Big Data for SDGs

Investing in Global and Enabling Public Goods

Olivier Dupriez, Senior Adviser
World Bank, Development Data Group
March 9, 2021

Mobilizing the data revolution (Big Data) for sustainable development

A slow and long process of modernization for NSOs in low-income countries

- Many criteria to be satisfied for use in official statistics: affordable, technically doable, sustainable, auditable, legal, ethical, safe

- Need a more enabling environment
  - Access to data (legal, trust, procurement, IT issues)
  - Tools, methods, IT infrastructure
  - Expertise
  - Funding

- Safeguards against accidental/intentional misuse of data

NSOs as enablers of Big Data use

- Without using Big Data, NSOs can have a big impact by supporting users of Big Data
  - NSO data needed to fit, calibrate, validate models
  - But issues of access, fitness-for-purpose
- NSOs to foster data integration and use by:
  - Adapting their data collection instruments/methods
    - How can data sources augment each other?
  - Disseminating more (micro)data, responsibly
- International community to contribute:
  - Data documentation and cataloguing tools to foster data discoverability and usability
  - Data protection tools and guidelines
  - Data dissemination tools (bulk, API)
  - Training, funding

NSOs as users of Big Data

- Multiple constraints
  - Legal and institutional environment (obstacles to data access)
  - Technical capacity (IT infrastructure, expertise)
  - Cost (storage, transfer, compute)
- Global public goods
  - Prioritize scalable, safe, sustainable solutions
  - Provide open, analysis-ready data to enable broader audiences
    - Free, formatted, documented, protected data
    - With open analytical tools and tutorials
  - Central catalog of relevant resources for improved discoverability
  - Training programs (incl. improvement of curricula)

Supporting the use of Big Data by providing analysis-ready open data

Example

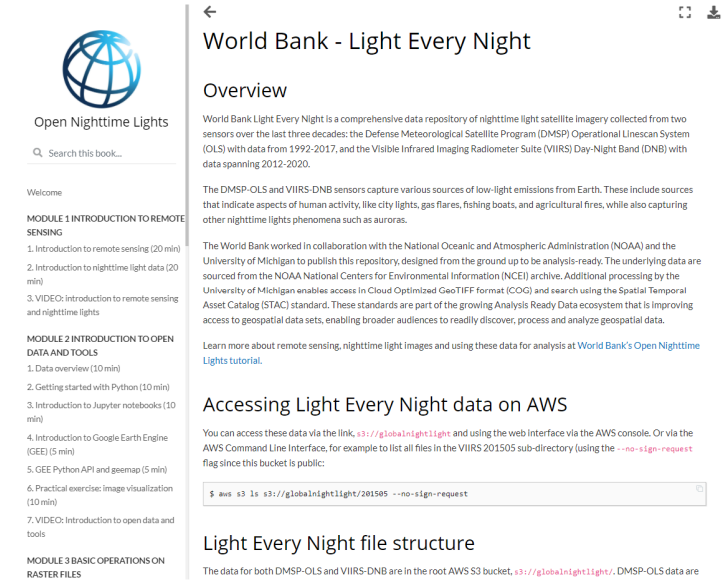
SDG 7: Sustainable, reliable, affordable energy for all

- **Data**: Light Every Night (1992-2020 – to be cont'd)
  - With the National Oceanic and Atmospheric Administration (NOAA) and the University of Michigan
  - Tbs of free data in Cloud Optimized GeoTIFF format
- **Tutorial**: in Python

Tutorial organization

- Module 1: Introduction to remote sensing
- Module 2: Introduction to open source tools
- Module 3: Basic operations on raster files
- Module 4: Charting
- Module 5: Data analysis and intercalibration
- Module 6: Exercise on analyzing economic activity in Nepal
- World Bank Light Every Night Archive: docs on accessing this new archive
- Applications of nighttime lights: assessing the power grid with nighttime lights

