Re-imagining the role of a national mapping agency to support spatially enabled governance

Mark TABOR, Karen HUNT, Andy WILSON, Anthony BECK, Juliet EZECHIE, United Kingdom

Key words: Surveyor 4.0, Informal Settlements, Land Management, Low Cost Technology, Commission 3 - Spatial Information Management

SUMMARY

Effective policy and decisions can only be made when supported by complete, accurate and timely evidence. However, we are surrounded by information and data, but don't necessarily know what is relevant for our needs. When we do find appropriate data, is it accessible, accurate, authoritative, or assured? Has it been collected at the right time, does it cover our area of interest? Finally, is the data licensed in a manner that allows us to use it effectively?

With such an array of constraints it is easy to see why "analysis paralysis" may set in. Ordnance Survey (OS) has been providing spatial data that supports evidence-based policy and decision-making and helping governments and business improvse efficiencies for many decades. This includes applications in land use change and (cadastral) land administration, but also across many public services including: emergency services, environment, planning, healthcare provision, transportation and local government. We have seen that spatial data forms the critical foundation for effective evidence-based decision making.

To answer this need for foundational spatial data we have developed an automated digital mapping solution based on machine learning and artificial intelligence algorithms to extract digital map features and land cover classes from either satellite, aerial or drone imagery. Our automated digital mapping solution, rapidly captures, integrates, and enables the visualisation of spatial data across the landscape. This provides the evidence upon which to collaborate, model, plan, implement and monitor policy and critical public services. The ability to rapidly review and change the modes of collection, interpretation and re-use are 'built-in' to the architectural framework allowing the evolution of workflows that deliver truly 'fit-for-purpose' outputs. This results in a more nuanced relationship between policy and practice as more appropriate tools are now available to measure and evaluate policy impact more effectively.

The use of geospatial information is also critical for good land governance. Fitness-for-purpose and timeliness are also critical characteristics: ideally one should be able to provide appropriate data to support decision making in a timely manner. The challenge is that

different decisions need different supporting data. Our automated digital mapping solution aims to address this challenge.

These innovative automated mapping technologies were used in Lusaka, Zambia giving city planners a much clearer picture of the city's rapid urbanisation and expanse and density of informal settlements. By bringing together national government with local government, academia, built environment professionals and the private sector, the Zambia mapping pilot provides an excellent example of a structured multi-stakeholder approach. It collectively addresses social and policy challenges, while deploying a configurable and dynamic technology stack which is both scalable and replicable. The resultant map is now being used to identify gaps in critical public services such as sanitation, health and transportation allowing resources to be targeted where they are needed most. The map is also being used to support the forthcoming census and also being assessed for use to support land registration and property taxation.

The map data now forms part of the National Spatial Data hub and is accessible by all parts of the Zambian government to assist with policy development as described above.

The partnership recognises that the provision of well-planned cities relies upon collaborative stakeholder evidence-based engagement, with supporting technology infrastructure which delivers information rich datasets that are so essential to the effective management of today's cities.

We are delighted this pilot won the AI Innovation of the Year award (Digital Leaders). We will present an overview of the technology, how the technology can support spatially enabled governance, and examples from the award-winning pilot from Lusaka.

FOOTER
Please do NOT include any footer text. Page numbers must be added above the footer. FIG will insert the correct footer text when the paper will be published on the web site.

Re-imagining the role of a national mapping agency to support spatially enabled governance

Mark TABOR, Karen HUNT, Andy WILSON, Anthony BECK, Juliet EZECHIE, United Kingdom

1. Introduction

Governments around the world face new challenges, from population growth and urbanisation, to climate change as well as pandemics. Geospatial Information is a crucial tool to help nations navigate these challenges effectively, providing insights and information to underpin effective policy and economic decisions by making it possible to monitor, measure, predict and adapt effectively. This in turn improves transparency, drives operational efficiencies, ensures compliance, and helps manage risk.

