Unequal impacts of urban industrial land expansion on economic growth and carbon dioxide emissions

Check for updates

Cheolhee Yoo [1,2,3], Huijuan Xiao[4], Qing-wei Zhong[5] & Qihao Weng [1,2,3] ✉

Industrial land drives economic growth but also contributes to global warming through carbon dioxide emissions. Still, the variance in its impact on economies and emissions across countries at different development stages is understudied. Here, we used satellite data and machine learning to map industrial land at 30 m resolution in ten countries with substantial industrial value-added, and analyzed the impact of industrial land expansion on economic growth and emissions in 216 subnational regions from 2000 to 2019. We found that industrial land expansion was the leading factor for economic growth and emissions in developing regions, contributing 31% and 55%, respectively. Conversely, developed regions showed a diminished impact (8% and 3%, respectively), with a shift towards other economic growth drivers like education. Our findings encourage developing regions to consider the adverse effects of climate change during industrial land expansion and that developed regions prioritize human capital investment over further land expansion.

Industrial land is a fundamental component of urban landscapes and the primary location for manufacturing facilities, warehouses, and other industrial operations. Industrial land supports the growth of industrial sectors and thus serves as a driving force of economic development[1–3]. Due to the accelerated pace of urbanization and industrialization across the world, the demand for industrial land is increasing at a substantial rate[4]. However, the rapid expansion of industrial land poses threats to the environment related to the associated increases in energy use and fossil fuel consumption[5–9]. Sectors closely related to industrial land use, including energy use in industrial activities and direct industrial processes, accounted for approximately 29.6% of global greenhouse gas emissions in 2016[10], translating to about 14.5 billion tonnes of carbon dioxide (CO_2) equivalent.

Most previous studies on the impact of industrial land on economic growth or CO_2 emissions have been limited to single-country or regional analyses, particularly in China. Many studies focusing on China have demonstrated that industrial land expansion has a positive interaction with economic growth[4,11,12] and is a major influencing factor in CO_2 emissions[13–16]. Regional case studies suggest that the impact of industrial land on economic growth and CO_2 emissions may differ in regions depending on socioeconomic achievements[17,18]. However, there is little understanding of how these impacts differ across multiple countries with varying levels of development over time. One important reason for the insufficiency of multi-country research regarding industrial land expansion is the lack of comparable and consistent datasets to identify industrial land in wide coverage areas.

Another research gap is that studies that concurrently explore the impacts of industrial land expansion on economic growth and CO_2 emissions are scarce. As interest in sustainable development and efficient land use continues to grow rapidly, we need a more comprehensive understanding of the interactions among industrial land expansion, economic growth, and CO_2 emissions. Such understanding can inform effective industrial land management, balancing economic development with minimizing environmental damage.

In this study, we developed a methodology for high-resolution (30 m) mapping of urban industrial land (IND) areas. Utilizing a synergistic approach involving multiple satellite-based datasets and machine learning techniques, we comprehensively mapped IND across ten countries with large industrial value-added —China, the United States, Japan, Germany,

[1]JC STEM Lab of Earth Observations, Department of Land Surveying and Geo-Informatics, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong SAR. [2]Research Centre for Artificial Intelligence in Geomatics, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong SAR. [3]Research Institute for Land and Space, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong SAR. [4]Department of Civil and Environmental Engineering, The Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong SAR. [5]School of Nursing, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong SAR. ✉e-mail: qihao.weng@polyu.edu.hk

India, South Korea, Italy, France, Vietnam, and Bangladesh. This methodological innovation elevates the precision and accessibility of monitoring detailed IND area distribution, marking a notable advancement in industrial land use studies.

We explored the impacts of IND expansion on economic growth and CO_2 emissions among 216 subnational regions in ten countries with varying levels of development, covering a twenty-year period from 2000 to 2019. Expanding our scope beyond just IND areas, we also compared the effects of other factors such as urban nonindustrial land (NIND), education level, population density, and soil properties on economic growth and CO_2 emissions. Our findings indicated that while IND expansion was a crucial driver for economic growth and emissions in developing regions, its contribution diminished in developed regions. We observed a shift in developed regions towards other drivers of economic growth, particularly education, underscoring the importance of investing in human capital. Our study culminated with proposing sustainable industrial land management strategies tailor-made to the specific developmental stage. These suggestions provide crucial guidance for policymakers seeking to harmonize economic growth with environmental considerations, enriching the global dialog on sustainable development and climate change mitigation.

Results

### Satellite-derived IND areas for ten countries

Satellite-derived maps of IND (built-ups used for industrial land use over impervious surfaces) effectively depict the spatial distribution and growth of dense IND regions across ten countries from 2000 to 2019 (Fig. 1a, b). Our mapping analysis revealed that IND areas exhibited a more localized, clustered distribution than NIND (remaining built-ups not classified as IND over an impervious surface) distributions did (Fig. 1 and Supplementary Fig. 1). Specifically, Asian countries such as China and India displayed a considerably broader area with a higher proportion of NIND relative to IND, as evidenced in northern India and northeast and east China. However, the expansion of IND areas in these countries from 2000 to 2019 was larger than that of NIND areas.

During this period, China observed the most substantial increase in IND areas, particularly in coastal regions. We found prominent expansions of IND areas around Shanghai (Fig. 1d), for example, where these areas dynamically expanded through the formation of clusters (e.g., Industrial Park). Conversely, developed countries, such as the United States, Japan, and three European nations, experienced relatively minor changes in IND areas from 2000 to 2019. Even major industrial cities in the United States
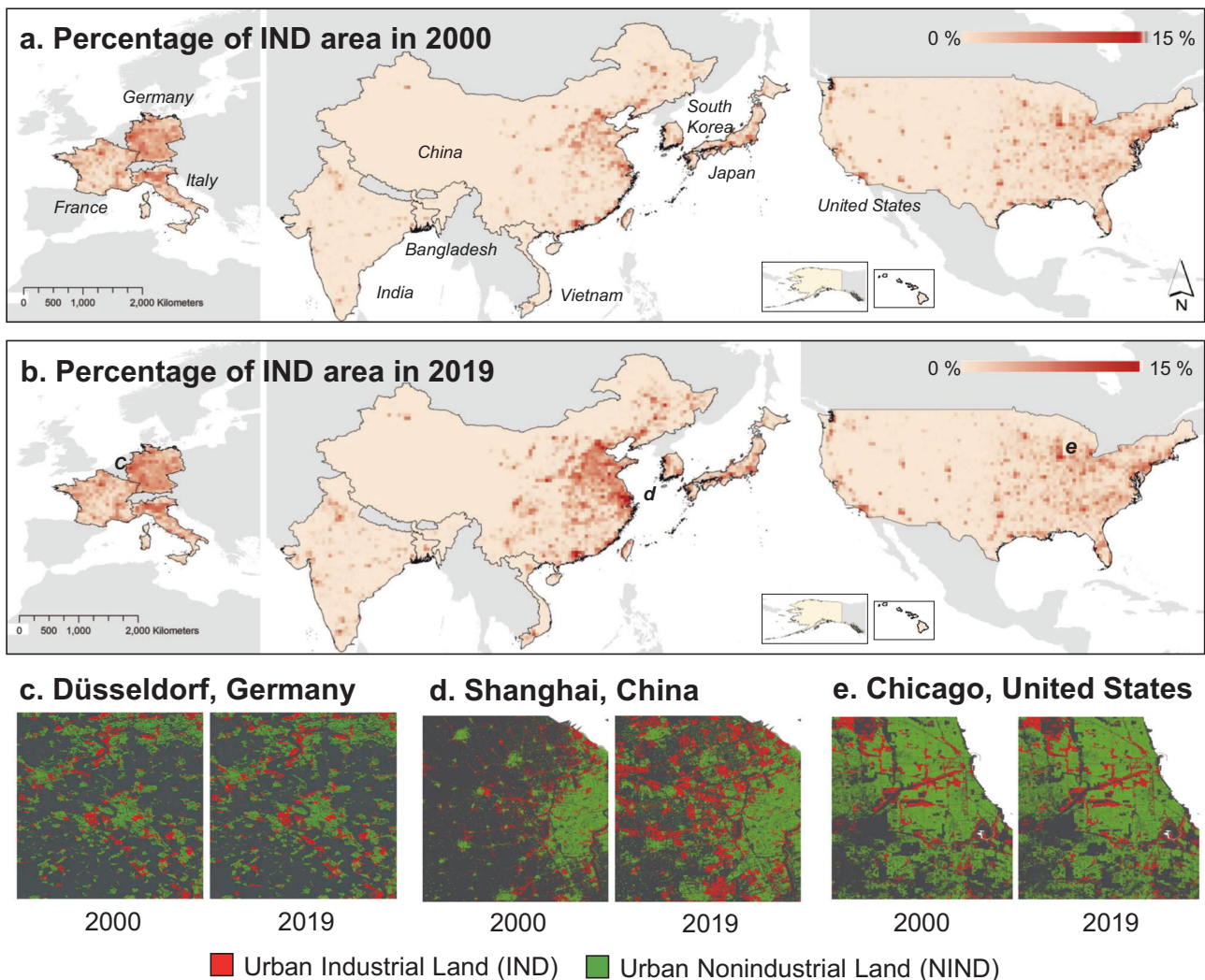
Fig. 1 | Distribution of Urban Industrial Land (IND). a IND area percentage in 2000 for ten countries; (**b**) IND area percentage in 2019 for ten countries. Note: the percentage was calculated by total IND area within a 0.5° grid; (**c**) case example of the distribution of IND and NIND areas in Düsseldorf, Germany (center coordinate as 51.25° N, 6.75° E); (**d**) case example of the distribution of IND and NIND areas in Shanghai, China (grid center coordinate as 31.25° N, 121.25° E); (**e**) case example of the distribution of IND and NIND areas in Chicago, the United States (grid center coordinate as 41.75° N, 87.75° W). Maps were generated by the Google Earth Engine platform (https://earthengine.google.com).
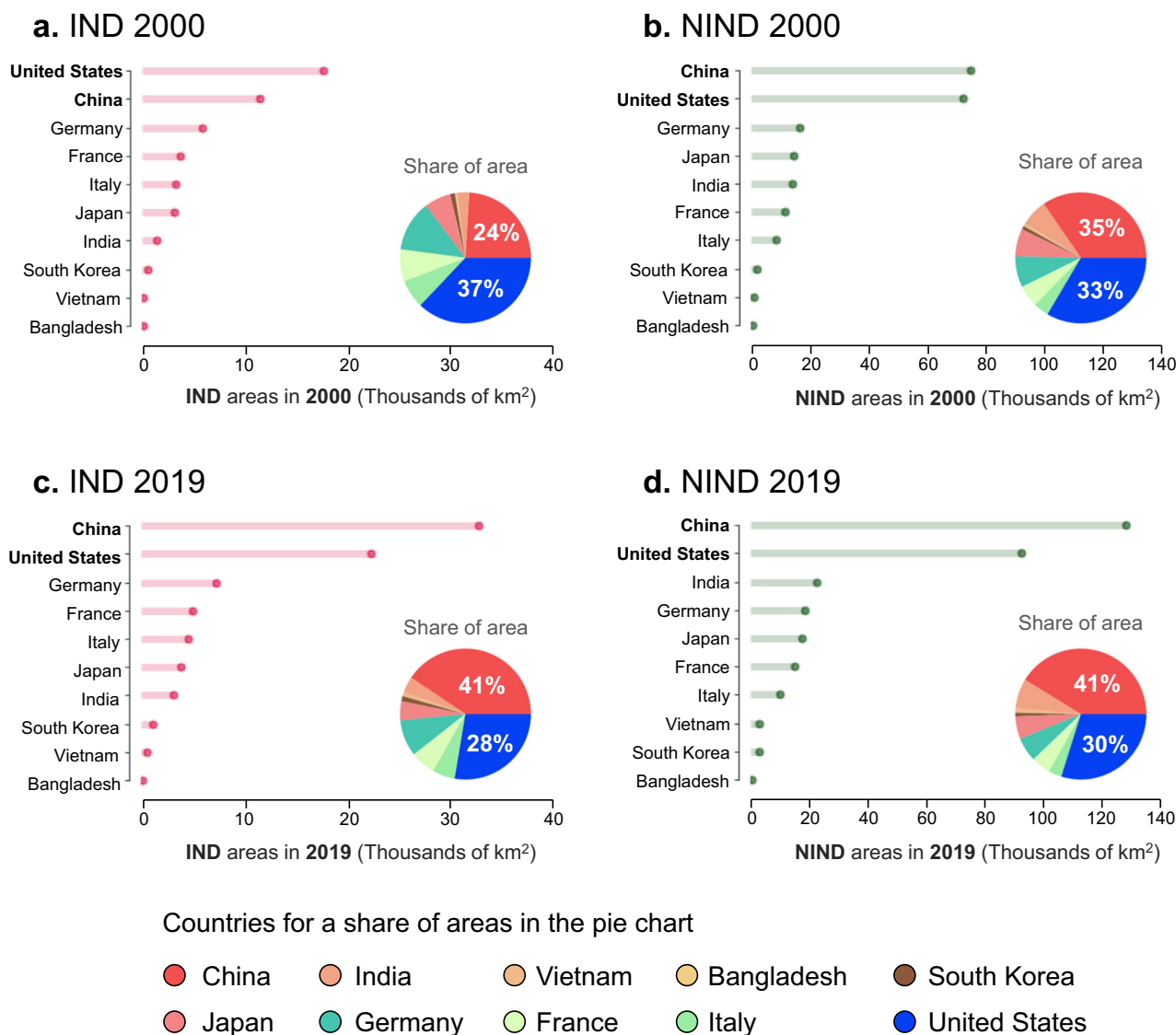
Fig. 2 | Total areas of Urban Industrial Land (IND) and Urban Nonindustrial Land (NIND) for the years 2000 and 2019 across ten countries. a IND area in 2000;
(**b**) NIND area in 2010; (**c**) IND area in 2019; (**d**) NIND area in 2019 for ten countries. The proportion of these areas is depicted as pie charts for each respective year.

