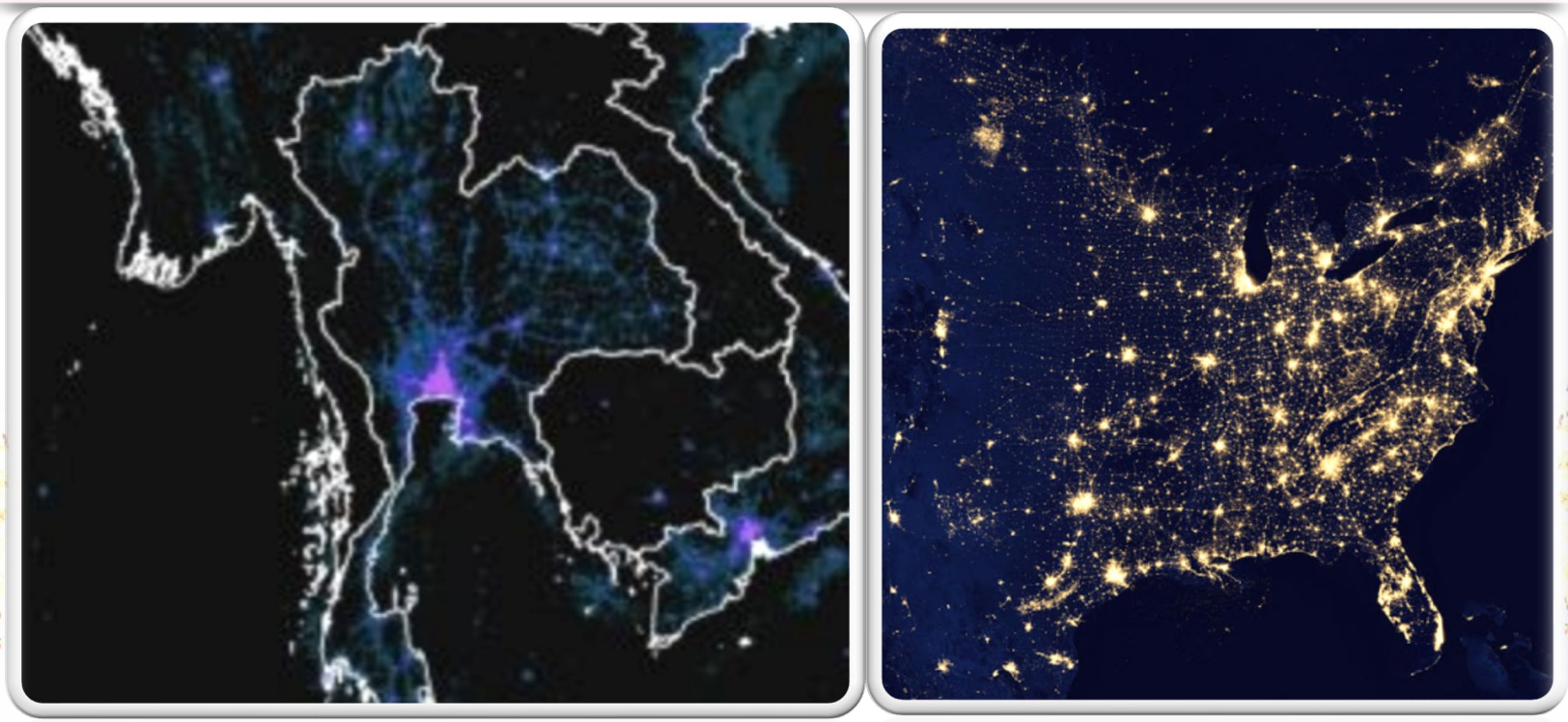
PROJECT ON USING BIG DATA TO MEASURE THE NEXUS BETWEEN ENVIRONMENT AND GENDER

Findings and Lessons Learned



Asia-Pacific Stats Cafe, 10th April 2023
Prepared by Buoy Somethea, Director Department of Statistical Information Services,
National Institute of Statistics, email: bsomethea@gmail.com