Ten of the seventeen United Nations sustainable development goals are land focused and link people to place. There are fundamental geospatial data themes that are crucial for informed social, economic, and environmental decision making – the three pillars of sustainable development required to run a country. The use of geospatial information is critical for good land governance and provides a mechanism for monitoring indicators relating to land access, property rights, ownership, land degradation, rapid urbanization, climate change, sustainable development and overall environmental, economic and social wellbeing and to demonstrate national progress globally.

Geospatial information serves an important role in definitively linking different data sets that would be otherwise unconnected, through their shared location and place. Enhancing and linking other data-driven initiatives, geospatial information is a critical enabler for nations seeking to grow their economies, drive sustainable development and enable new and innovative uses of government data to solve problems and provide new services.

Effective policy and decisions can only be made when supported by complete, accurate and timely evidence. We are surrounded by information and data, but don't necessarily know what is relevant for our needs. When we do find appropriate data, is it accessible, accurate, authoritative, or assured? Has it been collected at the right time, does it cover our area of interest? Finally, is the data licensed in a manner that allows us to use it effectively?

As the National Mapping agency of Great Britain these are issues that are important and critcal for Ordnance Survey in supporting the British Government, Businesses and Citizens.

2. Ordnance Survey: supporting evidence-based policy and decision making

Ordnance Survey has been providing spatial data that supports evidence-based policy and decision-making and helping governments and business improve efficiencies for many decades. This includes applications in land use change and (cadastral) land administration, but also across many public services including: emergency services, environment, planning, healthcare provision, transportation and local government. We have seen that spatial data forms the critical foundation for effective evidence-based decision making, by combining all the different data together for a specific property, neighborhood, area or region to ensure the right policies and resources are deployed.

The need for relevant, accurate and timely data that directly supports policy and decision making is unquestioned. However, as the challenges faced become more specific and nuanced, the data requirements equally become more challenging. In turn, it becomes more difficult for traditional national mapping data to respond to the dynamic nature of socio-economic or modern governance systems. Hence, fitness-for-purpose and timeliness are also critical characteristics: ideally one should be able to provide appropriate data to support decision making in a timely manner. The challenge is that different decisions need different supporting data. In summary data needs to be Authoritative and Accessible (A), Comprehensive and Consistent (C) and Timely (T) enabling organisations to ACT.

Ordnance Survey have recognized the need to fundamentally re-evaluate mechanisms for collecting foundational spatial data that reflect the need for a data baseline (providing longitudinal consistency) and change. We have developed a state of the automated digital mapping solution based on machine learning and artificial intelligence algorithms to flexibly and quickly extract digital map features and land cover classes.

3. Automated Feature Extraction

Our automated digital mapping solution utilises satellite, aerial or drone imagery to rapidly capture, integrate, to enable the visualisation of spatial data across the landscape. This provides the evidence upon which to collaborate, model, plan, implement and monitor policy. The ability to rapidly review and change the modes of collection, specification, interpretation and re-use are ‘built-in’ to the architectural framework allowing the evolution of workflows that deliver truly ‘fit-for-purpose’ outputs. This results in a more nuanced relationship between policy and practice as more appropriate tools are now available to measure and evaluate policy impact more effectively.

OS Automated Map Creation is powered by a configurable Automated Feature Extraction (AFE) process; allowing OS to rapidly collect data in a fit-for- purpose manner at City, Regional, Federal and National levels. AFE enables the rapid capture of features, land parcels or land cover to create base maps to meet a range of policy or commercial agendas. A base map can be defined as either a land parcel map, topographic land cover map, orthophoto map or a combination depending on the requirement.

OS has built supporting capabilities over a range of different environmental global conditions with appropriate classifiers and flexible workflows to produce accurate geospatial information to suit specific applications. Mapping features can be captured with configurable associated attribution and metadata, and tailored to specific end-use scenarios.

4. Lusaka – A case study

We received a request from the Zambian Ministry of Local Government to better understand the extent and location of their informal settlements, which will aid the implementation of the new National Urbanisation Policy (NUP). The NUP aims to promote prosperous, inclusive and resilient urban settlements, and ensure Zambia's towns and cities become engines of national growth. In order to implement the plan, a critical gap in knowledge is the need for a better understanding of the city's informal settlements and gaps in critical public services such as sanitation, health care and transportation.
1
Currently, the base mapping in Zambia is not sufficiently current or detailed enough to provide a meaningful spatial data infrastructure foundation, but Zambia Ministry of Lands and Natural Resources has made available their 20cm resolution aerial imagery of Lusaka to afford the opportunity for OS to utilise its advanced automated process to generate a new detailed base map.