and Germany, including Chicago (Fig. 1e) and Düsseldorf (Fig. 1c), exhibited limited IND growth during this time.

Chinese IND areas exceeded that of the United States (Fig. 2). In 2000, the United States had the largest IND area (17,571 km², 37%) and China had the second-largest (11,400 km², 24%), but by 2019, China's IND area had expanded to 32,868 km² (41%) and surpassed the United States' 22,366 km² (28%). Regarding NIND areas, China maintained the highest percentage (2000: 74,930 km², 35%; 2019: 127,843 km², 41%), followed by the United States (2000: 72,440 km², 33%; 2019: 92,506 km², 30%). The area gap between China and the United States consistently expanded throughout the study period. Among the ten countries analyzed, China, India, Vietnam, and South Korea—all in the Asian region—experienced increased proportions of both IND and NIND areas.

Spatial and temporal associations of per capita IND, GDP, and CO_2 in subnational regions

The spatial distributions of per capita IND correlated closely with the spatial distribution of per capita gross domestic product (GDP) across subnational regions in the ten countries (Fig. 3a, b). Specifically, the per capita IND maintained a strong spatial association with per capita GDP, presenting correlation coefficients of 0.90 for 2000 and 0.88 for 2019 (Supplementary

Fig. 2a, b). This pronounced spatial association was driven primarily by the higher values of per capita IND and GDP observed in regions within developed nations (the United States, three European countries, South Korea, and Japan). Conversely, regions within developing nations (China, India, Bangladesh, and Vietnam) showed lower comparative values. From 2000 to 2019, a considerable elevation in per capita GDP, particularly in regions within developing nations, corresponded with large growth in per capita IND (Fig. 4a, b). This link was further affirmed by the remarkably high temporal correlation between per capita IND and per capita GDP over the twenty years. A trend was particularly evident in most developing regions whose average correlation coefficient surpasses 0.8 (Supplementary Fig. 3a). Regions with high per capita IND values also exhibited greater per capita CO_2 emissions (Fig. 3c). The correlation coefficient for per capita IND and CO_2 emissions began at 0.91 in 2000 and then decreased slightly to 0.84 by 2019 (Supplementary Fig. 2c, d). From 2000 to 2019, notable surges (exceeding 100%) in per capita CO_2 emissions were clearly observable in the regions within developing nations (China, India, Vietnam, and Bangladesh; Fig. 4c). The temporal correlation between per capita IND and per capita CO_2 was notably high within these developing regions, where the average correlation coefficient exceeds 0.8 (Supplementary Fig. 3b).
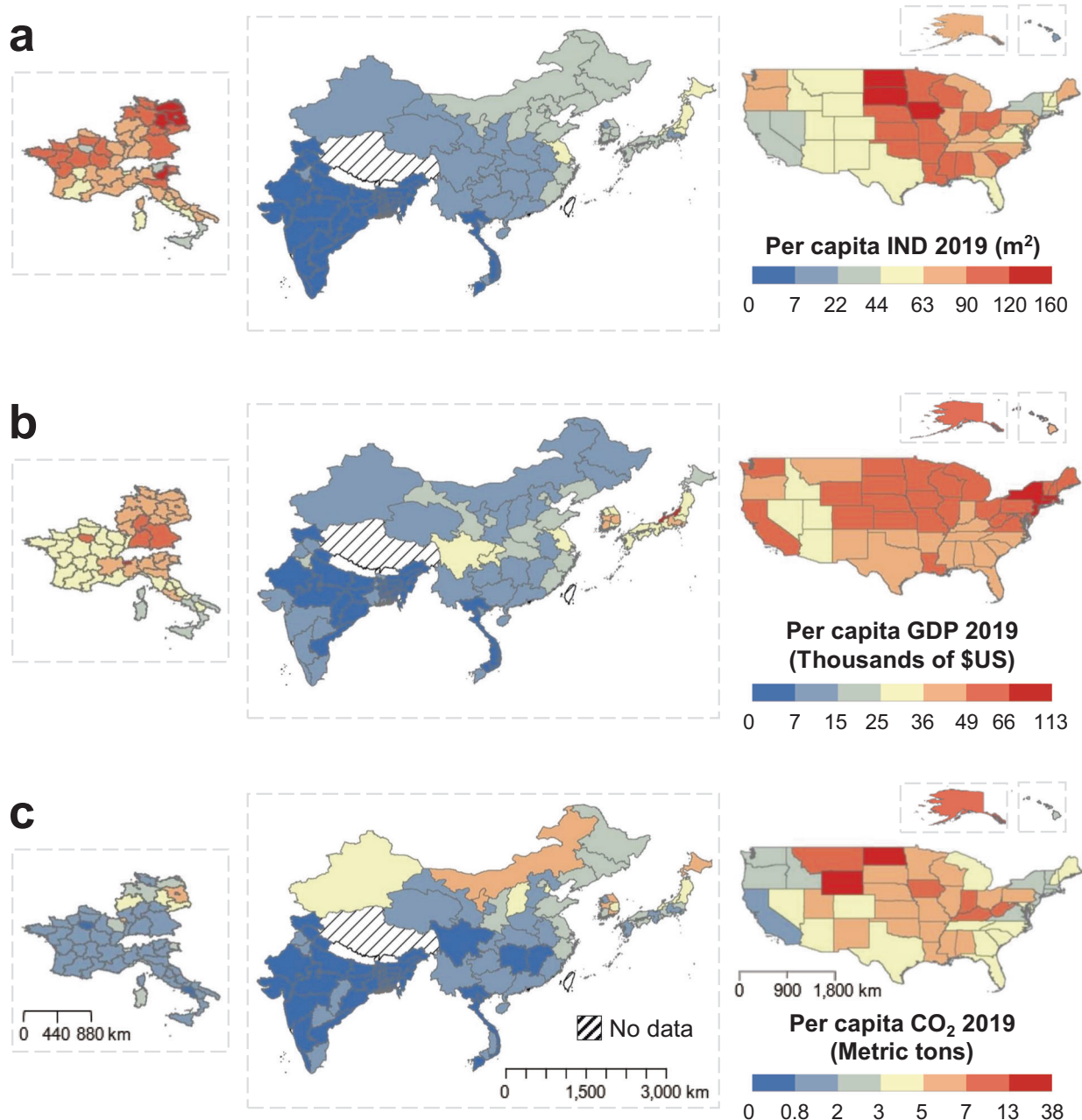
Fig. 3 | Distribution of per capita IND, GDP, and CO_2 for subnational regions. a Per capita IND in 2019; (**b**) per capita GDP in 2019; and (**c**) per capita CO_2 in 2019 for 216 subnational regions in ten countries. Note: the hatched area was excluded from the analysis due to a lack of data. Maps were created using ArcGIS v10.4.1 software.

