Local GDP Estimates Around the World*

Esteban Rossi-Hansberg
University of Chicago

Jialing Zhang
Princeton University

December 10, 2025

Appendix

Contents

*Esteban Rossi-Hansberg: earossih@uchicago.edu. Jialing Zhang: zhang.jialing@princeton.edu

1 Details of Data Source

1.1 GIS data

- Worldwide country level, province-equivalent level and county-equivalent level maps in Geopackage (".gpkg") format are downloaded from GADM version 4.1 "geopackage" action in the sentence "You can also download this version as six separate layers (one for each level of subdivision/aggregation), as a geopackage database" from the following website: https://gadm.org/download_world.html. Save the downloaded file with the name "gadm_410—levels.gpkg" in the folder "step1_obtain_gis_data/inputs /CGAZ_ADM1/".

- Another version of the province-level geometry file is the CGAZ dataset. Download "ADM1" global map in "geojson" format from the website: https://www.geoboundaries.org/globalDownloads.html (Runfola et al., 2020). Save the downloaded file with the name "geoBoundariesCGAZ_ADM1.geojson" in the folder "step1_obtain_gis_data/inputs/CGAZ_ADM1/".

- Worldwide large lakes are excluded from the maps. The data used are "Global Lakes and Wetlands Database: Large Lake Polygons (Level 1)" from the website: https://www.worldwildlife.org/publications/global-lakes-and-wetlands-database-large-lake-polygons-level-1 (Lehner and Döll, 2004). One with higher computer resources can even exclude permanent open water bodies with a surface area greater and equal to 0.1 km^2 and wetlands using other types of data from the website: https://www.worldwildlife.org/pages/global-lakes-and-wetlands-database. Save the downloaded files with filenames starting with "glwd_1.xxx" in the folder "step1_obtain_gis_data/inputs/large_inland_waters_geom_GLWD_level1/".

- We also need the spatial data from the paper named "DOSE – Global data set of reported sub-national economic output" (Wenz et al., 2023). Use the link: https://doi.org/10.5281/zenodo.7659599 and download the files from the folder "DOSE replication files/Data/spatial data/". We also need to download files from https://gadm.org/download_world36.html by clicking the action "shapefiles" in the sentence "You can also download this version as six separate layers (one for each level of subdivision/aggregation), as a geopackage database or as shapefiles). Save the downloaded files in the folder "step1_obtain_gis_data/inputs/DOSE_spatial_data/ ".

- China city level geometry shape files are downloaded from the website: https://github.com/GaryBikini/ChinaAdminDivisonSHP (GaryBikini, 2024). Download all files in the folder named "3.City". Save the downloaded files in the folder "step1_obtain_gis_data/inputs/china_city/".

1.2 GDP data

1.2.1 Regional GDP data

Table 1 lists the sources of regional GDP data by country. Table 2 lists countries in the training sample and their average area of regional units used to construct cell-level GDP data for model training. If a country is included in the training sample, the most granular administrative level data collected are used to construct the cell-level GDP. The following details provide information on where to find the regional GDP datasets referenced in Table 1:

- Obtain data from “DOSE – Global data set of reported sub-national economic output” (Wenz et al., 2023):
  - Click link `https://zenodo.org/records/13773040`, which is their paper’s recommended data download link
  - Download the file “DOSE_V2.11.csv”, and save it to folder “ step2 _obtain_gdp_data /inputs/gdp_data/regional/DOSE/”
- Obtain data from OECD iLibrary - Regional economy[1] (OECD, 2024b):
  - Click link `https://www.oecd-ilibrary.org/urban-rural-and-regional-development/data/oecd-regional-statistics_region-data-en#archive`, which refers to OECD iLibrary - OECD Regional Statistics
  - Click the “Archive 2023” under the “Datasets Archives”
  - Click the “csv” bottom beside “Regional economy (Edition 2023)”
  - Save the file with the name “REGION_ECONOM−2023−1−EN−20240216T100059 2.csv” to folder “step2_obtain_gdp_data/inputs/gdp_data/regional/oecd/”
  - Please note that for the following countries: ”AUT”, ”BEL”, ”CAN”, ”CHL”, ”DNK”, ”ESP”, ”FIN”, ”FRA”, ”GBR”, ”HUN”, ”ITA”, ”LVA”, ”NLD”, ”NOR”, ”PRT”, ”ROU”, and ”SWE”, the OECD provides some unregionalized data. We have removed this unregionalized data and rescaled the regional data to align with national GDP data from the World Bank or IMF.
- Obtain data from “OECD Data Explorer - Regions”[2] (OECD, 2024a):
  - Use the link `https://data-explorer.oecd.org/`, which refers to OECD Data Explorer main page.
  - Click “Regions” under “Regions, cities and local areas”
  - Choose the dataset “Gross domestic product - Regions”

[1]This dataset will also not be updated; therefore, we only retrieve data for the years 2012–2020. Data from 2021 and 2022 should be obtained from ”OECD Data Explorer - Regions”.

[2]The data was downloaded on Sept 22, 2025, and may have been updated since. For replication, the latest files should be downloaded, and the corresponding R scripts must be updated to process the new data correctly.

- Click “Download / Unfiltered data in tabular text (CSV)”, and save the file with the name “OECD.CFE.EDS,DSD_REG_ECO@DF_GDP,2.0+all.csv” to the folder “step2_obtain_gdp_data/inputs/gdp_data/regional/oecd/”
- Also click “OECD Territorial correspondence table (xlsx)” , and save the file with the name “OECD Territorial correspondence – TL2021.xlsx” to the folder “step2_obtain_gdp_data/inputs/gdp_data/regional/oecd/”. We need this file to understand the heritage relationships between different id names.

• Obtain Russia’s regional data after 2019[3] (*Russian Statistical Yearbook* 2022; *Russian Statistical Yearbook* 2023):

- Click the website link https://eng.rosstat.gov.ru/Publications/document/74811, which refers to Russian Statistical Yearbook
- Download the “Russian Statistical Yearbook 2023” and “Russian Statistical Yearbook 2022” in “RAR” format. Downloading the “pdf” format is also acceptable because we have to manually collect the data. Save the folders with name “russian_statistical_yearbook_20xx” in the folder ‘‘ step2_obtain_gdp_data/inputs/gdp_data/regional/RUS/’’
- Match the regions’ names from the yearbooks with OECD’s definitions of id names. Follow the matches in the file “RUS.xlsx” in the folder “step2_obtain_gdp_data/inputs/gdp_data/regional/RUS/”. When updating to new years, add the new data to the file “RUS.xlsx”.

• Obtain Brazil’s regional data in 2021 and 2022 (Geografia e Estatística (IBGE), 2024):[4]

- Click the website link https://www.ibge.gov.br/en/statistics/economic/national-accounts/16855-regional-accounts-of-brazil.html, which refers to the dataset “SCR - System of Regional Accounts” on the Instituto Brasileiro de Geografia e Estatística (IBGE)
- Download the “xls” file for the “GDP under the point of view of Production (2010-2022)”. Save the file with the name “PIB_Otica_Renda_UF.xls” to the folder “step2_obtain_gdp_data/inputs/gdp_data/regional/BRA/”.

• Obtain China’s regional data (Statistics of China, 2024):

- Click the website link https://data.stats.gov.cn/english/index.htm, which refers to the main page of National Bureau of Statistics of China.
- Click “Regional / Annual by Province” and choose the following:

[3] Russia’s regional GDP data for 2012-2019 are obtained directly from “OECD iLibrary - Regional economy” as described above. Since the OECD no longer updates Russia’s data, regional GDP data for future years must be obtained from alternative sources.

[4] As of the OECD data download date (Sept 22, 2025), Brazil’s regional data had not yet been updated. If, during future updates, OECD publishes Brazil’s province data, this step can be skipped by directly downloading the updated data using the same procedure as for the ”OECD Data Explorer.” This step serves as a backup in case Brazil’s data remains unpublished.

    * Year: LATEST20
    * Indicators: Gross
  - Download the "xls" format and save it with the name "AnnualbyProvince.xls" to the folder "step2_obtain_gdp_data/inputs/gdp_data/regional/CHN/".

- Obtain India's regional data (India, 2023):
  - Use the website link `https://www.rbi.org.in/Scripts/AnnualPublications.aspx?head=Handbook%20of%20Statistics%20on%20Indian%20States`, which refers to the "Handbook of Statistics of Indian States" on the Reserve Bank of India.
  - Choose the year 2024, download the excel format of "Table 27: Gross State Domestic Product (Current Prices)" and save the file with the name "T28_09122024E699603AE68F445FB6E485839CCB697B.XLSX" in the folder "step2_obtain_gdp_data/inputs/gdp_data/regional/IND/".

- Obtain Kazakhstan's regional data (National Statistics Agency for Strategic Planning and Republic of Kazakhstan, 2023):
  - Use the link `https://stat.gov.kz/en/industries/economy/national-accounts/`, which refers to the datasets of "Statistics of national accounts" on the Bureau of National Statistics Agency for Strategic Planning and Reforms of the Republic of Kazakhstan.
  - Find the "Gross regional product)", scroll down, and find the "Gross regional product" under the section of "Dynamic tables". Download the file and save it with the name "1. Gross regional product.xlsx" to the folder "step2_obtain_gdp_data/inputs/gdp_data/regional/KAZ/".

- Obtain USA's regional data (Economic Analysis U.S. Department of Commerce, 2023):
  - Use the link `https://apps.bea.gov/histdatacore/Regional_Accounts_new.html`, which refers to the "Regional Economic Accounts - Previously Published Estimates" of the Bureau of Economic Analysis U.S. Department of Commerce.
  - Click "Gross Domestic Product by County and Metro Area / December 7, 2023 / CAGDP2: GDP in Current Dollars by County and MSA". Download the folder and save it with the name "CAGDP2" to the folder "step2_obtain_gdp_data/inputs/gdp_data/regional/USA/".

- Obtain Philippine's regional data (Authority, 2023):
  - Use the link `https://psa.gov.ph/statistics/grdp/data-series`, which refers to the dataset "GRDP Data Series" on the Philippine Statistics Authority
  - Find the "2000-2024 Gross Regional Domestic Product" in the "Title" section. Download the file "GRDP by Region" and save it with the name "GRDP_Reg_2018PSNA_2000—2023.xlsx" to the folder "step2_obtain_gdp_data/inputs/gdp_data/regional/PHL/".

- Obtain Kyrgyzstan's regional data (Kyrgyz Republic, 2023):
  - Use the link https://www.stat.kg/en/statistics/nacionalnye-scheta/, which refers to the dataset "National accounts" on the National Statistical Committee of the Kyrgyz Republic.
  - Download the file "1.01.00.09 Gross regional product (GRP) at current prices" under the section "Dynamic tables". Save the file with the name "`1.01.00.09 Валовои региональныи продукт (ВРП) в текущих ценах..xlsx`" to the folder "step2_obtain_gdp_data/inputs/gdp_data/regional/KGZ/".

Table 1: Regional GDP data source by country

ISO	Available Years	Data Level	Source	Date
ALB	2012-2019	Third	DOSE – Global data set of reported sub-national economic output	Sept 22nd, 2025
AUS	2012-2022	Second	OECD iLibrary - Regional economy (2012-2020); OECD Data Explorer - Regions (2021-2022)	Sept 22nd, 2025
AUT	2012-2022	Second, Third	OECD iLibrary - Regional economy (2012-2020); OECD Data Explorer - Regions (2021-2022)	Sept 22nd, 2025
BEL	2012-2022	Second, Third	OECD iLibrary - Regional economy (2012-2020); OECD Data Explorer - Regions (2021-2022)	Sept 22nd, 2025
BGR	2012-2020	Second, Third	OECD iLibrary - Regional economy (2012-2020);	Sept 22nd, 2025
BIH	2012-2019	Second	DOSE – Global data set of reported sub-national economic output	Sept 22nd, 2025
BLR	2012-2019	Second	DOSE – Global data set of reported sub-national economic output	Sept 22nd, 2025
BRA	2012-2022	Second	OECD iLibrary - Regional economy (2012-2020); Instituto Brasileiro de Geografia e Estatística - System of Regional Accounts (2021-2022)	Sept 22nd, 2025

Continued on next page

Table 1: Regional GDP data source by country (Continued)

ISO	Available Years	Data Level	Source	Date
CAN	2012-2022	Second	OECD iLibrary - Regional economy (2012-2020); OECD Data Explorer - Regions (2021-2022)	Sept 22nd, 2025
CHE	2012-2022	Second, Third	OECD iLibrary - Regional economy (2012-2020); OECD Data Explorer - Regions (2021-2022)	Sept 22nd, 2025
CHL	2012-2022	Second	OECD iLibrary - Regional economy (2012-2020); OECD Data Explorer - Regions (2021-2022)	Sept 22nd, 2025
CHN	2012-2022	Second	National Bureau of Statistics of China - Regional - Annual by province	Sept 22nd, 2025
COL	2012-2022	Second	OECD iLibrary - Regional economy (2012-2020); OECD Data Explorer - Regions (2021-2022)	Sept 22nd, 2025
CZE	2012-2022	Second, Third	OECD iLibrary - Regional economy (2012-2020); OECD Data Explorer - Regions (2021-2022)	Sept 22nd, 2025
DEU	2012-2022	Second, Third	OECD iLibrary - Regional economy (2012-2020); OECD Data Explorer - Regions (2021-2022)	Sept 22nd, 2025
DNK	2012-2022	Second, Third	OECD iLibrary - Regional economy (2012-2020); OECD Data Explorer - Regions (2021-2022)	Sept 22nd, 2025
ECU	2012-2019	Third	DOSE – Global data set of reported sub-national economic output	Sept 22nd, 2025
ESP	2012-2022	Second, Third	OECD iLibrary - Regional economy (2012-2020); OECD Data Explorer - Regions (2021-2022)	Sept 22nd, 2025
EST	2012-2022	Second, Third	OECD iLibrary - Regional economy (2012-2020); OECD Data Explorer - Regions (2021-2022)	Sept 22nd, 2025

Continued on next page

Table 1: Regional GDP data source by country (Continued)

ISO	Available Years	Data Level	Source	Date
FIN	2012-2022	Second, Third	OECD iLibrary - Regional economy (2012-2020); OECD Data Explorer - Regions (2021-2022)	Sept 22nd, 2025
FRA	2012-2022	Second, Third	OECD iLibrary - Regional economy (2012-2020); OECD Data Explorer - Regions (2021-2022)	Sept 22nd, 2025
GBR	2012-2022	Second, Third	OECD iLibrary - Regional economy (2012-2020); OECD Data Explorer - Regions (2021-2022)	Sept 22nd, 2025
GRC	2012-2022	Second, Third	OECD iLibrary - Regional economy (2012-2020); OECD Data Explorer - Regions (2021-2022)	Sept 22nd, 2025
HRV	2012-2020	Second, Third	OECD iLibrary - Regional economy (2012-2020);	Sept 22nd, 2025
HUN	2012-2022	Second, Third	OECD iLibrary - Regional economy (2012-2020); OECD Data Explorer - Regions (2021-2022)	Sept 22nd, 2025
IDN	2012-2020	Second	OECD iLibrary - Regional economy (2012-2020)	Sept 22nd, 2025
IND	2012-2022	Second	Reserve Bank of India - Handbook of Statistics on Indian States (2012-2022)	Sept 22nd, 2025
IRL	2012-2022	Second, Third	OECD iLibrary - Regional economy (2012-2020); OECD Data Explorer - Regions (2021-2022)	Sept 22nd, 2025
ISL	2012-2022	Second, Third	OECD iLibrary - Regional economy (2012-2020); OECD Data Explorer - Regions (2021-2022)	Sept 22nd, 2025
ISR	2012-2022	Second	OECD iLibrary - Regional economy (2012-2020); OECD Data Explorer - Regions (2021-2022)	Sept 22nd, 2025

Continued on next page

Table 1: Regional GDP data source by country (Continued)

ISO	Available Years	Data Level	Source	Date
ITA	2012-2022	Second, Third	OECD iLibrary - Regional economy (2012-2020); OECD Data Explorer - Regions (2021-2022)	Sept 22nd, 2025
JPN	2012-2021	Second	OECD iLibrary - Regional economy (2012-2019); OECD Data Explorer - Regions (2020-2021)	Sept 22nd, 2025
KAZ	2012-2022	Third	National Statistics Agency for Strategic Planning and Republic of Kazakhstan	Sept 22nd, 2025
KEN	2013-2017	Third	DOSE – Global data set of reported sub-national economic output	Sept 22nd, 2025
KGZ	2012-2022	Third	National Statistical Committe of the Kyrgyz Republic	Sept 22nd, 2025
KOR	2012-2022	Second, Third	OECD iLibrary - Regional economy (2012-2020); OECD Data Explorer - Regions (2021-2022)	Sept 22nd, 2025
LKA	2013-2019	Third	DOSE – Global data set of reported sub-national economic output	Sept 22nd, 2025
LTU	2012-2022	Second, Third	OECD iLibrary - Regional economy (2012-2020); OECD Data Explorer - Regions (2021-2022)	Sept 22nd, 2025
LUX	2012-2022	Second, Third	OECD iLibrary - Regional economy (2012-2020); OECD Data Explorer - Regions (2021-2022)	Sept 22nd, 2025
LVA	2012-2022	Second, Third	OECD iLibrary - Regional economy (2012-2020); OECD Data Explorer - Regions (2021-2022)	Sept 22nd, 2025
MEX	2012-2022	Second	OECD iLibrary - Regional economy (2012-2020); Instituto Nacional de Estadística y Geografía - National Accounts System (2021-2022)	Sept 22nd, 2025
MOZ	2012-2019	Third	DOSE – Global data set of reported sub-national economic output	Sept 22nd, 2025

Continued on next page

Table 1: Regional GDP data source by country (Continued)

ISO	Available Years	Data Level	Source	Date
NLD	2012-2022	Second, Third	OECD iLibrary - Regional economy (2012-2020); OECD Data Explorer - Regions (2021-2022)	Sept 22nd, 2025
NOR	2012-2021	Second, Third	OECD iLibrary - Regional economy (2012-2020); OECD Data Explorer - Regions (2021)	Sept 22nd, 2025
NZL	2012-2022	Second	OECD iLibrary - Regional economy (2012-2020) OECD Data Explorer - Regions (2021-2022)	Sept 22nd, 2025
PER	2012-2020	Second	OECD iLibrary - Regional economy (2012-2020)	Sept 22nd, 2025
PHL	2012-2022	Second	Philippine Statistics Authority	Sept 22nd, 2025
POL	2012-2022	Second, Third	OECD iLibrary - Regional economy (2012-2020); OECD Data Explorer - Regions (2021-2022)	Sept 22nd, 2025
PRT	2012-2022	Second, Third	OECD iLibrary - Regional economy (2012-2020); OECD Data Explorer - Regions (2021-2022)	Sept 22nd, 2025
ROU	2012-2020	Second, Third	OECD iLibrary - Regional economy (2012-2020)	Sept 22nd, 2025
RUS	2012-2022	Second	OECD iLibrary - Regional economy (2012-2019); Federal State Statistics Service - Russian Statistical Yearbook (2020-2022)	Sept 22nd, 2025
SRB	2012-2018	Third	DOSE – Global data set of reported sub-national economic output	Sept 22nd, 2025
SVK	2012-2022	Second, Third	OECD iLibrary - Regional economy (2012-2020); OECD Data Explorer - Regions (2021-2022)	Sept 22nd, 2025
SVN	2012-2022	Second, Third	OECD iLibrary - Regional economy (2012-2020); OECD Data Explorer - Regions (2021-2022)	Sept 22nd, 2025

Continued on next page

Table 1: Regional GDP data source by country (Continued)

ISO	Available Years	Data Level	Source	Date
SWE	2012-2022	Second, Third	OECD iLibrary - Regional economy (2012-2020); OECD Data Explorer - Regions (2021-2022)	Sept 22nd, 2025
THA	2014-2018	Third	DOSE – Global data set of reported sub-national economic output	Sept 22nd, 2025
TUR	2012-2022	Second, Third	OECD iLibrary - Regional economy (2012-2020); OECD Data Explorer - Regions (2021-2022)	Sept 22nd, 2025
USA	2012-2022	Second, Third	Bureau of Economic Analysis - Regional Economic Accounts	Sept 22nd, 2025
UZB	2012-2019	Second	DOSE – Global data set of reported sub-national economic output	Sept 22nd, 2025
VNM	2012-2018	Third	DOSE – Global data set of reported sub-national economic output	Sept 22nd, 2025

1.2.2 National GDP data

Table 3 lists the sources of national GDP data by country. The following details provide information on where to find the national GDP datasets referenced in Table 3:

- Obtain data from World Bank (Bank, 2024d; Bank, 2024c; Bank, 2024b; Bank, 2024a):
  - Click link `https://data.worldbank.org/indicator/SP.POP.TOTL` for World Bank Population dataset with file name "API_SP.POP.TOTL_DS2_en_excel_v2_19296 .xls"; link `https://data.worldbank.org/indicator/NY.GDP.MKTP.CD` for World Bank GDP in current US$ with file name "API_NY.GDP.MKTP. CD_DS2_en_excel_v2_19310.xls"; link `https://data.worldbank.org/indicator/NY.GDP.MKTP.PP.CD` for World Bank GDP in current PPP adjusted international$ with file name "API_NY.GDP.MKTP.PP.CD_DS2_en_excel_v2_19548 .xls"; link `https://data.worldbank.org/indicator/NY.GDP.MKTP.PP.KD` for World Bank GDP in constant 2021 PPP adjusted international$ with file name "API_NY.GDP.MKTP.PP.KD_DS2_en_excel_v2_20528.xls".
  - Download as "Excel" files and save it in the corresponding folders in "step2_obtain_gdp_data /inputs/gdp_data/national/world_bank_data/xxx".
- Obtain the "Annual Index Value" data from US Census Bureau (Bureau, 2024):

Table 2: Average area of subnational units used for constructing cell-level GDP data by country

Developed Country	**Average Subnational Area (km^2)**	**Developing Country**	**Average Subnational Area (km^2)**
AUT	2383	ALB	2353
BEL	695	BIH	16988
BGR	3961	BLR	34328
CHE	1548	CHL	48669
CZE	5618	COL	34504
DEU	887	ECU	10695
DNK	3894	IDN	55053
ESP	8548	KEN	12276
EST	8641	KGZ	21134
FIN	16717	LKA	7314
FRA	6305	MOZ	70691
GBR	1356	PER	51699
GRC	2524	PHL	17006
HRV	2667	SRB	3122
HUN	4607	THA	6652
ITA	2756	UZB	29715
JPN	7832	VNM	5240
KOR	5635		
LTU	6459		
LVA	10699		
NLD	877		
NOR	29668		
NZL	21908		
POL	4254		
PRT	3668		
ROU	5642		
SWE	20600		
SVK	6107		
SVN	1692		
TUR	9500		
USA	2514		

- Use the link `https://www.census.gov/en.html`, which refers to main page of United States Census Bureau.
- Click "Topics / Income and Poverty / Income / Guidance for Data Users / Current versus Constant (or Real) Dollars".
- Download the excel file "Annual Index Value and Annual Percent Change in Price Series: 1947 to 2023" and save it to the folder "step2_obtain_gdp_data/inputs/gdp_data/national/US_census_bureau_data".