4.1 Methodology

The AFE processing flowline is described in the Figure below and includes the following components:

- Data preparation: Capture, Ingest and Ready the data for processing
- Pre-processing: Using the prepared data to generate a set of patches and tiles for training the model, validating our results, and testing accuracy in later processes.
- Train the model: Training a semantic segmentation Computer Neural Network using the created patches to produce a model optimised for classifying a set of feature classes
- Inference: Using the trained model to classify imagery from untrained regions, to produce a classified land cover output where every pixel within the imagery is classified to a feature class. Embedding within an ESRI ArcPro environment
- Assessment: Measures the success of our process against a completely unseen test block and indicates if further training or optimisation is required.
- Post Processing: Enhances the quality of the classified land cover output through cleaning up geometry, adding attribution.

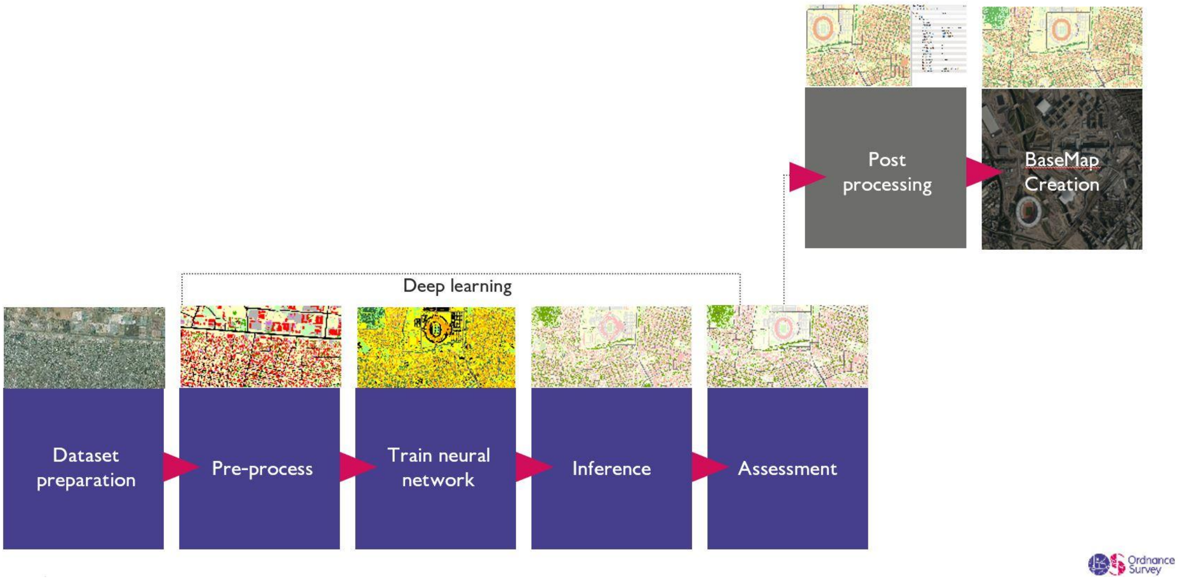

Figure 2 The AFE processing flowline

Data ingest

The first stage involves a technical analysis of the available 2017 aerial imagery (20cm 3 band - Red, Green, Blue (RGB colour)) to ensure that the required specification is achievable. This analysis ensures that the imagery dataset is free from any blurring and that the key features, such as buildings are clearly identifiable by the human eye from the imagery. Occasionally imagery can suffer from excessive shadow which can obscure vital features if the imagery was flown when the sun angle was too low.