In contrast, many regions within developed countries (the United States and the European countries) demonstrated a declining trend in per capita CO_2 emissions; changes fell below zero percent and in certain regions plunged below −25%. Many of these developed regions displayed a negative temporal correlation between per capita IND and per capita CO_2 over the two-decade span, with an average correlation coefficient of −0.39 (Supplementary Fig. 3b). Considering the generally positive changes in per capita IND and GDP in these regions, we can infer that developed regions made considerable efforts to mitigate CO_2 emissions from 2000 to 2019[19,20].

The impact of per capita IND on economic growth and CO_2 emissions

Our results, derived from a Mixed Effect Random Forest (MERF) machine learning model, revealed unequal contributions of per capita IND to per capita GDP, depending on the level of development across regions (Fig. 5a, c). The contribution of per capita IND to economic growth was larger in

developing regions (approximately 31%) than in developed regions (approximately 8%). In developing regions, per capita IND exerted the most pronounced impact on per capita GDP, followed by educational level and per capita NIND. Conversely, in developed regions, the most substantial impact came from the educational level, which accounts for approximately 35% of the influence. This suggests that the urban land level (for both IND and NIND) is less markedly associated with the economic level in developed regions than in developing regions.

In developing regions, the SHapley Additive exPlanations (SHAP) plot revealed that dominant contributing factors such as per capita IND and education level exhibit a pronounced positive correlation with economic level (Fig. 5b, d). This suggests that larger input factor levels (represented by the color pink in the plot) are associated with substantial economic growth. But in developed regions, factors other than the education level showed less or nondirectional influence (samples were clustered around zero SHAP value).
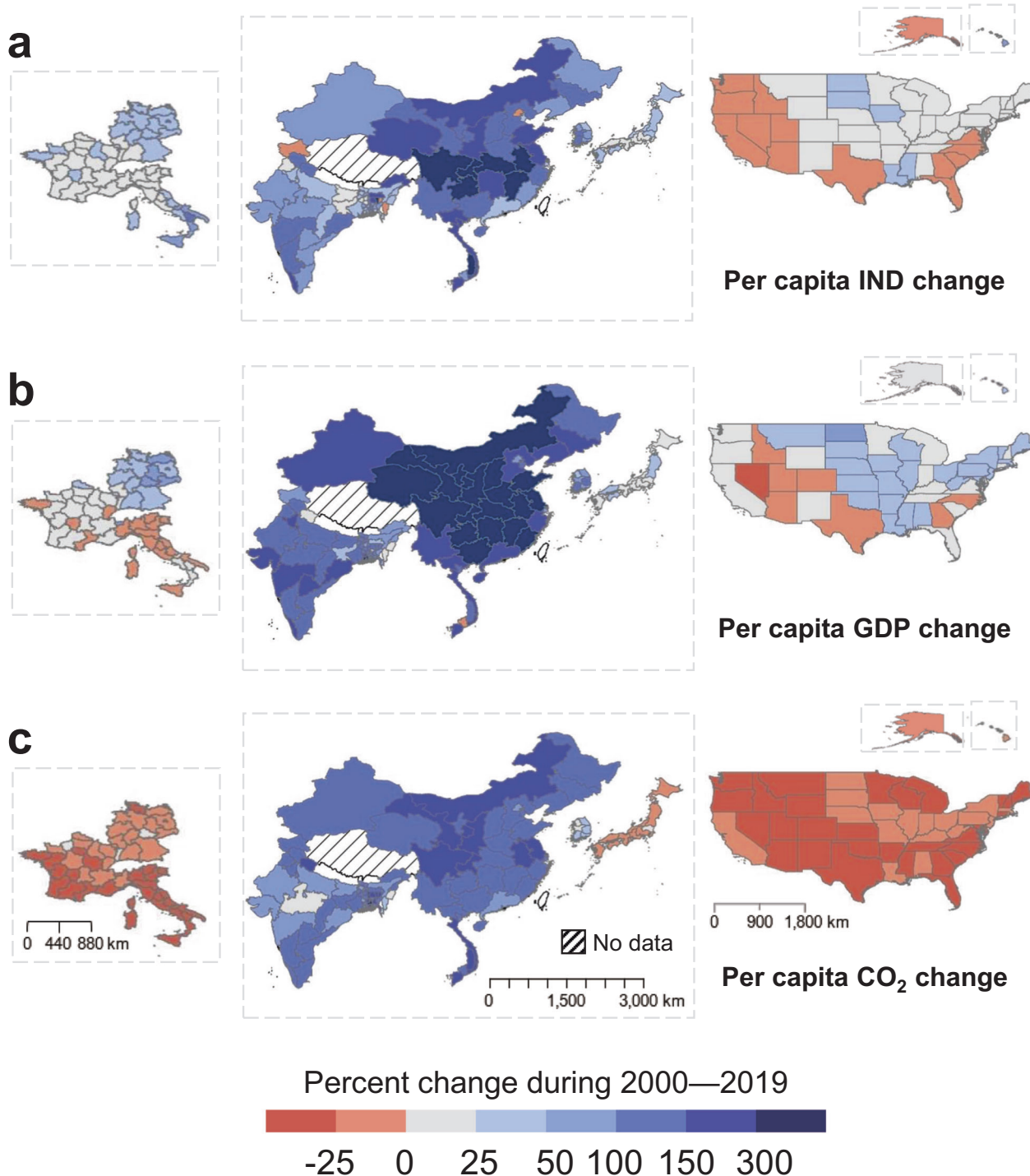
Fig. 4 | Percent change in per capita IND, GDP and CO$_2$ during 2000–2019 for subnational regions. a Percent change in per capita IND; (**b**) percent change in per capita GDP; and (**c**) percent change in Per capita CO$_2$ during 2000 to 2019 for 216 subnational regions in ten countries. Note: the hatched area was excluded from the analysis due to a lack of data. Maps were created using ArcGIS v10.4.1 software.

In developing regions, the impact of per capita IND on CO$_2$ emissions was overwhelmingly high (Fig. 6a). In developed regions, however, the impact of per capita IND was not as obviously evident (Fig. 6b). For developing regions, the impact of per capita IND (55%) presented a very distinct difference compared to other driving factors—such as per capita NIND (10%) and air temperature (8%)—all showing an impact of lower than 10%. However, in developed regions, population density emerged as the most influential impacting factor at 35%, followed by

soil pH at 22%, with the impact of per capita IND reduced to a mere 3%.

We observed an apparent positive direction for the impact of per capita IND on CO$_2$ emissions in developing regions (Fig. 6c). In contrast, some regions with larger per capita IND were found to produce less per capita CO$_2$ emissions (See Fig. 6d where the high feature values of per capita IND, represented in pink, appear under zero SHAP value). This could be attributed to the notable decreasing trend in per capita CO$_2$ from 2000 to 2019 observed in developed regions (Fig. 4d).