- Obtain data from IMF - WEO (Fund, 2024):[5]
  - Use the link https://www.imf.org/en/Publications/SPROLLS/world-economic-outlook-databases#sort=%40imfdate%20descending, which refers to the World Economic Outlook Databases (WEO) in International Monetary Fund (IMF).
  - Choose "World Economic Outlook Database, April 2025". Click "By Countries / ALL COUNTRIES / Continue".
  - Choose datasets "Gross domestic product, current prices U.S. DOLLARS", "Gross domestic product, current prices PURCHASING POWER PARITY; INTERNATIONAL DOLLARS", "Gross domestic product per capita, constant prices PURCHASING POWER PARITY; 2021 INTERNATIONAL DOLLARS", and "Population PERSONS".
  - Click "Continue" and choose "2012" as "Start Year", "2022" as "End Year". Click "ISO Alpha-3 Code".
  - Download the file and save the file with name "WEO_Data.xlsx" in the folder "step2_obtain_gdp_data/inputs/gdp_data/national/IMF_data".
- Obtain data from UNdata (UNdata, 2024):
  - Use the link https://data.un.org/Data.aspx?q=gdp&d=SNAAMA&f=grID%3a101%3bcurrID%3aUSD%3bpcFlag%3a1#SNAAMA, which refers to the dataset "Per capita GDP at current prices - US dollars" in UNdata.
  - Choose "Cuba", "Democratic People's Republic of Korea", "Eritrea", and years 2012 to 2023.
  - Download the files in csv format, value separated in comma. Save it with the name "UNdata_Export_20250426_211303453.csv" to the folder "step2_obtain_gdp_data/inputs/gdp_data/national/UN_data".

[5] The primary data source is the IMF, as it converts GDP reported on a fiscal year basis into a calendar year format, aligning with predictors that are primarily satellite-based. When updating to future years, care must be taken in selecting the dataset's end year, as some entries may be estimates rather than actual values. Additionally, note that data for certain countries may change during updates.

Table 3: National GDP and population data source by country

ISO	Data	Source	Date Retrieved
BMU, CYM, CUW, GRL, KXK, LIE, MCO, SXM, SYR, TCA, PSE	Population	World Bank - Population	Sept 15th, 2025
	GDP in current US dollars	World Bank - GDP (current US$)	Sept 15th, 2025
	GDP in constant 2021 US dollars	Calculated using World Bank - GDP (current US$) and Annual Index Value from US Census Bureau	Sept 15th, 2025
	GDP in current international dollars PPP adjusted	World Bank - GDP, PPP (current international $)	Sept 15th, 2025
	GDP in constant 2021 international dollars PPP adjusted	Calculated using World Bank - GDP, PPP (current international $) and World Bank - GDP, PPP (constant 2021 international $)	Sept 15th, 2025
CUB, ERI, PRK	Population	World Bank - Population	Sept 15th, 2025
	GDP in current US dollars	Calculated using UNdata - Per capita GDP at current prices - US dollars and World Bank - Population	Sept 15th, 2025
	GDP in constant 2021 US dollars	Calculated using UNdata - Per capita GDP at current prices - US dollars, World Bank - Population, and Annual Index Value from US Census Bureau	Sept 15th, 2025
	GDP in current international dollars PPP adjusted	NA	Sept 15th, 2025
	GDP in constant 2021 international dollars PPP adjusted	NA	Sept 15th, 2025
Other countries	Population	IMF - WEO - Population	Sept 15th, 2025
	GDP in current US dollars	IMF - WEO - Gross domestic product, current prices, US dollars	Sept 15th, 2025
	GDP in constant 2021 US dollars	Calculated using IMF - WEO - Gross domestic product, current prices, US dollars and Annual Index Value from US Census Bureau	Sept 15th, 2025
	GDP in current international dollars PPP adjusted	IMF - WEO - Gross domestic product, current prices PURCHASING POWER PARITY; INTERNATIONAL DOLLARS	Sept 15th, 2025
	GDP in constant 2021 international dollars PPP adjusted	Calculated using IMF - WEO - Gross domestic product per capita, constant prices PURCHASING POWER PARITY; 2021 INTERNATIONAL DOLLARS and IMF - WEO - Population	Sept 15th, 2025

1.2.3 Calculate Different Measures of GDP

There are four measures of GDP in our study: current USD, constant 2021 USD, current PPP-adjusted international dollars, and constant 2021 PPP-adjusted international dollars. Each type of GDP measure serves a specific purpose and requires different calculations to provide accurate economic comparisons. Below, we will explain what these different GDP measures mean and how to calculate them.

- GDP in Current USD:

  - This is the Gross Domestic Product (GDP) measured in current U.S. dollars. It represents the total value of all goods and services produced within a country in a given year, converted to USD using the current exchange rates. This measure is influenced by inflation, exchange rate fluctuations, and changes in the price level. GDP in current USD is useful for understanding the nominal size of an economy and comparing it to others using the same currency (USD).

- GDP in Constant 2021 USD:
  - This is the GDP adjusted for inflation and expressed in constant 2021 U.S. dollars. By using a base year (in this case, 2021), it removes the effects of price level changes over time, allowing for comparison of economic output across different years in real terms. This measure helps to isolate the actual growth in economic activity.
  - In our study, this measure is calculated using:

$$\text{Constant 2021 USD} = \text{Current USD} \times \left(\frac{\text{2021 Price Index}}{\text{Current Year Price Index}} \right) \tag{1}$$

  - The price index is the Chained Consumer Price Index for All Urban Consumers (C-CPI-U) obtained from US Census Bureau.

- GDP in Current PPP-Adjusted International $:
  - This GDP measure adjusts for purchasing power parity (PPP) and is expressed in current international dollars. Purchasing Power Parity (PPP) is a method of measuring the relative purchasing power of different countries' currencies over the same types of goods and services. International dollars are a hypothetical currency that has the same purchasing power over GDP as the U.S. dollar has in the United States. By using PPP adjustments, this measure accounts for differences in price levels between countries, providing a more accurate comparison of economic output and living standards. Unlike GDP in current USD, which is influenced by exchange rates, the PPP-adjusted measure reflects the value of goods and services in terms of what they can actually buy in each country.

- GDP in Constant 2021 PPP-Adjusted International $
  - This is the GDP adjusted for both purchasing power parity (PPP) and inflation, expressed in constant 2021 international dollars. By using a base year (2021) and adjusting for PPP, it allows for comparison of economic output across different years and countries, accounting for both inflation and differences in price levels. This measure provides a consistent basis for comparing real economic growth and living standards over time and across countries, isolating real growth from both inflation and price level differences.
  - In our study, this measure is calculated using:

$$\begin{array}{c}\text{Const 2021}\\ \text{PPP prices}\end{array} = \begin{array}{c}\text{Const 2022}\\ \text{PPP prices}\end{array} \times \left(\frac{\text{2021 GDP in Current PPP Prices}}{\text{2021 GDP in Const 2022 PPP Prices}} \right) \tag{2}$$

1.2.4 China city level GDP data

Please download the following files and place each one into its respective subfolder under "step2_obtain_gdp_data/inputs/regional/CHN/province_yearbook".

Table 4: City-Level GDP Data Sources for Seven Major Provinces in China

Province	Source	Dataset Name
Guangdong	http://stats.gd.gov.cn/gdtjnj/	Guangdong Province Statistical Yearbook 2024: 2-14 Gross Domestic Product by City
Henan	https://tjj.henan.gov.cn/tjfw/tjcbw/tjnj/	Statistical Yearbook 2024: 2-9 Gross Domestic Product by City (2022)
		Statistical Yearbook 2023: 2-9 Gross Domestic Product by City (2021)
		Statistical Yearbook 2022: 2-9 Gross Domestic Product by City (2020)
		Statistical Yearbook 2021: 2-9 Gross Domestic Product by City (2019)
		Statistical Yearbook 2020: 2-8 Gross Domestic Product by City (2018)
		Statistical Yearbook 2018: 2-9 Gross Domestic Product by City (2017)
		Statistical Yearbook 2017: 3-9 Gross Domestic Product by City (2016)
		Statistical Yearbook 2016: 3-9 Gross Domestic Product by City (2015)
		Statistical Yearbook 2015: 3-9 Gross Domestic Product by City (2014)
		Statistical Yearbook 2014: 3-10 Gross Domestic Product by City (2013)
		Statistical Yearbook 2013: 3-10 Gross Domestic Product by City (2012)
Hubei	https://tjj.hubei.gov.cn/tjsj/sjkscx/tjnj/qstjnj/	Statistical Yearbook 2023: 0115-Gross Domestic Product of Cities and Prefectures (2022).xls
		Statistical Yearbook 2022: 0115-Gross Domestic Product of Cities and Prefectures (2021).xls
		Statistical Yearbook 2021: 0115-Gross Domestic Product of Cities and Prefectures (2020).xls
		Statistical Yearbook 2020: 0115-Gross Domestic Product of Cities and Prefectures (2019).xls
		Statistical Yearbook 2019: 0115-Gross Domestic Product of Cities and Prefectures (2018).xls
		Statistical Yearbook 2018: 0119-Gross Domestic Product of Cities and Prefectures (2017).xls
		Statistical Yearbook 2017: 0119-Gross Domestic Product of Cities and Prefectures (2016).xls
		Statistical Yearbook 2016: 0119-Gross Domestic Product of Cities and Prefectures (2015).xls
		Statistical Yearbook 2015: 0119-Gross Domestic Product of Cities and Prefectures (2014).xls
		Statistical Yearbook 2014: 0119-Gross Domestic Product of Cities and Prefectures (2013).xls
		Statistical Yearbook 2013: 0119-Gross Domestic Product of Cities and Prefectures (2012).xls

Province	Source	Dataset Name
Jiangsu	https://www.jiangsu.gov.cn/col/col84736/index.html	Jiangsu Province Statistical Yearbook 2024: 2-12 Gross Domestic Product by Region
		Jiangsu Province Statistical Yearbook 2023: 2-12 Gross Domestic Product by Region
		Jiangsu Province Statistical Yearbook 2022: 2-12 Gross Domestic Product by Region
		Jiangsu Province Statistical Yearbook 2021: 2-12 Gross Domestic Product by Region
		Jiangsu Province Statistical Yearbook 2020: 2-12 Gross Domestic Product by Region
		Jiangsu Province Statistical Yearbook 2019: 2-12 Gross Domestic Product by Region
		Jiangsu Province Statistical Yearbook 2018: 2-12 Gross Domestic Product by Region
Shandong	http://tjj.shandong.gov.cn/col/col6279/index.html	Shandong Province Statistical Yearbook 2024: 2-6 Gross Domestic Product by Region (2022)
		Shandong Province Statistical Yearbook 2023: 2-6 Gross Domestic Product by Region (2021)
		Shandong Province Statistical Yearbook 2022: 2-6 Gross Domestic Product by Region (2020)
		Shandong Province Statistical Yearbook 2021: 2-6 Gross Domestic Product by Region (2019)
		Shandong Province Statistical Yearbook 2020: 2-6 Gross Domestic Product by Region (2018)
		Shandong Province Statistical Yearbook 2019: 2-9 Gross Domestic Product by Region (2017)
		Shandong Province Statistical Yearbook 2018: 2-9 Gross Domestic Product by Region (2016)
		Shandong Province Statistical Yearbook 2017: 2-9 Gross Domestic Product by Region (2015)
		Shandong Province Statistical Yearbook 2016: 2-8 Gross Domestic Product by Region (2014)
		Shandong Province Statistical Yearbook 2015: 2-8 Gross Domestic Product by Region (2013)
		Shandong Province Statistical Yearbook 2014: 2-9 Gross Domestic Product by Region (2012)
Sichuan	https://oversea.cnki.net/knavi/YearbookDetail?pcode=CYFD&pykm=YSCTN	Sichuan Province Statistical Yearbook 2024: 2-7 Gross Domestic Product by Region
Zhejiang	https://tjj.zj.gov.cn/col/col1525563/index.html	Zhejiang Province Statistical Yearbook 2023: 17-2 Major Indicators of National Economy by City and County (2022)
		Zhejiang Province Statistical Yearbook 2022: 17-2 Major Indicators of National Economy by City and County (2021)
		Zhejiang Province Statistical Yearbook 2021: 17-2 Major Indicators of National Economy by City and County (2020)
		Zhejiang Province Statistical Yearbook 2020: 17-2 Major Indicators of National Economy by City and County (2019)
		Zhejiang Province Statistical Yearbook 2019: 17-2 Major Indicators of National Economy by City and County (2018)
		Zhejiang Province Statistical Yearbook 2018: 17-2 Major Indicators of National Economy by City and County (2017)
		Zhejiang Province Statistical Yearbook 2017: 17-2 Major Indicators of National Economy by City and County (2016)
		Zhejiang Province Statistical Yearbook 2016: 17-2 Major Indicators of National Economy by City and County (2015)
		Zhejiang Province Statistical Yearbook 2015: 17-2 Major Indicators of National Economy by City and County (2014)

Province	Source	Dataset Name
		Zhejiang Province Statistical Yearbook 2014: 17-2 Major Indicators of National Economy by City and County (2013)
		Zhejiang Province Statistical Yearbook 2013: 17-2 Major Indicators of National Economy by City and County (2012)

1.3 Predictors data

- Population data:
  - The population data for our study is sourced from the LandScan Global Metadata, developed by the Oak Ridge National Laboratory (Bright et al., n.d.). The link is https://landscan.ornl.gov/metadata.
  - Choose "LandScan Global Metadata" and download for years 2000 to 2022. Save the downloaded files as "landscan-global-20xx.tif" in the folder "step3_obtain_cell_level_GDP_and_predictors_data/inputs/population/".
- Nighttime Light (NTL) data:
  - The Nighttime Light (NTL) data utilized in our study is the VIIRS VNP46A4 product from NASA's Black Marble suite (Román et al., 2018). The link is https://ladsweb.modaps.eosdis.nasa.gov/archive/allData/5000/VNP46A4/.
  - Data for years 2012 to 2022 are all downloaded and processed. Save the corresponding year files in the folder "step3_obtain_cell_level_GDP_and_predictors_data/inputs/NTL_VNP46A4/20xx/001".
- Net Primary Productivity (NPP) data
  - The Net Primary Productivity (NPP) data is the MOD17A3HGF Version 6.1 product from the MODIS suite (Running and Zhao, 2021). The link is https://search.earthdata.nasa.gov/search?q=C2565791034-LPCLOUD.
  - Data for years 2012 to 2022 are all downloaded and processed. Save the corresponding year files in the folder "step3_obtain_cell_level_GDP_and_predictors_data/inputs/NPP_V061/20xx/".
- Landcover data:
  - The landcover data that our model uses is the MCD12Q1 version 6.1 product from the MODIS suite (Friedl and Sulla-Menashe, 2022). The link is https://search.earthdata.nasa.gov/search?q=C2484079608-LPCLOUD.
  - Data for years 2012 to 2022 are all downloaded and processed. Save the corresponding year files in the folder "step3_obtain_cell_level_GDP_and_predictors_data/inputs/landcover_MCD12Q1V061/20xx/".
- Carbon dioxide (CO2) emissions data:

- The carbon dioxide (CO2) emissions data in our study are sourced from the EDGAR - Emissions Database for Global Atmospheric Research community. Data version "v8.0_GHG 1970-2022 (CO2, CH4, N2O, F-gases)". The link is `https://edgar.jrc.ec.europa.eu/dataset_ghg80`.
- They separate the CO2 emissions into "IEA-EDGAR CO2" and "EDGAR_CO2bio". For both of them, we aggregate sector-specific emissions to three main categories: manufacturing combustion, heavy industry, and transportation.
- For manufacturing combustion category, go to the section "Annual sector-specific gridmaps (1970-2022)", click "IEA-EDGAR CO2" and "EDGAR_CO2bio" under the subsection of "Combustion for manufacturing". Then download the files named "COMPLETE TIMESERIES [2022-1970] ENETCDF (xxx_emi_nc,zip)".
- For heavy industry category, it includes the subsection "Power industry", "Oil refineries & Transformation industry", "Fuel Exploitation", "Non-metallic minerals production", "Iron and steel production", and "Non-ferrous metals production". Note for "EDGAR_CO2bio", only subsections "Power industry", "Oil refineries & Transformation industry", and "Fuel Exploitation" have the data.
- For transportation category, it includes the subsection "Road transportation", and "Shipping".
- Save the corresponding year files in the folders "step3_obtain_cell_level_GDP_and_predictors_data/inputs/CO2_bio_specific_sectors/sector name/" and "step3_obtain_cell_level_GDP_and_predictors_data/inputs/CO2_non_org_specific_sectors /sector name/".

• Gas flare data:

- The gas flare data in our study are sourced from the Global Gas Flaring Data dataset in the Global Flaring and Methane Reduction Partnership (GFMR) community (Zhizhin et al., 2021; Elvidge, Zhizhin, Hsu, et al., 2013; Elvidge, Zhizhin, Baugh, et al., 2016). The link is `https://www.worldbank.org/en/programs/gasflaringreduction/global-flaring-data`.
- Click the "Dataset: 2012 - 2022 Flare Volume Estimates by Individual Flare Location" to download and save the corresponding file in the folder "step3_obtain_cell_level_GDP_and_predictors_data/inputs/gas_flare_data/". The file should be named "2012-2023-Flare-Volume-Estimates-by-individual-Flare-Location.xlsx".

• Terrain Ruggedness Index data:

- We use the Terrain Ruggedness Index produced in Nunn and Puga, 2012, calculated as a 30 arc-second grid across the earth. The link is `https://diegopuga.org/data/rugged/`.
- The dataset was downloaded on 01/20/2023. Save the corresponding file "step3_obtain_cell_level_GDP_and_predictors_data/inputs/ruggedness/" and name it "tri.txt".

The figure 1 shows the locations of gas flare spots used in this study.

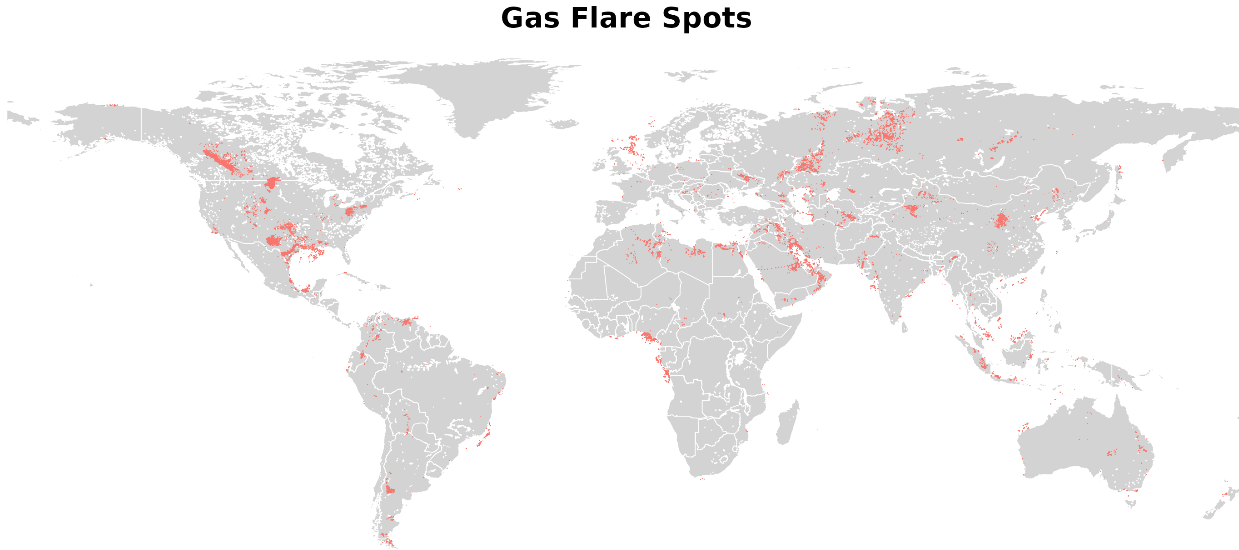

Figure 1: Locations of gas flare spots used in this study

1.4 Data Processing for Predictors

To prepare the geometry units for aggregating predictor data, the world is divided into grids with resolutions of 1°, 0.5°, and 0.25° in the EPSG:4326/WGS84 coordinate system. These grids are intersected with world country maps (as described in Section 1.1) to create country-cell-level geometry units. For specific countries, the geometry is at the state-equivalent level: Australia (8 units), Brazil (27 units), Canada (13 units), China (31 units), India (33 units), Kazakhstan (16 units), Mexico (32 units), Russia (83 units), and USA (51 units). Large inland water areas are then removed from the maps. This creates the final geometry units used for both GDP and predictor data. For NTL data aggregation only, an additional geometry unit is created by dividing the country/state-cell-level geometry into three land use categories: urban, cropland, and others by intersecting it with corresponding land use data.