CONTENTS

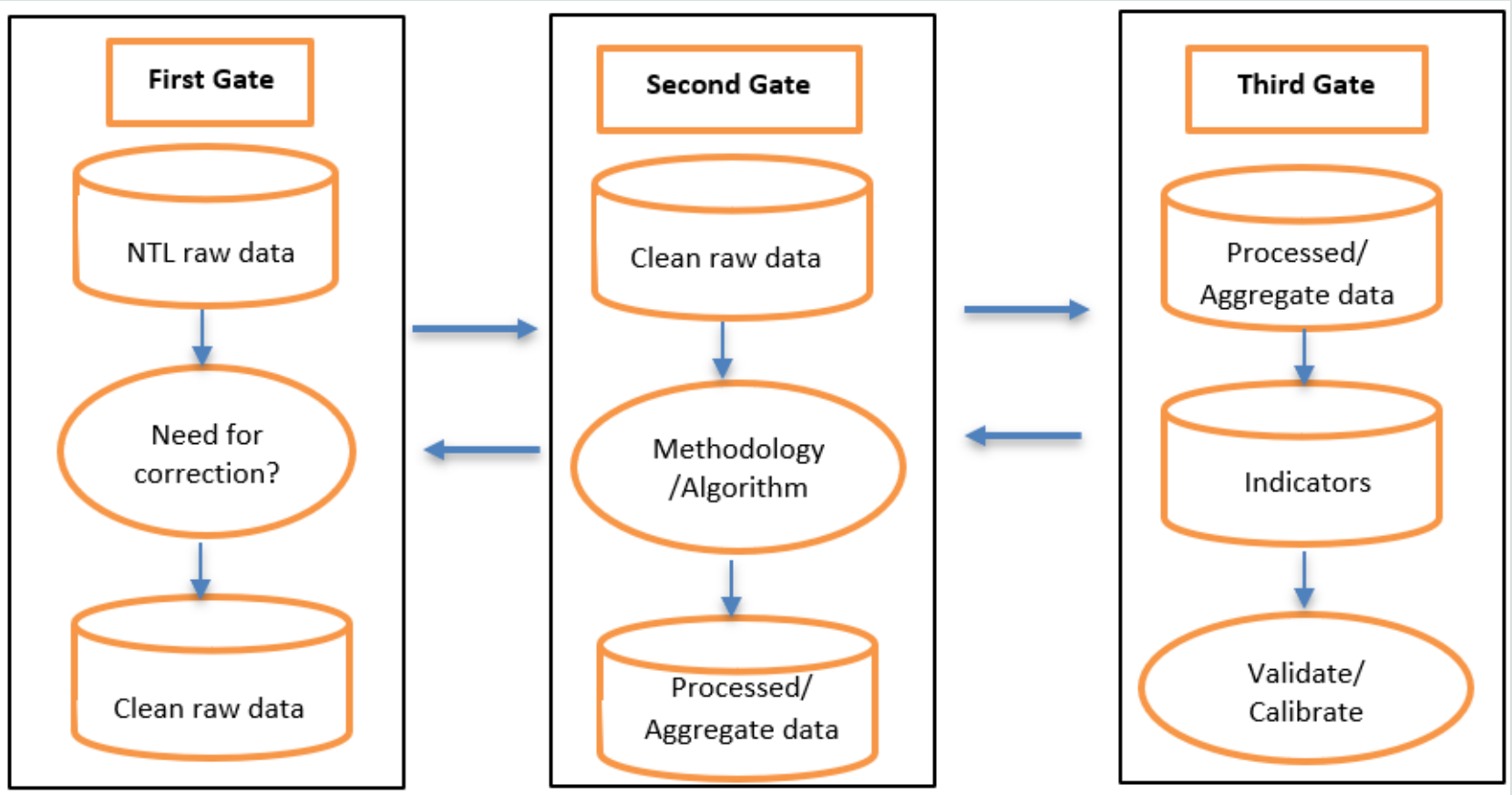
1. Project Background
2. Methodology
3. Hands-on
4. Findings
5. Challenges
6. Ways forward

1. PROJECT BACKGROUND

Purpose	Indicator	Expected Outcome
Measure the nexus between Environment and Gender	Population Access to electricity by Sex at National Level and Provinces	Testing Methodology National Guideline