Article

Developing regions (a–b)
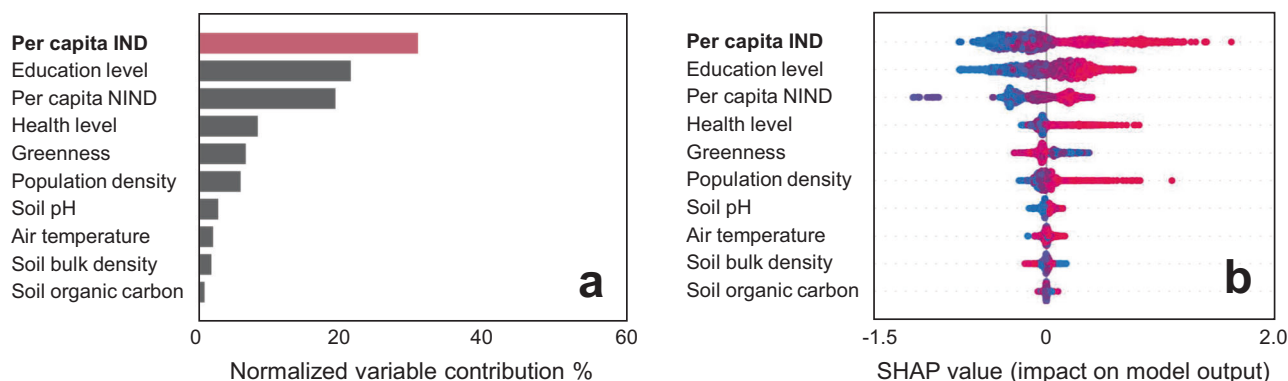
Developed regions (c–b)
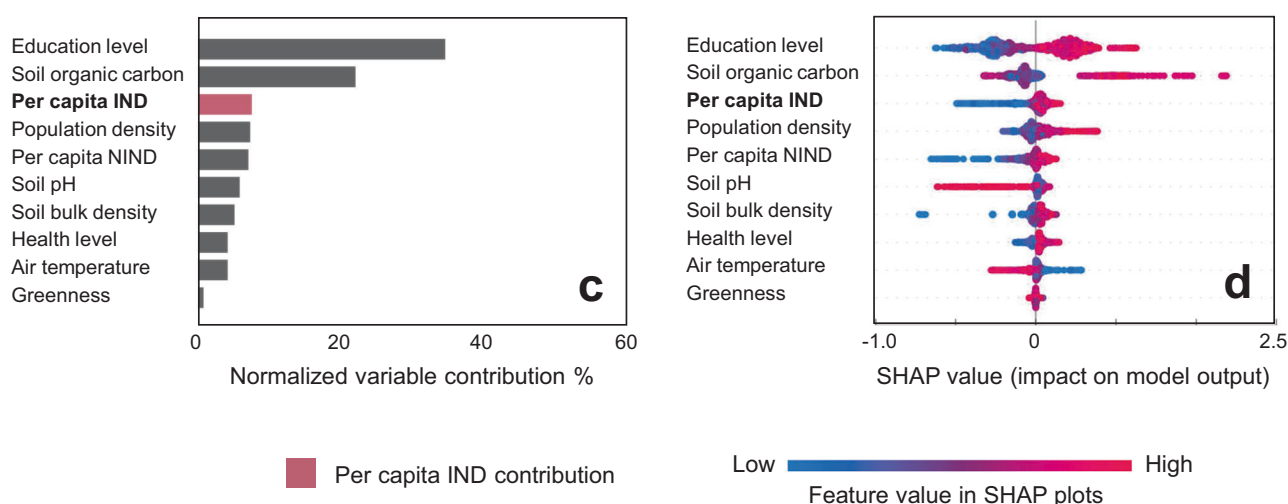
Fig. 5 | Impact of driving factors on per capita GDP in different development regions. a Normalized variable impact for per capita GDP in developing regions; (**b**) SHAP value of each driving factor for modeling per capita GDP in developing regions; (**c**) normalized variable impact for per capita GDP in developed regions; (**d**) SHAP value of each driving factor for modeling per capita GDP in developed regions. The magnitude of the variable impact was calculated using absolute SHAP values, which were then normalized as a percentage share of the total impact of all driving factors. In the SHAP plot, the color indicates the feature values from low (blue) to high (red). Each data point represents one SHAP value for a given prediction and feature. The distribution of red and blue dots provides the directionality of each feature's impact.

In both developing and developed regions, the impact of per capita IND was higher than that of per capita NIND on CO_2 emissions. The difference was more than 45% for developing regions and about 2% for developed regions. For developed regions, the impact of per capita NIND was contributing only 1.5%. These were consistent with those of per capita GDP modeling (Fig. 5), given that per capita IND also exhibited a higher impact on economic growth for developing and developed regions than per capita NIND does.

We conducted a further sensitivity analysis through a leave-one-region-out modeling approach for both developed and developing regions, in terms of per capita GDP and per capita CO_2 modeling (Supplementary Fig. 4). In developing regions, per capita IND consistently surfaced as the preeminent variable for both per capita GDP and CO_2 across all models in this analysis (Supplementary Fig. 4a, c). For developed regions, per capita IND held the third rank in contribution to per capita GDP, and the sixth for per capita CO_2, a ranking that aligned with the original model's results (Supplementary Fig. 4b, d). The median contribution values of per capita IND in the sensitivity analysis mirrored those of the original model encompassing all regions (Figs. 5a–c, 6a–c).

Discussion

Satellite-based historical IND maps show that China, a representative developing economy, exhibited the most dramatic increase in IND from 2000 to 2019 among the ten countries. Several factors could account for China's rapid IND expansion, such as rapid economic growth, low cost of land and labor, and government policies to promote manufacturing[7,21]. In contrast, mature industrial economies in developed countries (the United States, Japan, and three European nations) demonstrated relatively little change in their IND areas over the same period. Slower IND growth in these economies reflects structural economic changes, such as the shift toward service sectors, the offshoring of labor-intensive manufacturing, and the redevelopment of existing industrial zones, and limiting IND expansion[22–24].

Our findings highlight the importance of IND in driving economic growth, especially in developing regions. IND encompasses manufacturing, production, and logistical operations that directly bolster economic output by generating income, creating employment opportunities, and producing essential goods for economic growth[25,26]. The expansion of IND facilitates the transition of labor from less productive agricultural sectors to more productive manufacturing and industrial positions[21]. This "structural

Developing regions (a–b)
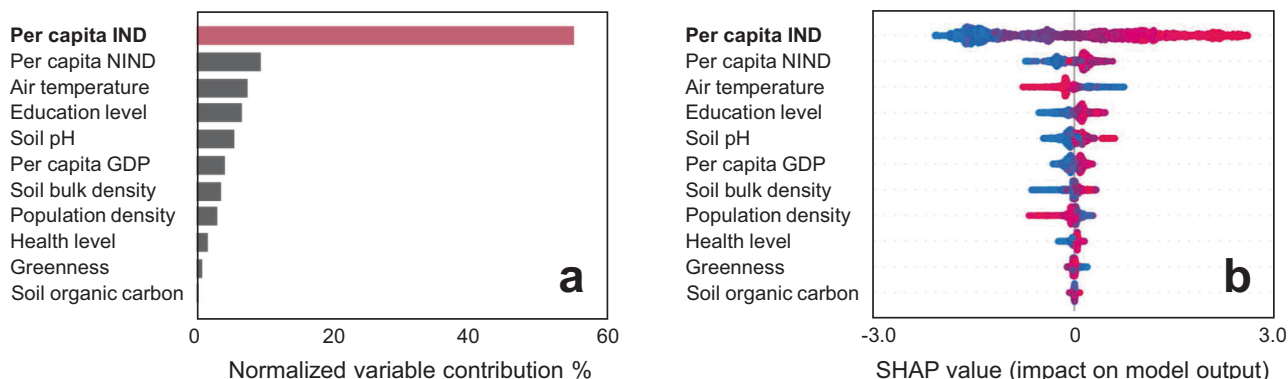
Developed regions (c–b)
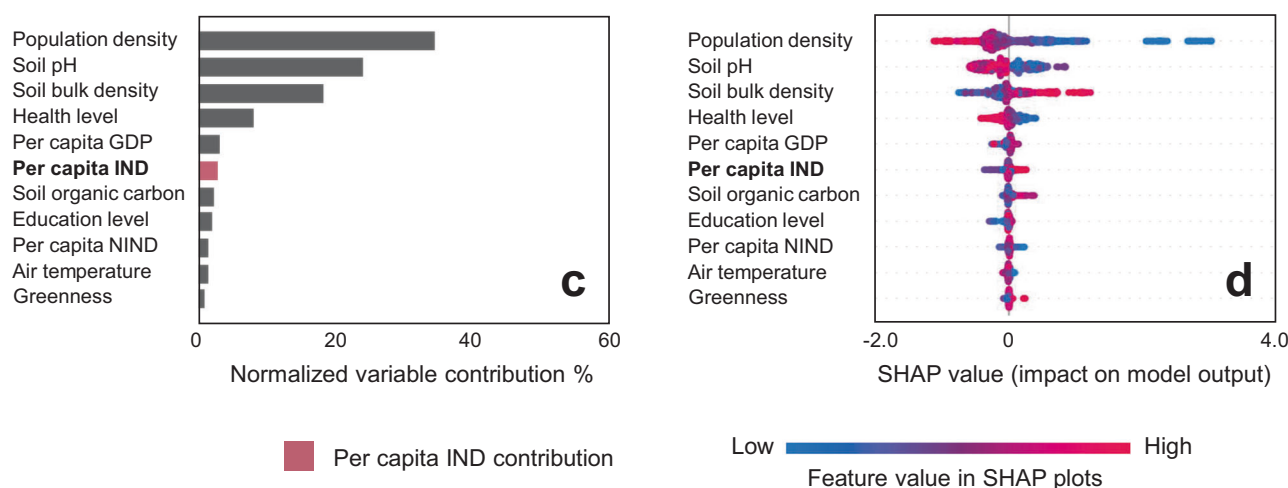
Fig. 6 | Impact of driving factors on per capita CO$_2$ in different development regions. a Normalized variable impact for per capita CO$_2$ in developing regions; (**b**) SHAP value of each driving factor for modeling per capita CO$_2$ in developing regions; (**c**) normalized variable impact for per capita CO$_2$ in developed regions; (**d**) SHAP value of each driving factor for modeling per capita CO$_2$ in developed regions. The magnitude of the variable impact was calculated using absolute SHAP values, which were then normalized as a percentage share of the total impact of all driving factors. In the SHAP plot, the color indicates the feature values from low (blue) to high (red). Each data point represents one SHAP value for a given prediction and feature. The distribution of red and blue dots provides the directionality of each feature's impact.

transformation" exerted a greater impact on GDP growth in developing regions, where a larger proportion of the workforce is initially engaged in agriculture and other activities in the primary production sector.