Before aggregating predictors' data into the geometry unit created above, the following processing steps are needed:

The original NASA VIIRS VNP46A4 Black Marble NTL dataset separates data by satellite angles and into snow-covered and snow-free periods. For this study, we use the snow-free layer covering all satellite angles. The dataset has already been pre-processed to adjust for atmospheric and lunar reflectance conditions and to remove noise sources such as stray light, twilight, clouds, scan edge artifacts, and temporary lights (e.g., lightning, wildfires, fireworks). However, persistent lights from gas flares and fishing boats still remain. To address this, we further process it by excluding the lights within a 0.2deg radius around gas flaring locations with positive gas flaring volumes. Gas flaring locations, provided by the Global Gas Flaring Data from the GFMR community, include annual point locations with estimated flaring volumes. Positive flaring locations for each year are identified and corresponding exclusion zones are applied. The original NTL data are in a sinusoidal coordinate system and are reprojected to EPSG:4326 using bilinear interpolation. This ensures spatial consistency but can introduce slight changes to original pixel values. The processed data are aggregated to year-country/state-cell-landuse levels.

The NPP dataset is also in a sinusoidal coordinate system and undergoes the same reprojection process as the NTL data.

The original land use data classify land into 11 categories. We select and aggregate them into five categories: urban, water, cropland (combining cropland, forest cropland, and herbaceous cropland), forest (combining open and dense forest), and snow ice. Recall that only large inland waters are removed, so smaller water bodies can still appear in the data.

The original CO_2 emissions data categorizes emissions into multiple sectors. We select and aggregate into six categories: fossil CO_2 manufacturing combustion, biofuel CO_2 manufacturing combustion, fossil CO_2 heavy industry (power industry + oil refineries transformation industry + fuel exploitation + non-metallic minerals production + iron and steel production + non-ferrous metals production), biofuel CO_2 heavy industry (power industry + oil refineries transformation industry + fuel exploitation), fossil CO_2 transportation (road transportation + shipping), biofuel CO_2 transporation (road transportation + shipping). EDGAR generates gridded emissions data by downscaling national emissions to each 0.1deg cell using sector-specific spatial proxies. In our study, we will convert all predictors to share terms rather than level terms, so we only care about EDGAR's downscaling method. EDGAR allocates industry emissions based on the locations and emissions of energy and manufacturing facilities, while transportation emissions are distributed according to the length and intensity of road networks and shipping routes. This distribution process relies mainly on statistical calculations, not machine learning or economic models. And their input data for each sector are primarily sourced from national government agencies, company reports, and international organizations. Excluding these datasets has minimal impact on GDP level predictions but can reduce the accuracy of growth rate predictions.

The LandScan population dataset, when aggregated to the national level, slightly differs from the IMF WEO and World Bank national population datasets used in this study. To ensure consistency, LandScan data are rescaled to match IMF and World Bank population values.

After preprocessing the predictor data, the values for each cell's predictors are then extracted and converted into shares, ready for training the random forest model.

2 Training Countries and Reference for GDP Share Prediction

Our model is trained using data from countries that provide GDP information at the county-equivalent level for developed nations and at a slightly broader administrative level for developing ones. These training countries span Northern America, Southern America, Europe, Africa, and Asia. Figure 2 shows the training countries.

Our model predicts cell GDP share which by definition is the proportion of GDP attributed to a specific cell relative to the aggregate GDP of its higher administrative entity. Correspondingly, our predictors are expressed as shares; for instance, the fraction of population, the proportion of nighttime light intensity, and the percentage of urban areas. The shares are calculated at the national level for most countries, indicating, for example, the share of national nighttime lights that a particular cell occupies. However, to improve the accuracy of our predictions by incorporating more detailed subnational GDP information, we treat province-equivalent units as “parent countries” in nations such as Australia, Brazil, Canada, China, India, Kazakhstan, Mexico, Russia, and the United States, where such province-level GDP data are available. In these cases, both the predictor shares and the predicted cell GDP shares are relative to the provincial level. Figure 2 illustrates the boundaries of “parent countries” in this context.

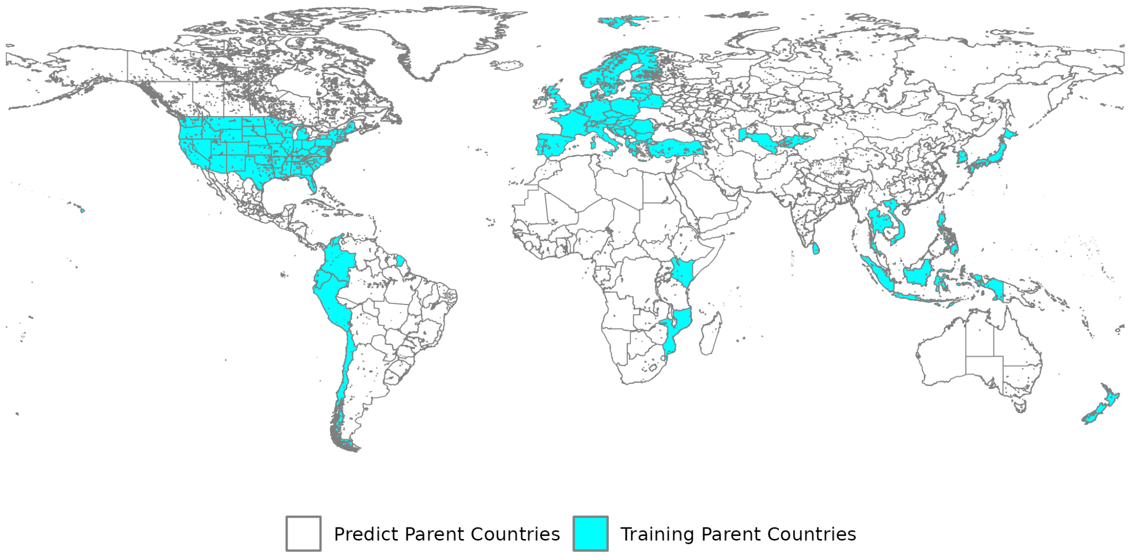

Figure 2: Boundaries of parent countries in the model

3 Method for Constructing Variable Importance Scores

Recall that the random forest models we use for training and prediction adopt cell GDP shares (expressed as a fraction of the parent state or country's GDP) as the target variable and draw on predictors that are also expressed as shares of the parent state or country. If we were to simply derive the variable importance scores from the prediction process (e.g., using a built-in R function), the scores generated would encompass a *scaling effect*: imagine a cell in Houston that has 1,000 times the population of a cell in West Texas. Since our prediction process concerns the distribution of state or country GDP (which is a known quantity) to cells, we could achieve great strides in prediction by simply distributing 1,000 times more GDP to the cell covering Houston than to the cell in West Texas. Here, population explains most of the margin in cross-cell variation in GDP shares (1,000 times); but, naturally, the variation in per capita GDP across cells within the same state or country is of a much smaller magnitude (e.g., less than 2 or 3 times). This warrants the construction of new variable importance scores that remove this scaling effect, in order to mitigate the bias given to cell population as the predominant predictor.

Our chosen approach considers the variation in $r_i = \log(y_i) - \log(p_i)$, where y_i is the GDP share and p_i is the population share of cell i. Essentially, we are asking the question: how well does each predictor explain the deviation of a cell's GDP share from what its population share may suggest? This approach does not require re-training or re-estimating the model. We are simply transforming our estimates of cell GDP shares and constructing variable importance scores based on a new form of variation.

We focus on within-country-year variation as we are interested in assessing how well each predictor explains the variation derived from our prediction pipeline which distributes a known country-year GDP to cells within the country-year. However, as shown in Tables 5 to 7, using overall variation does not qualitatively alter our findings.

Let $r_i = \log(y_i) - \log(p_i)$ denote the true log GDP-to-population ratio for cell i, where y_i is the true GDP share and p_i is the population share. Let $\hat{r}_i = \log(\hat{y}_i) - \log(p_i)$ denote the predicted log GDP-to-population ratio, where $\hat{y}_i$ is the model's predicted GDP share. Meanwhile, let $\bar{r}_{ct}$ denote the mean of r_i for country c in year t, and let $\bar{\hat{r}}_{ct}$ denote the corresponding mean of $\hat{r}_i$. Define the demeaned values as:

$$\begin{aligned}\tilde{r}_i &= r_i - \bar{r}_{c(i),t(i)},\\ \tilde{\hat{r}}_i &= \hat{r}_i - \bar{\hat{r}}_{c(i),t(i)},\end{aligned}$$

where $c(i)$ and $t(i)$ denote the country and year to which cell i belongs.

To assess the importance of predictor k, we compare the baseline model performance to performance after replacing predictor X_k with its global mean. Specifically, for each cell i, we replace $X_{k,i}$ with $\bar{X}_k = \frac{1}{N}\sum_j X_{k,j}$, where the sum is over all cells j in the dataset and N is the total number of cells. We then generate new predictions $\hat{y}_i^{(-k)}$ using the model with this replaced predictor, and compute $\hat{r}_i^{(-k)} = \log(\hat{y}_i^{(-k)}) - \log(p_i)$.

We calculate three within-country-year metrics on both the baseline predictions and after each variable replacement. The within R^2 compares prediction errors to within-country-year variance:

$$R^2_{\text{within}} = 1 - \frac{\sum_i (r_i - \hat{r}_i)^2}{\sum_i (r_i - \bar{r}_{c(i),t(i)})^2}. \tag{3}$$

Meanwhile, the within *Corr* is the Pearson correlation computed on demeaned values:

$$\text{Corr}_{\text{within}} = \text{Corr}(\tilde{r}_i, \tilde{\hat{r}}_i). \tag{4}$$

Furthermore, the within MSE measures the mean squared error of demeaned predictions:

$$\text{MSE}_{\text{within}} = \frac{1}{N}\sum_i (\tilde{r}_i - \tilde{\hat{r}}_i)^2. \tag{5}$$

Variable importance for predictor k is defined as the drop in each metric relative to baseline: $\text{Importance}_k = \text{Baseline metric} - \text{Metric after replacing } X_k$. Larger positive values indicate that within-country-year variation in X_k contributes more to model performance.

The variable importance scores constructed from these three metrics (and their counterparts that capture overall variation) for the baseline model are presented in Tables 5 to 7.

Table 5: Variable Importance: 1deg Resolution

	R^2 drop		*Corr* drop		MSE increase	
Variable	Within	Overall	Within	Overall	Within	Overall
Lag population (urban)	32.21	27.23	0.2020	0.1865	2.83	4.28
Population (Urban)	21.29	18.00	0.1933	0.1811	2.00	2.83
Population (other)	20.23	17.10	0.1815	0.1711	1.98	2.69
Lag population (other)	4.47	3.78	0.1433	0.1422	0.49	0.59
NTL (cropland)	3.55	3.00	0.1365	0.1361	0.39	0.47
Lag NTL (cropland)	2.70	2.28	0.1278	0.1285	0.30	0.36
Population (cropland)	1.30	1.10	0.1038	0.1052	0.14	0.17
Population (total)	0.76	0.59	0.1226	0.1096	7.71	15.68
NTL (urban)	0.37	0.32	0.1021	0.1217	0.02	0.05
Lag NTL (urban)	0.18	0.15	0.0528	0.0648	0.01	0.02
Lag cropland	0.14	0.12	0.0323	0.0356	0.02	0.02
Cropland	0.11	0.10	0.0241	0.0263	0.01	0.02
Lag population (cropland)	0.07	0.06	0.0165	0.0147	0.01	0.01
CO2 bio (manuf. combust.)	0.07	0.06	0.0196	0.0198	0.00	0.01
Lag CO2 non-org (manuf. combust.)	0.05	0.05	0.0202	0.0179	0.00	0.01
CO2 non-org (manuf. combust.)	0.05	0.04	0.0184	0.0166	0.00	0.01
Lag CO2 non-org (heavy industry)	0.05	0.04	0.0099	0.0092	0.00	0.01
Lag CO2 bio (manuf. combust.)	0.03	0.03	0.0055	0.0075	0.00	0.00
NTL (other)	0.02	0.02	0.0096	0.0090	0.00	0.00
Lag CO2 bio (transport)	0.02	0.02	0.0019	0.0032	0.00	0.00

Notes: This table reports importance scores after replacing each variable with its global mean: drop in R^2, drop in *Corr*, and increase in MSE. Variables shown are those ranking in the top 20 based on within-country-year R^2 importance. See Section 3 for the methodology used for constructing these scores.

Table 6: Variable Importance: 0.5deg Resolution

	R^2 drop		*Corr* drop		MSE increase	
Variable	Within	Overall	Within	Overall	Within	Overall
Population (urban)	41.02	34.74	0.1999	0.1854	3.31	6.21
Lag population (rrban)	40.46	34.27	0.1969	0.1828	3.25	6.13
Population (other)	19.61	16.61	0.1805	0.1724	2.02	2.97
Lag CO2 non-org (heavy industry)	16.12	13.65	0.1810	0.1758	1.79	2.44
CO2 non-org (heavy industry)	14.77	12.50	0.1795	0.1747	1.64	2.24
Population (cropland)	8.34	7.06	0.1608	0.1558	0.93	1.26
Lag population (cropland)	7.09	6.00	0.1567	0.1517	0.78	1.07
Lag population (other)	6.96	5.90	0.1559	0.1533	0.81	1.05
NTL (cropland)	6.01	5.09	0.1534	0.1562	0.77	0.91
Lag NTL (urban)	4.30	3.64	0.1433	0.1421	0.50	0.65
NTL (urban)	3.48	2.95	0.1363	0.1351	0.40	0.53
Lag NTL (cropland)	2.92	2.48	0.1292	0.1326	0.38	0.44
Lag CO2 bio (manuf. combust.)	2.08	1.76	0.1253	0.1296	0.25	0.32
NTL (other)	2.02	1.71	0.1177	0.1180	0.24	0.31
CO2 bio (manuf. combust.)	0.98	0.83	0.0995	0.1040	0.09	0.15
Population (total)	0.69	0.48	-0.1073	-0.1090	5.93	13.29
CO2 non-org (manuf. combust.)	0.29	0.24	0.0429	0.0452	0.03	0.04
Lag CO2 nonorg (manuf. combust.)	0.21	0.17	0.0356	0.0389	0.02	0.03
Lag NTL (other)	0.04	0.04	0.0106	0.0101	0.00	0.01
Cropland	0.04	0.04	0.0089	0.0079	0.00	0.01

Notes: This table reports importance scores after replacing each variable with its global mean: drop in R^2, drop in *Corr*, and increase in MSE. Variables shown are those ranking in the top 20 based on within-country-year R^2 importance. See Section 3 for the methodology used for constructing these scores.

Table 7: Variable Importance: 0.25deg Resolution

	R^2 drop		*Corr* drop		MSE increase	
Variable	Within	Overall	Within	Overall	Within	Overall
Population (urban)	27.13	22.82	0.1682	0.1632	2.20	4.26
Population (other)	18.72	15.75	0.1599	0.1581	1.79	2.94
Lag population (urban)	11.71	9.85	0.1569	0.1568	1.14	1.84
CO2 non-org (heavy industry)	11.66	9.81	0.1589	0.1606	1.22	1.83
Lag CO2 non-org (heavy industry)	10.25	8.62	0.1580	0.1609	1.10	1.61
Population (cropland)	8.17	6.87	0.1489	0.1505	0.86	1.28
Lag population (other)	6.52	5.48	0.1464	0.1500	0.72	1.02
Lag population (cropland)	5.48	4.61	0.1416	0.1450	0.60	0.86
NTL (cropland)	4.79	4.03	0.1429	0.1491	0.57	0.75
NTL (urban)	1.95	1.64	0.1152	0.1196	0.22	0.31
Lag NTL (cropland)	1.81	1.52	0.1115	0.1161	0.22	0.28
Lag NTL (urban)	1.58	1.33	0.1068	0.1113	0.18	0.25
Population (total)	0.97	0.61	-0.1295	-0.1108	5.13	12.95
Lag CO2 bio (manuf. combust.)	0.97	0.82	0.1029	0.1131	0.12	0.15
CO2 bio (manuf. combust.)	0.34	0.28	0.0581	0.0662	0.03	0.05
NTL (other)	0.26	0.22	0.0418	0.0436	0.03	0.04
Lag CO2 nonorg (manuf. combust.)	0.10	0.08	0.0189	0.0199	0.01	0.02
CO2 non-org (manuf. combust.)	0.08	0.07	0.0144	0.0148	0.01	0.01
Lag CO2 bio (transport)	0.01	0.01	0.0033	0.0040	0.00	0.00
Cropland	0.01	0.01	0.0030	0.0035	0.00	0.00

Notes: This table reports importance scores after replacing each variable with its global mean: drop in R^2, drop in *Corr*, and increase in MSE. Variables shown are those ranking in the top 20 based on within-country-year R^2 importance. See Section 3 for the methodology used for constructing these scores.

4 Post-adjustments

This section presents the number of cells affected by the post-adjustment process for each censoring threshold at each resolution.

Table 8: 1-degree cells affected by post-adjustments: threshold 0 pop per km2 land areas

Year	Total number of cells	Number of cells affected	Affected cells with zero pop	Percentage of inhabited cells affected	Affected cells with positive GDP
2012	21,194	4,112	3,713	2%	399
2013	21,194	3,888	3,466	2%	420
2014	21,194	3,901	3,461	2%	439
2015	21,194	3,928	3,498	2%	429
2016	21,194	3,812	3,366	3%	446
2017	21,194	3,813	3,349	3%	464
2018	21,194	3,924	3,468	3%	456
2019	21,194	4,210	3,726	3%	484
2020	21,194	4,612	4,086	3%	526
2021	21,194	4,420	3,885	3%	534
2022	20,416	3,608	3,071	3%	536

Table 9: 1-degree cells affected by post-adjustments: threshold 0.01 pop per km2 land areas

Year	Total number of cells	Number of cells affected	Affected cells with zero pop	Percentage of inhabited cells affected	Affected cells with positive GDP
2012	21,194	5,283	3,713	9%	415
2013	21,194	5,030	3,466	9%	423
2014	21,194	5,124	3,461	9%	436
2015	21,194	5,119	3,498	9%	432
2016	21,194	5,188	3,366	10%	454
2017	21,194	5,208	3,349	10%	454
2018	21,194	5,253	3,468	10%	436
2019	21,194	5,431	3,726	10%	445
2020	21,194	5,968	4,086	11%	450
2021	21,194	5,848	3,885	11%	461
2022	20,416	5,023	3,071	11%	464

Table 10: 1-degree cells affected by post-adjustments: threshold 0.02 pop per km2 land areas

Year	Total number of cells	Number of cells affected	Affected cells with zero pop	Percentage of inhabited cells affected	Affected cells with positive GDP
2012	21,194	5,692	3,713	11%	419
2013	21,194	5,480	3,466	11%	425
2014	21,194	5,550	3,461	12%	430
2015	21,194	5,548	3,498	12%	426
2016	21,194	5,616	3,366	13%	449
2017	21,194	5,633	3,349	13%	445
2018	21,194	5,706	3,468	13%	431
2019	21,194	5,884	3,726	12%	448
2020	21,194	6,318	4,086	13%	444
2021	21,194	6,244	3,885	14%	451
2022	20,416	5,440	3,071	14%	450

Table 11: 1-degree cells affected by post-adjustments: threshold 0.05 pop per km2 land areas

Year	Total number of cells	Number of cells affected	Affected cells with zero pop	Percentage of inhabited cells affected	Affected cells with positive GDP
2012	21,194	6,383	3,713	15%	420
2013	21,194	6,212	3,466	15%	422
2014	21,194	6,295	3,461	16%	440
2015	21,194	6,327	3,498	16%	443
2016	21,194	6,376	3,366	17%	454
2017	21,194	6,387	3,349	17%	445
2018	21,194	6,411	3,468	17%	428
2019	21,194	6,555	3,726	16%	448
2020	21,194	6,902	4,086	16%	443
2021	21,194	6,877	3,885	17%	458
2022	20,416	6,074	3,071	17%	460

Table 12: 0.5-degree cells affected by post-adjustments: threshold 0 pop per km2 land areas

Year	Total number of cells	Number of cells affected	Affected cells with zero pop	Percentage of inhabited cells affected	Affected cells with positive GDP
2012	73,728	18,327	17,390	2%	919
2013	73,728	17,098	16,071	2%	1,010
2014	73,728	17,211	16,159	2%	1,032
2015	73,728	17,291	16,250	2%	1,020
2016	73,728	17,291	16,204	2%	1,057
2017	73,728	17,370	16,293	2%	1,055
2018	73,728	17,616	16,545	2%	1,043
2019	73,728	18,686	17,601	2%	1,066
2020	73,728	20,570	19,432	2%	1,114
2021	73,728	20,071	18,925	2%	1,129
2022	70,826	16,912	15,750	2%	1,142

Table 13: 0.5-degree cells affected by post-adjustments: threshold 0.01 pop per km2 land areas

Year	Total number of cells	Number of cells affected	Affected cells with zero pop	Percentage of inhabited cells affected	Affected cells with positive GDP
2012	73,728	21,911	17,417	8%	916
2013	73,728	20,655	16,080	8%	972
2014	73,728	20,942	16,165	8%	987
2015	73,728	20,995	16,253	8%	979
2016	73,728	21,221	16,211	9%	1,003
2017	73,728	21,384	16,301	9%	1,000
2018	73,728	21,520	16,569	9%	988
2019	73,728	22,357	17,632	8%	995
2020	73,728	24,452	19,460	9%	962
2021	73,728	24,239	18,952	10%	987
2022	70,826	20,991	15,806	9%	988

Table 14: 0.5-degree cells affected by post-adjustments: threshold 0.02 pop per km2 land areas

