Transport Starter Data Kit: Historical socio-transport data for Colombia

Authors

Naomi Tan[1,2], Robert Ambunda[3], Nikola Medimorec[3] Angel Cortez[3], Agustina Krapp[3], Erin Maxwell[1], John Harrison[1], Mark Howells[1,2]

Affiliations

1. Centre for Sustainable Transitions: Energy, Environment and Resilience, Loughborough University
2. Centre for Environmental Policy, Imperial College London
3. SLOCAT Partnership on Sustainable, Low Carbon Transport

Corresponding author(s)

Naomi Tan (n.tan@lboro.ac.uk)

Abstract

Colombia is the third-most populated country in Latin America with 50 million people. Three quarters of the population live in cities. It is an upper middle-income country with a strong GDP growth rate in the past two decades. The GDP per capita is over USD 5,300 (as of 2020). However, Colombia also has the second-highest inequality score among Latin American countries.[1]

Colombia's transport system is well developed with an extensive road and highway network and railways spanning through the most populated areas.[2] The Nationally Determined Contribution by Colombia covers the introduction of 600,000 zero-emission vehicles. Freight transport will shift from road transport to inland waterways. Infrastructure improvements as well as urban development to increase density and accessibility are intended by Colombia among many other activities to ensure a more efficient and low carbon transport system.[3]

Transport demand modelling can be used to assess the implications of different scenarios and support improved policymaking. Data on transport activity is an important element for the development of national transport decarbonisation strategies. By having freight and passenger transport information, the impacts on vehicle and fuel consumption changes from replacing internal combustion engine vehicles with electric vehicles can be calculated. The development of a national decarbonisation strategy requires significant efforts. However, access to data is often a barrier to starting transport system modelling in developing countries, thereby causing delays. This article provides data that can be used to support a model for Colombia, which may act as a starting point

[1] https://www.mobiliseyourcity.net/sites/default/files/2022-05/Colombia.pdf
[2] https://en.wikipedia.org/wiki/Transport_in_Colombia
[3] https://changing-transport.org/ndc_country/colombia/

for further model development and scenario analysis. The data are collected entirely from publicly available and accessible sources, focusing on national reports, statistical yearbooks and academia.

Keywords

U4RIA

Transport data

Transport modelling

MAED

Colombia

Specifications Table

Subject	Transport
Specific subject area	Transport Data
Type of data	Tables Graphs
How data were acquired	Literature survey (databases and reports from international organisations; journal articles)
Data format	Raw and analysed
Parameters for data collection	Data collected based on inputs required to create an energy system model for Colombia
Description of data collection	Data were collected from the websites, annual reports and databases of international organisations, as well as from academic articles and existing modelling databases.
Data source location	Not applicable
Data accessibility	With the article and in a repository. Repository name: Zenodo. Direct URL to data: https://doi.org/10.5281/zenodo.7998839

Value of the Data

- These data can be used to develop national transport demand models to inform national transport investment outlooks and policy plans, as well as provide insights on the evolution of total final energy demand.
- The data are useful for country analysts, policy makers, and the broader scientific community, as a zero-order starting point for model development.
- These data could be used to examine a range of possible transport pathways, in addition to the examples given in this study, to provide further insights into the evolution of the country's power system.
- The data can be used for conducting an analysis of transport activity and capacity-building activities. Additionally, the methodology of translating the input data into modelling assumptions for a demand projection tool is presented in this article, which is useful for developing a zero-order national transport demand model. This is consistent with the U4RIA goals.

- The data can also be used as a call to action in addressing transport data gaps and establishing parameters for data collection to improve the consistency of transport-climate research in these countries.

1. Data Description

The data provided in this paper can be used as input data to develop transport demand models for Colombia. The data provided in this paper can be used to support the development of a transport model for Colombia. The data provided were collected from publicly available sources, including statistical yearbooks, transport ministry reports, statistics from national authorities and affiliated research institutions, academia and journal articles. Global datasets (primarily from the World Bank) were only consulted if severe data gaps existed. The dataset includes parameters on passenger and freight transport activity, disaggregated by transport mode (road, rail, aviation etc.) and geographic scale (inter-city or inner-city), if available. The dataset also covers the size of the vehicle fleet, disaggregated by vehicle types. The data coverage and sub-types vary among the parameters. The overall ambition is to include the most recent available year(s).