We note, however, that per capita IND also produced a substantially high impact on CO$_2$ emissions in developing regions. Developing regions generally possess more carbon-intensive industrial sectors and less efficient production processes, whereas developed regions have transitioned away from emissions-intensive industries and toward more advanced technologies[27,28]. By examining the decoupling relationship between per capita GDP and per capita CO$_2$ from 2000 to 2019 (Supplementary Fig. 5), we observed that strong decoupling states—economic growth while decreasing CO$_2$ emissions—predominate in developed regions (the United States, several European countries, and Japan). This supports the notion that developed economies can achieve economic growth from industrial activity without a commensurate increase in CO$_2$ emissions.

Educational level appeared to be the most influencing factor for economic growth in developed regions (Fig. 5c). A population with a high level of education embodies a more substantial human capital, which is a crucial driver of economic development[29]. Further, education cultivates innovation

by equipping individuals with the necessary skills and knowledge to pioneer new technologies. This can lead to new industries in developed regions, serving as a potent engine for economic growth.

Population density also emerged as a major influence on CO$_2$ emissions in developed regions. For developed regions, the impact of population density was much larger than that of per capita IND on per capita CO$_2$. In developed regions in which urban expansion has reached saturation, population dynamics appeared to predominantly dictate per capita CO$_2$; this underscores the critical role of demographic changes in shaping the complex interplay between land use and environmental repercussions[30,31].

In developed regions, particular soil properties have proven crucial contributing variables, specifically soil organic carbon for per capita GDP, and soil pH for per capita CO$_2$. This could be linked to the spatial distribution of soil properties: regions rich in organic carbon generally display higher per capita GDP, while regions with lower soil pH tend to correlate with higher CO$_2$ emissions, as indicated by SHAP values (Figs. 5d and 6d). This aligned with previous studies suggesting that soil acidification, potentially leading to lower pH, could be induced by increased atmospheric CO$_2$ levels[32]. Furthermore, soils enriched with organic carbon often drive

high agricultural productivity, thereby contributing markedly to economic growth[33].

A key methodological contribution lies in the extensive temporal and spatial analysis of IND area changes across both developed and developing regions. This was accomplished by directly generating high-resolution (30 m) IND areas for ten countries, utilizing a combination of satellite-based multiple datasets and machine learning techniques. Previous research has relied on national statistical data or point-source–based data[12,14,17], which inherently limited their scope to highly localized areas, such as individual countries. This study, by contrast, constructed spatially continuous IND areas at a high resolution, allowing us to trace IND distribution and its expansion over two decades across 216 subnational regions in ten countries. This led to the crucial discovery of an inequality in the contribution of industrial land among development levels across a multitude of regions.

Another noteworthy innovation in this study is our application of machine learning–based MERF for longitudinal modeling. This enabled us to model the complex impacts among variables while addressing the multicollinearity between factors[34]. Furthermore, the SHAP values for each input variable facilitated us to identify the directionality of various socioeconomic and environmental variables on economic growth and CO_2 emissions.

Compared to conventional linear mixed-effects models (LMM) in longitudinal analysis (see the Supplementary Note 1), MERF exhibited superior performance in both developed and developing regions for per capita GDP and CO_2 (Supplementary Table 1). In developed regions, for instance, MERF yielded an R-value of 0.987 and a root mean square error (RMSE) of 2.26 for per capita GDP modeling. In contrast, LMM produced an R-value of 0.965 and a higher RMSE of 3.64, thus highlighting the lower error rate associated with MERF. Moreover, the sensitivity analysis of the MERF model affirms its reliability (as shown in Supplementary Fig. 4), emphasizing its proficiency in robustly discerning the variegated impacts of industrial land.

This study is the first to separate IND and NIND areas from urban land and to analyze their impacts on both economic growth and CO_2 emissions. Compared to previous studies which treated urban land as a single built-up area[31,35], the use of IND and NIND as separate factors in our longitudinal model leads to higher modeling accuracy (Supplementary Table 1). In particular, using the urban land area as a single factor to interpret the impacts on economic growth and CO_2 emissions (Supplementary Fig. 6) revealed the potential of overestimating the impact of the NIND area.

We carried out an extensive analysis of the expansion of IND areas by using long-term impervious surface data to map their distribution in 2019 and simulate historical changes (2000–2019). However, this method did not address the demolition of existing IND or the conversion from IND to NIND. Certain input variables required for mapping IND distributions—synthetic aperture radar (SAR) data, nighttime light data, and local climate zone (LCZ) maps—were available only in recent years. This limited the development of historical IND mapping models on an annual basis, particularly for years prior to 2010.

In the longitudinal study, we compiled numerous variables at the subnational level across ten countries from 2000 to 2019. Potential errors existed in each variable. Our IND and NIND data, for instance, had an accuracy of 91% (see the method detail), so there is still room for improvement. The WorldPop population data, employed in our per capita calculations, could also introduce errors, stemming from source data quality and assumptions made during the modeling process[36]. Other variables, such as education and health levels, might contain errors due to their dependence on statistical data. Likewise, observation-based data, such as satellite Normalized Difference Vegetation Index (NDVI) or soil properties, possess inherent measurement inaccuracies. Lastly, the gridded GDP data and gridded CO_2 emissions data possessed uncertainties due to errors in interpolation and extrapolation (GDP[37]), and discrepancies linked to emission factors and point source estimates (CO_2[38]).

Future research could broaden its geographic and temporal scope by including a more diverse set of countries, particularly those with emerging economies or smaller industrial bases. Moreover, should data become available, extending the analysis to cover the period beyond 2020 would enable us to capture the latest trends and dynamics in industrial expansion and its economic and environmental impacts.

Further studies should also consider the impact of IND expansion within individual regions on the economic growth of other regions (e.g., offshoring). The economic growth of the United States, for example, is assisted by US companies offshoring parts of their operations to Asian countries[39]. By investigating the effects of IND expansions on economic growth in the context of multinational corporations and foreign investment, the analysis can offer a more thorough understanding of the intricate relationships between regional developments and worldwide economic dynamics.

The framework proposed in this study can be extended to evaluate the impact of industrial land on other environmental factors, such as air pollutants. We further attempted to apply our framework to model sulfur dioxide (SO_2) emissions and fine particulate matter ($PM_{2.5}$) concentrations (refer to Supplementary Note 2 for details). Notably, per capita IND did not appear as a major contributor to SO_2 and $PM_{2.5}$ in either developing or developed regions, contrasting with its notable influence on CO_2 emissions in developing regions (Supplementary Figs. 7 and 8). CO_2 emissions primarily emanate from the combustion of fossil fuels, making CO_2 the most consequential greenhouse gas emitted from industrial activities, as compared to SO_2 and $PM_{2.5}$. Future research could deepen this explanation by potentially integrating insights into the health impacts and climate change consequences of industrial land expansion.

Conclusions

We mapped IND areas for ten industrialized countries from 2000 to 2019 and compared the impact of per capita IND on economic growth and CO_2 emissions across subnational regions with different development levels. The main finding is that industrial land exerted an unequal impact on economic growth and CO_2 emissions across varying levels of development. Specifically, in developing regions, a large and positive directional impact from the IND level for economic growth and CO_2 emissions was observed. This impact was diminished in developed regions, with no discernible trends in CO_2 emissions. Another key finding is that the level of education emerges as the primary driver of economic growth in developed regions. This could be associated with the emphasis on human capital investment in these advanced stages of development. We also observed that per capita IND surpassed that of per capita NIND in economic growth and CO_2 emissions, regardless of a region's development status.

This study demonstrated that while IND promoted economic growth, developing regions continued to experience substantial CO_2 emissions from IND. Policymakers should strive for a balanced approach between fostering economic growth and mitigating emissions when developing sustainable IND management strategies. In developing regions, the facilitation of IND expansion can enhance productivity, job creation, and economic growth. Alongside these efforts, policy interventions must address environmental consequences by prioritizing cleaner energy sources and low-carbon technology to reduce CO_2 emissions from IND activities[40,41]. Meanwhile, in developed regions, especially where CO_2 emissions have decoupled from economic growth, the optimization of existing industrial land use and the promotion of technical innovation driven by human capital can maximize output gains[42]. When multinational corporations from developed countries build IND facilities in developing regions, they should conduct thorough environmental impact assessments and implement measures to minimize local emissions. Enabling the transfer of best management practices and clean technologies for IND operations from developed to developing regions can improve land use efficiency, lower emissions, and support global carbon neutrality goals. Government policies and international development initiatives should actively encourage the sharing of such knowledge and technology.

Methods

Study area

Ten countries, such as China, the United States, Japan, Germany, India, South Korea, Italy, France, Vietnam, and Bangladesh, were selected for the study area. Excluding Bangladesh and Vietnam, the eight countries are the top eight for a share of global manufacturing output based on the United Nations Statistics Division in 2019[43]. We chose Bangladesh and Vietnam because they achieved rapid industrial development and are Asia's most promising textile-producing nations[44]. The selected study period for the analysis was from 2000 to 2019, considering the availability of the data in use. In addition, this period is notable because it was found that economic growth contributed to global urban expansion at a higher rate after the year 2000 than pre-2000[31].

The IND mapping framework

We devised a framework to map the IND areas for the ten countries in 2000 and 2019 at 30 m spatial resolution. First, the IND map for the year 2019 was produced using multisource satellite-based datasets and machine learning. Then, the distribution of IND areas from 2000 to 2019 was extracted from historical impervious surface data. The overall flow of the framework is shown in Supplementary Fig. 9.

Human settlement built-ups and impervious surfaces, which are fundamental to defining the spatial distribution of IND and NIND areas, were obtained from the World Settlement Footprint (WSF) and global impervious surface area dataset (GISA), respectively. The WSF provided a binary masked representation of human settlement built-ups on a global scale and was generated by the German Aerospace Center (DLR). We obtained WSF2019, which was produced at a spatial resolution of 10 m for the year 2019 using multitemporal Sentinel-1/2 imagery. To delineate historical impervious surfaces, we used the most recent version of the GISA dataset, GISA2.0, which was generated based on Landsat imagery covering the period from 1972 to 2019, with a spatial resolution of 30 m. GISA 2.0 is known to be the most accurate and stable compared to the existing historical global urban datasets[45]. To construct reference data for modeling the identification of IND and NIND regions, we utilized the OpenStreetMap (OSM) land use data, which was downloaded in Polygon shapefile format for ten countries.