Year	Total number of cells	Number of cells affected	Affected cells with zero pop	Percentage of inhabited cells affected	Affected cells with positive GDP
2012	73,728	23,365	17,418	11%	935
2013	73,728	22,165	16,080	11%	958
2014	73,728	22,494	16,165	11%	970
2015	73,728	22,582	16,253	11%	965
2016	73,728	22,820	16,212	11%	984
2017	73,728	22,910	16,301	12%	979
2018	73,728	23,075	16,569	11%	964
2019	73,728	23,888	17,632	11%	971
2020	73,728	25,717	19,460	12%	918
2021	73,728	25,648	18,952	12%	937
2022	70,826	22,503	15,807	12%	945

Table 15: 0.5-degree cells affected by post-adjustments: threshold 0.05 pop per km2 land areas

Year	Total number of cells	Number of cells affected	Affected cells with zero pop	Percentage of inhabited cells affected	Affected cells with positive GDP
2012	73,728	25,744	17,418	15%	960
2013	73,728	24,710	16,080	15%	931
2014	73,728	25,107	16,165	16%	926
2015	73,728	25,156	16,253	15%	942
2016	73,728	25,331	16,212	16%	953
2017	73,728	25,471	16,301	16%	956
2018	73,728	25,649	16,569	16%	946
2019	73,728	26,301	17,632	15%	931
2020	73,728	27,745	19,460	15%	896
2021	73,728	27,830	18,952	16%	909
2022	70,826	24,774	15,807	16%	914

Table 16: 0.25-degree cells affected by post-adjustments: threshold 0 pop per km2 land areas

Year	Total number of cells	Number of cells affected	Affected cells with zero pop	Percentage of inhabited cells affected	Affected cells with positive GDP
2012	271,692	85,793	83,535	1%	2,204
2013	271,692	79,412	76,976	1%	2,365
2014	271,692	80,405	77,957	1%	2,372
2015	271,692	80,784	78,331	1%	2,387
2016	271,692	81,597	79,084	1%	2,435
2017	271,692	82,404	79,855	1%	2,470
2018	271,692	83,015	80,499	1%	2,437
2019	271,692	87,240	84,741	1%	2,436
2020	271,692	94,604	92,134	1%	2,396
2021	271,692	94,284	91,786	1%	2,429
2022	260,590	82,298	79,797	1%	2,439

Table 17: 0.25-degree cells affected by post-adjustments: threshold 0.01 pop per km2 land areas

Year	Total number of cells	Number of cells affected	Affected cells with zero pop	Percentage of inhabited cells affected	Affected cells with positive GDP
2012	271,692	94,485	83,827	6%	2,138
2013	271,692	88,746	77,046	6%	2,213
2014	271,692	89,796	78,006	6%	2,232
2015	271,692	90,074	78,374	6%	2,232
2016	271,692	90,909	79,147	6%	2,284
2017	271,692	91,779	79,920	6%	2,302
2018	271,692	92,196	80,664	6%	2,285
2019	271,692	95,766	84,889	6%	2,284
2020	271,692	103,621	92,268	6%	2,189
2021	271,692	103,731	91,902	7%	2,266
2022	260,590	91,423	80,132	6%	2,280

Table 18: 0.25-degree cells affected by post-adjustments: threshold 0.02 pop per km2 land areas

Year	Total number of cells	Number of cells affected	Affected cells with zero pop	Percentage of inhabited cells affected	Affected cells with positive GDP
2012	271,692	99,391	83,830	8%	2,112
2013	271,692	93,698	77,048	9%	2,144
2014	271,692	94,984	78,008	9%	2,185
2015	271,692	95,406	78,376	9%	2,185
2016	271,692	96,230	79,148	9%	2,208
2017	271,692	97,131	79,920	9%	2,221
2018	271,692	97,528	80,664	9%	2,202
2019	271,692	100,781	84,889	9%	2,180
2020	271,692	108,263	92,269	9%	2,062
2021	271,692	108,850	91,903	9%	2,120
2022	260,590	96,681	80,138	9%	2,146

Table 19: 0.25-degree cells affected by post-adjustments: threshold 0.05 pop per km2 land areas

Year	Total number of cells	Number of cells affected	Affected cells with zero pop	Percentage of inhabited cells affected	Affected cells with positive GDP
2012	271,692	107,285	83,835	12%	2,030
2013	271,692	102,482	77,053	13%	2,014
2014	271,692	104,076	78,010	13%	2,054
2015	271,692	104,417	78,377	13%	2,059
2016	271,692	105,125	79,151	13%	2,066
2017	271,692	106,154	79,921	14%	2,066
2018	271,692	106,446	80,667	13%	2,054
2019	271,692	109,311	84,890	13%	2,010
2020	271,692	115,781	92,272	13%	1,939
2021	271,692	116,742	91,905	14%	1,979
2022	260,590	104,882	80,144	14%	1,994

5 Model Within Training Sample Fit

Recall that in the paper Section 2.1 and 2.2, we present the results of models trained using data from 2012 to 2022 for available countries (excluding China) under the optimal hyper-parameters. Here we present the within-sample fit of the models on the training sample. We report the the R^2 of log GDP levels and annual changes for each country in the training sample. Note that the predicted values are out-of-bag (OOB) estimates. The out-of-bag estimate is calculated using predictions from trees that excluded the cell from their training.

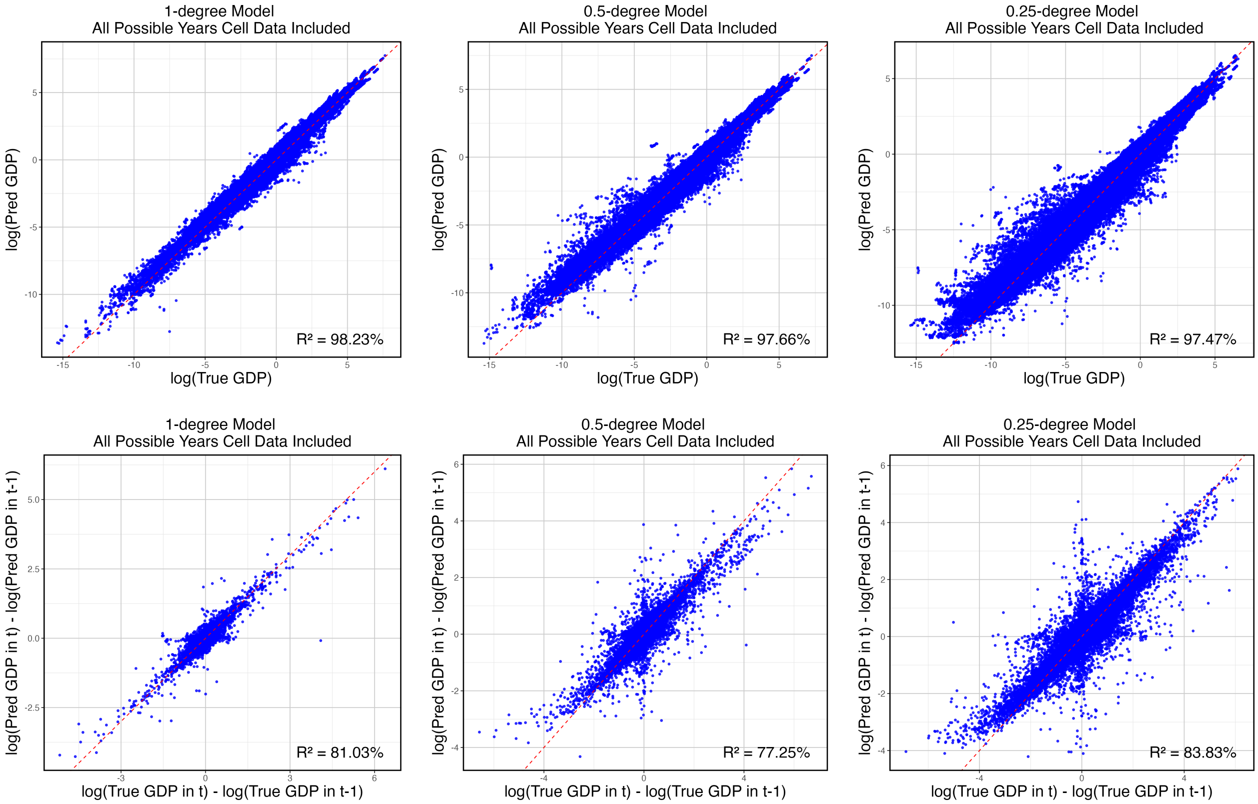

Figure 3: Predicted vs. True GDP Across Different Spatial Resolutions for Training Data

Table 20: R^2 Results for the 1-Degree Model: Comparing Predicted vs. Actual Log Cell GDP

Iso: Developed	R2: Developed	Iso: Developing	R2: Developing
AUT	99.58%	ALB	96.40%
BEL	98.07%	BGR	98.53%
CHE	97.02%	BIH	92.40%
CZE	99.24%	BLR	98.72%
DEU	98.28%	CHL	98.29%
DNK	99.13%	COL	94.89%
ESP	97.72%	ECU	94.50%
EST	96.16%	HUN	97.78%
FIN	99.30%	IDN	96.33%
FRA	98.63%	KEN	91.09%
GBR	99.31%	KGZ	96.73%
GRC	98.28%	LKA	94.68%
HRV	96.24%	MOZ	95.36%
ITA	97.69%	PER	97.60%
JPN	99.12%	PHL	98.38%
KOR	98.93%	POL	96.79%
LTU	98.94%	ROU	98.46%
LVA	97.24%	SRB	95.39%
NLD	98.54%	THA	85.01%
NOR	99.52%	TUR	92.17%
NZL	98.96%	UZB	98.81%
PRT	99.47%	VNM	93.40%
SVK	98.24%		
SVN	97.84%		
SWE	99.66%		
USA	98.09%		

Table 21: R^2 Results for the 0.5-Degree Model: Comparing Predicted vs. Actual Log Cell GDP

Iso: Developed	R2: Developed	Iso: Developing	R2: Developing
AUT	99.30%	ALB	97.08%
BEL	98.70%	BGR	97.60%
CHE	98.71%	BIH	94.75%
CZE	99.43%	BLR	98.09%
DEU	97.22%	CHL	96.93%
DNK	99.37%	COL	94.88%
ESP	98.28%	ECU	90.46%
EST	96.27%	HUN	98.73%
FIN	99.27%	IDN	95.37%
FRA	98.22%	KEN	90.14%
GBR	98.59%	KGZ	96.28%
GRC	98.45%	LKA	98.70%
HRV	98.59%	MOZ	92.41%
ITA	96.19%	PER	97.32%
JPN	98.63%	PHL	97.92%
KOR	99.02%	POL	96.38%
LTU	96.57%	ROU	96.90%
LVA	95.08%	SRB	91.33%
NLD	99.11%	THA	91.27%
NOR	99.21%	TUR	88.39%
NZL	97.81%	UZB	96.69%
PRT	99.08%	VNM	94.62%
SVK	97.95%		
SVN	99.55%		
SWE	99.18%		
USA	97.17%		

Table 22: R^2 Results for the 0.25-Degree Model: Comparing Predicted vs. Actual Log Cell GDP

Iso: Developed	R2: Developed	Iso: Developing	R2: Developing
AUT	98.51%	ALB	97.56%
BEL	98.43%	BGR	96.13%
CHE	98.56%	BIH	91.76%
CZE	99.37%	BLR	96.74%
DEU	95.57%	CHL	97.55%
DNK	99.33%	COL	94.40%
ESP	98.28%	ECU	90.85%
EST	96.84%	HUN	98.07%
FIN	99.56%	IDN	94.36%
FRA	98.00%	KEN	88.93%
GBR	98.52%	KGZ	94.69%
GRC	97.88%	LKA	98.92%
HRV	98.08%	MOZ	88.65%
ITA	96.11%	PER	97.17%
JPN	98.64%	PHL	96.56%
KOR	98.28%	POL	95.86%
LTU	95.94%	ROU	96.43%
LVA	90.10%	SRB	86.10%
NLD	99.27%	THA	88.32%
NOR	98.86%	TUR	87.18%
NZL	98.34%	UZB	93.86%
PRT	97.62%	VNM	90.66%
SVK	97.91%		
SVN	99.04%		
SWE	99.50%		
USA	97.20%		

Table 23: R^2 Results for the 1-Degree Model: Comparing Predicted vs. Actual Log Differences in Cell GDP (log(cell GDP in t) - log(cell GDP in t-1))

Iso: Developed	R2: Developed	Iso: Developing	R2: Developing
AUT	66.58%	ALB	79.95%
BEL	88.33%	BGR	53.74%
CHE	-62.28%	BIH	78.21%
CZE	84.13%	BLR	83.27%
DEU	76.24%	CHL	95.05%
DNK	79.32%	COL	73.36%
ESP	30.13%	ECU	17.16%
EST	84.53%	HUN	72.11%
FIN	92.54%	IDN	66.41%
FRA	87.31%	KEN	55.12%
GBR	81.21%	KGZ	73.53%
GRC	75.37%	LKA	46.14%
HRV	60.37%	MOZ	86.26%
ITA	88.85%	PER	84.17%
JPN	65.10%	PHL	55.20%
KOR	77.20%	POL	82.71%
LTU	84.67%	ROU	68.96%
LVA	6.08%	SRB	71.72%
NLD	24.08%	THA	12.61%
NOR	86.70%	TUR	8.60%
NZL	53.18%	UZB	82.95%
PRT	66.85%	VNM	-19.93%
SVK	90.70%		
SVN	92.29%		
SWE	91.47%		
USA	74.03%		

Table 24: R^2 Results for the 0.5-Degree Model: Comparing Predicted vs. Actual Log Differences in Cell GDP (log(cell GDP in t) - log(cell GDP in t-1))

Iso: Developed	R2: Developed	Iso: Developing	R2: Developing
AUT	67.66%	ALB	89.90%
BEL	-522.63%	BGR	60.95%
CHE	31.87%	BIH	89.74%
CZE	84.46%	BLR	87.73%
DEU	69.99%	CHL	84.04%
DNK	93.36%	COL	77.88%
ESP	65.29%	ECU	-22.73%
EST	73.60%	HUN	77.86%
FIN	92.19%	IDN	46.19%
FRA	87.95%	KEN	64.93%
GBR	56.49%	KGZ	68.09%
GRC	71.38%	LKA	71.48%
HRV	76.09%	MOZ	69.39%
ITA	49.77%	PER	86.07%
JPN	69.70%	PHL	62.05%
KOR	67.72%	POL	88.03%
LTU	85.50%	ROU	70.84%
LVA	51.30%	SRB	60.32%
NLD	75.60%	THA	65.80%
NOR	88.01%	TUR	5.05%
NZL	63.26%	UZB	78.90%
PRT	75.24%	VNM	44.88%
SVK	90.02%		
SVN	92.32%		
SWE	91.49%		
USA	74.43%		

Table 25: R^2 Results for the 0.25-Degree Model: Comparing Predicted vs. Actual Log Differences in Cell GDP (log(cell GDP in t) - log(cell GDP in t-1))

Iso: Developed	R2: Developed	Iso: Developing	R2: Developing
AUT	78.50%	ALB	89.12%
BEL	-24.86%	BGR	70.49%
CHE	18.34%	BIH	89.74%
CZE	88.05%	BLR	87.13%
DEU	80.51%	CHL	90.67%
DNK	92.00%	COL	85.64%
ESP	73.29%	ECU	-216.44%
EST	66.09%	HUN	65.89%
FIN	96.76%	IDN	69.84%
FRA	80.28%	KEN	82.19%
GBR	57.01%	KGZ	67.81%
GRC	88.34%	LKA	72.68%
HRV	88.52%	MOZ	80.20%
ITA	85.81%	PER	87.31%
JPN	72.83%	PHL	63.49%
KOR	67.28%	POL	79.52%
LTU	88.39%	ROU	64.78%
LVA	76.96%	SRB	70.81%
NLD	72.51%	THA	49.65%
NOR	96.03%	TUR	42.81%
NZL	74.71%	UZB	73.58%
PRT	80.71%	VNM	-117.47%
SVK	84.74%		
SVN	91.69%		
SWE	96.32%		
USA	83.72%		

6 Model Performance Under COVID Shock

In the paper, we evaluate the model's performance under the COVID-19 shock using China's data, which is excluded from the training sample. Here, we further demonstrate that the model's strong performance on COVID-affected data is not due to prior exposure to COVID years in the training sample. To verify this, we train the models using data from 2012 to 2019 for all available countries (excluding China) and assess whether the model's performance remains consistent on out-of-sample data. Note that all years of data for China are treated as out-of-sample and only the years 2020, 2021 and 2022 are considered out-of-sample for the training countries.

The results of the China test, as shown in Figure 4, align closely with the findings presented in the paper. Similarly, the tests conducted on the training countries, presented in Figures 5, 6, and 7, for the years 2020, 2021 and 2022 also demonstrate strong performance. These results highlight the model's ability to generalize effectively, even in the absence of specific shocks in the training data.

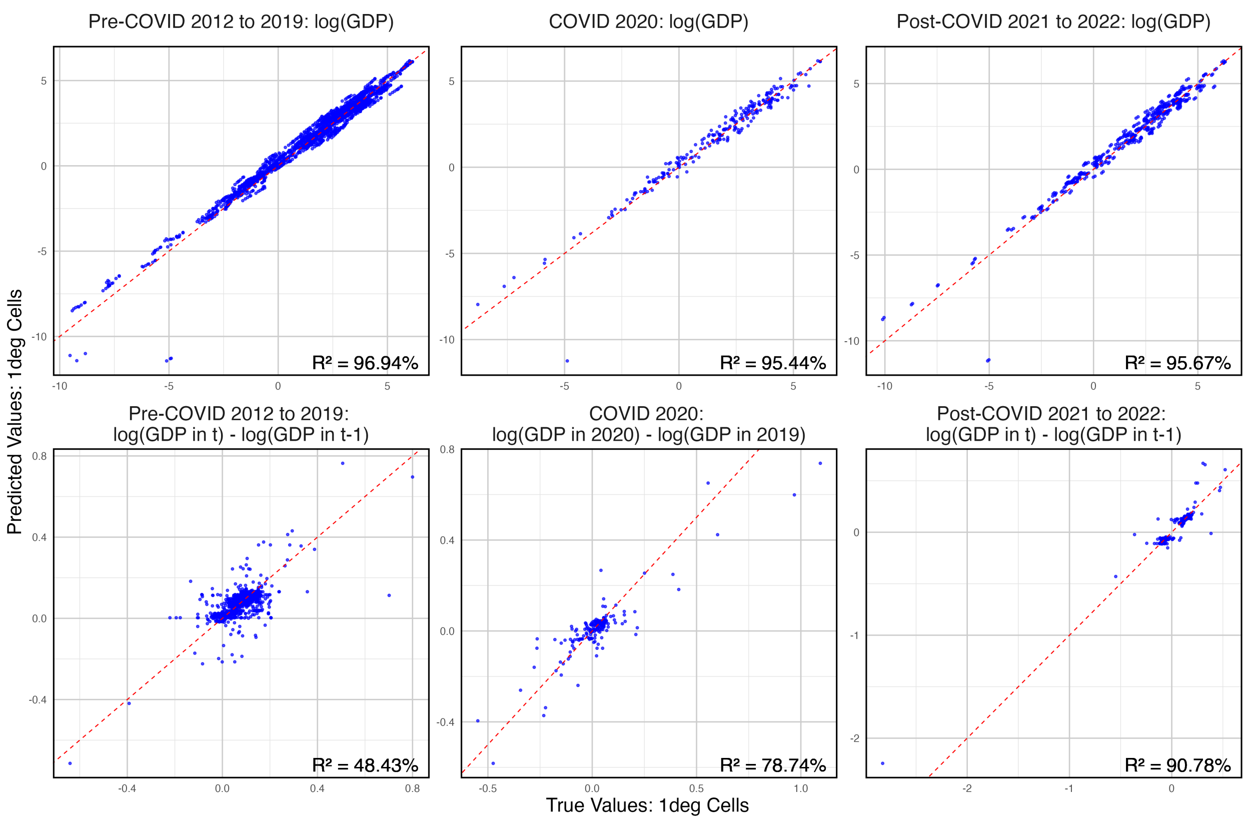

Figure 4: Model Predictions Against Actual Values in Billion Constant 2021 USD for Seven Leading Provinces in China

Note: The red dashed line represents the 45-degree line. Cells with a GDP value of zero are omitted to enable the calculation of logarithmic values.