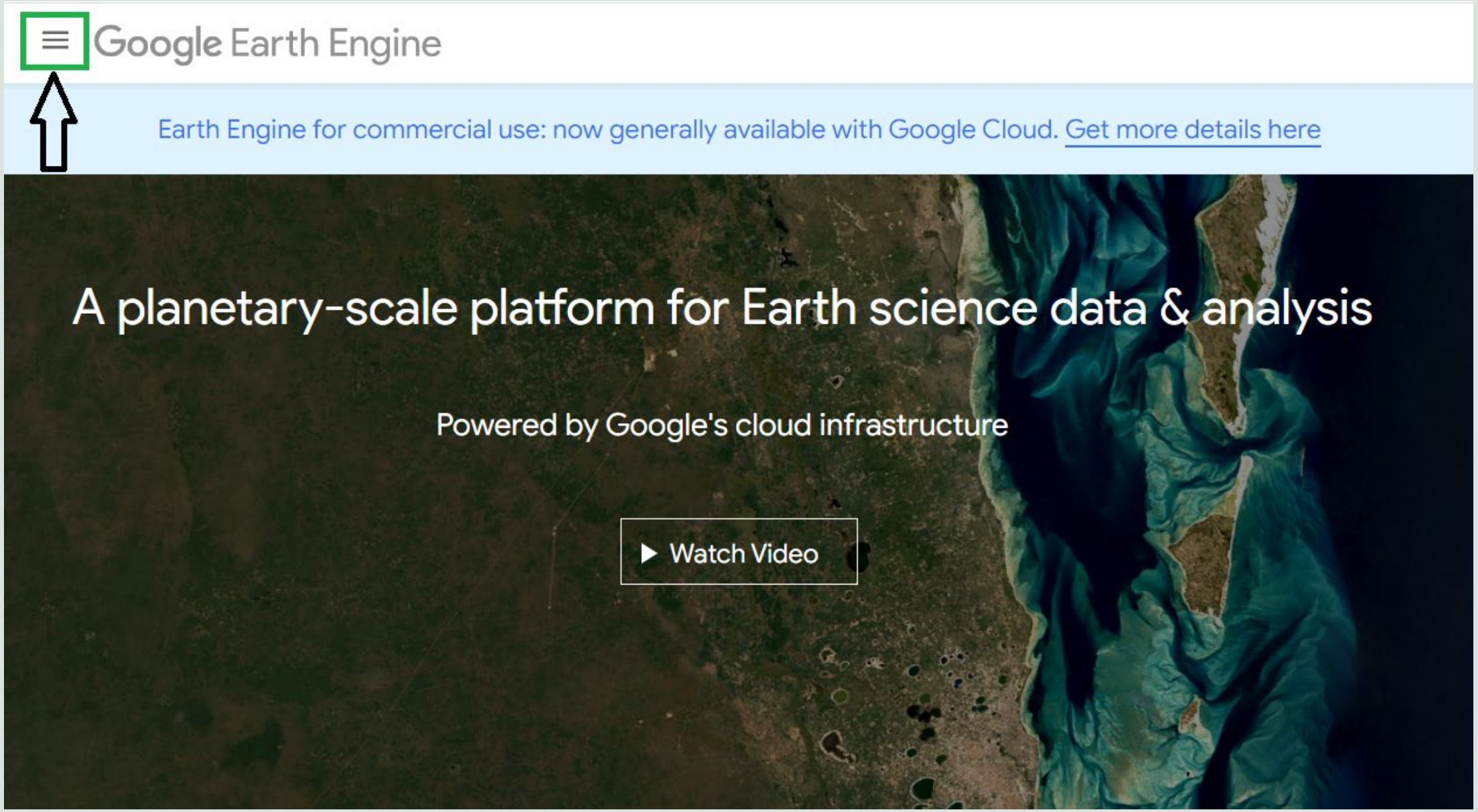
2. METHODOLOGY

2. 1. MODEL FOR THIS PROJECT-OVERVIEW

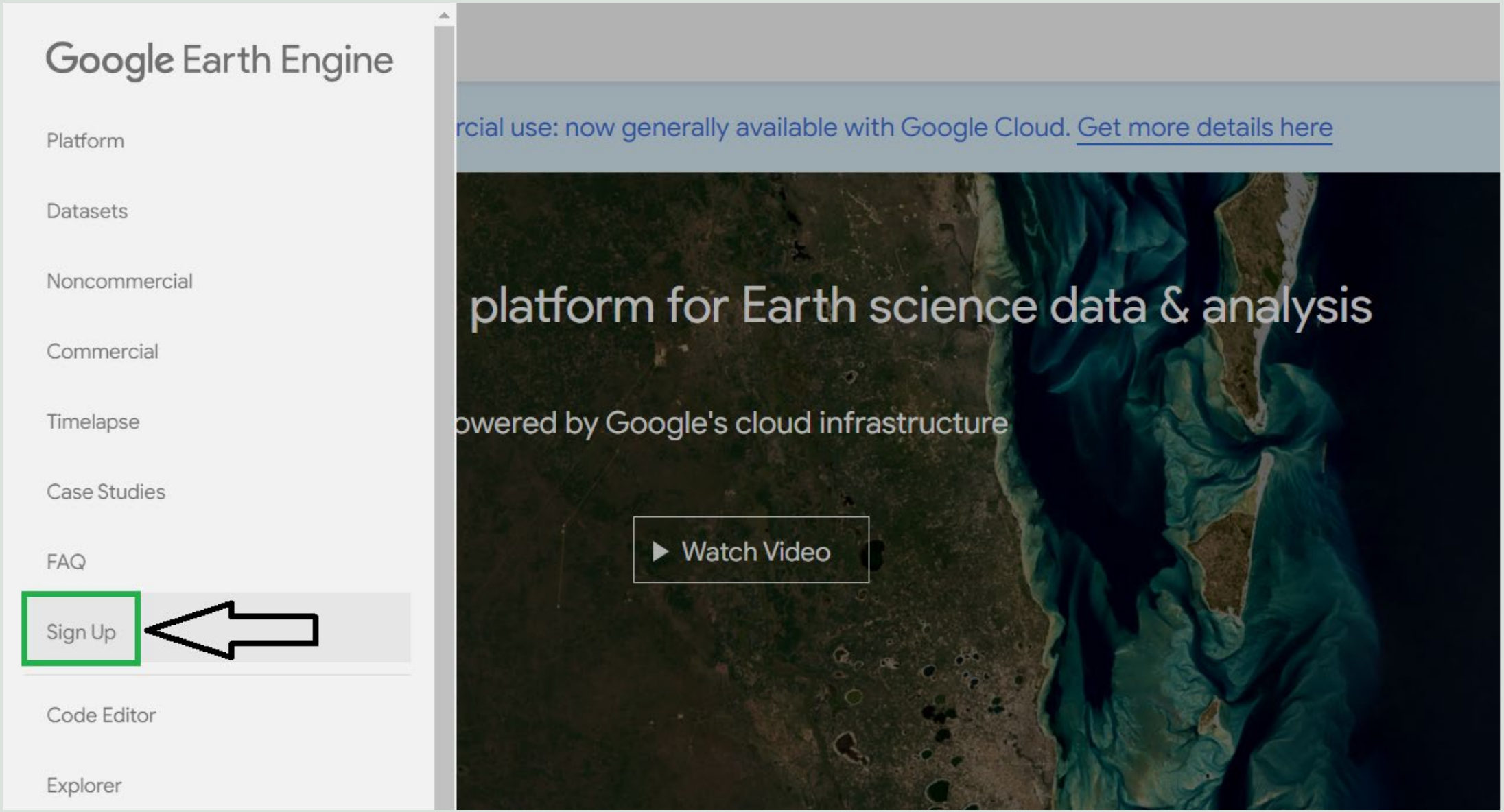


- **The first gate:** Google Earth Engine (GEE), system will provide Night-time Light (NTL) data
- **The second gate:** Setting thresholds, process data against thresholds, and aggregate data
- **The third gate:** Quality assurance, sensitivity analysis, and validation