The area of WSF2019 located on the impervious surface was extracted using GISA2.0. Then, human settlement built-ups over impervious surface pixels distributed within the polygon boundary corresponding to the industrial category of OSM land use data were selected, and IND samples were constructed by converting them into point data. In the case of NIND, OSM land uses polygons corresponding to non-industrial classes under the 'developed' land category (e.g., commercial, residential, and institutional) were extracted. The human settlement built-ups over impervious surface pixels distributed within the non-industrial polygon boundary were extracted and converted into points to obtain NIND reference samples.

Several multi-source data, mainly satellite data that can represent the characteristics of industrial land, were used as input data for IND mapping. A total of six surface reflectance (three visible bands (blue, green, and red), one near-infrared band, and two short-wave infrared bands) at 30 m spatial resolution were used from the Landsat-8 surface reflectance data. This surface reflectance was widely used as major input data for satellite-based land cover and land use mapping studies due to the unique characteristics of each wavelength band. Landsat-8 surface temperature, which is 30 m resolution resampled from 100 m data, was also used as input data to account for the relatively large heat emission characteristics of industrial buildings[46]. The USGS Landsat-8 Level 2 collection was used to obtain the surface reflectance and temperature, and all images available covering study regions during the period 2019-2020 were acquired.

We used the SAR data as an input feature, which is known as a suitable source of building feature extraction[47]. VV and VH dual polarizations with 10 m spatial resolution were obtained from Sentinel-1 SAR Ground Range Detected (GRD) data. Considering that some industrial buildings are actively operated at night, satellite-based nighttime light was also considered

as input data. Monthly average radiance composites using nighttime data from the Visible Infrared Imaging Radiometer Suite (VIIRS) Day/Night Band (DNB) with 500 m spatial resolution were used. All available data for the period 2019-2020 were obtained for the Sentinel SAR and VIIRS DNB, respectively.

LCZ was also selected as an input feature based on its standard for characterizing urban landscapes. The LCZ consists of 17 classes where ten classes for distinguishing urban forms based on building height and density and seven for natural type. An open-source global LCZ with 100 m resolution produced by machine learning on a 2018-year basis was obtained[48]. We used some human impact variables, such as population density and human modification measures, as input data. The population density was selected with the assumption that the people are relatively less distributed in the industrial area than in the residential and commercial areas. WorldPop data with 100 m spatial resolution was used for the population density. WorldPop is produced by disaggregating the admirative unit's population data through machine learning, and data has been provided on a yearly basis from 2000 to the present. We also used the global Human Modification dataset (gHM) data, which can cumulatively measure how much terrestrial lands have been modified due to human activity, including energy production and electrical infrastructure. gHM is a 1 km resolution data provided by Conservation Science Partners, produced in 2016 by integrating 13 individual datasets, including electrical infrastructure, mining, and energy production. The gridded gHM has a value between 0 and 1 by the proportion of human modification at each pixel.

Since Landsat-8 surface reflectance and surface temperature are vulnerable to the cloud, only scenes under a 30 % cloud ratio were selected from all images from 2019 to 2020. Cloud areas were masked using the quality assessment information. One single image for each band was obtained by median calculation from the remaining clear-sky scenes for the six surface reflectance.

The surface temperature could vary largely even for the same type of industrial buildings due to geographical and weather conditions. To reduce this effect, we extracted the temperature value only for the impervious surface, and the min-max normalized over the selected areas for each clear-sky scene. Then, the median produced one single surface temperature image for the normalized images.

Furthermore, for a total of seven Landsat-8 multispectral bands, we devised a method of extracting the surrounding information of each pixel location and using it as an additional input feature, considering the characteristics of industrial buildings, mainly clustering[11]. Here, the mean, maximum, minimum, and standard deviation values of the neighboring pixels were obtained using a circle-shaped kernel. At this time, 5, 10, 15, and 20 pixels were set as candidates for the size of the radius of the kernel, and the ideal size was selected for each study area to learn through empirical experiments. A total of 28 neighboring information layers from seven Landsat-8 multispectral bands (four for each band) with 30 m resolution were constructed as input features.

The median values of all available images in 2019-2020 were obtained for the VV and VH bands from Sentinel SAR and monthly average VIIRS DNB data. The population of 2019 was chosen from the yearly WorldPop. Because LCZ and gHM are single-period data, they were used as input data without specific period selection. All input features were resampled to 30 m, which is the target spatial resolution in this study.

We developed a separate model for each subnational region of ten countries. We adopted the subnational region delineation provided by the Global Data Lab (GDL)[49]. Such administrative divisions may include states, provinces, or regions, depending on the country's administrative structure. Here, some of the original subnational regions were merged with adjacent regions because the number of references to IND samples was insufficient (described in Supplementary Fig. 10). The training and test sets were separated from the entire reference data. If samples included in the same OSM polygon are used for training and testing, overestimates of accuracy may occur[50]. To deal with this, we randomly divided the OSM polygons into 8:2 for IND and NIND, respectively. Then the reference samples originating

from the polygons corresponding to 80% were used as training sets, and samples from the polygons corresponding to 20% were used as test sets. For the increasing modeling efficiency, only 10,000 IND samples and 10,000 NIND samples were extracted from each unit's full training set as the final training samples. For the final test samples, a total of 10,000 IND and 10,000 NIND were extracted from the full test sets per country. Here, stratified random sampling was performed considering the ratio of the impervious surface of each subnational unit in each country (Supplementary Fig. 10).

Machine learning random forest (RF) was used for modeling. RF is an ensemble tree-based statistical model and is actively used in global land cover classification due to its high accuracy, stability, and relatively fast processing speed[51]. The tree size, which is the main factor of RF, was selected as 200, and even if the tree size becomes larger here, it was confirmed that the accuracy does not change substantially due to sufficient ensemble. The model was trained using training samples, and the performance was evaluated using test samples. Overall Accuracy (OA) and F1 Score were used in accuracy evaluation indices. A final map of ten countries was produced by classifying entire human settlements built-ups over impervious surfaces into IND and NIND for 2019. The GOOGLE EARTH ENGINE platform was used for modeling and mapping. The accuracy of the 2019 IND map was found to range between 88.5% and 96.6% for ten countries, with an area-weighted overall accuracy of approximately 91% (Supplementary Table 2).

The framework used to map the IND in 2019 is difficult to apply to the year 2000 due to the limited availability of relevant input features over that period. Assuming a high association between the expansion of IND and NIND areas and the growth of their underlying impervious surface areas, we overlaid the 2019 IND (with NIND) map—generated using machine learning—onto the GISA2.0 impervious surface area data for each year between 2000 and 2018. This procedure led to the creation of historical IND and NIND maps spanning the years 2000 to 2019.

Economic growth and CO_2 emissions analysis

This study set the subnational region (defined by GDL, Supplementary Fig. 11) as the basis for analyzing economic growth and CO_2 emissions. We classified the subnational regions of the ten countries into two groups based on the development level, using the human development index (HDI) provided by the Subnational Human Development Database. We selected 2010, the median year of the research period, as the base year. The regions with an HDI lower than 0.8 were classified as developing regions, while those with a higher HDI were classified as developed regions. The 0.8 HDI threshold follows the United Nations Development Programme's definition of high human development.

Among all the GDL's subnational regions in the ten countries, we excluded regions where the total area of IND and NIND in 2019 (the last year in the study period) was less than 5 km^2. This helps minimize the impact of high omission errors inherent in the datasets we utilized in IND and NIND area extraction, such as GISA and WSF. Furthermore, regions for which GDL did not provide HDI values were excluded. Finally, 216 subnational regions across ten countries were used in the analysis (Supplementary Fig. 11).

To derive the GDP data as an economic indicator, we used the global gridded GDP dataset developed by[37]. This dataset provides GDP values in 2011 US dollars, covering 1992 to 2015 at a 5-arc-minute spatial resolution. We extracted the total summed GDP within each subnational region annually from 2000 to 2015. Since this GDP data was only available by 2015, we predicted GDP from 2016 to 2019 using Gross National Income (GNI) values provided by the Subnational Human Development Database. Based on the assumption that GDP growth is highly correlated with GNI growth over time[52], we performed a linear regression between GNI and GDP from 2000 to 2015 for each subnational region. Using the regression models, we extrapolated GDP from 2016 to 2019 to construct a time series of total GDP for each subnational region from 2000 to 2019.

As a CO_2 emissions indicator, we employed the Open-Data Inventory for Anthropogenic Carbon Dioxide (ODIAC) dataset[38]. ODIAC offers high

spatiotemporal resolution data on CO_2 emissions from fossil fuel combustion, spanning from 2000 to the present. The ODIAC product computes CO_2 emissions resulting from the combustion of coal, oil, and natural gas by integrating fuel consumption statistics with spatial allocation factors that reflect the dispersion of emission sources. Specifically, we utilized the ODIAC2020b version, which provides monthly CO_2 emissions estimates covering January 2000 to December 2019 at a 0.01° resolution. We generated annual CO_2 emissions by summing the monthly data.

To facilitate appropriate comparisons between regions with differing population sizes, we computed metrics such as per capita GDP and per capita CO_2 for each subnational region. Using WorldPop population data, we determined the total population of each subnational region from 2000 to 2019. Subsequently, we calculated the per capita indices (per capita GDP and per capita CO_2) for each subnational region in the ten countries under study from 2000 to 2019.