Figure 3 Aerial imagery supplied by Zambian Land Authority - 20cm Ground Sample Distance, 3 band RGB colour

Training data creation

Labelling of every pixel of a representative sample is required to generate the data needed to train a deep learning system. Ordnance Survey evaluated the total geography of Lusaka and selected a sample of 9% of the total data which equated to 37kms of imagery for training data creation. As the training model improves the perecentage of training data required reduces significantly. This selection covered the variety of feature classes required so that the best possible automated pre-processing results can be achieved. As the priority was the capture of buildings for informal settlements, the sample specifically included eleven different informal settlement types based on a localised typology.

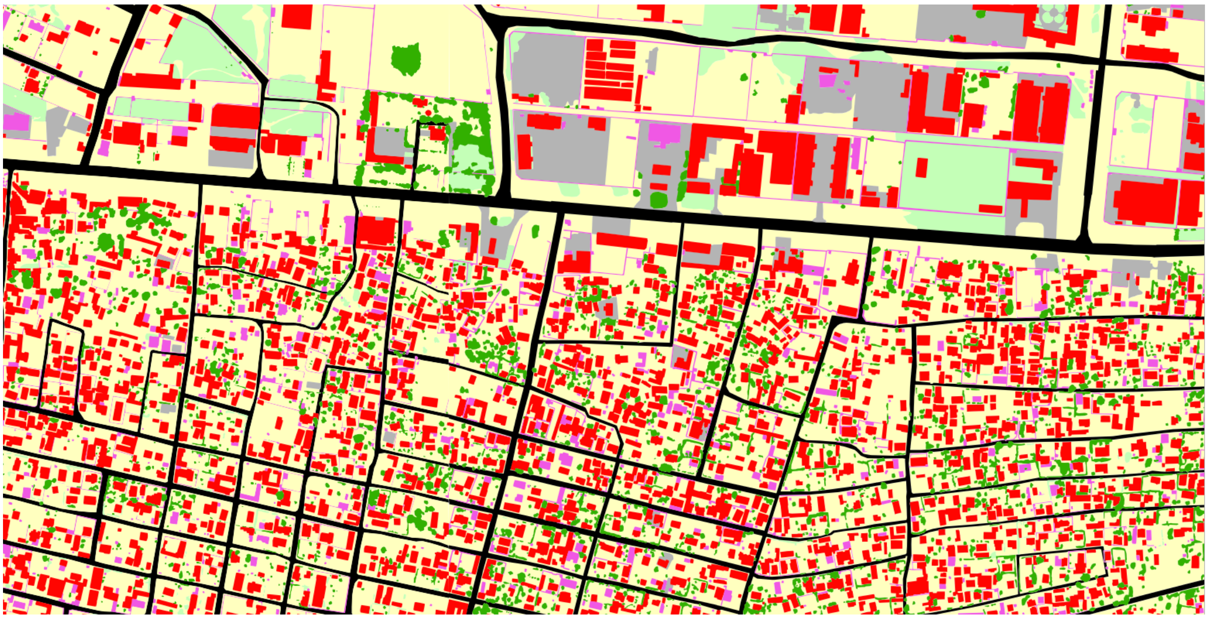

Figure 4 Training data labelled - every pixel is classified

Pre-processing

The steps taken to prepare the imagery prior to training include:

- Colour balancing to ensure that the image colour is consistent throughout the whole block which will ensure a consistent output from the pre-processing
- The choice of distribution of training data to cover all feature classes
- Creation of dataset statistics to confirm that all classes are covered

This process was built and run within the Feature Manipulation Engine (FME) which is a platform that streamlines the translation of spatial data between geometric and digital formats. It is intended especially for use with geographic information system (GIS), computer-aided design (CAD) and raster graphics software. FME was adapted specifically for this project.

Classification – Train and infer

Classification was performed over a two-week period. Classification covers the deep learning element of the project in which labelled data is used to train a model.

The model architecture used by Ordnance Survey is based on the popular semantic image segmentation U-Net algorithm (an end to end fully convolutional network) and has been customised to incorporate geographic data. A model was successfully created following the training process which uses the above architecture with training data. The overall accuracy for the model itself covering all predictions was at 78%. This is the validation accuracy we received following the model training and covers all classes.

