Implementing Earth Observation Data for Official Statistics

Considerations for National Statistical Offices (NSOs)

Presented by: Jacinta Holloway
EO data for official statistics

In this session:

- How can EO data be used for official statistics?
- Cost benefit criteria for implementing EO data
- Collaboration between NSOs and policy makers
Who am I? Not Ric...

- **Co-author and editor**: of UN Satellite Imagery and Geospatial Data Task Team Handbook (author of chapter 5, co-author of chapters 3 and 4).
- **Presenter**: of hands on UN workshops introducing use of EO for official statistics.
- **Researcher**: developing new methods for analysing EO data to measure environment (reforestation).
- **Public servant**: Worked at ABS for 7 years, including in Tourism Satellite Account and Methodology.

Twitter: @thejholloway
EO data for official statistics

How can EO data be used for official statistics?

Possible applications of EO and other big data sources:

- sample frame or register creation
- partial data substitution for a subgroup of a population
- partial data substitution for some required data items
- imputation of missing data items
- editing
- linking to other data
- data confrontation
- generating new analytical insights (Tam & Clark 2015, p.442).

Example: Mexico

<table>
<thead>
<tr>
<th>Satellite Imagery</th>
<th>Other sources</th>
<th>National Uses</th>
<th>SDG / other applications</th>
</tr>
</thead>
<tbody>
<tr>
<td>SPOT ERMEX</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Very high resolution (0.5 m)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GEOEYE EVISMAR</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Medium resolution (5-30m)</td>
<td>Natural resources &amp; topographic charts, Forestry &amp; water data, In situ validation</td>
<td>Land Use &amp; Vegetation map series, Deforestation, land use changes, Monitoring crops</td>
<td>2. Zero hunger, 16. Clean water and sanitation, 13. Climate action, 14. Life below water, 15. Life on land</td>
</tr>
<tr>
<td>RAPIDEYE LANDSAT</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low resolution (250 m)</td>
<td>Topographic maps, Land use &amp; vegetation</td>
<td>Disaster monitoring Fires, large flooding</td>
<td>Sendai Framework, Climate action</td>
</tr>
<tr>
<td>MODIS</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Radar RADARSAT</td>
<td></td>
<td>Disaster monitoring Flooding, digital models in foggy areas</td>
<td>Sendai Framework, Climate action</td>
</tr>
</tbody>
</table>

Source: Presentation by Mr. Rolando Ocampo, Vice President, INEGI, Mexico. Meeting organized by UNSD on Geospatial Information and Earth Observations: Supporting Official Statistics in Monitoring the SDGS, New York, 7th March 2016.
EO data for official statistics

Conducting a cost benefit analysis

According to Tam & Clarke (2015) the costs and benefits of using a new data source need to be assessed in terms of:

- Costs
- Reduction in provider load;
- Sustainability of new source;
- Accuracy, relevance, consistency, interpretability, and timeliness of outputs stipulated in Data Quality Framework;
- Availability of knowledge for processing of EO data adequately
EO data for official statistics

Conducting a cost benefit analysis

- Costs
- Reduction in provider load;
- Sustainability of new source (access and ownership);
- Accuracy, relevance, consistency, interpretability, and timeliness of outputs stipulated in Data Quality Framework;
- Availability of knowledge for processing of EO data adequately
- Policy implications
- Ground truth and uncertainty
The importance of policy drivers

Additional cost benefit criteria: Policy implications
Are there clear policy drivers for producing these statistics? What are these?

What kinds of policy decisions rely on these statistics? How serious are these?

Do you need an indicator or a specific figure?

Are you as an NSO and your stakeholders comfortable with estimates informing this policy that are not directly collected via survey or census?
Collaboration between methodologists and policy makers

Development of methods and official statistics for environmental accounting needs to be iterative and collaborative process.

Need a clear and defined policy requirement.

Communication throughout process – not methodologists working in a silo for months and return with ‘the solution’.
Collaboration between methodologists and policy makers

For national accounts ongoing communication is key. Not a ‘set and forget’ approach to producing statistics.

E.g. ABS Tourism Satellite Account determining how to factor Air Bnb into the National Accounts, adjustments made to phone expenses at hotels when mobiles became popular.
Using EO data: CSIRO Decision Tree

**DECISION TREE**

- Define the nature of the issues to be monitored
- Define the status of existing observation networks
- Do you NEED to use satellite EO data?
  - YES
  - Can satellite EO potentially provide the required data?
    - YES
    - Determine minimum required data characteristics
    - Can the EO product meet the data requirements?
      - YES
      - USE EO PRODUCTS
      - NO
    - NO
  - NO

**QUESTIONS - Rationale**

- Which issues to be addressed?
  - Policy and/or regulatory drivers
  - Who are the stakeholders and beneficiaries?

