

Project on Using Big Data to Measure the Nexus between Environment and Gender in Asia and the Pacific

Inception Workshop

Date and Time: 9:00 – 12:00 (Phnom Penh time), Tuesday, 6th December 2022

Venue: Raffles Hotel Le Royal, Phnom Penh, Cambodia, and [MS Teams](#)

Introduction

The nexus between the environment and gender has been of interest in recent years, with the 2030 Agenda for Sustainable Development providing renewed impetus to the discussion. However, the critical links between gender and environment are not well understood and gaps in data availability impede progress assessment. To better understand the environment-gender nexus, UN-Women and United Nations Economic and Social Commission for Asia and the Pacific (ESCAP) partnered with United Nations Environment Programme (UNEP) and the International Union for Conservation of Nature (IUCN) to put forward a proposal for an Environment-Gender Indicators Set that could capture issues of relevance for Asia-Pacific countries in the form of a 2019 working paper¹.

In response to the challenge on data availability, big data could be used as a data source for producing appropriate indicators relevant to the environment-gender nexus in a cost-effective and sustainable manner. Many National Statistical Offices (NSOs) in Asia and the Pacific region are increasingly experimenting with the application of non-traditional data sources including big data for the production of official statistics.

In line with efforts by NSOs in the region and with the aim of developing a better understanding of the nexus between environment and gender, ESCAP and the National Institute of Statistics (NIS) jointly developed a pilot project to explore the use of big data for the production of an indicator from the Indicator Set. *GE28 proportion of population with access to electricity, by sex* is the selected indicator for this project.

A team of UNESCAP staff and national and international consultants will work with NIS until March 2023 to implement the project. The team will work closely with NIS to produce a national strategy or strategic document detailing the key considerations, procedures, challenges and recommendations for using big data to produce the indicator. A key expectation is that the project may provide some significant insights that will support NIS and other NSOs in the region in developing appropriate overarching strategies for the use of big data in the context of official statistics.

Relevant data sources in scope of the project, given the timetable and resources available have been identified as:

Data source	Potential use and benefits	Challenges to be addressed
Nighttime Light (NTL) (Average day-night-band (DNB) values)	<ul style="list-style-type: none">Techniques exist to use this data as a proxy for understanding patterns of access to electricity on a highly frequent basisRelevant data can be obtained freely from sources such as: https://eogdata.mines.edu/products/vnl/	<ul style="list-style-type: none">Data science and geospatial skills are required to process and disseminate the dataData storage issues need to be resolved (eg Data lakes, the Cloud)

¹ ESCAP, UN Women, UN Environment, and IUCN, [Mainstreaming gender in environment statistics for the SDGs and beyond: Identifying priorities in Asia and the Pacific](#)

		<ul style="list-style-type: none"> Techniques for combining with other data sources are needed to understand the gender nexus
Open Street Map And/or Mapping data from Mapping Agency	<ul style="list-style-type: none"> Options for aligning NTL data with Local Administrative Units (LAU) Open Street Map freely available Establishing data sharing with other Government bodies may be relatively easy 	<ul style="list-style-type: none"> Software requirements for processing: QGIS or ArcGIS Data access from Mapping Agency (MoU, Agreement) Useful to use both sources if possible and compare quality
Population and Housing Census/Household surveys	<ul style="list-style-type: none"> Assuming gender patterns are relatively stable, can provide ready access to data on population and gender that can be combined with NTL data Potentially useful for ground-truthing results 	<ul style="list-style-type: none"> Managing confidentiality/ data storage issues Matching granularity of data to target geographies Identifying most useful datasets
World Bank's Multi-Tier Framework (MTF) Global Survey on access to electricity²	<ul style="list-style-type: none"> World Bank's 2017 MTF global survey for Cambodia provides details on different levels of electricity access for Cambodian households, offering information beyond simple access 	<ul style="list-style-type: none"> World Bank's 2017 MTF global survey data is less recent and might require integration with more recently produced datasets to yield up-to-date results

However, other data sources could be considered for future investigation as part of the national strategy, particularly where they have the potential to add significant insights based on sufficient data quality into the gender aspects of the indicator.

Aims of the workshop

The workshop represents the finalisation of an inception period where ESCAP and NIS have worked closely in building up cornerstone knowledge to support the use of relevant types of big data sources for the indicator. The workshop aims to:

1. Establish target data sources and methodologies to be investigated and on which to focus forward learning activities
2. Establish priorities for the strategic focus of the project in terms of user needs and the National Strategy for the Development of Statistics (NSDS)
3. Review and (re)align a work plan with the above and plan in detail for actions required

Project work plan

The general phases of the work plan to be discussed and adapted during and after the inception workshop is:

Time period	Actions
December 2022 - January 2023	Investigate and experiment with downloading and combining data sources; reporting on learning and future needs
January 2023	Finalisation of proposed methodology
January 2023	Additional training
January – February 2023	Liaise with potential users and partners to guide the development of the national strategy

² <https://documents1.worldbank.org/curated/en/141011521693254478/pdf/Cambodia-Beyond-connections-energy-access-diagnostic-report-based-on-the-multi-tier-framework.pdf>

February 2023	Drafting National Strategy
March 2023	National closing workshop
March 2023	Regional workshop and webinar (Stats Café)

National strategies target contents

The proposed content for the national strategies is currently as follows, with sections i to iii leading to the development of recommendations in sections iv to vii:

- i. User needs and rationale for using big data to support the monitoring of the indicator
- ii. Main production stakeholders, data partnership models and data flow
- iii. Proposed methodology:
 - a. Overview in relation to the GSBPM
 - b. Collection methods including approach to partnerships, data access, and to data confidentiality
 - c. Methodology for data processing and finalisation, including approaches to data storage
 - d. Requirements for dissemination
 - e. Data quality assurance and risk mitigation
 - f. Roles and responsibilities
- iv. Enabling regulatory framework and legal basis governing the use and storage of data
- v. Statistical indicators and quality assurance considerations
- vi. Limitations and risk mitigation
- vii. Strategic action plan:
 - a. Technical developments matched to a resourcing and training plan
 - b. User engagement plan
 - c. Leveraging wider gains for the National Statistical System

Agenda

Time	Program
8:30 – 9:00 (30 mins)	Registration
9:00 – 9:20 (20 mins)	Opening remarks <ul style="list-style-type: none"> Ms. Rachael Beaven, Director, Statistics Division, ESCAP H.E. Heang Kanol, Under Secretary of State, Ministry of Planning, Cambodia
9:20 – 9:45 (25 mins)	Photos and coffee break
9:45 – 10:00 (15 mins)	Presentation – Project overview: Objective, design, and methodology
10:00 – 10:15 (15 mins)	Live Demonstration – Big data collection using Nighttime Light (NTL) data from Google Earth
10:15 – 10:30 (15 mins)	Q&A
10:30 – 10:45 (15 mins)	Break
10:45 – 11:30 (45 mins)	Stakeholder consultation – Aligning the project to the National Strategy for the Development of Statistics (NSDS) and evolving data ecosystem, and generating inputs from stakeholders on project implementation and national strategy development.
11:30 – 11:45 (15 mins)	Feedback from participants, discussion, review of work plan, and next steps
11:45-12:00 (15 mins)	Closing remarks