Unseen imagery is then classified using the model via a process known as inference to produce the automated output. Each 1km² image was then processed in approximately 4 minutes. The specification created is covered in detail in the metadata section, but included is:

- Roofed Structures – Buildings deemed to be permanent (over 12m² in area)
- Water – Rivers, Lakes, Ponds.
- Natural Surface – Dirt, Gravel, Sand
- Sealed Surface – Paved, asphalt, concrete
- Roads – Roads if definable
- Railways
- Grass/Shrubland – Low/Medium Vegetation
- Tree/Forest - High Vegetation with visible canopy
- Other Structures – Buildings under construction, Walls, Jetties, etc
- Contour dataset at 5m intervals

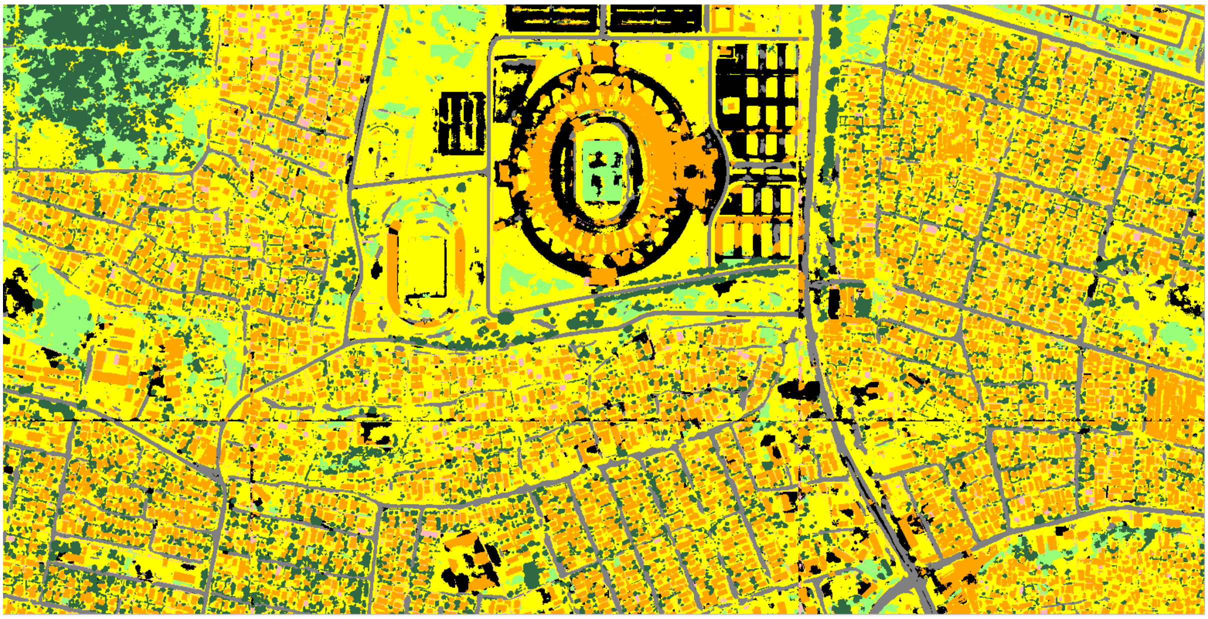

Figure 5 Raw deep learning classified output with overlap

Post Processing

Post processing adds geometric improvements to the classification output. These improvements include: Squaring of buildings, Smoothing of vegetation classification, and the automatic calculation of roads widths. Existing ESRI ArcGIS Pro coding was adapted for this use case.

As the nature of the road network in Lusaka does not lend itself to consistent automated results due to the number of unsealed roads which look identical to non-road features, it was decided to capture the majority of roads manually. The roads were captured at this stage.

Figure 6 Post processing pre-building edits

Manual editing

This process manually edits any discrepancies with the data that cannot be resolved within the automated process. Additional editing of roads, water, buildings, and trees were required. This is needed because classification and post processing will often introduce error which can't be resolved automatically. This requires an experienced operator to assess the resulting data and identify any significant deviations that have occurred and apply edits.

Cartographic Styling and creation of attribution.