Item	*Description of Content*
Figure 1	A graph showing total population (million people), as well as the share of urban and rural population in Colombia.
Figure 2	A graph showing total GDP (million USD in 2015), as well as the share of the different sectors contributing to GDP in Colombia: agriculture, construction, mining, manufacturing, service, and energy.
Table 1	An additional table showing passenger transport activity in Colombia based on UN DESA Statistics Division data (see explanation below). The data feature information for 2018.
Table 2	A table showing freight transport activity in Colombia for the most recent year data was available. The data are curated from national statistics agencies or other government-affiliated agencies.
Table 3	A table showing freight transport activity in Colombia based on UN DESA Statistics Division data (see explanation below). The data feature information for 2018.
Table 4	A table showing global estimates for transport energy intensity levels (MJ per passenger-km) in 2018.
Table 5	A table showing load factors (average number of people per vehicle) for urban transport in Bogota.
Table 6	A table showing vehicle fleet data in Colombia for the most recent year data was available.

For the parameters on passenger and freight transport activity, an additional dataset was included in Table 1 and Table 3. The UN DESA Statistics Division modelled passenger activity and freight activity

for every country in support of SDG Indicator 9.1.2[4]. Passenger activity data provide information for road, rail and air transport. Freight data cover road, rail and inland water, aviation. The passenger-km and tonnes-km data originate from the Open SDG Data Hub. In this dataset only the data for International Transport Forum (ITF) (representing mostly OECD countries) and UNECE countries (mostly European countries) are based on national reporting. For non-ITF/UNECE countries, the data are estimated using the ITF model, which uses several covariates such as GDP, population and transport network coverage. A description of the model can be found in the ITF Transport Outlook 2017.

1.1 Population

Population data including total population, population growth, and split by rural or urban was gathered from The World Bank Open Data platform[5]. Figure 1 displays the total population disaggregated by urban and rural in Colombia.

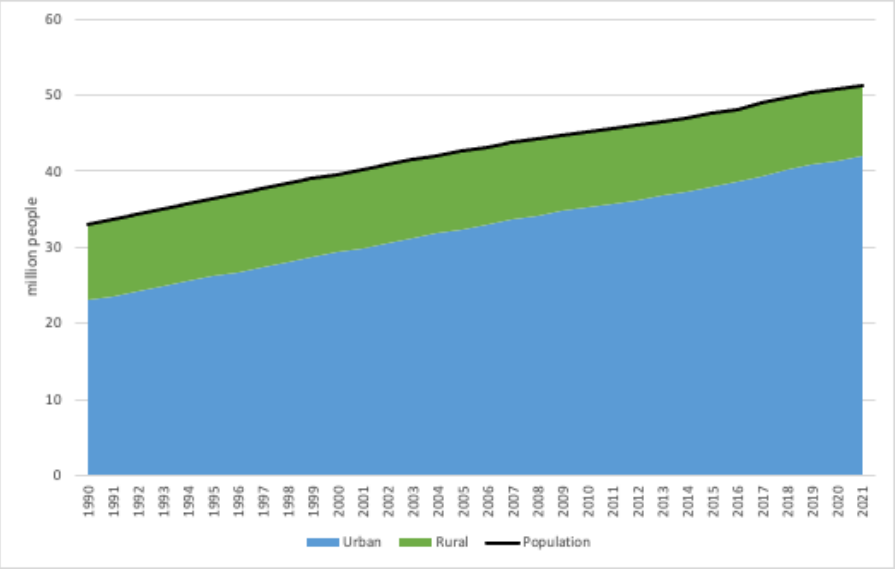

Figure 1: Total population (million people) disaggregated by urban and rural in Colombia

1.2 Gross domestic product (GDP)

GDP data including total GDP, GDP growth, and GDP share by sector (agriculture, manufacturing, service) was collected from The World Bank Open Data platform[5]. Where data was not available, data processing was done. Figure 2 shows the total GDP, as well as the share by sector, in Colombia.