GitHub



https://worldbank.github.io/OpenNightLights/wb-light-every-night-readme.html

Supporting the use of Big Data by providing analysis-ready open data

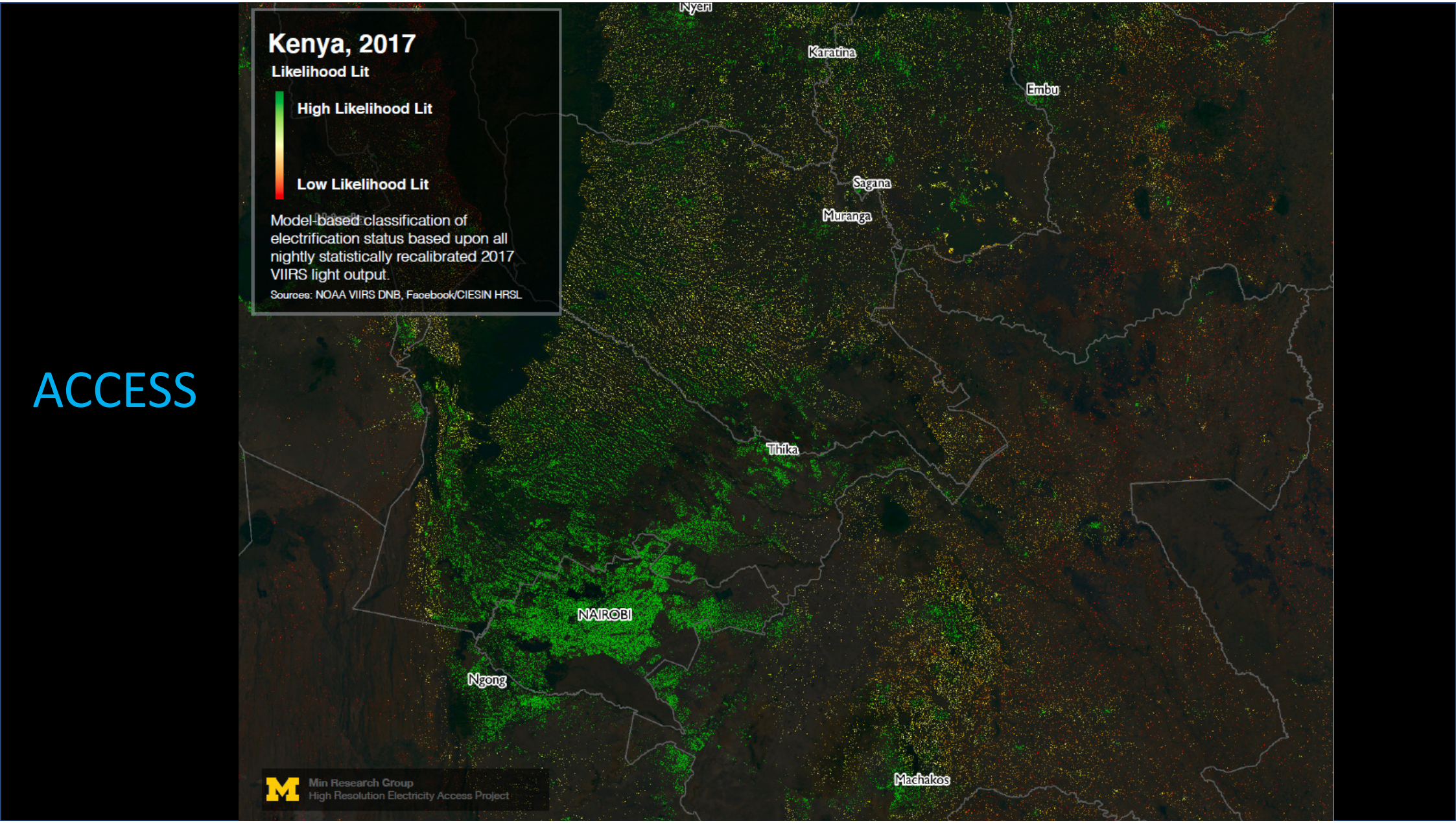
Example

- **Tool**: High Resolution Electricity Access (HREA) model by University of Michigan
  - Combines night light data and high-resolution population density data
  - Estimates at settlement level
    - Access: Likelihood lit
    - Reliability: Frequency lit
    - Use: Brightness z-scores
  - Results validated using Sustainable Energy for All (SE4ALL) data as gold standard