Key information, such as the date of the aerial imagery used, the size of the building, and a unique GIS identifier were incorporated into the attribution of each building.

Two additional datasets were also incorporated:

- The first was part of a complimentary project on characterising informal settlements in Lusaka, conducted by Patrick Lamson-Hall and Valectus Ltd. This included a shape file in which residential settlements were automatically classified into eleven different

categories. Buildings could then be categorised as being located in ‘informal’ or ‘formal’ townships.
- Administrative boundary information of Lusaka was provided by the Zambian Lands authority.

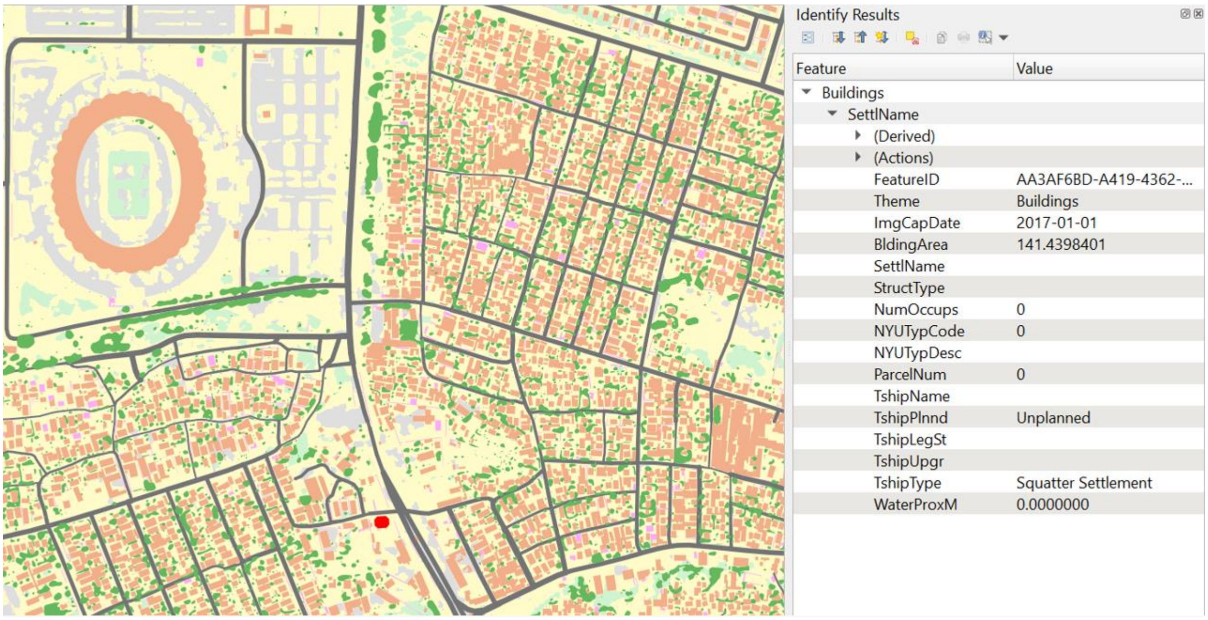

Figure 7 Post processed post building edit output (the atribute box details the more refined attribution in the building layer)

The information of the two datasets was applied to the attribution of the building class. Additional building attributes were created and deliberately left blank to await additional data after project delivery. Attribute themes included the following:

- Date of acquisition (Date of imagery capture)
- Type of structure
- Number of occupants
- Proximity to drinking water
- Different types of land Use / Informal Settlement – Typology
- Parcel number
- Name of informal settlement
- District / constituency / ward
- Declared / undeclared settlement

Results

Our automated process generated a new base map across 410 km² of Lusaka, in one tenth of the time that manual techniques would have taken. it took just 140 hours (approx. 6 days). Our goal was to prove how quickly a key policy need, in this case the identification of informal settlements could be directly met using automated repeatable machine learning approaches. Computers were taught to look for features in images using training data, which was applied to source imagery to recognise buildings, roads, water, and natural features. The technology then automatically created a map quickly and accurately, therefore proving that these features can be automatically extracted.