[4] UN DESA (2020), Indicator 9.1.2: Freight volume by mode of transport (tonne kilometres), https://unstats-undesa.opendata.arcgis.com/datasets/4a5d7189e27148c48f045729ef9e40c8_0/about;
UN DESA (2020), Indicator 9.1.2: Passenger volume (passenger kilometres) by mode of transport, https://hub.arcgis.com/datasets/undesa::indicator-9-1-2-passenger-volume-passenger-kilometres-by-mode-of-transport-5/about

[5] https://data.worldbank.org/

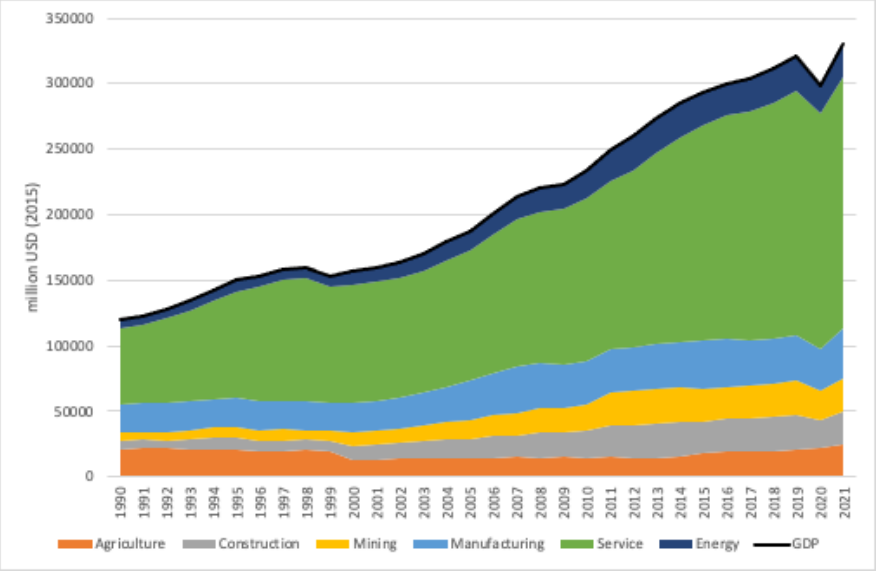

Figure 2: Total GDP (million USD in 2015) disaggregated by share in Colombia

1.3 Passenger transport activity

Colombia has very detailed transport data but it lacks the parameters in need for this research.[6] There is no information available on passenger activity from country reporting or any dataset.

According to the UN DESA modelled data, it is estimated that the passenger activity in Colombia recorded over 137161 million passenger-km for road, 10220 million passenger-km for rail and 20431 million passenger-kms for aviation in 2019. The large majority of passenger activity is conducted through road transport.

Table 1: Modelled passenger transport activity (million passenger-km) in Colombia

Mode	2019
Aviation	20431.79361
Rail	10220.37143
Road	137161.1143

[6] Transport ministry by Colombia provides several data points here: https://www.mintransporte.gov.co/documentos/15/estadisticas/

Source: UN DESA (2020), Indicator 9.1.2: Passenger volume (passenger kilometres) by mode of transport, https://hub.arcgis.com/datasets/undesa::indicator-9-1-2-passenger-volume-passenger-kilometres-by-mode-of-transport-5/about, last accessed April 2022.

1.4 Freight transport activity

Information on freight activity for Colombia has been retrieved for surface freight transport (rail, road and inland water transport). IRF World Road Statistics 2022 recorded that road recorded 246 million tonnes-km, rail 50 million tonnes and inland water 4.0 million tonnes-km in 2019.

Table 2: Freight transport activity (million tonnes-km) in Colombia

Mode	2019
Inland Water	4.857
Rail	50
Road	246.9896

Source: IRF (2023), World Road Statistics 2022, International Road Federation, https://worldroadstatistics.org/

The UN DESA modelled data for 2019 estimates that freight activity through roads surpasses 92701 million tonnes-km, rail 16542 million tonnes-km, aviation 1547 million tonnes-km and inland water transport at 14064 million tonnes-km.

Table 3: Modelled freight transport activity (million tonnes-km) in Colombia

Mode	2019
Aviation	1547.7
Rail	16542.2
Road	92701
Inland Water	14064.2

Source: UN DESA (2020), Indicator 9.1.2: Freight volume by mode of transport (tonne kilometres), https://unstats-undesa.opendata.arcgis.com/datasets/4a5d7189e27148c48f045729ef9e40c8_0/about, last accessed April 2022.

1.5 Energy intensities for transport

To further understand the efficiency of the transport system, information on the transport energy intensity is relevant. It is together with load factors (see *1.4 Load Factors*) inputs to MAED. However, such information is difficult to retrieve and there were no values available for this country. Global estimates by IEA can be found below in Table 4.

