Informal settlements – opportunities, risks and challenges of using new information technologies

Report of Online Workshop held on March 20, 2020

Summary

On March 20, 2020, the Austrian Foundation for Development Research (ÖFSE) in cooperation with the Helmholtz Centre for Environmental Research (UFZ), Stanford University, and the FUSE project, organized the online workshop: "Informal settlements – opportunities, risks and challenges of using new technologies." The workshop – made possible by funding from the Austrian Development Agency and Volkswagen Stiftung – was motivated by the assumption that many critical problems of rapid urbanisation can be addressed by recent technological innovations. Rapid urbanisation entails challenges for cities, particularly in informal settlements where human well-being and living conditions are poor. To address these challenges, novel remote sensing and digital technologies offer means to improve infrastructure and basic services through tools aimed at planning well functioning and sustainable urban spaces.

However, to make sure that the use of these technologies is beneficial for urban poor communities, a range of issues must be addressed. These include privacy, access to and ownership of data, potential for participation, and respective power relations of vested interests. Bearing in mind this constellation, the workshop looked into opportunities, risks and challenges of the use of new technologies in deprived areas.

The workshop brought together 20 researchers from both the technical and social sciences as well as representatives from NGOs. Participants came from Johannes Kepler University of Linz, Joanneum Research Graz, University of Salzburg, International Institute for Applied Systems Analysis (IIASA, Laxenburg), Technical University Munich (TUM), German Aerospace Center (DLR, Munich) and Helmholtz Centre for Environmental Research (UFZ, Leipzig) as well as from Medicins sans Frontières (MSF, Vienna).

In her keynote, Monika Kuffer from the University of Twente, Netherlands, one of the leading international experts in the field of earth observation (EO) and deprived areas, linked the debates about the need and collection of data with ethical challenges. Three comments by researchers from different fields complemented her keynote. Finally, the overall discussion confirmed the importance of bringing together different disciplines and perspectives. Participants exchanged technical as well as social and ethical considerations and their interrelations.

In the course of the discussion, participants identified the need for better coordination in order to harmonize data and to avoid parallel work. Additionally, the need was expressed for best practice examples, which allow others to learn and better understand the complex issues in this field. There was also agreement that the needs and wishes from the people affected must be understood in order to be able to put research at the service of a better life for their communities.

For further questions, please contact: Karin Küblböck
k.kueblboeck@oefse.at.

ÖFSE Austrian Foundation for Development Research

UFZ HELMHOLTZ Centre for Environmental Research

Program

10.00 AM	**Welcome and introduction of participants**
10.30 AM	**Keynote: Monika Kuffer, University of Twente**
11.30 AM	**Comments:** Ian McCallum, IIASA: Citizen Science applications Daniel Klotz, Kepler University Linz: The role of EO/AI for hydrologic modelling Manuela Hirschmugl, Joanneum Research: EO for Sustainable Development
12.00 PM	**Discussion**
13.00 PM	**End of Workshop**

Speakers:

Monika Kuffer, Assistant Professor at the Faculty of Geo-Information Science and Earth Observation (ITC) of the University of Twente. Her main research focus is on urban remote sensing and modelling linked to the global development agenda (e.g., SDGs). She is co-chairing of a large multi-disciplinary network of researchers and experts across the Global North and South on urban deprivation related research topics

Ian McCallum, International Institute for Applied Systems analysis (IIASA). Ian has a background in terrestrial environmental monitoring, employing techniques including earth observation and citizen science to address issues around land use, environmental degradation, natural disasters and inequality.

Daniel Klotz, Kepler University Linz, Institute for Machine Learning. Daniel has a background in environmental engineering and works on rainfall-runoff modelling (focusing on floods and water management). He is currently developing AI based simulation models with the long-term goal to use them to improve decision making.

Manuela Hirschmugl, senior researcher at Joanneum Research in Graz. She has a background in environmental sciences and remote sensing and has been working with satellite data analysis in different fields of application from vegetation monitoring to risk mapping. For the past three years, she participated in the EO4SD-urban project of the European Space Agency (ESA).

Facilitation and Documentation:

Karin Küblböck, Ines Omann, Hannes Grohs (ÖFSE)

VolkswagenStiftung

With funding from
Austrian Development Cooperation

A. Keynote: Monika Kuffer – Deprived urban areas: What role can digital technologies play to improve living conditions?

In her keynote, Monika Kuffer linked debates concerning the need and collection of data with ethical challenges within the field of earth observation. The following questions guided her presentation:

What role can digital technologies play to improve living conditions informal settlements/deprived urban areas?