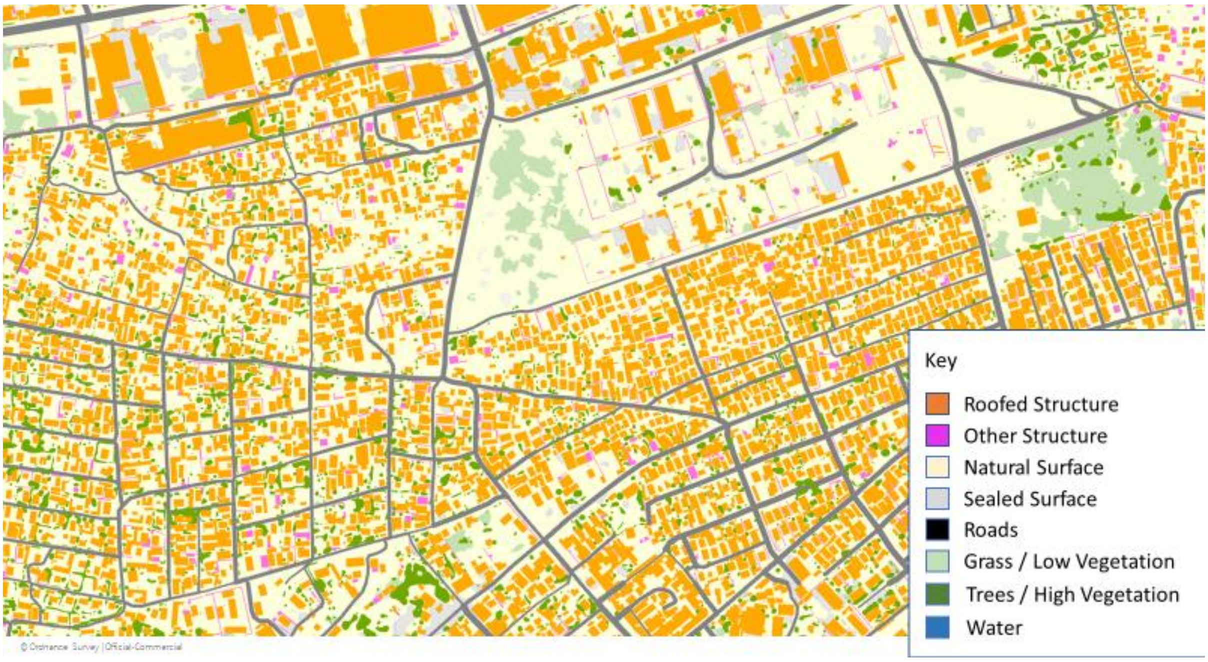

Figure 8 Final basemap

By leveraging the basemap created by Ordnance Survey, the Ministry of Local Government can undertake the following analyses:

- Identifying and locating informal settlements, their size, and the number of built structures
- How the formal and informal neighbourhoods are served by roads and public spaces
- Assessing the density of population based upon the number of buildings
- Prediction of where informal settlements may continue to grow, their potential capacity and extents

The map is also being assessed for use to support land registration and property taxation.

The map data now forms part of the National Spatial Data hub and is accessible by all parts of the Zambian government to assist with policy development as described above.

The basemap provides a robust foundation for the integration of future census data and for a range of other functions including land administration. Together, this will enable the Ministry to better target investment in critical infrastructure and services, upgrading informal settlements in order to provide for the most vulnerable city residents. It will also assist in better planning for urban expansion, which reduces the overall cost of infrastructure investment, limits informality, and enables more resilient and sustainable urban futures.

The pilot is working closely with local stakeholders, including the Ministry of Local Government and the Survey and Mapping department in Zambia, as well as Lusaka city council, to ensure it is tailored to meet their needs. It is also running alongside another IGC project with New York University's Marron Institute, which is analysing satellite data to create a typology of informal settlements using neighbourhood characteristics and building types. They are also utilising this information to better understand their density and expansion, as well as evaluate their access to water and sanitation infrastructure. The NYU project builds off previous work that we have done with the Commonwealth Association of Planners and the Commonwealth Association of Architects, and will provide useful training data and assist in the classification of settlements for the base map.