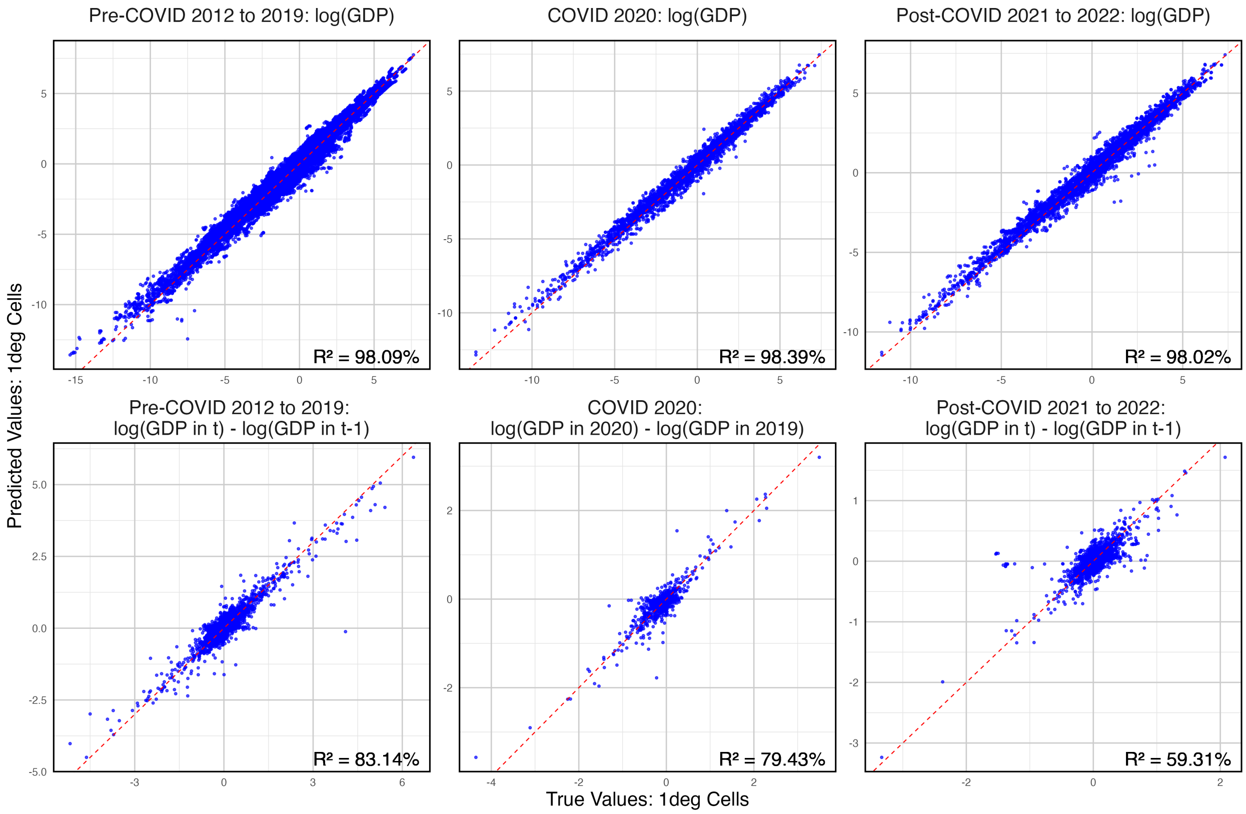

Figure 5: 1deg Model Predictions Against Actual Values in Billion Constant 2021 USD for All Training Countries

Note: The red dashed line represents the 45-degree line. Cells with a GDP value of zero are omitted to enable the calculation of logarithmic values. Data for 2012 to 2019 are within training sample, so use out-of-bag predictions. Data for years 2020 to 2022 are predictions from the model.

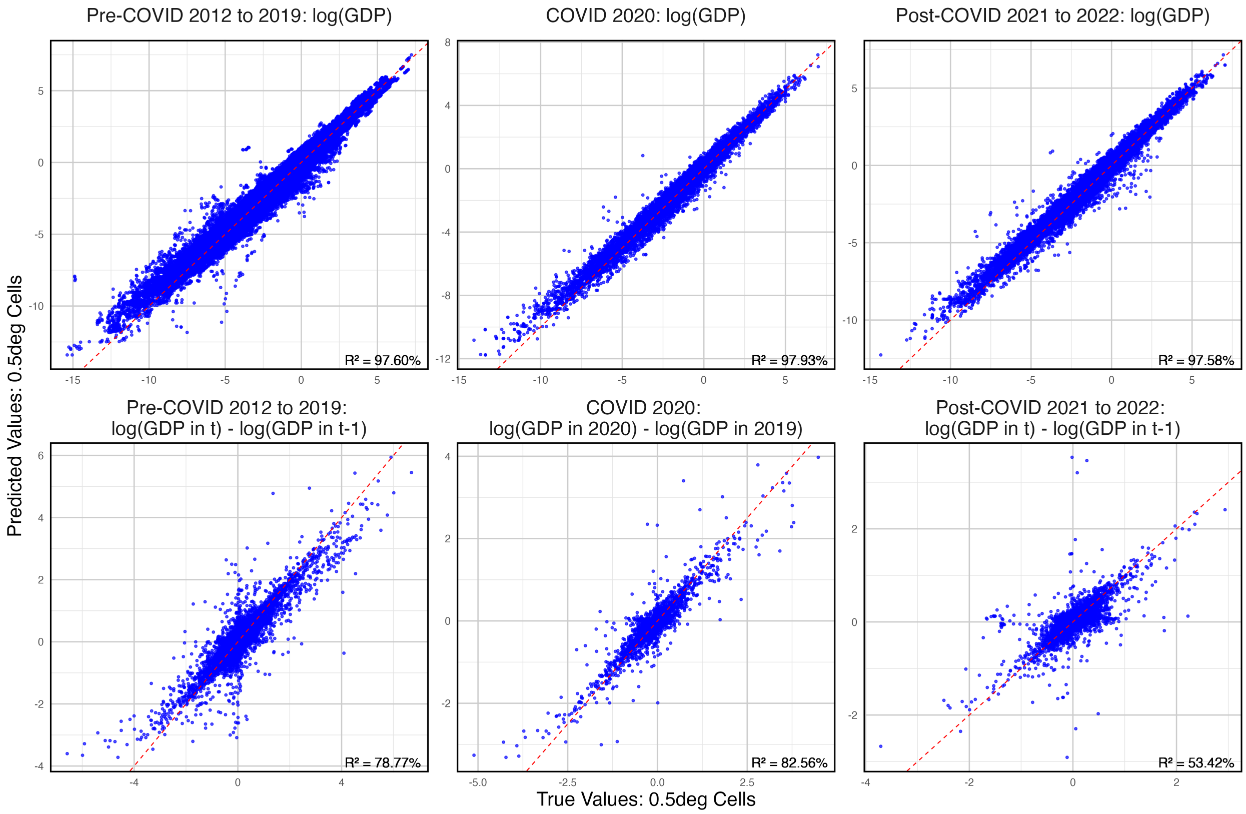

Figure 6: 0.5deg Model Predictions Against Actual Values in Billion Constant 2021 USD for All Training Countries

Note: The red dashed line represents the 45-degree line. Cells with a GDP value of zero are omitted to enable the calculation of logarithmic values. Data for 2012 to 2019 are within training sample, so use out-of-bag predictions. Data for years 2020 to 2022 are predictions from the model.

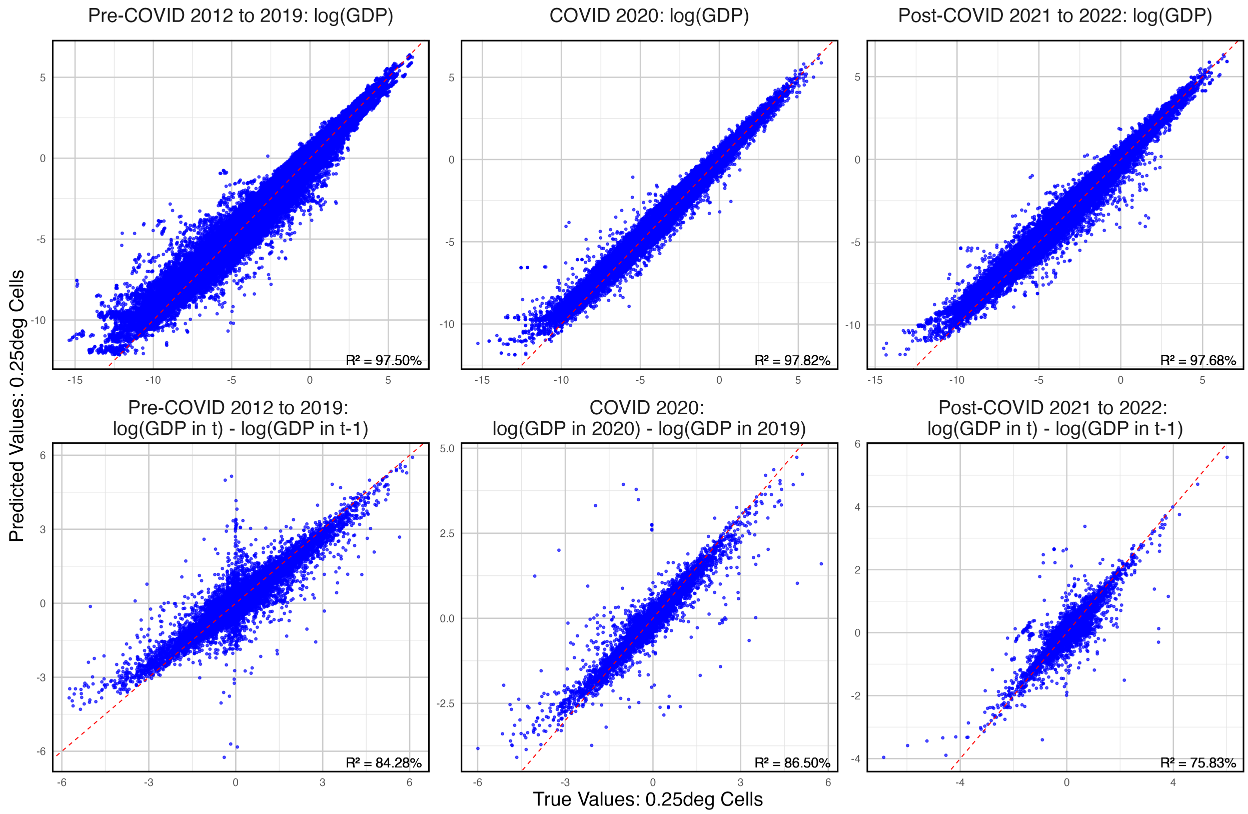

Figure 7: 0.25deg Model Predictions Against Actual Values in Billion Constant 2021 USD for All Training Countries

Note: The red dashed line represents the 45-degree line. Cells with a GDP value of zero are omitted to enable the calculation of logarithmic values. Data for 2012 to 2019 are within training sample, so use out-of-bag predictions. Data for years 2020 to 2022 are predictions from the model.

7 Robustness Check

7.1 Compare Benchmark Models with Models Tuned Based on Mean Square Error

In this section, we compare the results of two models: the model trained using data from 2012 to 2022 for all available countries (excluding China) with hyperparameters chosen to maximize the R^2 of annual log changes in GDP, and a model trained on the same data but with hyperparameters chosen to minimize the mean squared error (MSE) of GDP share. This comparison is essential because traditional hyperparameter tuning in machine learning is often based on minimizing MSE for the predicted variable - in our case, GDP share.

Table 26 shows the cross-validation performance of the model tuned to minimize MSE. Compare this table with the corresponding table in the paper, we show that while GDP level predictions are largely similar between the two models, tuning for maximizing annual log change R^2 improves the accuracy of year-over-year GDP changes, especially for 1deg model.

We then use the two models to predict all global cells and compare their predictions. Figure 8 demonstrates that the level predictions are highly consistent, while the annual changes exhibit slight differences especially at the 1deg and 0.25deg resolutions.

We also performed the same tests on the MSE-tuned model for comparison with the benchmark model, as shown in Figures 9, 10, and 11. The results indicate that the benchmark model outperforms the MSE-tuned model in capturing annual changes in GDP.

Table 26: Cross-Validated Performance Metrics Across Spatial Resolutions (MSE Objective)

	1-degree Model		0.5-degree Model		0.25-degree Model	
Panel A: Mean Square Error (MSE)						
MSE (Developed)	0.00046		0.00018		0.00001	
MSE (Developing)	0.00040		0.00009		0.00001	
MSE (All)	0.00033		0.00011		0.00001	
Weighted MSE	0.00042		0.00011		0.00001	
Panel B: R^2 of Log GDP Level						
R^2 (Developed)	97.81%		97.40%		97.87%	
R^2 (Developing)	95.61%		95.75%		93.83%	
R^2 (All)	97.66%		97.12%		96.96%	
Weighted R^2	96.22%		96.21%		94.96%	
Panel C: R^2 of $\log(GDP_t) - \log(GDP_{t-1})$						
R^2 (Developed)	66.71%		73.69%		82.97%	
R^2 (Developing)	76.27%		76.21%		64.58%	
R^2 (All)	71.38%		72.91%		77.05%	
Weighted R^2	73.60%		75.51%		69.72%	
Panel D: Variable Importance Scores						
	R^2	*Corr*	R^2	*Corr*	R^2	*Corr*
Lag population (urban)	17.32	0.0547	0.17	0.0457	4.19	0.2157
Population (urban)	12.54	0.0480	0.32	0.0955	15.89	0.2595
Lag CO2 non-org (heavy industry)	0.00	0.0000	15.88	0.1795	10.74	0.2548
Population (other)	14.82	0.0407	0.02	0.0038	11.36	0.2533
CO2 non-org (heavy industry)	0.00	0.0000	14.19	0.1783	12.03	0.2573
NTL (cropland)	0.00	0.0000	6.71	0.1562	4.75	0.2241
Population (cropland)	0.00	0.0000	0.02	0.0056	2.52	0.1896
Lag NTL (cropland)	0.00	0.0000	2.07	0.1169	0.79	0.1222
Lag population (other)	0.00	0.0000	0.06	0.0115	1.58	0.1669
Population	0.25	-0.1838	0.75	-0.0769	1.21	0.0641
Lag population (cropland)	0.00	0.0000	0.31	0.0548	0.91	0.1310
Lag CO2 bio (manuf. combust.)	0.00	0.0000	0.08	0.0193	0.62	0.1257
CO2 bio (manuf. combust.)	0.00	0.0002	0.15	0.0345	0.57	0.1202
CO2 bio (heavy industry)	0.00	0.0000	0.02	0.0041	0.14	0.0362
Lag CO2 bio (transport)	0.00	0.0000	0.00	0.0003	0.13	0.0368
Lag CO2 bio (heavy industry)	0.00	0.0000	0.03	0.0049	0.10	0.0275
CO2 bio (transport)	0.00	0.0000	0.00	0.0007	0.10	0.0300
Forest	0.00	0.0000	-0.00	0.0130	0.08	0.1074
NTL (urban)	0.06	0.0121	0.07	0.0155	0.07	0.0188
NTL other (snow-free period)	0.00	0.0008	0.07	0.0176	0.04	0.0122

Notes: Panels A, B, and C report cross-validated MSE and R^2 metrics. Panel D reports variable importance scores: drop in within-country R^2 and *Corr* when each variable is replaced with its global mean. Variables shown are the top 20 by their maximum R^2 importance score across the three degree levels.

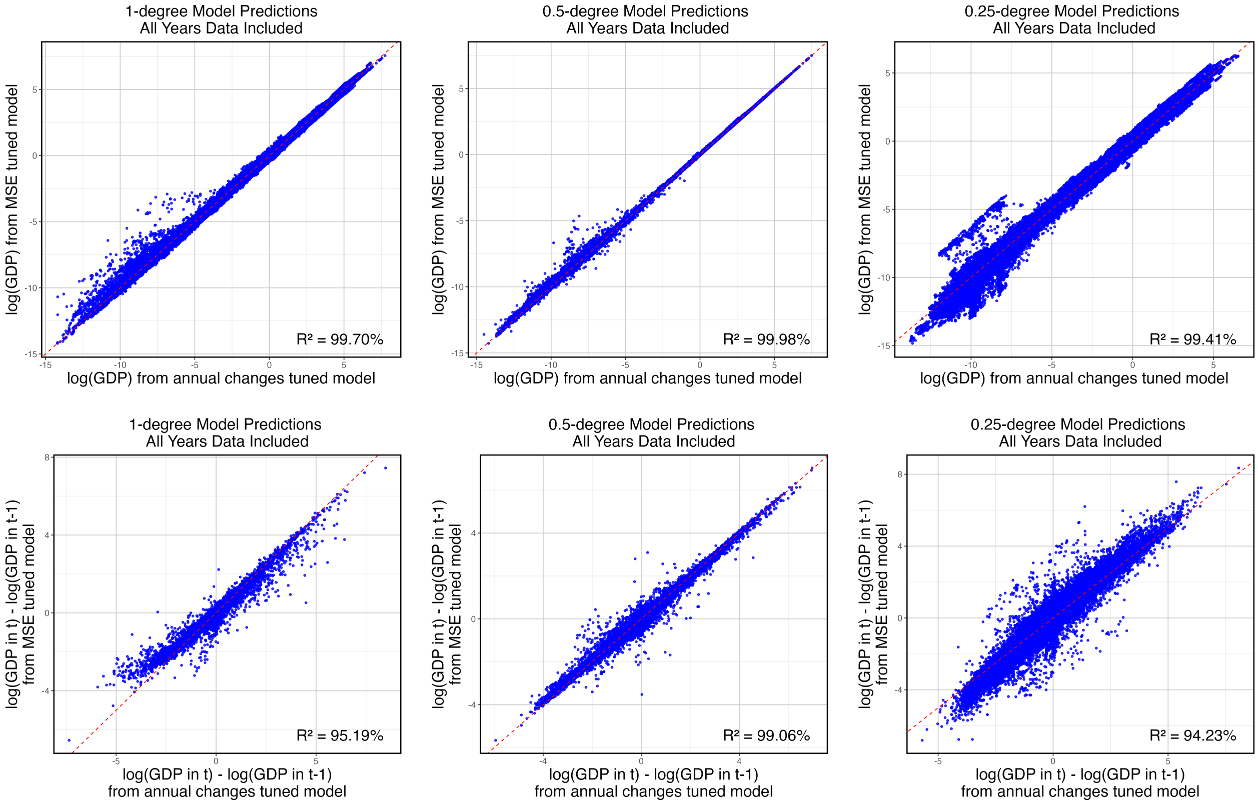

Figure 8: Comparison of Model Predictions: MSE-Tuned vs. Annual Change-Tuned Models

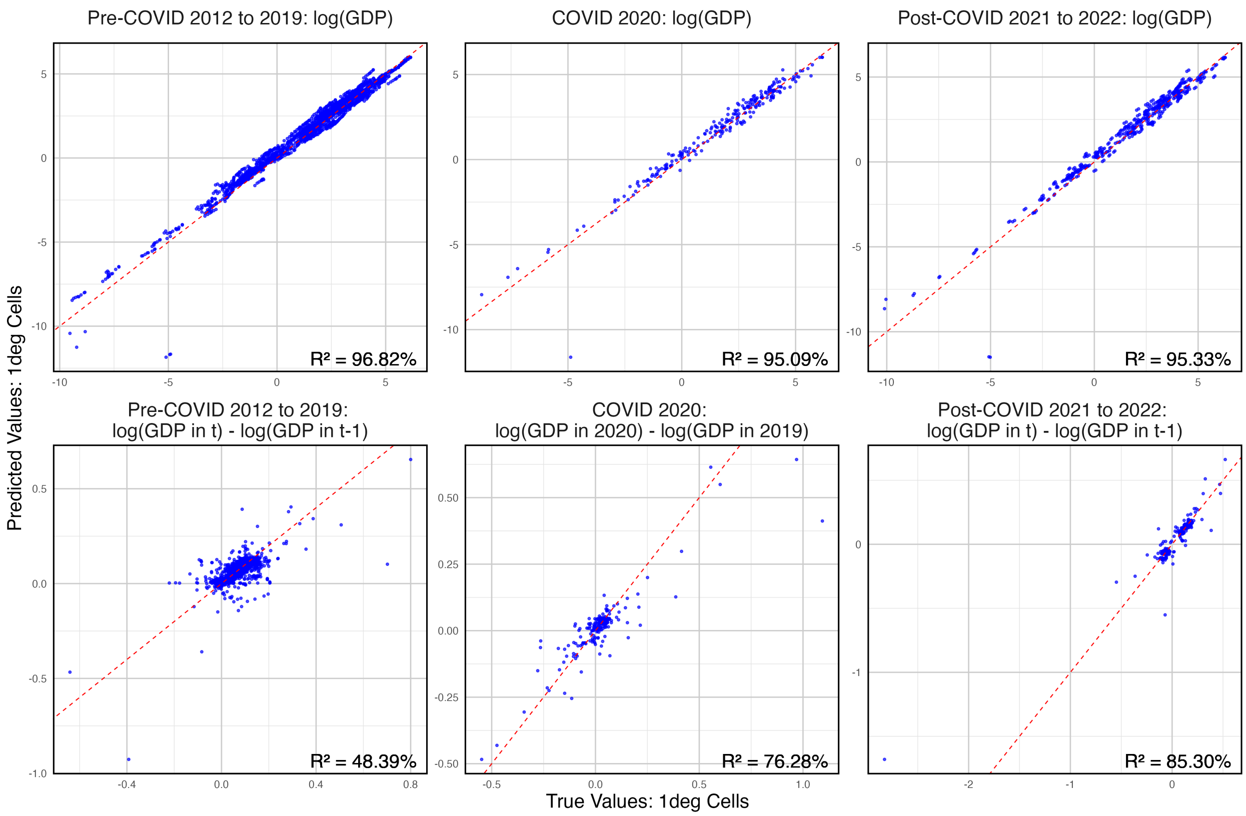

Figure 9: Model Predictions Against Actual GDP Values for China's Seven Leading Provinces, Using the Model Trained on years 2012 to 2022 Data and Tuned for Minimizing MSE

Note: This plot is comparable with the same plot in the paper. The only difference is the way hyperparameters are tuned.