What are the scope and limits of present technological developments? Is current research asking the right questions?

To answer these questions Monika Kuffer divided her input into 4 parts:

1. Introduction – defining urban deprivation
2. Why are data / and which data on the location of deprived areas needed?
   - What role can digital technologies play to improve living conditions?
3. Examples of EO mapping approaches
   - What are the scope and limits of present technological developments?
   - Is current research asking the right questions?
4. What are the main ethical challenges for mapping deprived areas?

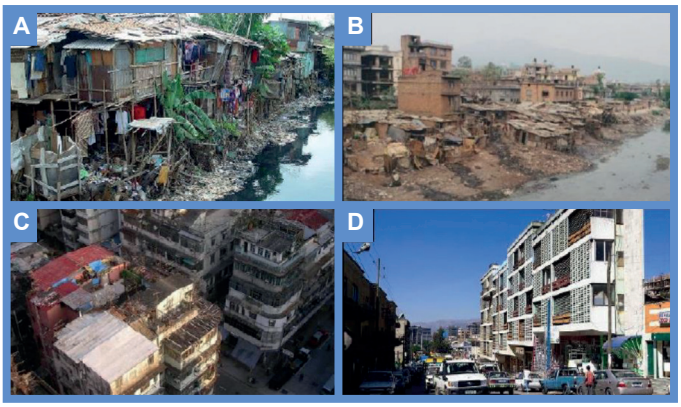
1. Defining urban deprivation

In the first part of her talk, Monika introduced the participants to the complexity of defining slum households and slums/deprived areas, and the consequences of those definitions for mapping exercises. The keynote's starting point was her observation that endeavours to choose the right terms and to define slums/deprived areas often entail more questions than answers. However, different terms and definitions have far-reaching implications on how to map these areas. Often, there are no clear-cut boundaries for defining slum and non-slum areas. UN Habitat produced a household-based **definition of slums** that includes five dimensions: water, sanitation, tenure security, overcrowding, endurable housing. This definition differs from an area-based definition of deprivation. When people live in deprived areas, they often face additional risks (e.g., flooding, health, security).

Definition of household deprivation versus area based deprivation

- A deprived "slum" neighbourhood reflects social, environmental, and ecological risk factors to health and wellbeing above and beyond household and individual characteristics.
- A "slum household" reflects household poverty risk factors to individual health and wellbeing.

Monika illustrated her definitions with examples (Picture 1): The picture shows slum households in deprived (A) and in not-deprived neighbourhoods (C), and of non-slum households in deprived (B) and in not-deprived neighbourhoods (D)



Picture 1: The complexity of defining deprivation[1].
Source: Thomson et al (2020)[2], in: Monika Kuffer, slide 5

1 Image credits:
A – Jonathan McIntosh – Own work, CC BY 2.0, https://commons.wikimedia.org/w/index.php?curid=53838
B – Peter van der Sluijs – Own work, CC BY-SA 3.0, https://comons.wikimedia.org/w/index.php?curid=18905096
C – J. Carmichael – Own work, Public Domain, https://comons.wikimedia.org/w/index.php?curid=6441162
D – The Chroniclers – Own work, CC BY-SA 4.0, https://comons.wikimedia.org/w/index.php?curid=51369584

2 Thomson, D. et al (2020): Critical Commentary: Need for an Integrated Deprived Area "Slum" Mapping System (IDeAMapS) in LMICs, available as preprint: https://www.preprints.org/manuscript/201910.0242/v2

As a next step, Monika introduced the participants to the definition of deprivation:

Deprivation is a multi-dimensional phenomenon, covering a wide range of socio-economic dimensions, including:

- Physical deprivation (e.g., quality of buildings)
- Environmental deprivation (e.g., heat islands)
- Security deprivation (e.g., no lighting)
- Health deprivation (e.g., no access to clean water)
- Services and access deprivation (e.g., lack of water and sanitation)

Taking into account different terms and definitions, it becomes obvious that a clear-cut demarcation of slums/ deprived areas is complex.

Criteria/parameters used for mapping slums are objects (roofs, sizes of houses, types of roads), settlements (measured by densities, irregular patterns), environment (e.g., settlements built on a slope). These parameters are also used to train models as part of machine learning (slide 11).