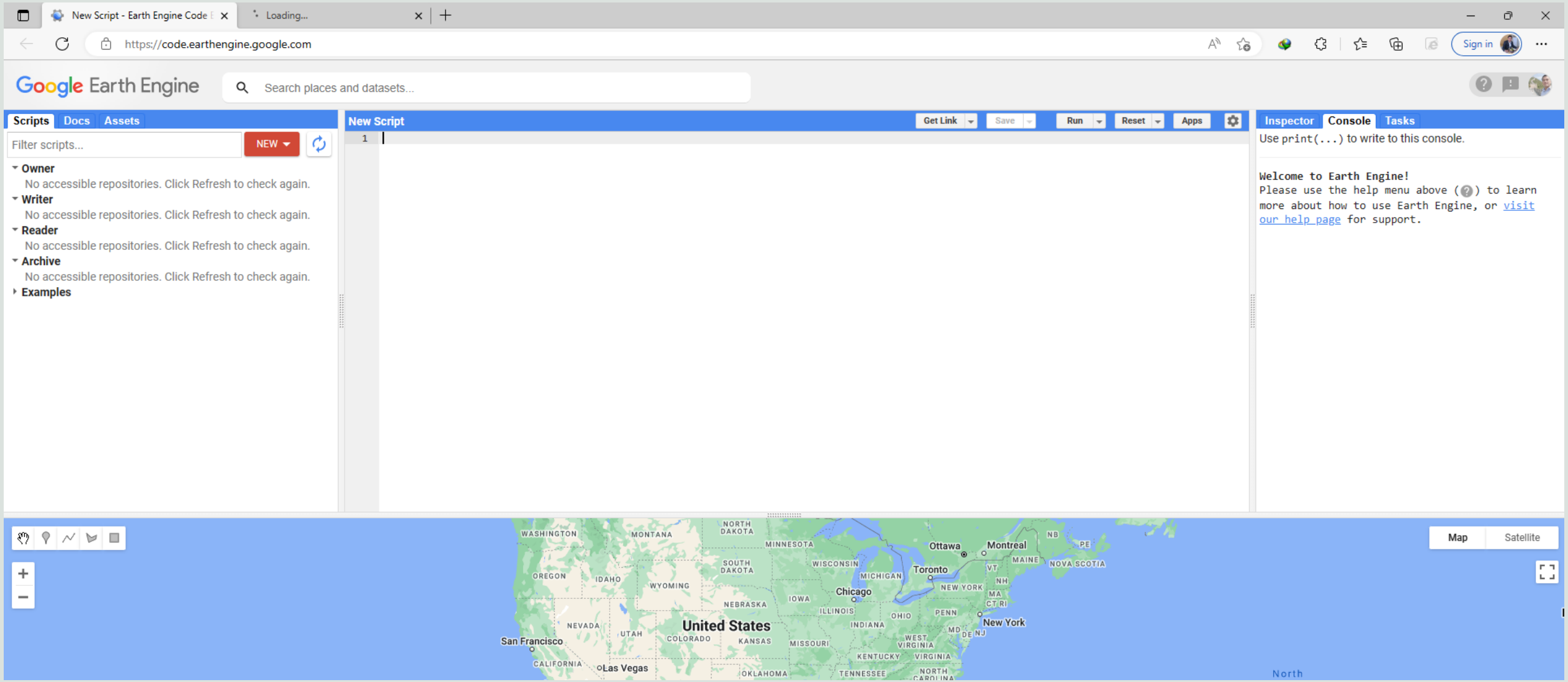
2.2 DETERMINING INPUT DATA AND PRE-PROCESSING - EARTH OBSERVATION DATA



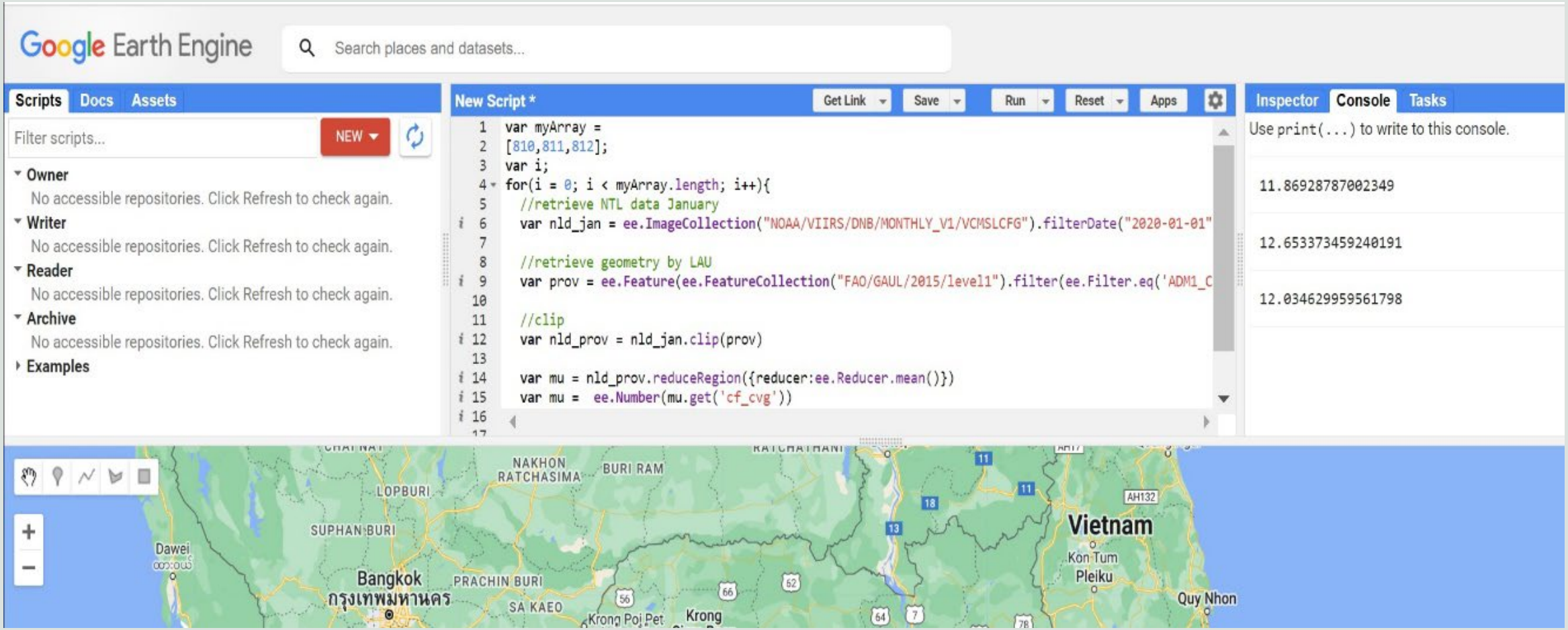
2.3 DETERMINING INPUT DATA AND PRE-PROCESSING - EARTH OBSERVATION DATA



2.4 DETERMINING INPUT DATA AND PRE-PROCESSING - EARTH OBSERVATION DATA



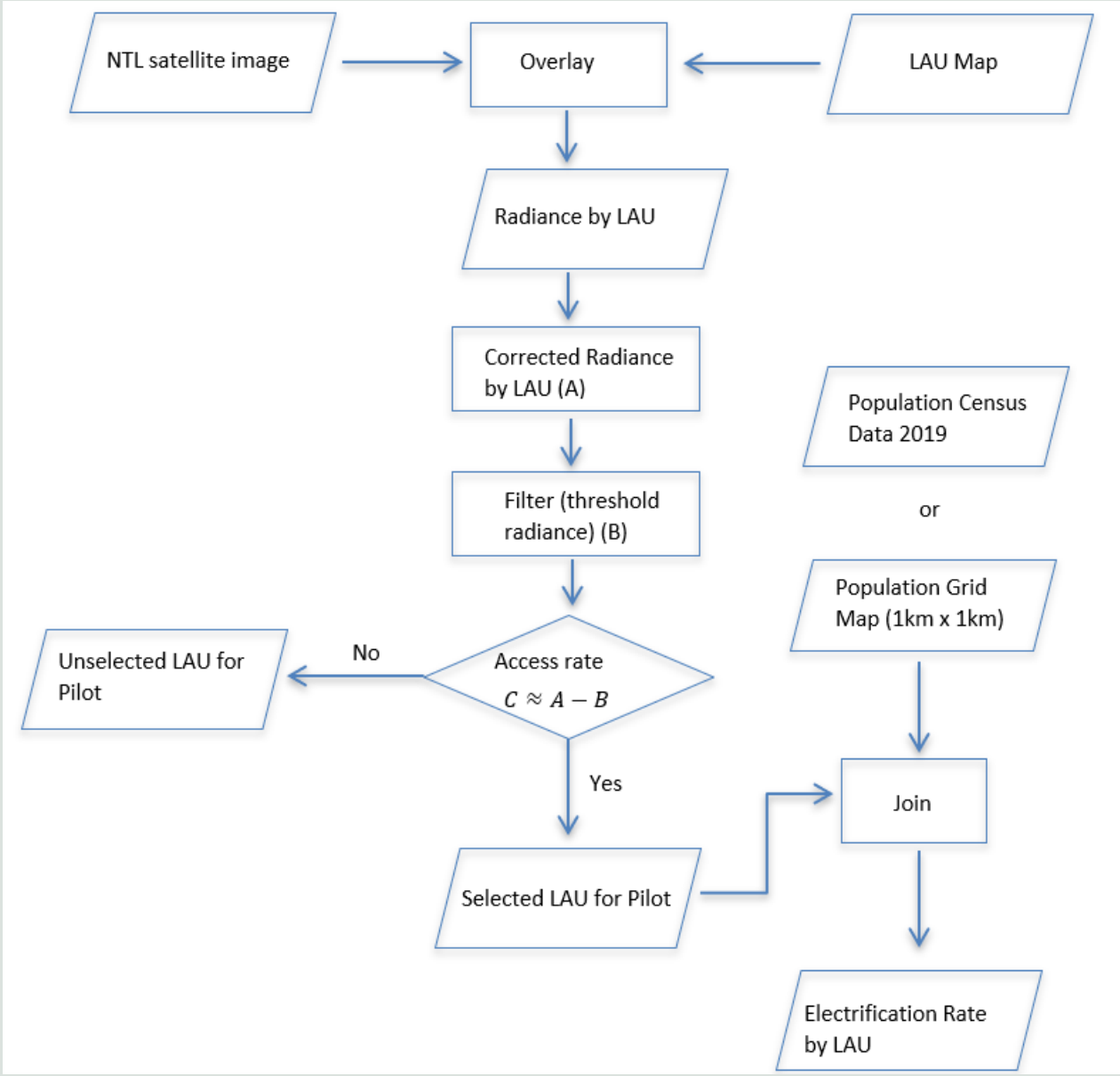
2.5 DETERMINING INPUT DATA AND PRE-PROCESSING - EARTH OBSERVATION DATA



- Writing/copy the codes/scripts to the GEE console.
- Download WFP code for specific areas_Country, Province/City, District
- Process the radiance/illumination of the NTL data on the GEE platform
- The results as csv or excel extension

Note:
GEE : Google Earth Engine
LAU : Local Administrative Unit