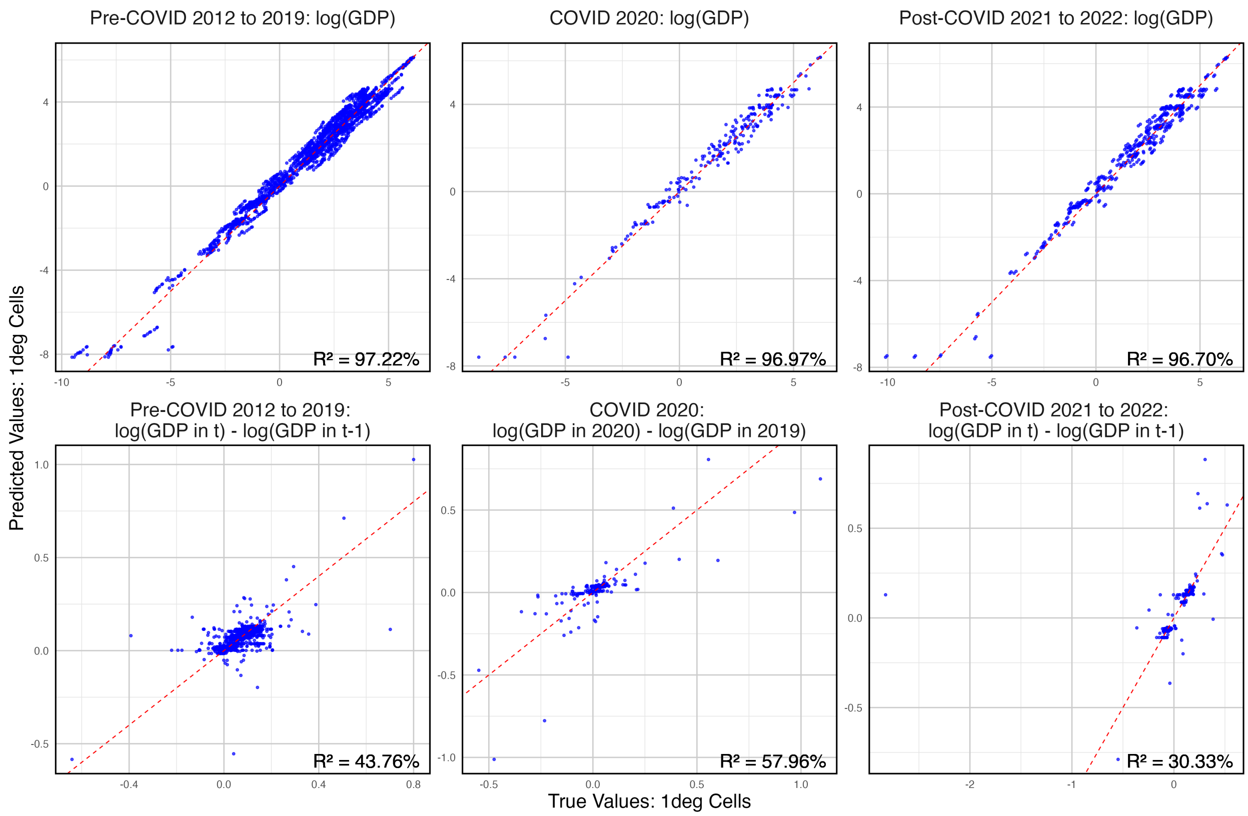

Figure 10: Model Predictions Against Actual GDP Values for China's Seven Leading Provinces, Using the Model Trained on years 2012 to 2019 Data and Tuned for Minimizing MSE

Note: This plot is comparable with Figure 4. The only difference is the way hyperparameters are tuned.

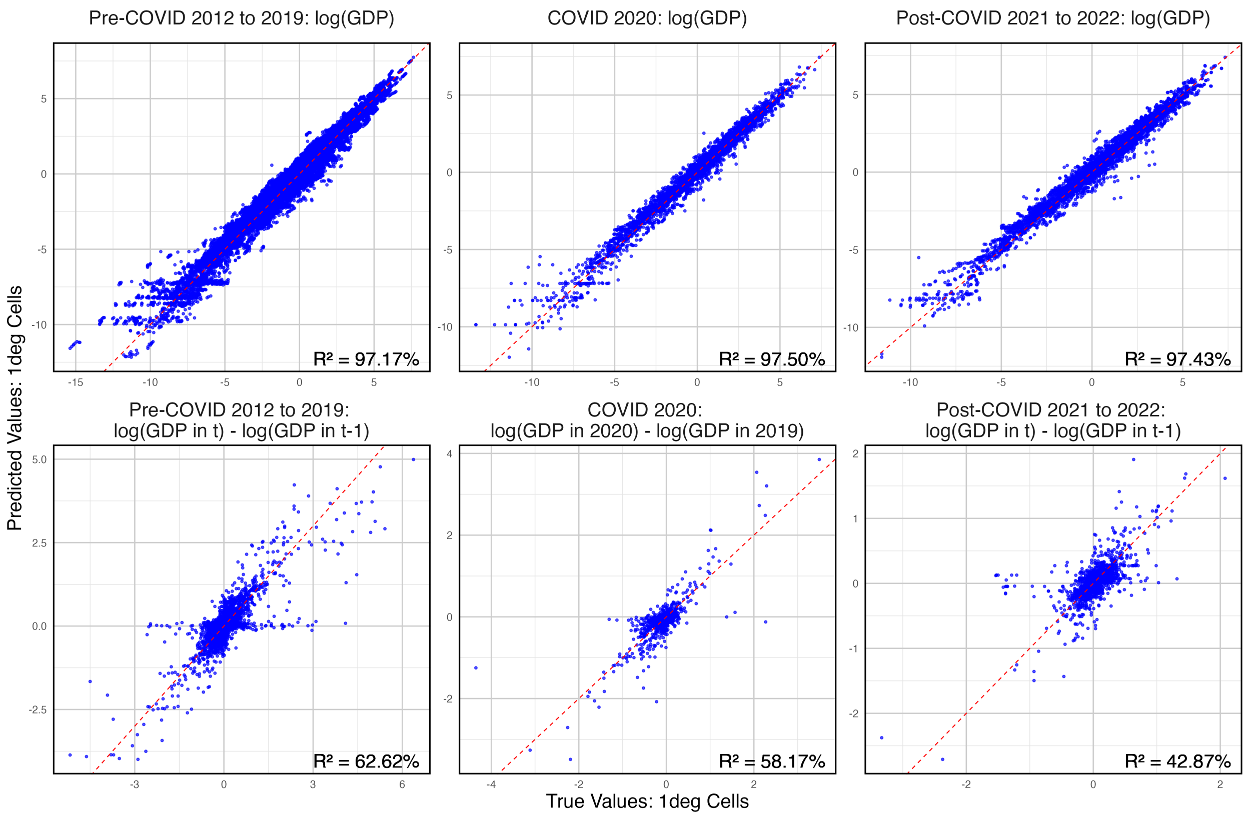

Figure 11: Model Predictions Against Actual GDP Values for All Training Countries, Using a Model Trained on years 2012 to 2019 Data and Tuned for Minimizing MSE

Note: Data for 2012 to 2019 are within training sample, so use out-of-bag predictions. Data for years 2020 to 2022 are predictions from the model. This plot is comparable with Figure 5. The only difference is the way hyperparameters are tuned.

7.2 Compare Benchmark Models with Models Trained Without Weights

The models used in the paper are trained with weights to address the imbalance between cells from developed and developing countries in the training sample. The weights are the rescaling factors of cell shares from developed and developing countries in the training sample, adjusted to match their real-world proportions. Cells can only be assigned one of the two weights: one assigned to cells from developed countries and the other to cells from developing countries, depending on their classification. These weights influence the probability of each cell being selected in the bootstrap sample used to build each decision tree. By assigning higher probabilities to underrepresented groups, the model ensures that both developed and developing regions are adequately represented during training.

Here we compare the predictions from models trained with and without weights. Table 27 and Figure 12 illustrate that the differences between the two approaches are minimal and the weights have a limited impact on the overall predictive performance of the methodology.

We also performed the same tests on the model trained without weights for comparison with the benchmark model (trained with weights), as shown in Figures 13, 14, and 15. The results continue to show the limited impact of weights on predictions.

Table 27: Cross-Validated Performance Metrics Across Spatial Resolutions

	1-degree Model		0.5-degree Model		0.25-degree Model	
Panel A: R^2 of Log GDP Level						
R^2 (Developed)	98.32%		97.45%		97.65%	
R^2 (Developing)	95.95%		95.84%		94.63%	
R^2 (All)	97.95%		97.16%		97.06%	
Weighted R^2	96.61%		96.29%		95.47%	
Panel B: R^2 of $\log(GDP_t) - \log(GDP_{t-1})$						
R^2 (Developed)	73.43%		74.50%		84.07%	
R^2 (Developing)	79.35%		76.58%		80.62%	
R^2 (All)	76.19%		73.52%		81.84%	
Weighted R^2	77.69%		76.00%		81.58%	
Panel C: Variable Importance Scores						
	R^2	*Corr*	R^2	*Corr*	R^2	*Corr*
Population (urban)	21.76	0.1924	40.94	0.1999	24.66	0.1743
Lag population (urban)	31.78	0.2002	40.48	0.1968	11.29	0.1629
Population (other)	20.43	0.1804	19.50	0.1803	16.75	0.1660
Lag CO2 non-org (heavy industry)	0.05	0.0100	16.13	0.1809	10.29	0.1644
CO2 non-org (heavy industry)	0.01	0.0026	14.79	0.1795	12.11	0.1662
Population (cropland)	1.27	0.1020	8.19	0.1602	7.35	0.1529
Lag population (cropland)	0.07	0.0164	7.16	0.1570	3.03	0.1316
Lag population (other)	4.76	0.1441	6.83	0.1553	5.08	0.1471
NTL (cropland)	3.42	0.1343	6.02	0.1533	4.86	0.1487
Lag NTL (urban)	0.18	0.0525	4.39	0.1439	0.01	0.0023
NTL (urban)	0.37	0.1002	3.41	0.1355	0.02	0.0054
lag NTL (cropland)	2.59	0.1256	2.90	0.1289	1.47	0.1080
Lag CO2 bio (manuf. combust.)	0.03	0.0056	2.04	0.1244	0.59	0.0879
NTL other (snow-free period)	0.02	0.0099	1.97	0.1167	0.15	0.0294
Population	0.75	0.1118	0.69	-0.1069	1.00	-0.1081
CO2 bio (manuf. combust.)	0.06	0.0196	0.96	0.0988	0.18	0.0355
CO2 non-org (manuf. combust.)	0.05	0.0174	0.32	0.0458	0.05	0.0096
Lag CO2 non-org (manuf. combust.)	0.06	0.0207	0.20	0.0354	0.05	0.0090
Lag cropland	0.14	0.0324	0.04	0.0076	0.04	0.0088
Cropland	0.09	0.0194	0.04	0.0090	0.01	0.0030

Notes: Panels A and B report cross-validated R^2 metrics. Panel C reports variable importance scores: drop in within-country R^2 and *Corr* when each variable is replaced with its global mean. Variables shown are the top 20 by their maximum R^2 importance score across the three degree levels.

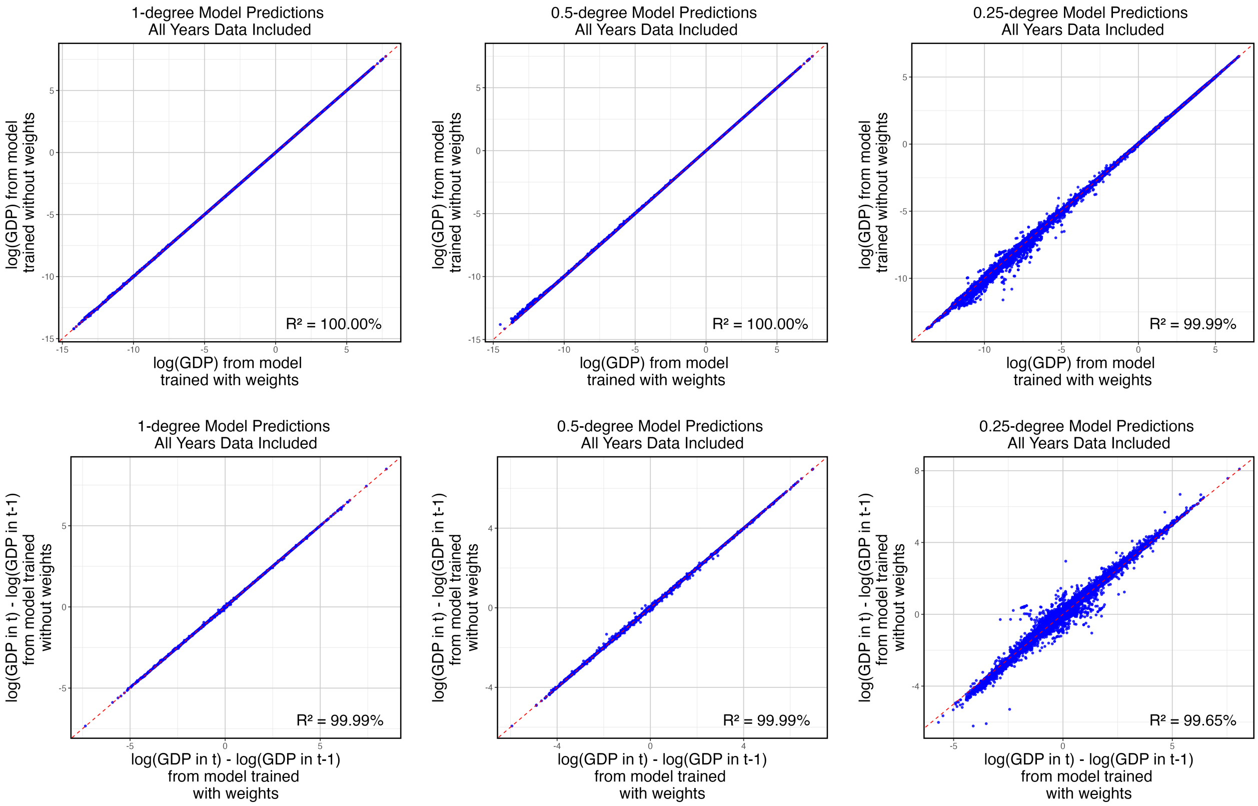

Figure 12: Comparison of Model Predictions: Models Trained With and Without Weights

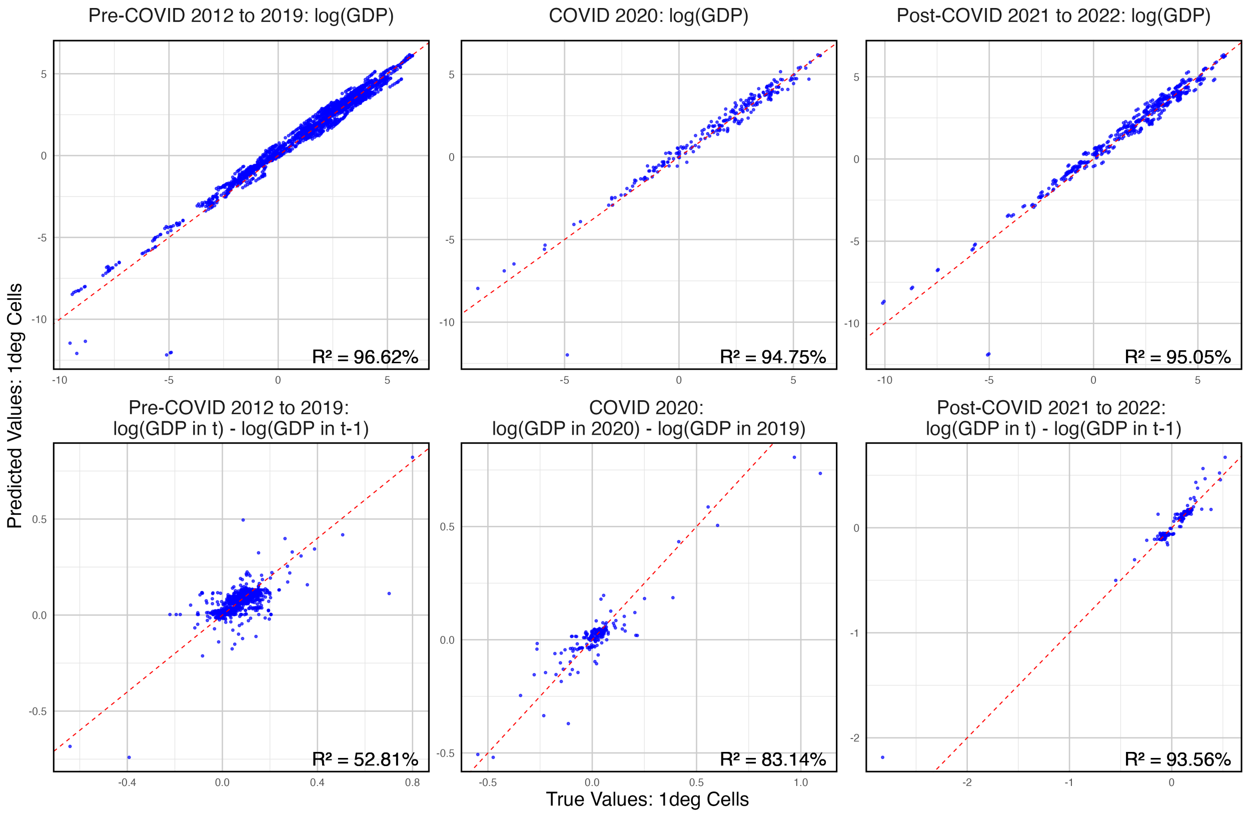

Figure 13: Model Predictions Against Actual GDP Values for China's Seven Leading Provinces, Using the Model Trained on years 2012 to 2022 Data and Without Weights

Note: This plot is comparable to the corresponding plot in the paper, with the only difference being whether the models were trained with or without weights.

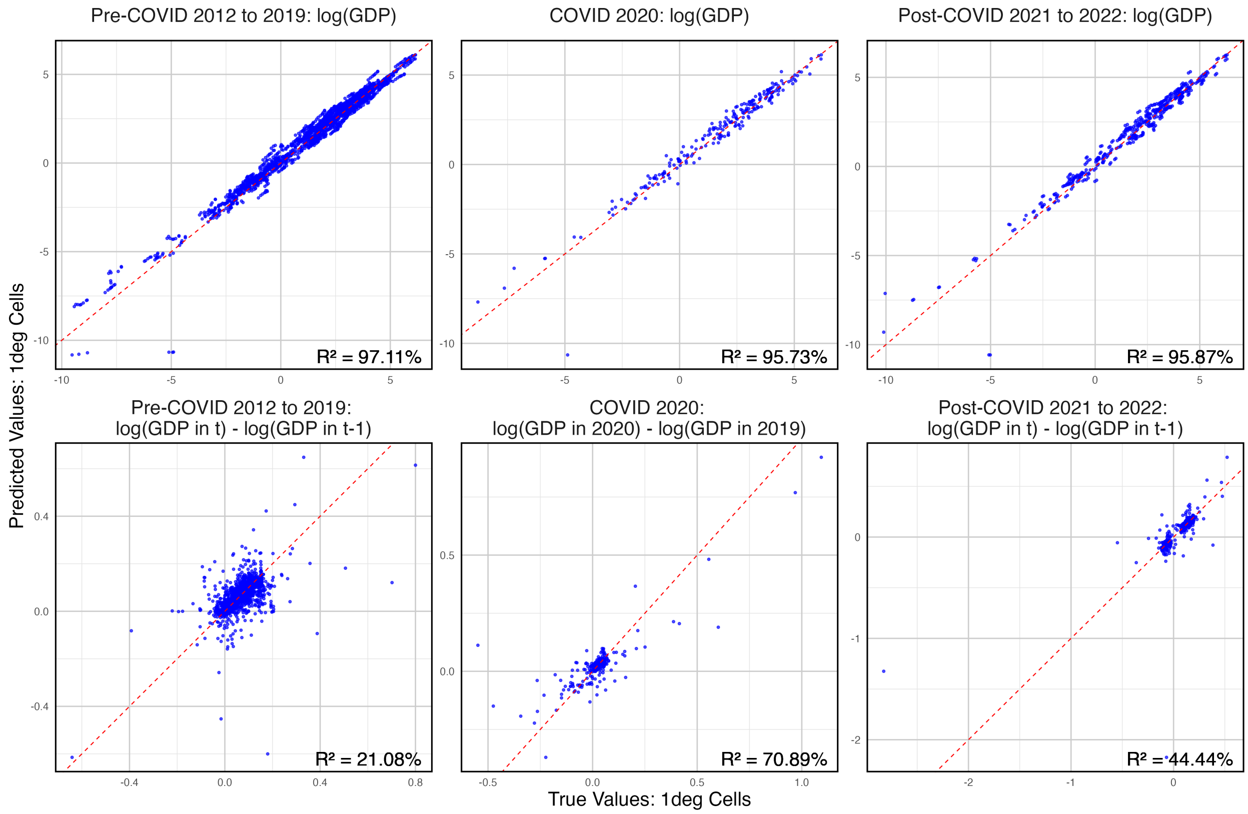

Figure 14: Model Predictions Against Actual GDP Values for China's Seven Leading Provinces, Using the Model Trained on years 2012 to 2019 Data and Without Weights

Note: This plot is comparable with Figure 4. The only difference lies in whether the models were trained with or without weights.

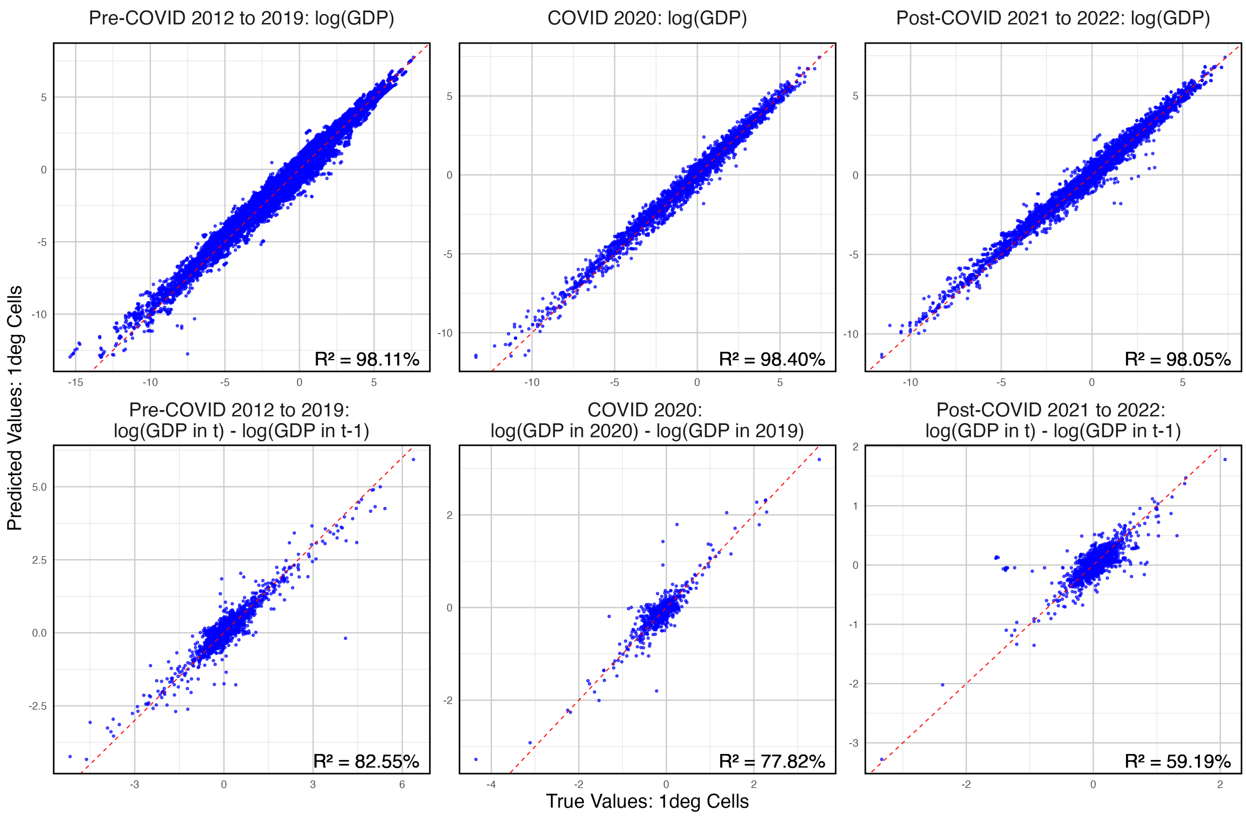

Figure 15: Model Predictions Against Actual GDP Values for All Training Countries, Using the Model Trained on years 2012 to 2019 Data and Without Weights

Note: Data for 2012 to 2019 are within training sample, so use out-of-bag predictions. Data for years 2020 to 2022 are predictions from the model. This plot is comparable with Figure 5. The only difference lies in whether the models were trained with or without weights.

