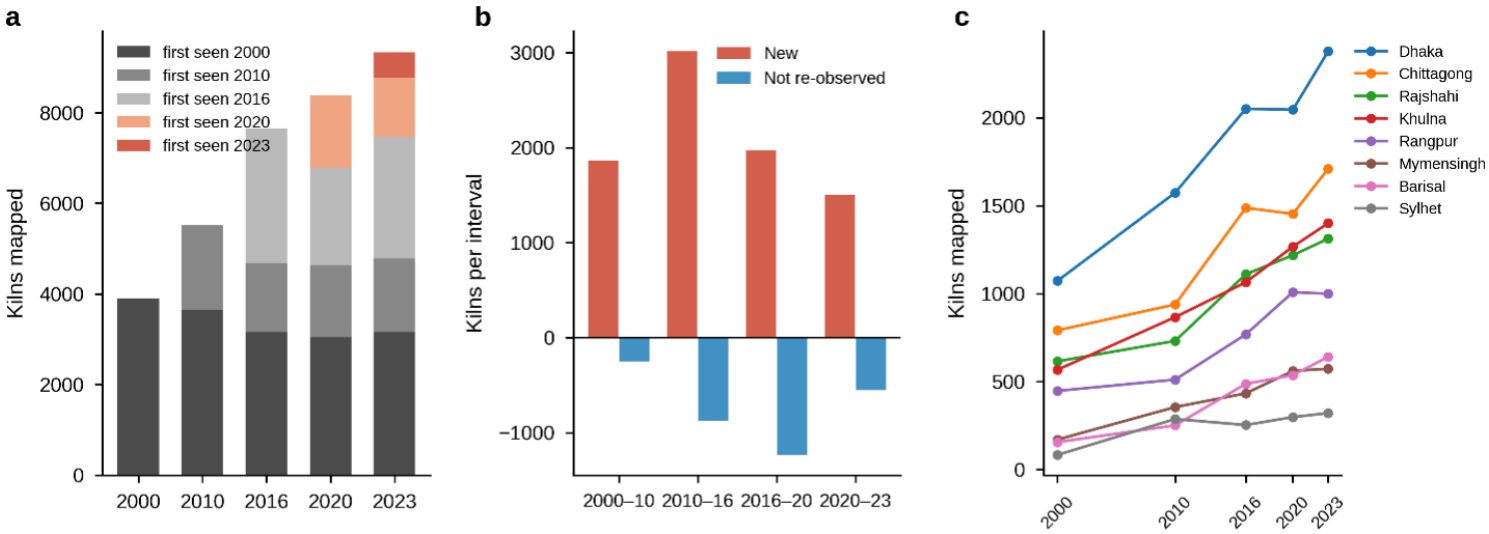


Figure 3

Composition and dynamics of the kiln stock. (a) Kilns mapped at each epoch, colored by the epoch in which the site was first observed. (b) Sites first appearing (new) and sites not re-observed in the following epoch, per interval. (c) Kilns mapped by division.

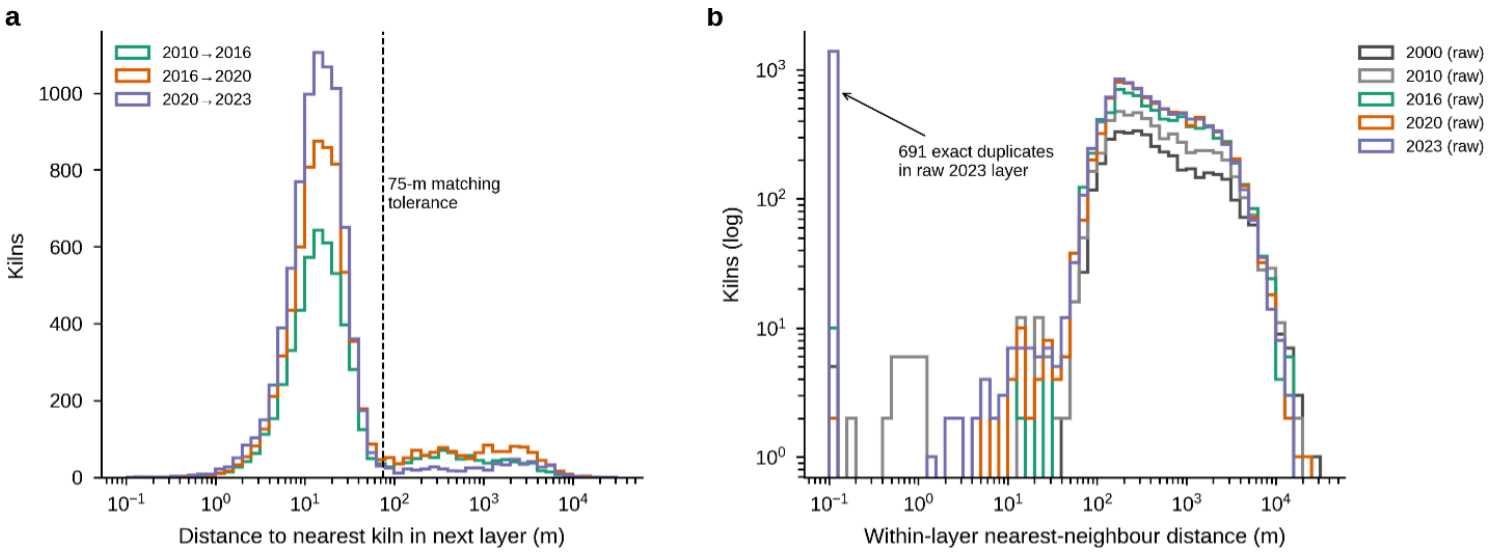


Figure 4

Technical validation. (a) Distance from each kiln to the nearest kiln in the following epoch for the independently digitized layers; the dashed line marks the 75-m matching tolerance. (b) Within-layer nearest-neighbor distances in the raw layers, showing the duplicate records in the raw 2023 layer.