NTL : Night Time Light

2.7 DETERMINING INPUT DATA AND PRE-PROCESSING - INSTRUCTIONS FOR PRE-PROCESSING THE DATA



2.8 PROCESSING AND OUTPUT - SETTING THE THRESHOLD

- A standardized threshold for using radiance to determine access to electricity is not yet available.
- For this project, it is proposed to experiment with different thresholds using medians from quartiles and conduct a sensitivity analysis comparing electricity access rates from radiance data with electricity access rates from the latest Census and surveys to see which is the most suitable.

2.9 PROCESSING AND OUTPUT –ELECTRICITY ACCESS RATE, DISAGGREGATED BY SEX

$$Total\ Pop_E = \sum_{i=1}^{n} Pop_{E,i}$$

$$Total\ Female\ Pop_E = \sum_{i=1}^{n} Female\ Pop_{E,i}$$

i = Commune with access to electricity

$Pop_{E,i}$ = Population with access to electricity for each commune "i"

$Total\ Pop_E$ = Total population with access to electricity

$Female\ Pop_{E,i}$ = Female population with access to electricity for each commune "i"

$Total\ Female\ Pop_E$ = Total female population with access to electricity

Data on the total number of women with access to electricity can be derived by summing all female population data for all communes/provinces with access to electricity. To produce the proportion of population with access to electricity by sex, the total number of women with access to electricity derived from the NTL data is divided by the total female population in Cambodia.