7.3 Compare Benchmark Models with Models Trained Without Developing Countries Data

This section presents the results of models trained without data from developing countries. Note that weights are also excluded when developing country data is not included. Table 28 summarizes the cross-validated performance across all three resolutions. No large differences are observed. The models were then used to predict data for cells in developing countries, and the results were compared with the truth. Tables 29, 30, and 31 demonstrate that the models perform well in predicting developing country data, even without having been trained on it. When applied to predict all world cells, Figure 16 reveals that level predictions are highly consistent, while slight differences are observed in annual changes.

We also performed the same tests on the model trained without developing countries for comparison with the benchmark model (trained with developing countries), as shown in Figures 17, 18, 19 and 20. The results show that the models continue to behave well even when developing countries are excluded.

Table 28: In-Sample Performance Metrics Across Spatial Resolutions

	1-degree Model		0.5-degree Model		0.25-degree Model	
Panel A: R^2 of Log GDP Level						
R^2 (All)	98.79%		98.95%		98.89%	
Panel B: R^2 of $\log(GDP_t) - \log(GDP_{t-1})$						
R^2 (All)	80.07%		86.31%		90.07%	
Panel C: Variable Importance Scores						
	R^2	Corr	R^2	Corr	R^2	Corr
Lag population (urban)	17.90	0.7095	66.74	0.4207	10.16	0.1983
Lag CO2 non-org (transport)	45.80	0.7750	10.14	0.3529	15.01	0.1996
Population (other)	44.79	0.7662	42.30	0.3965	36.18	0.2072
Population (urban)	13.49	0.6647	40.79	0.4103	7.34	0.1951
Lag population (other)	27.91	0.7179	12.82	0.3597	8.76	0.1921
CO2 non-org (transport)	27.83	0.7221	7.93	0.3433	14.39	0.1992
Population (cropland)	19.17	0.6730	16.06	0.3678	14.15	0.1985
Lag NTL (urban)	16.92	0.6713	0.01	0.0043	0.00	0.0014
Lag population (cropland)	10.94	0.6021	16.43	0.3687	5.10	0.1840
NTL from Urban	14.35	0.6562	0.01	0.0050	0.01	0.0040
CO2 bio (transport)	13.31	0.6346	0.01	0.0107	0.00	0.0008
Lag CO2 bio (transport)	11.70	0.6184	0.03	0.0238	0.00	0.0011
Lag NTL (cropland)	7.75	0.5626	0.01	0.0051	0.00	0.0006
Lag CO2 nonorg (manuf. combust.)	6.25	0.5511	1.67	0.2629	0.19	0.0699
NTL (cropland)	4.54	0.4867	0.00	0.0039	0.87	0.1342
CO2 non-org (manuf. combust.)	4.45	0.5035	1.39	0.2504	0.41	0.1040
Lag cropland	2.51	0.3985	0.02	0.0148	0.00	0.0011
Cropland	2.10	0.3703	0.02	0.0156	0.00	0.0009
Population	1.68	0.6198	0.90	0.1203	0.99	-0.0703
Lag CO2 bio (manuf. combust.)	1.18	0.2830	0.01	0.0108	0.00	0.0001

Notes: Panels A and B report in-sample R^2 metrics. Panel C reports variable importance scores: drop in within-country R^2 and Corr when each variable is replaced with its global mean. Variables shown are the top 20 by their maximum R^2 importance score across the three degree levels.

Table 29: 1-degree Model Performance Metrics for Developing Group

ISO	R^2 for log(GDP)	GDP Loss	R^2 for log(GDP in t) - log(GDP in t-1)
ALB	98.29%	8.70%	95.78%
BIH	97.50%	10.66%	80.66%
BLR	99.11%	6.12%	95.51%
CHL	95.46%	10.36%	92.58%
COL	95.27%	10.35%	96.10%
ECU	93.73%	6.22%	87.66%
IDN	95.59%	11.50%	91.00%
KEN	88.63%	13.07%	89.48%
KGZ	95.82%	24.20%	94.87%
LKA	90.75%	22.79%	80.18%
MOZ	95.58%	14.32%	92.27%
PER	98.03%	13.11%	97.45%
PHL	98.38%	6.59%	94.77%
SRB	95.43%	10.40%	93.64%
THA	86.72%	23.64%	92.73%
UZB	98.37%	9.56%	98.13%
VNM	91.52%	14.76%	90.44%

Table 30: 0.5-degree Model Performance Metrics for Developing Group

ISO	R^2 for log(GDP)	GDP Loss	R^2 for log(GDP in t) - log(GDP in t-1)
ALB	94.08%	23.19%	97.19%
BIH	95.27%	8.80%	89.03%
BLR	98.44%	10.00%	95.44%
CHL	96.89%	11.76%	94.87%
COL	95.30%	13.01%	97.56%
ECU	87.19%	12.27%	81.67%
IDN	95.06%	11.16%	91.39%
KEN	88.58%	20.77%	89.95%
KGZ	95.87%	29.92%	97.32%
LKA	98.45%	9.18%	94.34%
MOZ	92.71%	20.08%	89.79%
PER	97.85%	13.55%	97.13%
PHL	98.01%	9.87%	97.58%
SRB	78.77%	12.12%	76.64%
THA	89.67%	27.12%	93.71%
UZB	98.06%	8.91%	97.51%
VNM	92.82%	15.21%	93.54%

Table 31: 0.25-degree Model Performance Metrics for Developing Group

ISO	R^2 for log(GDP)	GDP Loss	R^2 for log(GDP in t) - log(GDP in t-1)
ALB	93.64%	20.36%	96.43%
BIH	92.30%	6.61%	90.84%
BLR	96.51%	14.67%	95.44%
CHL	96.40%	14.97%	96.59%
COL	94.44%	16.95%	97.67%
ECU	89.18%	16.50%	86.27%
IDN	94.84%	19.59%	93.42%
KEN	89.29%	17.23%	92.55%
KGZ	95.08%	32.36%	97.96%
LKA	98.38%	9.93%	97.21%
MOZ	89.10%	27.36%	90.28%
PER	97.19%	13.64%	97.64%
PHL	97.04%	15.07%	98.97%
SRB	81.14%	17.14%	80.88%
THA	86.32%	29.02%	91.18%
UZB	97.18%	16.57%	97.31%
VNM	91.23%	18.24%	92.59%

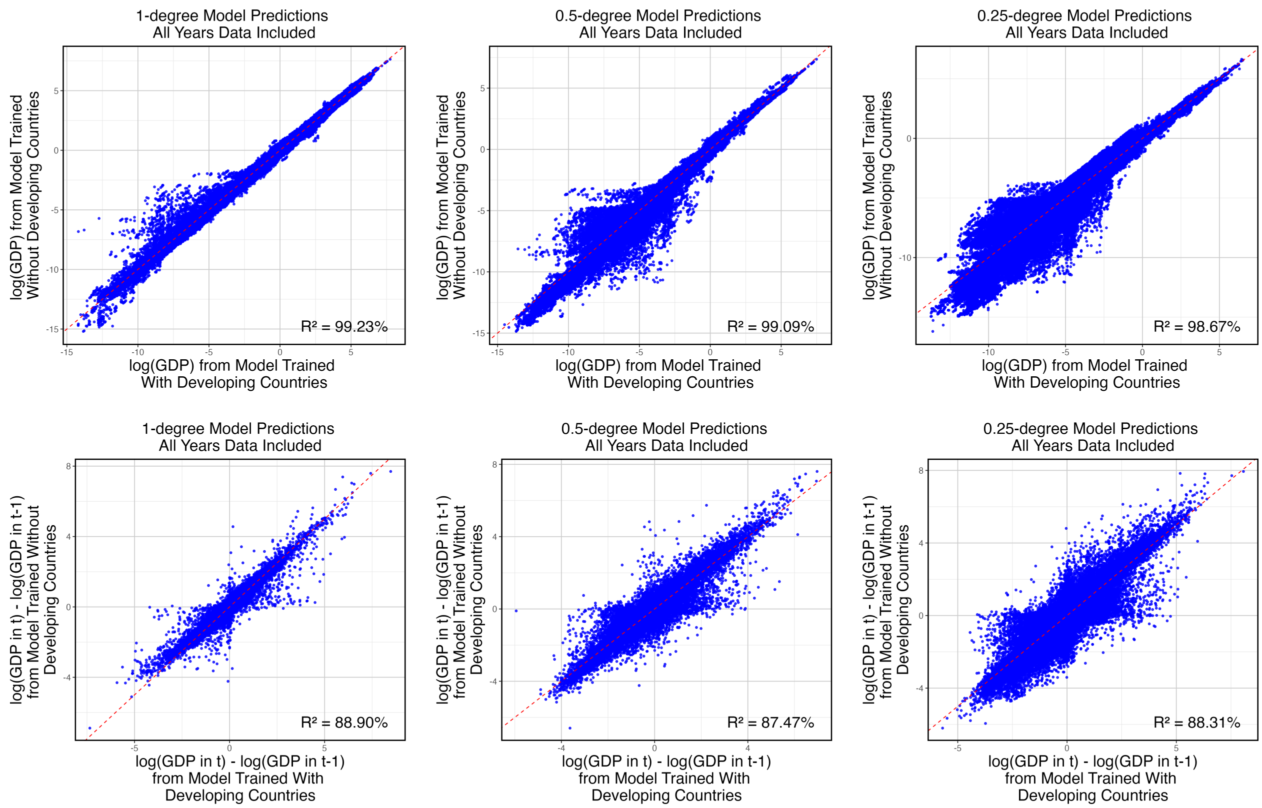

Figure 16: Comparison of Model Predictions: Models Trained With and Without Developing Countries Data

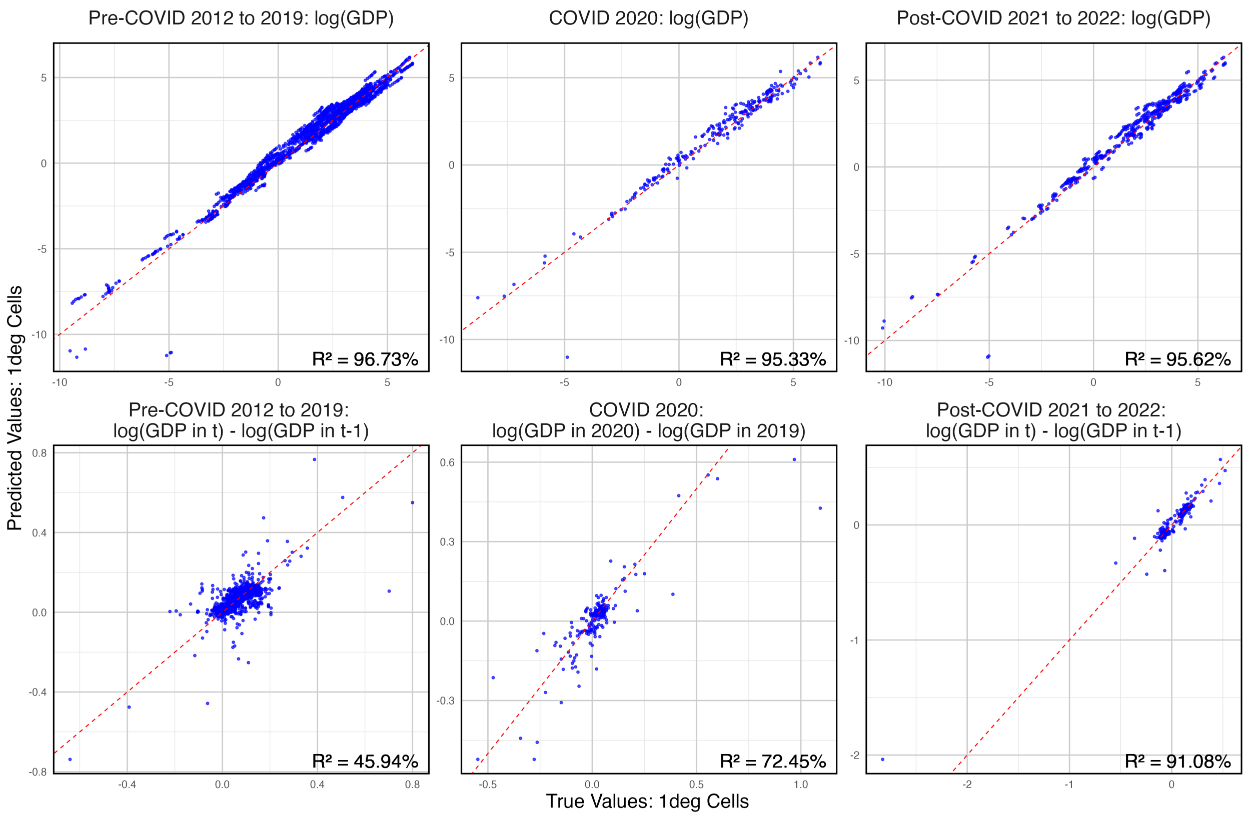

Figure 17: Model Predictions Against Actual GDP Values for China's Seven Leading Provinces, Using a Model Trained on years 2012 to 2022 Data and Without developing countries data

Note: This plot is comparable to the corresponding plot in the paper, with the only difference being whether the models were trained with or without developing countries data.

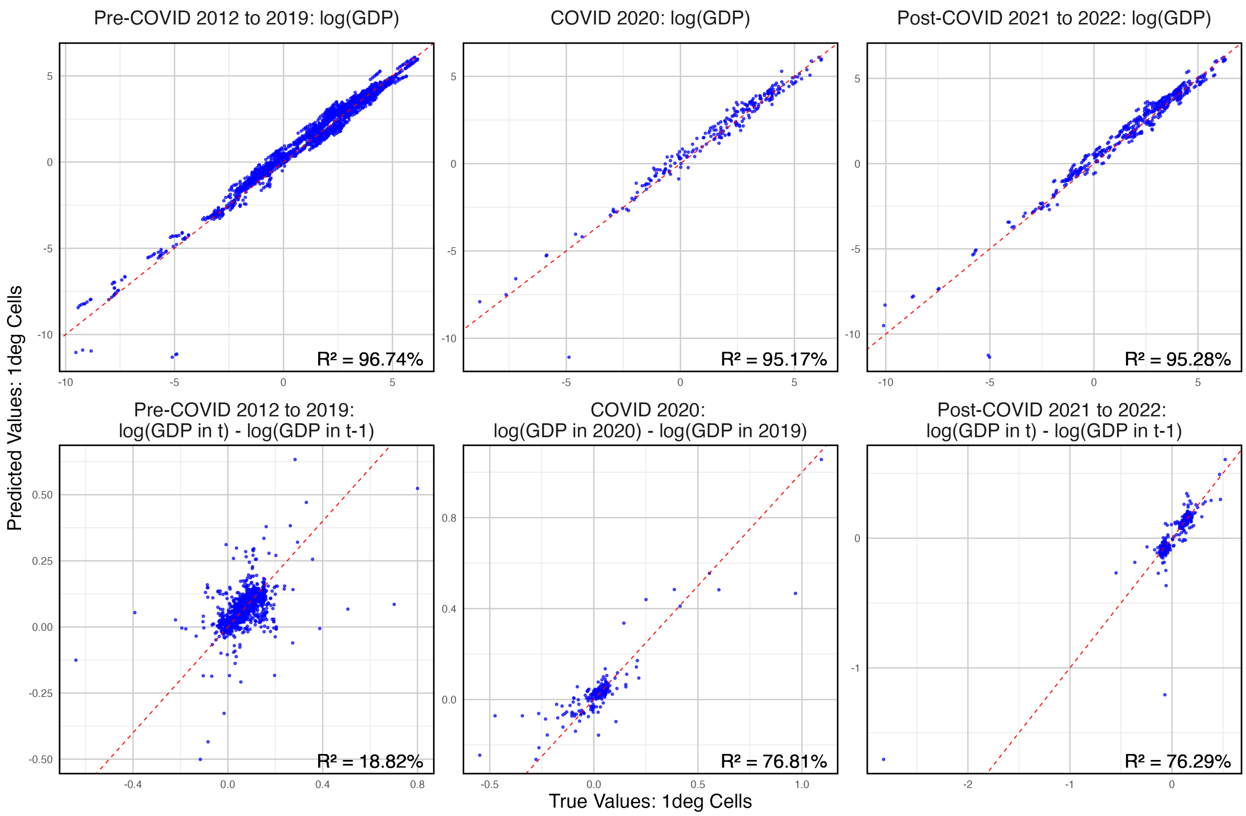

Figure 18: Model Predictions Against Actual GDP Values for China's Seven Leading Provinces, Using a Model Trained on years 2012 to 2019 Data and Without developing countries data

Note: This plot is comparable with Figure 4. The only difference lies in whether the models were trained with or without developing countries data.

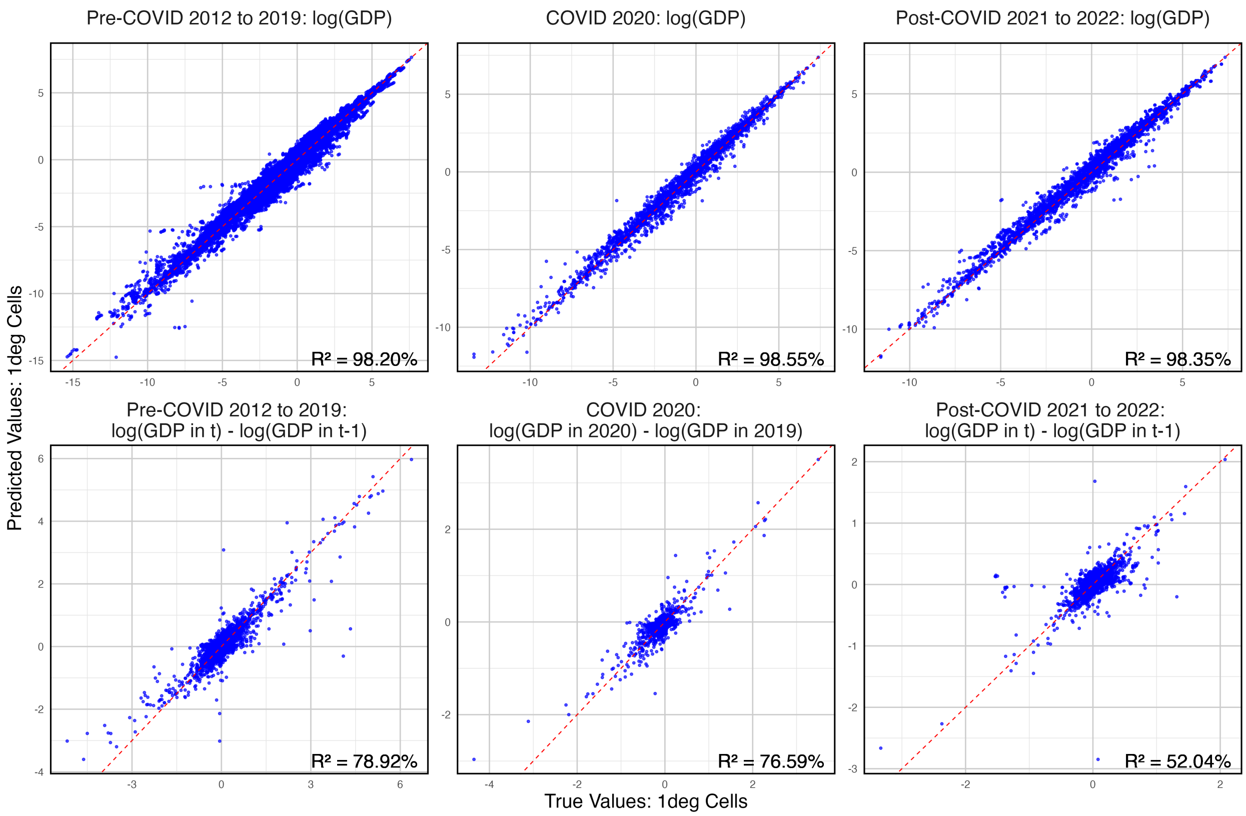

Figure 19: Model Predictions Against Actual GDP Values for All Training Countries, Using a Model Trained on years 2012 to 2019 Data and Without developing countries data

Note: Data for 2012 to 2019 are within training sample, so use out-of-bag predictions. Data for years 2020 to 2022 are predictions from the model. This plot is comparable with Figure 5. The only difference lies in whether the models were trained with or without developing countries data.