Table 4: Transport energy intensity for 2018 (in MJ / passenger-km)

Mode	Load factors
Rail	0.2

Two/three-wheelers	0.5
Buses and minibuses	0.7
Cars	1.8
Aviation	1.8
Large Cars	2.7

Source: IEA (2022), Energy intensity of passenger transport modes, 2018, https://www.iea.org/data-and-statistics/charts/energy-intensity-of-passenger-transport-modes-2018

1.6 Load factors

The load factors in the Starter Data Kits for Transport focuses on the average number of people transported by one unit in each transport mode. For example, for a bus, it is the average number of people per trip. In some cases, it might be also referred to as 'occupancy levels' for passenger transport. The values for Colombia are based on a study for Bogota, where the average load factor is 1.2 for cars, 1.3 for taxis and 36 for public buses for 2018.

Table 5: Load factors

Mode	Load factors
Private Car	1.2
Taxi	1.3
Public bus	36

Source: Oviedo, D., Granada, I. and Perez-Jaramillo, D. (2020), Ridesourcing and Travel Demand: Potential Effects of Transportation Network Companies in Bogotá, Sustainability, 12, 1732. https://doi.org/10.3390/su12051732

1.7 Vehicle fleet

Colombia has a total of 6.3 million road vehicles, as of 2020. Records cover information from 2015 to 2020. The vehicle fleet data is sourced from the IRF World Road Statistics 2022.

Table 6: Vehicle fleet in Colombia

Mode	2015	2016	2017	2018	2019	2020
Buses and Motor Coaches In Use	184172	188610	192458	195555	194456	192183
Lorries and Road Tractors In Use	352338	360910	367368	374091	378107	385048
Motorcycles and Mopeds In Use	6825884	7398986	7899676	8450799	8978337	9496550
Passenger Cars In Use	3019099	3165613	3299597	3439197	3548476	3625604
Total Vans, Pickups, Lorries and Road Tractors In Use	1983261	2086399	2187473	2302286	2400355	2492250
Vans and Pickups In Use	1630923	1725489	1820105	1928195	2022248	2107202

Total Vehicles In Use	5186532	5440622	5679528	5937038	6143287	6310037

Source: IRF (2023), World Road Statistics 2022, International Road Federation, https://worldroadstatistics.org/

2. Experimental Design, Materials, and Methods

The above data were gathered through extensive desk research. This included material from international organisations, journal articles, and media reports. The World Bank's data platform provided GDP share by sector for agriculture, manufacturing, and services. However, GDP share by construction, mining, and energy was also needed to align the data structure with the MAED tool. To address the lack of data available for these sectors, the authors assumed that construction, mining, manufacturing, and energy all fall within the industry sector. Thus, to obtain data for the three remaining sectors, the remaining percentage after considering agriculture, manufacturing, and services from The World Bank's data platform, was divided by three. It is therefore assumed that the GDP share of the construction, mining, and energy sectors are the same. The transport data was also compiled, presented, and discussed with local stakeholders to reach a consensus on the main data and assumptions.

3. Ethics Statement

Not applicable.

4. CRediT Author Statement

Naomi Tan: Investigation, Conceptualisation, Methodology; Data Collection; Visualization, Writing and Editing; **Robert Ambunda:** Data Collection; Investigation; Writing and Editing; **Nikola Medimorec:** Conceptualisation; Methodology; Data Collection; Investigation; Writing, Review & Editing; Supervision; **Angel Cortez:** Data Collection; **Agustina Krapp:** Data Collection; **Erin Maxwell:** Data Collection; **John Harrison:** Supervision; **Mark Howells:** Supervision

Acknowledgements

We would like to acknowledge the SLOCAT Partnership on Sustainable, Low Carbon Transport who helped make this and future iterations possible. We would also like to acknowledge the International Road Federation (IRF) and the International Union of Railways (UIC) for providing us with these data. The data are extracted from IRF World Road Statistics (WRS) and their use is subject to copyright and specific Terms and Conditions available on the WRS website. More WRS data are available for free on its Data Warehouse www.worldroadstatistics.org. Likewise, data was extracted from the UIC Statistics Rail Information System and Analyses (Railisa) and more can be found on its online tool https://uic-stats.uic.org/

Funding

As well as support in kind provided by the employers of the authors of this note, we also acknowledge core funding from the Climate Compatible Growth (CCG) Programme of the United Kingdom's Foreign, Commonwealth & Development Office (FCDO). The views expressed in this paper do not necessarily reflect the UK government's official policies.

Declaration of Competing Interests

The authors declare that they have no known competing financial interests or personal relationships which have or could be perceived to have influenced the work reported in this article.