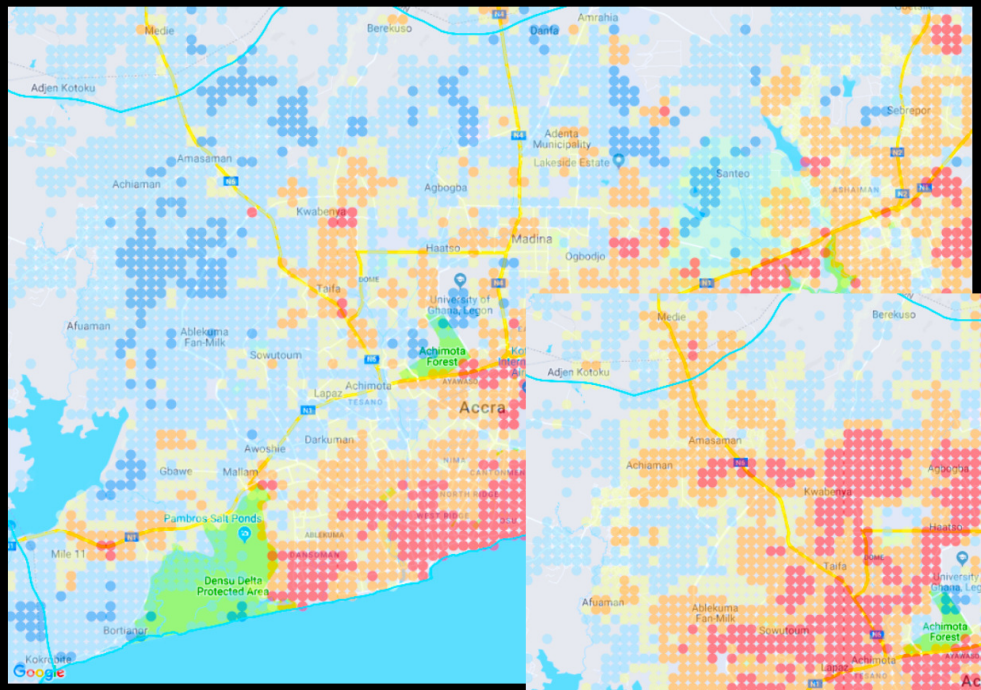
GitHub

https://github.com/zachokeeffe/nightlight_electrification

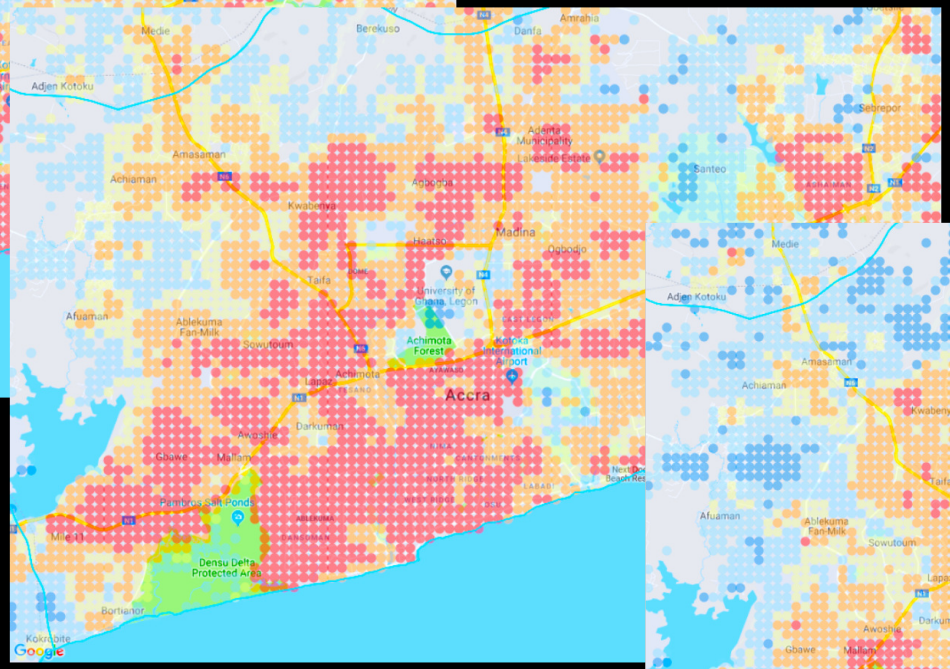
ACCESS


Kenya, 2017
Likelihood Lit
High Likelihood Lit
Low Likelihood Lit
Model-based classification of electrification status based upon all nightly statistically recalibrated 2017 VIIRS light output.