There are different visions/approaches used to map deprivation:

- Household level mapping (via census data and surveys) aggregated to an area level: if there is a large number of deprived households in one area, it is defined as a slum/deprived neighbourhood.
- Field-based mapping, using definitions based on the local context, amongst others done by Slum Dwellers International and also governments (e.g., government of Indonesia)
- Human visual interpretation (interpreting images – what features make a slum area different in an image)
- Machine imagery classification including via machine learning

2. Why do we need data on deprived areas?

Monika first started with a brainstorming exercise among the participants, asking why are data needed on deprived areas?

Why are data needed on deprived areas (informal areas/slums)?

To know the number of people in need, and the nature of their needs	To accurately document the areas for agencies, governments, etc.	for measuring population, for understanding ethnic distribution, for measuring health standards
for improved assessing need/scale of areas and for planning	provision of services (health eg. vacination, water, ...)	to make deprivation visible, to plan infrastructure to improve access to water, sanitation, electricity
To know about population density and their needs	better planning, improving services for the affected people	Because there is no formal planning so formal maps are difficult to draw
data help for planning, for instance for number of toilets, wastewater systems, how much water is needed, which form of infrastructure	health service epidemiology vector control potential land rights	official data not transparent

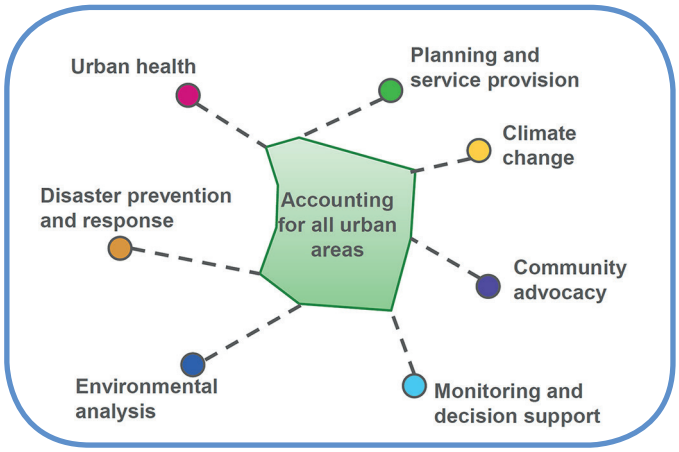
Picture 2: Results of Mentimeter brainstorming

Question from the audience. How good is the quality of Google Street View maps?

Answer: It varies, for example in Dar es Salam, Tanzania the map of high quality, because of a World Bank project that has supported mapping. However, in n most cites (e.g., Indian cities) you typically see boundaries, but individual buildings are not mapped. Even in cities where buildings are generally mapped individually, in informal settlements, this is not the case.

After the brainstorming session, Monika presented her view of the reasons why data are needed on deprived areas: Data on deprived areas are often inaccurate as the spatial dimension is missing. For example, there are no urban poverty maps. Population estimates are very uncertain (often underreported), and existing methods lack transferability and scalability. In Bangalore/India the "Slum board" recognised around 600 slums, while a local survey (based on images and ground-based mapping) mapped 1.500 slums (slide 17). A large part of the population in deprived areas is very vulnerable to climate change, e.g., when areas are below sea level. Underestimation of people living in those areas is therefore extremely problematic, to say the least.

Earth Observation can be used to obtain more accurate data, but there is a lack of scalability and transferability – methods are typically developed for one city/area, and cannot be transferred easily to other places. However, accurate data are essential for many domains, such as planning for urban health, disaster prevention, planning and service provision (see picture 3)



Picture 3: examples of domains for which data on deprived areas is needed, in: Monika Kuffer, slide 18

Monika mentioned some **existing initiatives/efforts** to improve data on deprived areas:

Amongst others, the UN is trying to fill data gaps, e.g., there is an initiative coordinated by **UN-Habitat**: Building the Climate Resilience of the Urban Poor (slide 19) – because of lacking data, it was difficult to establish a list of hotspots of urban poor areas most vulnerable to climate change.

Community based data. The NGO **Slum Dwellers International** (SDI) is producing community based maps in many cities. However, this only covers the cities where SDI is working in. http://knowyourcity.info/

Million neighbourhood project: – data production, also in co-operation with e.g., slum dwellers international https://millionneighborhoods.org

She stressed the importance of combining earth observation with community based mapping in order to scale it up to produce global data sets. She also pointed to ethical questions in mapping processes. As an example, Monika mentioned the mapping of Favelas in Brazil: the government started to collect data on slums and later used it to clear and demolish these areas with a view to the 2014 Football World Cup. The answer of the community was twofold: on one hand, they protested against the government mapping initiative, on the other hand, they engaged in their own mapping exercise, with the approach to not only map negative aspects – but also to map the assets that deprived communities have; eventually the community could prevent the threat of their slums being demolished.