5. Conclusion

Our AFE solution is highly flexible and can be configured to apply to multiple use cases or applications including:

- **Government**: Census planning and analysis; improved cross-government working and e-government; monitoring progress against UN SDGs; critical infrastructure planning; provision of public services; revenue protection and identification of new revenue streams
- **Resilience and security:** Disaster response and preparedness, resilience planning
- **Commercial:** Effective construction planning; optimised mining operations; improved agricultural yields; telecoms networks planning and logistics, route optimisation
- **Smart cities:** Optimised roll out of IoT (sensor) technologies and automated vehicles
- **Sustainable urbanisation:** Settlement analysis; population and building density mapping; urbanisation monitoring
- **Environment monitoring:** Land use and land use change tracking; accurate deforestation monitoring
- **Land Governance**: Enhanced land administration
- **Insurance:** Risk analysis and modelling

As shown in the figure below, variations on the techniques deployed in Lusaka were trialed in other areas with different geographical issues but with equally impressive results.

Figure 9 Application of the AFE technique in different geographical settings.

Within Lusaka the key benefits of the Automated Map Creation are:

- Time: The automated process ensures that the volume of land parcels required can be captured and completed within the project timeframe.
- Consistency: The automated process ensures a consistent output that is machine driven. Under a purely manual process each human operator will make individual decisions in deriving the land parcels which will produce an inconsistent output. Process following a prescribed set of rules which will end up giving the same output. Data produced will be statistically similar because the same inputs and model are used.
- Quality: Process ensures the best quality. The machine-driven approach can be fully quality assured and quality controlled in near real-time, and adjustments to the process flow can be made as required with minimal effort, again saving time and thereby cost.
- Manual efficiency: The automated process and technology automatically identify the land parcels (10% of total) that cannot be captured automatically and will direct the manual operators to the specific area that requires manual completion.

However, the technology needs to be applied within a framework. By bringing together national and local government, academia, built environment professionals and the private sector, the Zambia mapping pilot provides an excellent example of a structured multi-stakeholder approach. A multi-stakeholder approach provides the necessary understanding from which tailored workflows can be created to provide key data to support evidence-based decision making. These collectively addresses social and policy challenges, while deploying a configurable and dynamic technology stack which is both scalable and replicable.

Rapid urbanisation, land use/land governance, sustainable cities and communities, food & water security, environmental monitoring, climate change, biodiversity and sdigital economic transformation are all complex challenges and can only be overcome by integrating policy agendas and spatial data. The governance of well-planned cities relies upon collaborative approaches. Collaboration is key between Geospatial Agencies, Cadastre Agencies, Ministries, Local Governments, Academia, Communities and the Private Sector. By developing a collaborative ecosystem the AFE tool can flexibly support the rapid delivery of critical spatial data & services required to underpin the socio-economic development of a country.

This case study has indicated that the AFE technology can deliver promising results. By judicially increasing the spatial, spectral, temporal and radiometric resolution of input data with appropriate training datasets then it is technologically possible to address a range of governance and policy challenges together with optimal implementation of evidence led critical public services.

REFERENCES

BIOGRAPHICAL NOTES

Mark TABOR,
Karen HUNT,
Andy WILSON,
Anthony BECK,
Juliet EZECHIE,

CONTACTS

Mr Mark TABOR
Ordnance Survey
Adanac Park
Southampton
SO16 0AS
mark.tabor@os.uk
www.os.uk/international

Karen HUNT
Ordnance Survey
Adanac Park
Southampton
SO16 0AS
karen.hunt@os.uk
www.os.uk/international

Dr Anthony BECK
Ordnance Survey
Adanac Park
Southampton
SO16 0AS
anthony.beck@os.uk
www.os.uk/international

Mr Andy WILSON
Ordnance Survey
Adanac Park
Southampton
SO16 0AS
andy.wilson@os.uk
www.os.uk/international

Juliet EZECHIE
Ordnance Survey
Adanac Park
Southampton
SO16 0AS
julliet.ezechie@os.uk
www.os.uk/international