Sources: NOAA VIIRS DNB, Facebook/CIESIN HRSL
Nyeri
Karatina
Embu
Sagana
Muranga
Thika
NAIROBI
Ngong
Machakos
Min Research Group
High Resolution Electricity Access Project

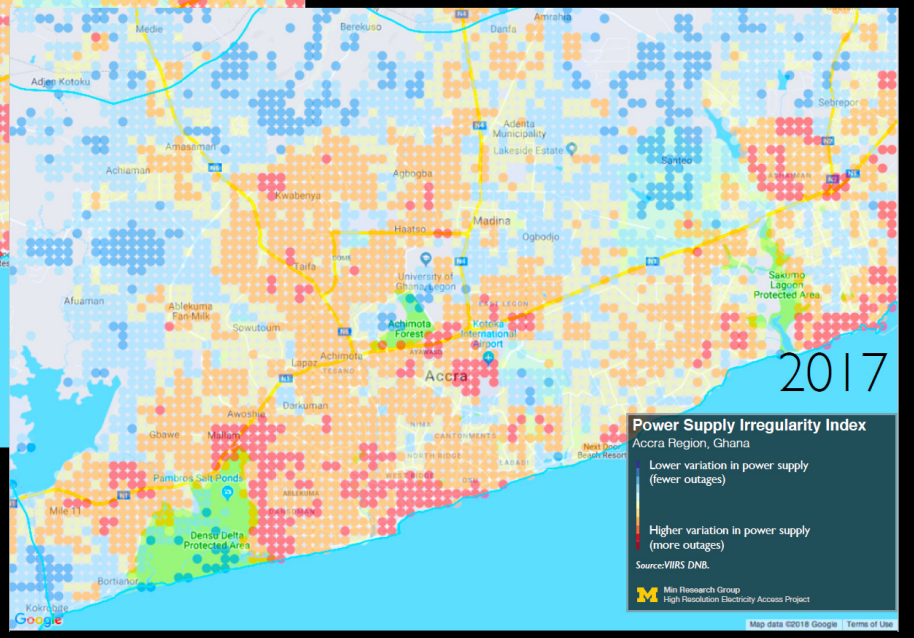
Power supply irregularity index
Accra region, GHANA