To assess the impact of industrial land on economic growth and CO_2 emissions, we used per capita IND as an input variable. Additionally, per capita NIND was obtained as an input to discern the effects of the remaining urban land areas, excluding industrial land (such as commercial and residential land). These comparisons are intended to aid in understanding the relative impacts of these areas on economic growth and CO_2 emissions. We derived these two input variables by aggregating data from the developed 30 m resolution IND and NIND area maps for each subnational region across the ten countries, spanning a 20-year period.

We compiled several additional socioeconomic variables at the subnational level for the period 2000–2019. We included health and educational level as input variables in our study because of their established associations with both economic growth and CO_2 emissions[53–55]. These indices were obtained from the Subnational Human Development Index (SHDI) Database, curated by GDL. The health level was defined by life expectancy at birth, while the education level was established based on an equal combination of two factors: mean years of schooling for adults aged 25 and over, and expected years of schooling. These SHDI-based values were obtained from Eurostat for European countries, and from national statistical offices for other countries. A comprehensive explanation of the SHDI data can be found in[49]. We also calculated the population density for each subnational region using WorldPop population data over a 20-year period.

Environmental factors were also incorporated as input variables. We collected air temperature data, known to interact with economic growth and CO_2 concentration[56,57]. We used TerraClimate's monthly climate datasets, which are generated at approximately a 4 km resolution globally. TerraClimate employs a method called climatically aided interpolation, where high-spatial resolution climatological normals from the WorldClim dataset are combined with temporally varying data from CRU Ts4.0 and the Japanese 55-year Reanalysis (JRA55) that have a coarser spatial resolution[58]. We first computed the monthly average of maximum and minimum air temperatures, and then calculated the annual mean value for each subnational unit to derive air temperatures for the 20-year period.

We used 'greenness' as a variable to represent the amount of vegetation in each region, which can influence the economic level from an agricultural productivity perspective and CO_2 emissions concerning its role in emissions mitigation[59,60]. The measure of 'greenness' was obtained from the NDVI, a satellite-based measure of vegetation greenness. Specifically, we utilized data from the Terra Moderate Resolution Imaging Spectroradiometer (MODIS) Vegetation Indices Monthly (MOD13A3) Version 6.1, which offers a spatial scale of 1 km. Subsequently, we computed the annual mean value of NDVI for each subnational region for the period of 2000 to 2019. Here, our calculation of the annual greenness only averaged the NDVI for the vegetation growing season (May to October), to mitigate potential distortion effects, such as snow cover during the cold season.

Finally, we used soil properties as input variables because they have been identified as contributing to economic growth by increasing agricultural yield[61,62], and to CO_2 concentrations due to the role of healthy soil physical properties in mitigating greenhouse emissions[63,64]. We selected key

soil attributes: soil bulk density (a physical property), soil pH, and soil organic carbon (both chemical properties) from the top 0-5 cm of soil depth, using the 250 m SoilGrids 2.0 dataset[65]. We then computed the average for these soil properties across all subnational units. The statistics of each variable used in the analysis comparing developing and developed regions were presented in Supplementary Table 3.

We carried out a longitudinal analysis over 20 years (2000–2019), examining the impact of per capita IND on per capita GDP and per capita CO_2 emissions across different regions. We employed the MERF machine learning model for the longitudinal analysis. MERF, an extension of the RF model, incorporates both fixed and random effects, combining the benefits of traditional linear mixed models with the flexibility of RF in handling non-linear relationships and interactions between predictors (See Supplementary Note 3). This makes MERF an ideal choice for analyzing complex longitudinal data[66].

For the per capita GDP model, the input variables included per capita IND, per capita NIND, health level, education level, population density, soil bulk density, soil pH, soil organic carbon, air temperature and greenness. For the per capita CO_2 model, per capita GDP was added as an additional input variable to the same ten driving factors. All input and target variables were transformed to log2-scale in order to address skewed data, as suggested by[67]. The MERF model was run separately for development-level groups (developing and developed regions) and for the two target factors (per capita GDP and per capita CO_2).

To interpret the output of the MERF model, we computed the SHAP values for each input variable. SHAP values quantify the contribution of each variable to the model's prediction for a specific instance, providing an explanation of the model's prediction[68]. The direction of impact for each variable on the model's prediction can also be determined using SHAP values. A positive SHAP value for a variable implies that the presence of the feature increases the model's prediction relative to the expected value, whereas a negative SHAP value suggests a decrease in the prediction.

To assess the impact of each driving factor, we calculated the mean absolute SHAP value for each input variable and normalized the values by dividing each impact by the total sum of all input impacts. The MERF model and SHAP values were computed using Python, utilizing the *merf* and *shap* packages.

Data availability
The datasets supporting the findings of this study are available in the Zenodo repository, accessible via: https://doi.org/10.5281/zenodo.10802879. The repository contains: A comprehensive list of public data sources with corresponding links for industrial land mapping; Detailed public datasets with links for conducting the longitudinal impact analysis on CO_2 emissions and economic growth; and the input datasets employed in MERF longitudinal modeling, which facilitate the examination of variable contributions to CO_2 emissions and economic growth across developing and developed regions, are provided separately.

Code availability
Scripts for MERF longitudinal modeling, calculation of SHAP values, and generation of figures for the analysis of CO2 emissions and economic growth are available in the Zenodo repository at https://doi.org/10.5281/zenodo.10802879. The Google Earth Engine code used in the industrial land mapping can be provided by Q.W. or C.Y. upon request.

References
1. Alvarez, S. A., Barney, J. B. & Newman, A. M. The poverty problem and the industrialization solution. *Asia Pac. J. Manag.* **32**, 23–37 (2015).
2. Kimura, F. & Chang, M. S. Industrialization and poverty reduction in East Asia: Internal labor movements matter. *J. Asian Econ.* **48**, 23–37 (2017).
3. McMillan, M. & Zeufack, A. Labor productivity growth and industrialization in Africa. *J. Econ. Perspect.* **36**, 3–32 (2022).
4. Kuang, W., Liu, J., Dong, J., Chi, W. & Zhang, C. The rapid and massive urban and industrial land expansions in China between 1990 and 2010: A CLUD-based analysis of their trajectories, patterns, and drivers. *Landscape Urban Plan.* **145**, 21–33 (2016).
5. Dong, J., He, J., Li, X., Mou, X. & Dong, Z. The effect of industrial structure change on carbon dioxide emissions: a cross-country panel analysis. *J. Syst. Sci. Inf.* **8**, 1–16 (2020).
6. Grubler, A. et al. A low energy demand scenario for meeting the 1.5 C target and sustainable development goals without negative emission technologies. *Nat. Energy* **3**, 515–527 (2018).
7. Li, Q., Chen, W., Li, M., Yu, Q. & Wang, Y. Identifying the effects of industrial land expansion on PM2. 5 concentrations: A spatiotemporal analysis in China. *Ecol. Indicat.* **141**, 109069 (2022).
8. Mishra, M., Sahu, S. K., Mangaraj, P. & Beig, G. Assessment of hazardous radionuclide emission due to fly ash from fossil fuel combustion in industrial activities in India and its impact on public. *J. Environ. Manag.* **328**, 116908 (2023).
9. Ye, L. et al. Effects of dual land ownerships and different land lease terms on industrial land use efficiency in Wuxi City, East China. *Habitat Int.* **78**, 21–28 (2018).
10. Ritchie, H., Roser, M. & Rosado, P. CO_2 and greenhouse gas emissions. *Our world in data* (2020).
11. Tian, J., Liu, W., Lai, B., Li, X. & Chen, L. Study of the performance of eco-industrial park development in China. *J. Cleaner Prod.* **64**, 486–494 (2014).
12. Zhou, L., Tian, L., Cao, Y. & Yang, L. Industrial land supply at different technological intensities and its contribution to economic growth in China: A case study of the Beijing-Tianjin-Hebei region. *Land Use Policy* **101**, 105087 (2021).
13. Ke, Y. et al. The carbon emissions related to the land-use changes from 2000 to 2015 in Shenzhen, China: Implication for exploring low-carbon development in megacities. *J. Environ. Manag.* **319**, 115660 (2022).
14. Li, Y.-N., Cai, M., Wu, K. & Wei, J. Decoupling analysis of carbon emission from construction land in Shanghai. *J. Cleaner Prod.* **210**, 25–34 (2019).
15. Xia, C. et al. Exploring potential of urban land-use management on carbon emissions—A case of Hangzhou, China. *Ecol. Indicat.* **146**, 109902 (2023).
16. Zhang, T., Chen, L., Yu, Z., Zang, J. & Li, L. Spatiotemporal evolution characteristics of carbon emissions from industrial land in Anhui Province. *China. Land* **11**, 2084 (2022).
17. Qi, J., Hu, M., Han, B., Zheng, J. & Wang, H. Decoupling Relationship between Industrial Land Expansion and Economic Development in China. *Land* **11**, 1209 (2022).
18. Wu, S., Hu, S. & Frazier, A. E. Spatiotemporal variation and driving factors of carbon emissions in three industrial land spaces in China from 1997 to 2016. *Technol. Forecasting Soc. Change* **169**, 120837 (2021).
19. Chen, J., Shi, Q., Shen, L., Huang, Y. & Wu, Y. What makes the difference in construction carbon emissions between China and USA? *Sustainable Cities Soc.* **44**, 604–613 (2019).
20. Ganda, F. The impact of innovation and technology investments on carbon emissions in selected organisation for economic Co-operation and development countries. *J. Cleaner Prod.* **217**, 469–483 (2019).
21. Xu, G. et al. The effect of industrial relocations to central and Western China on urban construction land expansion. *J. Land Use Sci.* **16**, 339–357 (2021).
22. Ellram, L. M. Offshoring, reshoring and the manufacturing location decision. *J. Supply Chain Manag.* **49**, 3 (2013).
23. Shahidul, M. & Syed Shazali, S. Dynamics of manufacturing productivity: lesson learnt from labor intensive industries. *J. Manuf. Technol. Manag.* **22**, 664–678 (2011).