3. Examples of Earth Observation (EO) mapping approaches – Scopes and limits

Monika started the next part with a brainstorming session on the main methodological challenges for EO mapping approaches/machine learning.

What do you see as the main methodolocial challenges?

lack of communication between earth observation experts and slum communities. lack of computer literacy of slum dwellers communities to critique data. lack of participatory systems

harmonizing disparate datasets

integration of EO analyses (only some aspects visible) with field/community-based/ socio-economic data

validation

population estimates, Quality assurance especially in dynamic evolution

comunication and coordination

Sometimes unknown visible indicators for deprived areas

Picture 4: Results of mentimeter brainstorming

Methodological challenges

After the brainstorming session, Monika presented a list of major challenges (slides 22ff):

1. **Scalability** – The main problem is consistency, data are often not spatially well recorded, and datasets often miss the spatial dimensions. One big problem of working with EO is the cost of high resolution images (they cost 20€/km2). One can explore whether cheaper, lower resolution images are suitable.
2. **Transferability** (spatial, sensor, temporal) is often not discussed, but resolving the issues involved is very important for producing global and multi-temporal data on deprived areas.
3. **Understanding the local context:** The main problem is that many researchers have good data access, but they often do not work with ground-truth data. So the EO community trains the algorithms with data but without having ground-truth data (e.g., historical centres in Indian cities can look similar to slum areas from EO (narrow streets, etc.)).
4. **Mapping has to go beyond physical characteristics and beyond binary maps** (slum/non-slum) – it makes more sense to look at degrees of deprivation. How to train models via deep learning to recognise different degrees of deprivation – can we predict variations of deprivation between settlements?
5. **Map Validation** is often difficult via image interpretation or expert judgements. Example: In an Indonesia study, five experts judged slum presence, yet only in one area of the city did they agreed that this is a slum.
6. **Privacy concerns + ethics:** see next part

4. What are the main ethical challenges for mapping deprived areas?

At the end of her talk, Monika discussed a range of ethical considerations, related amongst others to questions of data access, privacy issues, and the use of data.

One question in this context is how spatially detailed should deprived area, publicly available, maps and images be? What should be displayed on those maps? There are also local differences in what people feel should be displayed. These questions are even more relevant in the future, as the quality of images, computational power and algorithms are constantly improving. What does this mean for mapping slums? What are the downsides? Maybe it is sufficient to use fuzzy maps that do not show the exact boundaries, because of the danger of violating privacy of inhabitants.

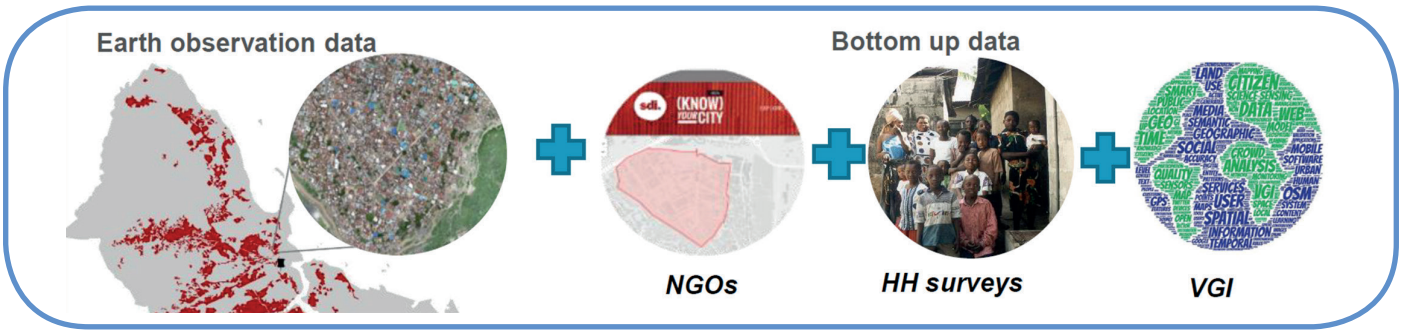
Besides privacy issues, there is the question of the benefits of the use of maps. In principle, maps should be produced for the benefit of communities, but they can also be used against communities (see the above-mentioned example of Favela-mapping in Brazil). Monika raised the question of the responsibility of researchers: What can they do to protect communities, to avoid that maps produced are used against them?