2013



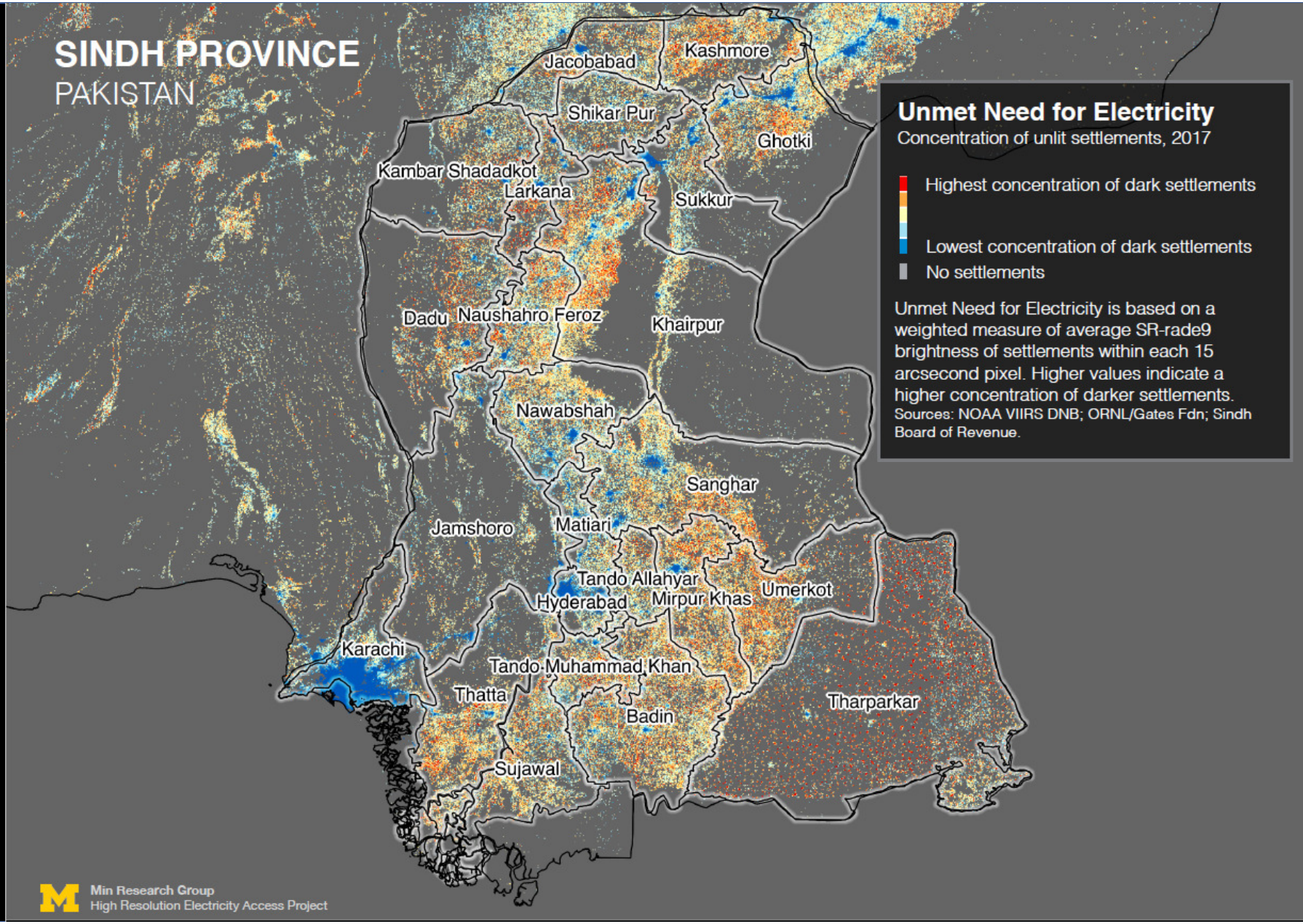
2015

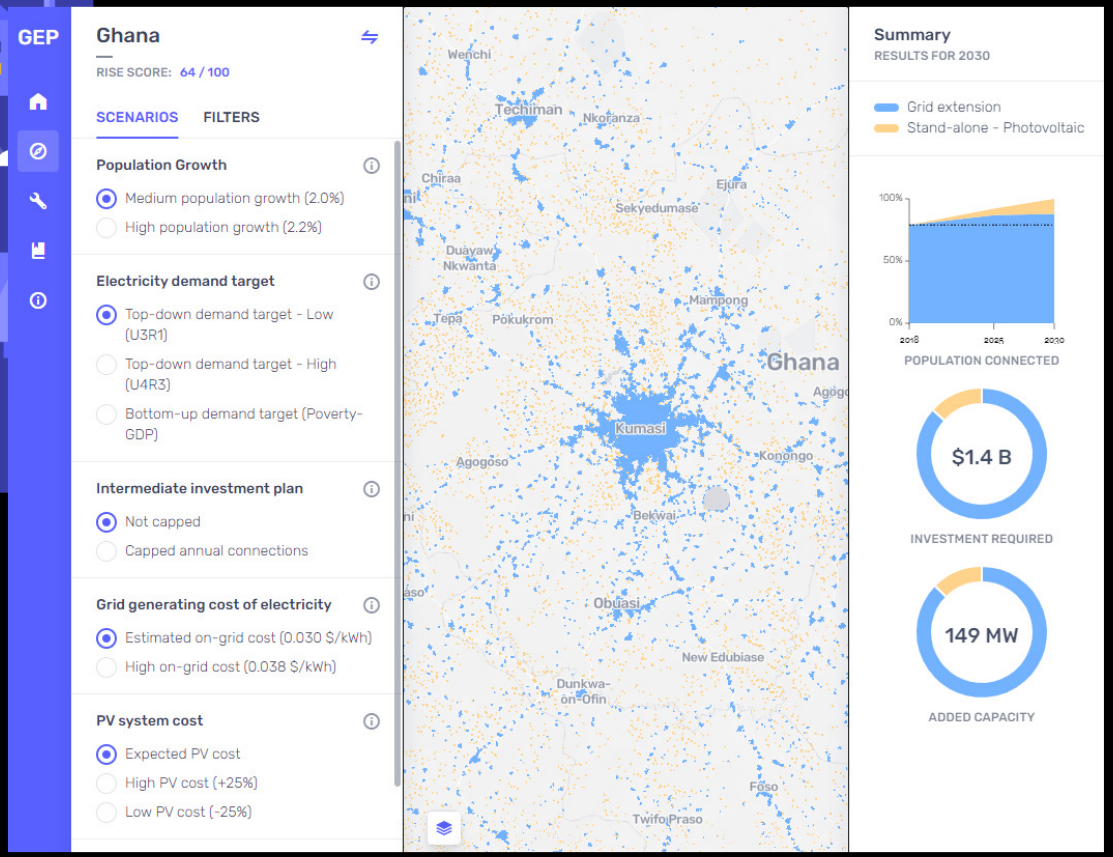


2017

RELIABILITY

UNMET NEED


SINDH PROVINCE
PAKISTAN
Jacobabad
Kashmore
Shikar Pur
Ghotki
Kambar Shadadkot
Larkana
Sukkur
Dadu
Naushahro Feroz
Khairpur
Nawabshah
Sanghar
Jamshoro
Matiari
Tando Allahyar
Hyderabad
Mirpur Khas
Umerkot
Karachi
Tando Muhammad Khan
Thatta
Badin
Tharparkar
Sujawal
Unmet Need for Electricity
Concentration of unlit settlements, 2017
Highest concentration of dark settlements
Lowest concentration of dark settlements
No settlements
Unmet Need for Electricity is based on a weighted measure of average SR-rade9 brightness of settlements within each 15 arcsecond pixel. Higher values indicate a higher concentration of darker settlements.
Sources: NOAA VIIRS DNB; ORNL/Gates Fdn; Sindh Board of Revenue.
Min Research Group
High Resolution Electricity Access Project



Then use other (big) data and model to act …

https://electrifynow.energydata.info/