- Metrics available?
  - Condition of the data networks?
  - Impediments of sharing, collecting, archiving the data?
  - Any papers/case-studies already written?
  - Any monitoring/modelling?

- Justification
  - Will complement ground-based monitoring networks or serve as the sole information source?
  - Will it be used in conjunction with modeling?
  - Are the EO data streams suitable for long-term decision-support?

- Suitability
  - What variables can EO provide?
  - Are data products readily available? Or will they be in future?

**Spatial resolution:** What is the appropriate pixel size?
**Temporal frequency:** How frequent do these observations need to be?
**Record-length:** How far back in time does your data record need to go?

- In-situ data requirements (how much in situ data are used in data product?)
- Reliability: What’s the certainty of supply of that product across space and through time?
- Accuracy: What’s the uncertainty associated with the data estimates?
- Maturity: How established is the data product?
- Complexity: What is the level of complication in the process of connecting the EO processed data into the data product?

Example: Measuring tree cover using EO data

Scenario: You are a manager with responsibility for monitoring deforestation and illegal clearing in South East Queensland. There is a known koala population in the area you are aiming to protect, and the preserved forest is also intended for bush walking and recreational use.

Can and should you use EO data to help you monitor clearing activities and improve the forest cover in this region?

Example: Measuring Tree Cover with EO data

Define the nature of the problem:

What management questions need to be answered?
What is the current state of the forest?
What changes to the forest cover have occurred in the past 5-10 years? Has clearing already occurred?
What changes can be made to improve and increase forest cover?
Would reforestation projects be helpful?
How should forest resources be monitored into the future?

What are the policy or regulatory drivers?
Preservation of forest for species and environmental protection, and human consumption.

Who are the stakeholders and beneficiaries?
Your employer, local residents, local government, environmental groups, forest management groups nationally and internationally, potentially the legal system if illegal clearing is identified and pursued.
Example: Measuring Tree Cover with EO data

Define the status of existing observation networks

What is the condition of the data networks?
Field measurements of koala presence and a sample of tree presence and type with latitude, longitude and tree height angle recorded.

What are the impediments to sharing, collating and archiving the data?
The size of required EO data can potentially make storage difficult. Storage options for NSOs include: storing the data onsite if they have server capacity, outsourcing the data storage or storing the information in a cloud.

What has been done in the past?
Some field measurements. There are maps of forest cover and forest regrowth available through the QLD Statewide Landcover and Trees Study (SLATS)
Any monitoring? Modelling?
Some historical data available from SLATS. No modelling of this area performed previously.
## Example: Measuring Tree Cover with EO data

<table>
<thead>
<tr>
<th>Question</th>
<th>Characteristic</th>
<th>Answer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Do you need to use EO data?</td>
<td>Justification</td>
<td>Yes, if there is no other source on forest cover that goes back 5-10 years.</td>
</tr>
<tr>
<td>Can EO provide the required data products?</td>
<td>Suitability</td>
<td>Yes, from EO data we can obtain NDVI, EVI, fractional cover from free products.</td>
</tr>
<tr>
<td>What is the appropriate pixel size?</td>
<td>Spatial resolution</td>
<td>Depends on computing capacity. Modis or Landsat (30m) would be suitable for identifying forest gain and loss.</td>
</tr>
<tr>
<td>What degree of accuracy is needed in the data products?</td>
<td>Accuracy</td>
<td>Capability to discern bare earth and other vegetation from forest cover.</td>
</tr>
<tr>
<td>How frequent do observations need to be?</td>
<td>Temporal frequency</td>
<td>It depends on how frequently you want to monitor land cover changes. Tree clearing should be identifiable quickly and tree growth is not fast enough to make it undetectable if checked on a monthly or seasonal basis.</td>
</tr>
</tbody>
</table>
**Example: Measuring Tree Cover with EO data**

<table>
<thead>
<tr>
<th>Question</th>
<th>Characteristic</th>
<th>Answer</th>
</tr>
</thead>
<tbody>
<tr>
<td>How far back in time does your data need to go?</td>
<td>Record length</td>
<td>10 years for this project.</td>
</tr>
<tr>
<td>Do you have guaranteed continuation of data supply in future?</td>
<td>Reliability</td>
<td>Yes, assuming sensor continues to operate. If not, another sensor could be substituted.</td>
</tr>
</tbody>
</table>