$$ER_{NTL} = \frac{Total\ Female\ Pop_E}{Total\ Female\ Pop}$$

3. HANDS-0N

	A	B	C	D	E	F	G	K	L	M	N	O	P	Q
1	WFP Dis_Code	WFP Dis_Name	Radiance from GEE	Adjusted Radiance	Cut-off Point by Quartile	Median by Quartile	Access to Electricity (yes=1, no=0)	# Households	# Female Population	# Total Population	% Households with access to electricity	% Population with access to electricity	% F pop with access to electricity per total female	% F pop with electricity popul
2	12378	Anlong Veaeng	9.7934	9.4434	11.51373		=IF(D2>=$F$25,1,0)	12,884	28,490	58,389				
3	12362	Thma Bang	9.8718	9.5218	=QUARTILE($D$2:$D$186,2)		0	1,646	3,433	7,077				
4	12277	Thma Puok	10.1416	9.7916	**12.6751**		0	18,885	39,276	77,356				
5	12359	Mondol Seima	10.2350	=IF(C5>0,C5-0.35,"")			1	3,357	7,587	15,236				
6	12379	Banteay Ampil	10.4097	10.0597			0	12,534	26,941	54,716				
7	12408	Preah Sdach	10.4537	10.1037			0	30,485	63,786	121,867				
8	12426	Veun Sai	10.4754	10.1254			0	5,391	11,274	22,969				
9	12339	Krong Kampot	10.5316	10.1816			0	7,216	17,556	33,762				
10	12425	Ta Veaeng	10.5362	10.1862			0	1,960	3,756	7,533				
11	12381	Krong Samraong	10.5759	10.2259			0	15,739	36,346	72,803				
12	12415	Veal Veaeng	10.7535	10.4035			0	6,706	13,479	28,357				
13	12420	Bar Kaev	10.9727	10.6227			0	6,633	14,165	28,454				
14	12386	Chamkar Mon	10.9887	10.6387			0	28,805	73,920	146,038				
15	12276	Svay Chek	11.0481	10.6981			0	18,774	41,055	81,338				
16	12419	Krong Ban Lung	11.0816	10.7316			0	6,877	15,800	32,388				
17	12388	Doun Penh	11.1537	10.8037			0	30,996	85,437	163,330				
18	12442	Siem Pang	11.1759	10.8259			0	5,113	12,711	26,538				
19	12380	Chong Kal	11.1966	10.8466			0	6,261	14,307	27,910				
20	12439	Varin	11.2685	10.9185			0	10,180	23,301	46,475				
21	12402	Kampong Trabaek	11.2960	10.9460			0	29,178	61,194	115,757				
22	12375	Ou Reang	11.3007	10.9507			0	1,465	3,323	6,645				
23	12440	Sesan	11.3016	10.9516			0	5,608	12,834	26,209				
24	12406	Peam Chor	11.3850	11.0350		=MEDIAN(D2:D48)	0	15,033	31,316	61,721				
25	**12382**	**Trapeang Prasat**	**11.4098**	**11.0598**		**11.0748**	0	51 998	56 306	108 304	**0**		**0**	

4. FINDING- ACCESS RATE AT DISTRICT LEVEL

District Case

Threshold	Access Rate (Pre-process Radiance)	Access Rate CSES2019/20	Difference C = A – B
	A	B	
11.0598	87.6%	86.4%	1.2%
11.8439	62.2%	86.4%	-24.2%
12.4577	37.3%	86.4%	-49.1%

4.1. RESULT- ACCESS RATE AT HOUSEHOLD AND POPULATION LEVEL

% Households with access to electricity	% Population with access to electricity	% F pop with access to electricity per total female population	Result Population Census_HH	Result CSES2021_HH
Big data - Year 2022			2019	2021
92.07%	**91.88%**	**91.96%**	**84.10%**	**90.80%**

5. CHALLENGES

1. The WFP district code is provided by Consultant
2. The current WFP Map does not allow to find radiance at the commune level or lower
3. The current WFP Map used in GEE needs to be updated
4. Administrative boundary code is different (NIS Vs WFP code)
5. Population projection is not compute up to commune level

WFP Com_Code	WFP Com_Name	NIS District_Code/Name
12270	Malai	109. Malai
12271	Mongkol Borei	102. Mongkol Borei
12272	Ou Chrov	105. Ou Chrov
12273	Phnum Srok	103. Phnum Srok
12274	Preah Netr Preah	104. Preah Netr Preah
12275	Krong Serei Saophoan	106. Krong Serei Saophoan
12276	Svay Chek	108. Svay Chek
12277	Thma Puok	107. Thma Puok
		110. Krong Paoy Paet
12386	Chamkar Mon	1201. Chamkar Mon
12388	Doun Penh	1202. Doun Penh
12385	Prampir Meakkakra	1203. Prampir Meakkakra
12391	Tuol Kouk	1204. Tuol Kouk
12387	Dangkao	1205. Dangkao
12389	Mean Chey	1206. Mean Chey
12390	Ruessei Kaev	1207. Ruessei Kaev
		1208. Sen Sok
		1209. Pou Senchey
		1210. Chrouy Changvar
		1211. Prek Pnov
		1212. Chbar Ampov

6. WAY FORWARD

1. Upload Cambodia's latest shapefile to google earth engine
2. Calculate up to commune level (every month/year)
3. More capacity building workshops/trainings for the NIS staff
   - ✓ to better understand the basic coding languages used to retrieve NTL data (R program/Phyton/ JavaScript)
   - ✓ to be able to apply NTL/GEE to estimate other data (poverty rate/crop cultivated areas, etc.)


thank
you