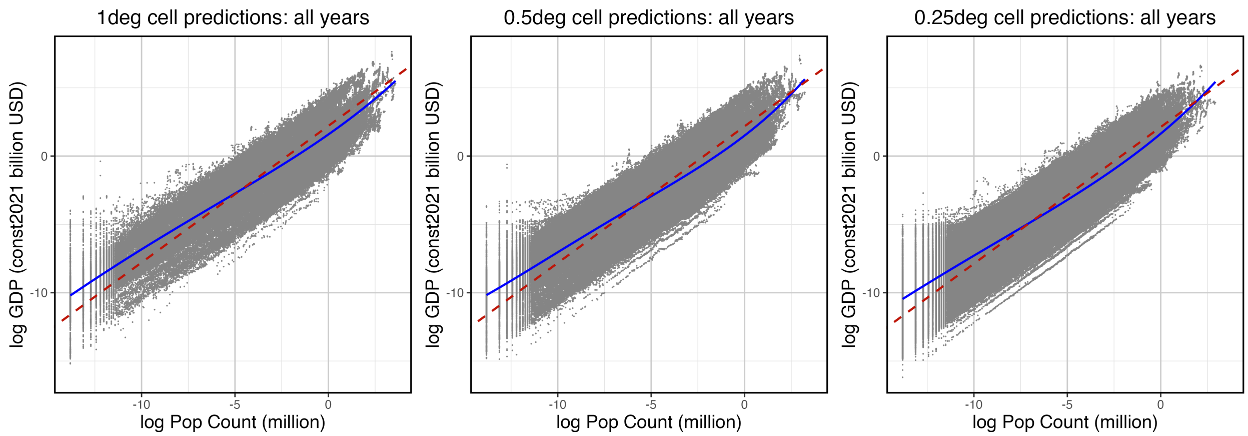

Figure 20: Cell Population Against Cell GDP, Using Models Trained Without developing countries data

Note: For the 1-degree resolution, the blue solid line is $y = 1.59 + 0.96x + 0.03x^2 + 0.002x^3 + 0 * x^4$ and the red dashed line is $y = x + 2.21$. For the 0.5-degree resolution, the blue solid line is $y = 1.5 + 1.05x + 0.05x^2 + 0.005x^3 + 0.0001x^4$ and the red dashed line is $y = x + 2.18$. For the 0.25-degree resolution, the blue solid line is $y = 1.64 + 1.14x + 0.05x^2 + 0.003x^3 + 0 * x^4$ and the red dashed line is $y = x + 2.13$.

8 Consistency of Predictions Across Resolutions

In this section, we demonstrate that the predictions from models at different spatial resolutions are consistent. Figure 21 compares the predictions generated by our benchmark models in the paper: trained using data from all available countries (excluding China) from 2012 to 2022 with optimized hyperparameters for each resolution. To evaluate consistency, we aggregate the predictions from finer resolutions to match the coarser resolution and compare the results. Figure 21 shows that the predictions remain relatively consistent and confirms the robustness of our models across different spatial scales.

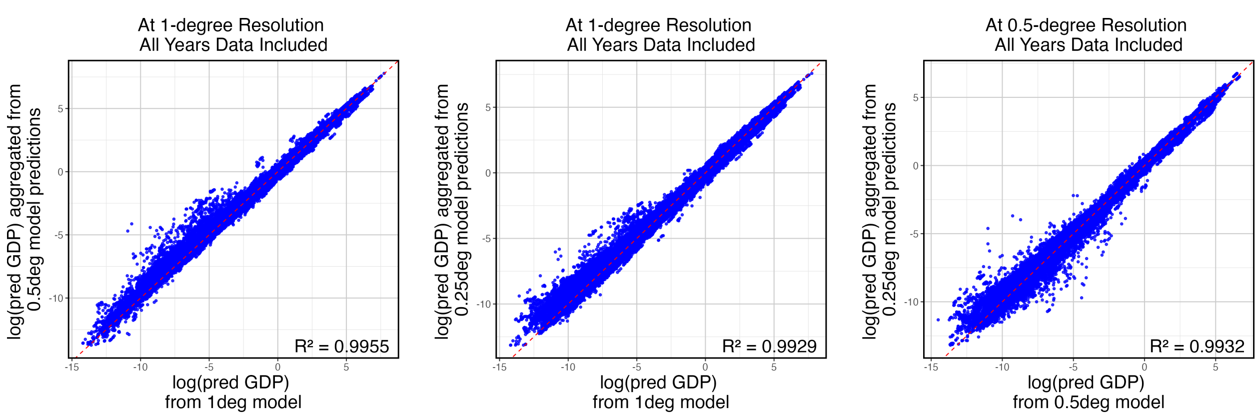

Figure 21: Comparison of Model Predictions Across Resolutions: Aggregated vs. Direct Predictions

9 Consistency of Predictions Across Dataset Versions

Our datasets are updated annually, and this process involves incorporating new training samples. As a result, predictions may change, not only for the new years but also for the previously published years. This highlights the importance of testing whether our models produce consistent predictions across different dataset versions. To conduct this test, we trained two versions of the model: 1) uses data from all available countries (excluding China) for the years 2012 to 2019 2) uses data from the same countries but extends the coverage to include years 2012 to 2022. We compared the predictions for both the newly updated years (2020, 2021 and 2022) and selected previous years (2018 and 2019). Figures 22 and 23 illustrate that, despite minor discrepancies, the models produce highly similar results across versions for both new and old years.

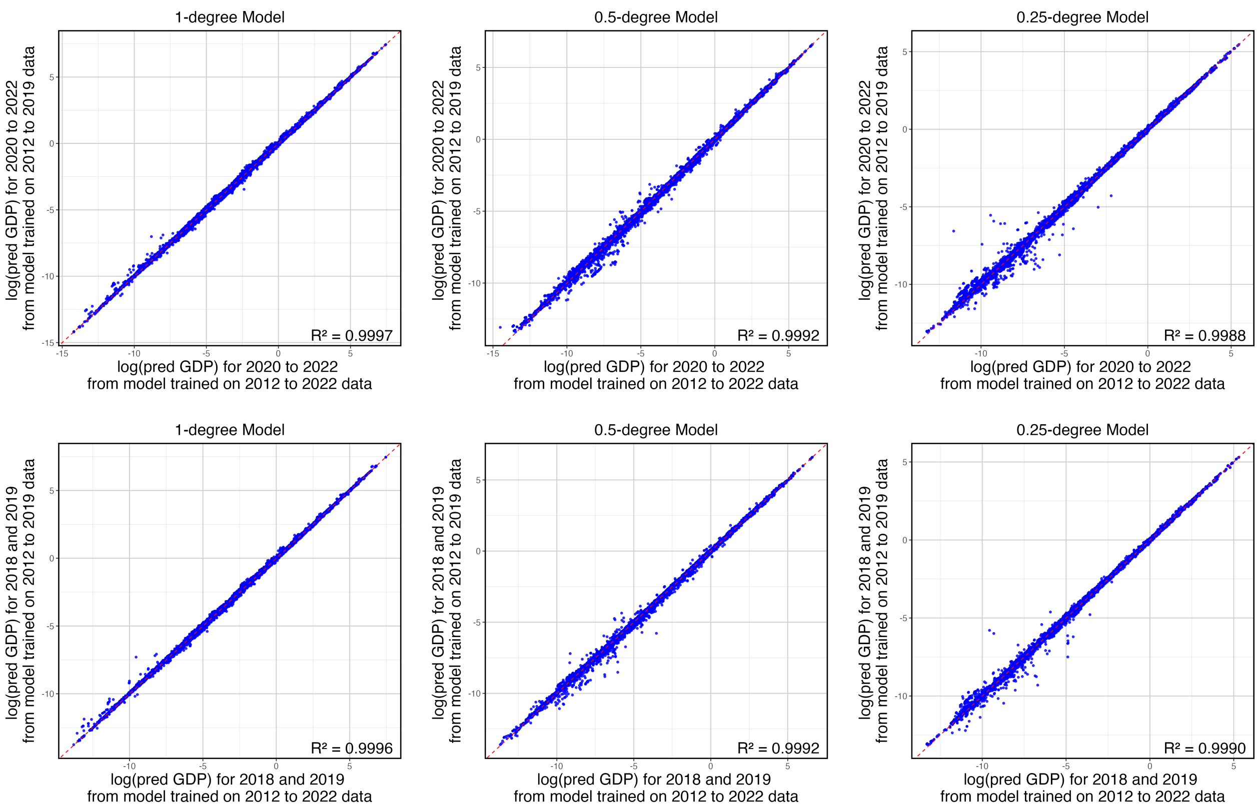

Figure 22: Comparison of Predicted Cell GDP Levels Across Model Versions

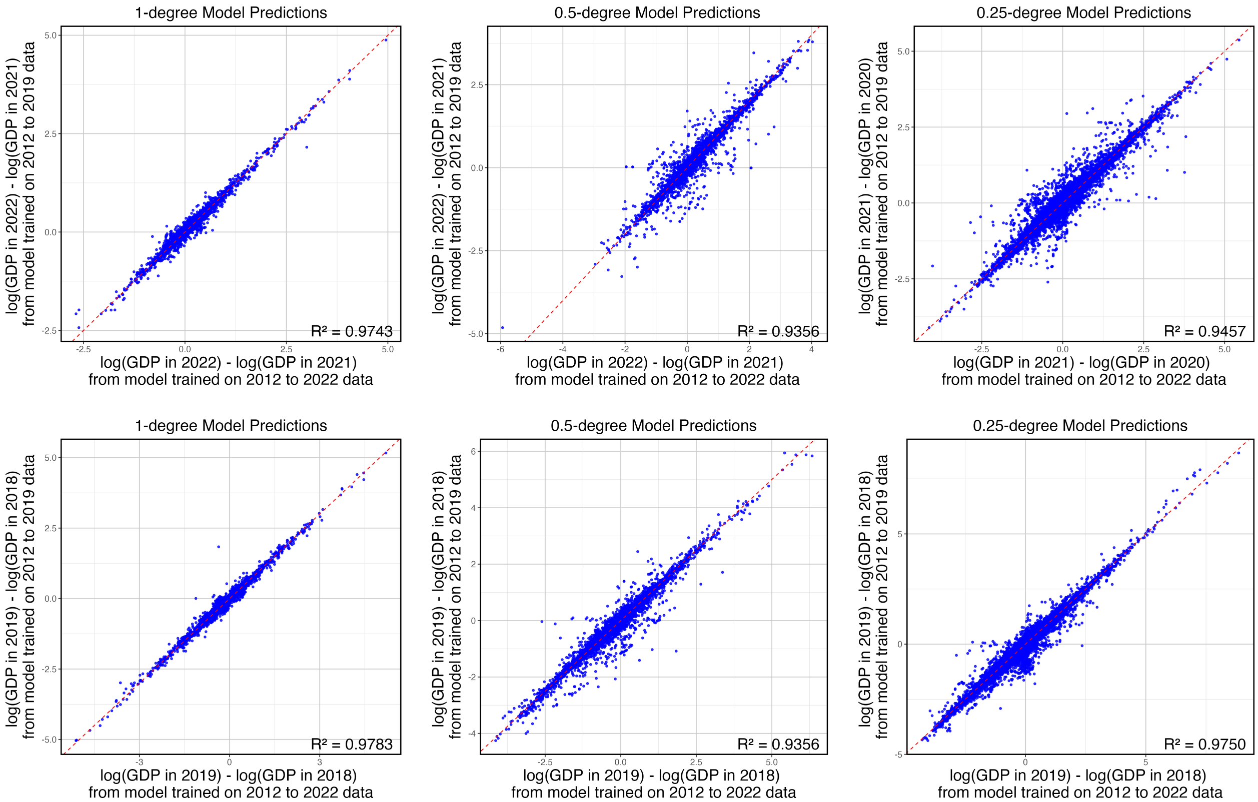

Figure 23: Comparison of Predicted Cell GDP Annual Changes Across Model Versions

References

Authority, Philippine Statistics (2023). *2000-2023 Gross Regional Domestic Product - GRDP by Region*. GRDP Data Series (Database). Accessed on 15th June 2024. URL: https://psa.gov.ph/statistics/grdp/data-series.

Bank, World (2024a). *GDP in Constant 2021 PPP Adjusted International$ (Dataset)*. The World Bank Group. Accessed on 15th June 2024. URL: https://data.worldbank.org/indicator/NY.GDP.MKTP.PP.KD.

— (2024b). *GDP in Current PPP Adjusted International$ (Dataset)*. The World Bank Group. Accessed on 15th June 2024. URL: https://data.worldbank.org/indicator/NY.GDP.MKTP.PP.CD.

— (2024c). *GDP in Current US$ (Dataset)*. The World Bank Group. Accessed on 15th June 2024. URL: https://data.worldbank.org/indicator/NY.GDP.MKTP.CD.

— (2024d). *Population (Dataset)*. The World Bank Group. Accessed on 15th June 2024. URL: https://data.worldbank.org/indicator/SP.POP.TOTL.

Bright, Edward, Patricia Coleman, Andrew King, Amy Rose, Michael Urban, Jarrod McKee, Kelsey Sims, and Andrew Reith (n.d.). *LandScan Global*. Oak Ridge National Laboratory. Accessed: 2023-12-20. URL: https://doi.org/10.48690/1524196,%20https://doi.org/10.48690/1524197,%20https://doi.org/10.48690/1524198,%20https://doi.org/10.48690/1524199,%20https://doi.org/10.48690/1524200,%20https://doi.org/10.48690/1524201,%20https://doi.org/10.48690/1524202,%20https://doi.org/10.48690/1524203,%20https://doi.org/10.48690/1524204,%20https://doi.org/10.48690/1524205,%20https://doi.org/10.48690/1524206,%20https://doi.org/10.48690/1524207,%20https://doi.org/10.48690/1524215,%20https://doi.org/10.48690/1524208,%20https://doi.org/10.48690/1524209,%20https://doi.org/10.48690/1524210,%20https://doi.org/10.48690/1524211,%20https://doi.org/10.48690/1524212,%20https://doi.org/10.48690/1524213,%20https://doi.org/10.48690/1524214,%20https://doi.org/10.48690/1523378,%20https://doi.org/10.48690/1527702.

Bureau, U.S. Census (2024). *Annual Index Value and Annual Percent Change in Price Series: 1947 to 2022*. U.S. Census Bureau Database. Accessed on 15th June 2024. URL: https://www.census.gov/en.html.

Economic Analysis U.S. Department of Commerce, Bureau of (2023). *CAGDP2: GDP in Current Dollars by County and MSA*. Regional Economic Accounts - Previously Published Estimates: Gross Domestic Product by County and Metro Area. Accessed on 15th June 2024. URL: https://apps.bea.gov/histdatacore/Regional_Accounts_new.html.

Elvidge, Christopher D, Maxim Zhizhin, Kimberly Baugh, Feng-Chi Hsu, and Tilottama Ghosh (2016). "Methods for Global Survey of Natural Gas Flaring from Visible Infrared Imaging Radiometer Suite Data". In: *Energies* 9.1, p. 14. DOI: 10.3390/en9010014.

Elvidge, Christopher D, Maxim Zhizhin, Feng-Chi Hsu, and Kimberly E Baugh (2013). "VIIRS Nightfire: Satellite Pyrometry at Night". In: *Remote Sensing* 5, pp. 4423–4449. DOI: 10.3390/rs5094423.

Friedl, Mark and Damien Sulla-Menashe (2022). *MODIS/Terra+Aqua Land Cover Type Yearly L3 Global 500m SIN Grid V061*. Distributed by NASA EOSDIS Land Processes

Distributed Active Archive Center. Accessed 2023-12-20. URL: https://doi.org/10.5067/MODIS/MCD12Q1.061.

Fund, International Monetary (2024). *World Economic Outlook Database, April 2024*. IMF World Economic Outlook Database. Accessed on 15th June 2024. URL: https://www.imf.org/en/Publications/SPROLLS/world-economic-outlook-databases#sort=%40imfdate%20descending.

GaryBikini (2024). *China Administrative Division Shapefiles*. Version 24.02.06. DOI: 10.5281/zenodo.10624971. URL: https://doi.org/10.5281/zenodo.10624971.

Geografia e Estatística (IBGE), Instituto Brasileiro de (2024). *GDP under the point of view of Income (2010-2021)*. SCR - System of Regional Accounts (Database). Accessed on 15th June 2024. URL: https://www.ibge.gov.br/en/statistics/economic/national-accounts/16855-regional-accounts-of-brazil.html.

India, Reserve Bank of (2023). *Table 27: Gross State Domestic Product (Current Prices)*. Handbook of Statistics of Indian States. Accessed on 15th June 2024. URL: https://www.rbi.org.in/Scripts/AnnualPublications.aspx?head=Handbook%20of%20Statistics%20on%20Indian%20States.

Kyrgyz Republic, National Statistical Committee of the (2023). *1.01.00.09 Gross regional product (GRP) at current prices*. National accounts (Database). Accessed on 15th June 2024. URL: https://www.stat.kg/en/statistics/nacionalnye-scheta/.

Lehner, Bernhard and Petra Döll (2004). "Development and validation of a global database of lakes, reservoirs and wetlands". In: *Journal of Hydrology* 296.1-4, pp. 1–22.

National Statistics Agency for Strategic Planning, Bureau of and Reforms of the Republic of Kazakhstan (2023). *Gross regional product*. Statistics of national accounts (Database). Accessed on 15th June 2024. URL: https://stat.gov.kz/en/industries/economy/national-accounts/.

Nunn, Nathan and Diego Puga (2012). "Ruggedness: The Blessing of Bad Geography in Africa". In: *Review of Economics and Statistics* 94.1, pp. 20–36.

OECD (2024a). *Gross domestic product - Regions. OECD Regional Economy (Database)*. Accessed on 15th June 2024. URL: https://data-explorer.oecd.org/vis?fs%5B0%5D=Topic%2C1%7CRegional%2C%20rural%20and%20urban%20development%23GEO%23%7CRegions%23GEO_REG%23&fs%5B1%5D=Topic%2C2%7CRegional%2C%20rural%20and%20urban%20development%23GEO%23%7CRegions%23GEO_REG%23%7CRegional%20economy%23GEO_REG_ECO%23&pg=0&fc=Topic&snb=39&df%5Bds%5D=dsDisseminateFinalDMZ&df%5Bid%5D=DSD_REG_ECO%40DF_GDP&df%5Bag%5D=OECD.CFE.EDS&df%5Bvs%5D=2.0&lo=5&lom=LASTNPERIODS&dq=A..AUS%2BAU1%2BAU2%2BAU3%2BAU4%2BAU5%2BAU6%2BAU7%2BAU8%2BAUT%2BAT11%2BAT111%2BAT112%2BAT113%2BAT12%2BAT121%2BAT122%2BAT123%2BAT124%2BAT125%2BAT126%2BAT127%2BAT13%2BAT130%2BAT21%2BAT211%2BAT212%2BAT213%2BAT22%2BAT221%2BAT222%2BAT223%2BAT224%2BAT225%2BAT226%2BAT31%2BAT311%2BAT312%2BAT313%2BAT314%2BAT315%2BAT32%2BAT321%2BAT322%2BAT323%2BAT33%2BAT331%2BAT332%2BAT333%2BAT334%2BAT335%2BAT34%2BAT341%2BAT342..GDP..Q.USD_PPP_PS&ly%5Brw%5D=REF_AREA%2CTERRITORIAL_LEVEL&ly%5Bcl%5D=TIME_PERIOD&to%5BTIME_PERIOD%5D=false.

— (2024b). *Regional Economy. OECD iLibrary - OECD Regional Statistics (Database)*. Accessed on 15th June 2024. URL: https://www.oecd-ilibrary.org/urban-rural-and-regional-development/data/oecd-regional-statistics_region-data-en.

Román, Miguel O, Zhuosen Wang, Qingsong Sun, Virginia Kalb, Steven D Miller, Andrew Molthan, Lyle Schultz, Jordan Bell, Erika C Stokes, Bhaskar Pandey, and Karen C Seto (2018). "NASA's Black Marble nighttime lights product suite". In: *Remote Sensing of Environment* 210, pp. 113–143. DOI: 10.1016/j.rse.2018.03.017.

Runfola, Daniel, Dylan Anderson, Daniel Mahr, Will R. Turner, and Martin Wegmann (2020). "geoBoundaries: A global database of political administrative boundaries". In: *PLoS ONE* 15.4, e0231866. DOI: 10.1371/journal.pone.0231866. URL: https://doi.org/10.1371/journal.pone.0231866.

Running, Steven and Maosheng Zhao (2021). *MODIS/Terra Net Primary Production Gap-Filled Yearly L4 Global 500m SIN Grid V061*. Distributed by NASA EOSDIS Land Processes Distributed Active Archive Center. Accessed 2023-12-20. URL: https://doi.org/10.5067/MODIS/MOD17A3HGF.061.

Russian Statistical Yearbook (2022). Federal State Statistics Service (Rosstat). Accessed on 15th June 2024.

Russian Statistical Yearbook (2023). Federal State Statistics Service (Rosstat). Accessed on 15th June 2024.

Statistics of China, National Bureau of (2024). *Annual by Province - Gross*. Regional (Database). Accessed on 15th June 2024. URL: https://data.stats.gov.cn/english/index.htm.

UNdata (2024). *Per Capita GDP at Current Prices - US Dollars*. UNdata Database. Accessed on 15th June 2024. URL: https://data.un.org/Data.aspx?q=gdp&d=SNAAMA&f=grID%3a101%3bcurrID%3aUSD%3bpcFlag%3a1#SNAAMA.

Wenz, Leonie, Robert Devon Carr, Noah Kögel, Maximilian Kotz, and Matthias Kalkuhl (2023). "DOSE–Global data set of reported sub-national economic output". In: *Scientific Data* 10.1, p. 425.

Zhizhin, Maxim, Andrey Matveev, Tilottama Ghosh, Feng-Chi Hsu, Mark Howells, and Christopher Elvidge (2021). "Measuring Gas Flaring in Russia with Multispectral VIIRS Nightfire". In: *Remote Sensing* 13.16, p. 3078. DOI: 10.3390/rs13163078.