Another aspect is the question of benefits from mapping exercises for inhabitants of deprived areas. They often feel that they are used to obtain data, but they do not see any consequent positive changes or improvements, which causes frustration.

> ***"When you participate in so many researches and you don't know what is happening [as a result] and it's not changing the environment that we are in, you feel like wasted. You keep on asking questions, but this data will go where?***
>
> ***The one who takes data will never come back to us like 'okay, we took this and these are the results.' So you feel wasted".***

Picture 5: Quote on benefits for inhabitants. Source: Nicera Wanjiru, Slum Dwellers International Federation, in: Monika Kuffer, slide 35

Monika also addressed the ability of EO data to **estimate the population of deprived areas**. Often, bottom up and top down population estimates reveal huge differences, as algorithms fail in areas with very high density in the Global South; e.g., the top down approach to estimate the population in Dar es Salam resulted in 3 million inhabitants, the bottom up estimation resulted in 5 million. Underestimation of the population can have severe consequences, e.g., for water provision or for estimating the number of people exposed to flooding. Different methods should be combined to get more accurate data estimates.



Picture 6: Top down versus bottom up approaches of population estimates: EO data, data from NGOs, household surveys, volunteered geographic information, in: Monika Kuffer, slide 37

5. New projects and final thoughts

At the end of her talk Monika presented some new projects in which she is involved:

The project IDeaMaps (Integrated Deprivation Area Mapping) is a project that combines crowd wisdom, space technology and different methods. It aims to show the degree of deprivation.

The Slumap project (Remote Sensing for Slum Mapping and Characterization in sub-Saharan African Cities) aims to provide an open-source framework that allows for the processing of remote sensing images to provide information on the location and extension of slums and to characterize the physical environment within slums at limited cost.

Finally, Monika presented some thoughts on how to advance with mapping deprived areas.

She stated that Sustainable Development Goals indicators related to slums need a spatial dimension in order to better estimate, for instance, vulnerability to climate change. At the same time, more data are needed for local planning purposes. Localised data could be used at aggregate levels, and employ the dimension of probability to estimate degrees of deprivation in different areas, while also communicating the uncertainties involved.

Question from the audience.

- ***Are there examples of benefits of those maps for the population? Did urban planners ever use research results / data to change something in their decisions?***

Answer:

> *Slum inhabitants want to be seen, they are often ignored, so maps help them to become visible, to take account for areas, to include them in service provision.*
> *In a project in Dar es Salaam, titles were given to inhabitants in the process of a mapping exercise.*
> *Slum dweller international has used it maps for court cases – for people to have evidence that they have been living in this area, so they cannot be expelled.*
> *It would actually be a good idea to build up archives with such cases.*

- ***Who should have access to data and how to avoid misuse?***

Answer:

> *This is an open question; in any case, data should be accessible for local communities. In the next years so much data is going to be produced – researcher have also a responsibility of the data collection process and at least to empower communities to know that they see the data and are aware of what data is produced about them.*

- ***To what extent do you use crowdsourcing data?***

Answer:

> *That's the way forward, but how to do this is still to be developed.*

B. Comments

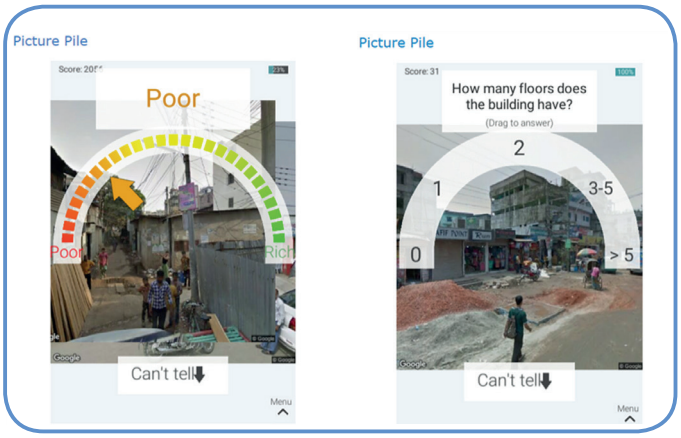
Comment 1: Ian McCallum, IIASA – Citizen Science applications

Within IIASA, the Centre for Earth Observation and Citizen Science is working on the merger of these two disciplines. The original motivation was to provide better data and to reduce the uncertainty around the global data sets that they are feeding into with their modelling groups. For example, they have a large land-use model, so they are interested in land cover, and biophysical datasets that go into these models. Over time they have also moved into socio-economic aspects, e.g., inequality, which is related to today's discussion.