24. Wirtz, J., Tuzovic, S. & Ehret, M. Global business services: Increasing specialization and integration of the world economy as drivers of economic growth. *J. Serv. Manag.* **26**, 565–587 (2015).

25. Li, Q., Chen, W., Zhang, S. & Shi, H. Achieving sustainable development by reducing income inequality: The differential impact of industrial land expansion in urban agglomerations. *Sustain. Dev.* **31**, 2770–2783 (2023).

26. Yu, Z., Yan, T., Liu, X. & Bao, A. Urban land expansion, fiscal decentralization and haze pollution: Evidence from 281 prefecture-level cities in China. *J. Environ. Manag.* **323**, 116198 (2022).

27. Avenyo, E. K. & Tregenna, F. Greening manufacturing: Technology intensity and carbon dioxide emissions in developing countries. *Appl. Energy* **324**, 119726 (2022).

28. Fankhauser, S. & Jotzo, F. Economic growth and development with low-carbon energy. *Wiley Interdiscip. Rev.: Clim. Change* **9**, e495 (2018).

29. Nowak, A. & Dahal, G. The contribution of education to economic growth: Evidence from Nepal. *Int. J. Econ. Sci.* **5**, 22–41 (2016).

30. Luqman, M., Rayner, P. J. & Gurney, K. R. On the impact of urbanisation on CO2 emissions. *npj Urban Sustain.* **3**, 6 (2023).

31. Mahtta, R. et al. Urban land expansion: The role of population and economic growth for 300+ cities. *Npj Urban Sustain.* **2**, 5 (2022).

32. Oh, N. H. & Richter, D. D. Jr Soil acidification induced by elevated atmospheric CO2. *Global Change Biol.* **10**, 1936–1946 (2004).

33. Ma, Y. et al. Global crop production increase by soil organic carbon. *Nat. Geosci.* **16**, 1159–1165 (2023).

34. Lindner, T., Puck, J. & Verbeke, A. Beyond addressing multicollinearity: Robust quantitative analysis and machine learning in international business research. *J. Int. Bus. Stud.* **53**, 1307–1314 (2022).

35. Chen, W., Gu, T., Fang, C. & Zeng, J. Global urban low-carbon transitions: Multiscale relationship between urban land and carbon emissions. *Environ. Impact Assessment Rev.* **100**, 107076 (2023).

36. Tatem, A. J. WorldPop, open data for spatial demography. *Sci. Data* **4**, 1–4 (2017).

37. Kummu, M., Taka, M. & Guillaume, J. H. Gridded global datasets for gross domestic product and Human Development Index over 1990–2015. *Sci. Data* **5**, 1–15 (2018).

38. Oda, T., Maksyutov, S. & Andres, R. J. The Open-source Data Inventory for Anthropogenic CO 2, version 2016 (ODIAC2016): a global monthly fossil fuel CO 2 gridded emissions data product for tracer transport simulations and surface flux inversions. *Earth Syst. Sci. Data* **10**, 87–107 (2018).

39. Moran, T. & Oldenski, L. *How offshoring and global supply chains enhance the US economy*. (Peterson Institute for International Economics, 2016).

40. Hao, L.-N., Umar, M., Khan, Z. & Ali, W. Green growth and low carbon emission in G7 countries: how critical the network of environmental taxes, renewable energy and human capital is? *Sci. Total Environ.* **752**, 141853 (2021).

41. Wang, J., Dong, X. & Dong, K. How renewable energy reduces CO2 emissions? Decoupling and decomposition analysis for 25 countries along the Belt and Road. *Appl. Econ.* **53**, 4597–4613 (2021).

42. Wang, Q. & Wang, S. Decoupling economic growth from carbon emissions growth in the United States: The role of research and development. *J. Clean. Prod.* **234**, 702–713 (2019).

43. Richter, F. China is the world's manufacturing superpower. *Statista* **4**, 2021 (2021).

44. Shen, B. & Mikschovsky, M. in *Fashion Supply Chain Management in Asia: Concepts, Models, and Cases* (eds Shen, B., Gu, Q. & Yang, Y.) 1–17 (Springer Singapore, 2019).

45. Huang, X. et al. Toward accurate mapping of 30-m time-series global impervious surface area (GISA). *Int. J. Appl. Earth Observ. Geoinf.* **109**, 102787 (2022).

46. Gourlis, G. & Kovacic, I. Passive measures for preventing summer overheating in industrial buildings under consideration of varying manufacturing process loads. *Energy* **137**, 1175–1185 (2017).

47. Li, X. et al. Progressive fusion learning: A multimodal joint segmentation framework for building extraction from optical and SAR images. *ISPRS J. Photogrammetry Remote Sens.* **195**, 178–191 (2023).

48. Demuzere, M. et al. A global map of Local Climate Zones to support earth system modelling and urban scale environmental science. *Earth Syst. Sci. Data Discussions* **2022**, 1–57 (2022).

49. Smits, J. & Permanyer, I. The subnational human development database. *Sci. Data* **6**, 1–15 (2019).

50. Yoo, C., Han, D., Im, J. & Bechtel, B. Comparison between convolutional neural networks and random forest for local climate zone classification in mega urban areas using Landsat images. *ISPRS J. Photogrammetry Remote Sens.* **157**, 155–170 (2019).

51. Rodriguez-Galiano, V. F., Ghimire, B., Rogan, J., Chica-Olmo, M. & Rigol-Sanchez, J. P. An assessment of the effectiveness of a random forest classifier for land-cover classification. *ISPRS J. Photogrammetry Remote Sens.* **67**, 93–104 (2012).

52. Sørensen, B. E., Wu, Y.-T., Yosha, O. & Zhu, Y. Home bias and international risk sharing: Twin puzzles separated at birth. *J. Int. Money Finance* **26**, 587–605 (2007).

53. Bedir, S. Healthcare expenditure and economic growth in developing countries. *Adv. Econ. Bus.* **4**, 76–86 (2016).

54. Dong, H., Xue, M., Xiao, Y. & Liu, Y. Do carbon emissions impact the health of residents? Considering China's industrialization and urbanization. *Sci. Total Environ.* **758**, 143688 (2021).

55. Hanushek, E. A. & Woessmann, L. Education and economic growth. *Econ. Educ.* **60**, 1 (2010).

56. Chaplot, V. Water and soil resources response to rising levels of atmospheric CO$_2$ concentration and to changes in precipitation and air temperature. *J. Hydrol.* **337**, 159–171 (2007).

57. Moore, F. C. & Diaz, D. B. Temperature impacts on economic growth warrant stringent mitigation policy. *Nat. Clim. Change* **5**, 127–131 (2015).

58. Abatzoglou, J. T., Dobrowski, S. Z., Parks, S. A. & Hegewisch, K. C. TerraClimate, a high-resolution global dataset of monthly climate and climatic water balance from 1958–2015. *Sci. Data* **5**, 1–12 (2018).

59. Fan, Z., Bai, X. & Zhao, N. Explicating the responses of NDVI and GDP to the poverty alleviation policy in poverty areas of China in the 21st century. *Plos One* **17**, e0271983 (2022).

60. Rahaman, Z. A. et al. Assessing the impacts of vegetation cover loss on surface temperature, urban heat island and carbon emission in Penang city, Malaysia. *Build. Environ.* **222**, 109335 (2022).

61. Cen, Y. et al. Using organic fertilizers to increase crop yield, economic growth, and soil quality in a temperate farmland. *PeerJ* **8**, e9668 (2020).

62. Selim, M. M. Introduction to the integrated nutrient management strategies and their contribution to yield and soil properties. *Int. J. Agronomy* **2020**, 2821678 (2020).

63. Lal, R. Soil health and carbon management. *Food Energy Secur.* **5**, 212–222 (2016).

64. Ogle, S. M. et al. Climate and soil characteristics determine where no-till management can store carbon in soils and mitigate greenhouse gas emissions. *Sci. Rep.* **9**, 11665 (2019).

65. Poggio, L. et al. SoilGrids 2.0: producing soil information for the globe with quantified spatial uncertainty. *Soil* **7**, 217–240 (2021).

66. Capitaine, L., Genuer, R. & Thiébaut, R. Random forests for high-dimensional longitudinal data. *Stat. Methods Med. Res.* **30**, 166–184 (2021).

67. Fox, J. *Regression diagnostics: An introduction*. (Sage publications, 2019).

68. Lundberg, S. M. et al. From local explanations to global understanding with explainable AI for trees. *Nat. Mach.Intell.* **2**, 56–67 (2020).

Author contributions

C.Y. was responsible for the conceptualization, development of the methodology framework, model coding, and analysis of the results. C.Y. also took the lead in writing the initial draft of the manuscript and contributed to subsequent revisions, as well as the visualization of data. H.X. participated in analyzing the results and provided critical feedback that helped shape the research, analysis and manuscript. Q.Z. contributed to the methodology and analysis of the results. Q.W. played a pivotal role in conceptualization and supervision, offered guidance throughout the research process, and contributed to manuscript writing and revisions. Q.W. was instrumental in funding acquisition.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s43247-024-01375-x.

Correspondence and requests for materials should be addressed to Qihao Weng.

Peer review information *Communications Earth and Environment* thanks Nikunj Patel, Yacouba Kassour and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Primary Handling Editors: Sadia Ilyas and Martina Grecequet. A peer review file is available.

Reprints and permissions information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.