A decade ago, they launched the **Geo-Wiki** to produce better environmental data, and to derive more information about various socio-economic aspects (http://geo-wiki.org).

Another tool that they developed is **Picture Pile**, which consists of feeding in piles of images into databases, typically satellite images, and providing this data to the crowd, combined with some training and asking people to classify these images. They also increasingly use Google Street View, and an open dataset called mapilary. At the beginning, they asked more binary questions, this has now become more elaborated.

Related to slums/deprived areas, they show e.g., snapshots of imagery to people who then help to classify those images, for instance to estimate the degree of poverty in an area or the number of floors of buildings. This can be used for physical and socio-economic mapping exercises. The data can also be used to develop new data sets, or for machine learning.



Picture 7 and 8: examples from Picture Pile, in: McCallum, slide 6

Question from the audience: **How to validate the data that you receive from the people?**

Answer: We use several techniques, including showing the same image to many different people to determine the crowd agreement and asking individuals to classify some images multiple times to infer consistency.

Comment 2: Daniel Klotz, Kepler University Linz: The role of EO/AI for hydrologic modelling

Daniel works at the Institute for Machine Learning at JKU Linz. The main focus of the institute is to develop new methods for **machine learning**. Given the topic of the workshop, he focused his talk on applications. His institute is building simulation models, for traffic or, more relevant to this workshop, **environmental simulations** (e.g., rainfall-runoff simulation that describe the generation of streamflow from meteorological data)

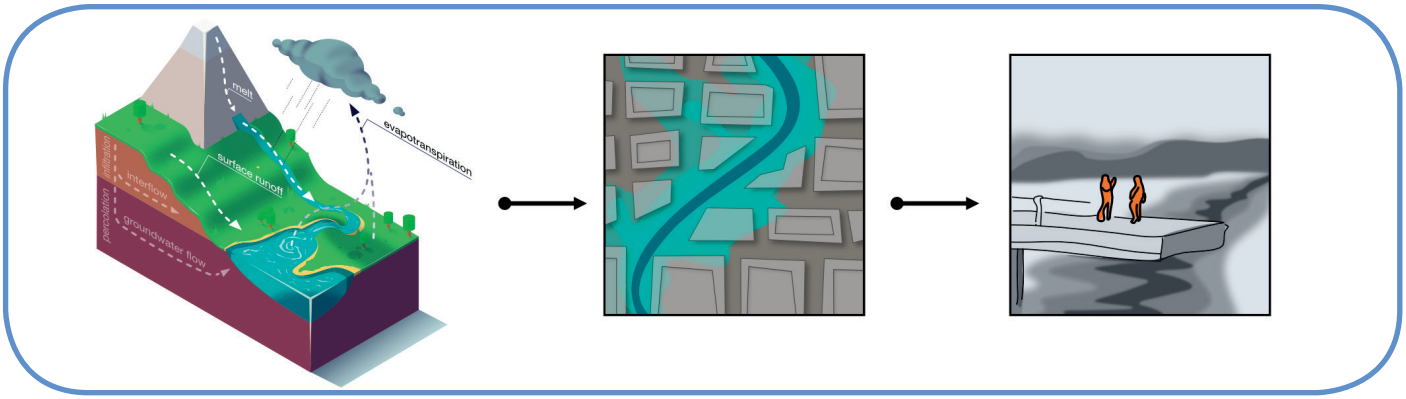
At the beginning of his comment, Daniel raised the issue of transferability of models: Statistical learning theory assumes that the models are trained and applied for independent samples of the same distribution. Thus, in theory those models are not suited for use outside of the domain of the observed data (keywords for this context are out-of-sample-testing, feature-drift, and concept-drift).

To remedy this, his group trained a single model with a large set of rivers and provided it with static catchment attributes (such as percentage of sand in the soil, or long-term annual precipitation) to learn the underlying process from the given data. They used the CAMELS dataset. It contains more than 600 basins within the contiguous United States, with almost no anthropogenic influences. So, while traditional models are usually only fed with meteorological variables, the machine learning model also has context information.

The reason for this particular setup is to be able to trade "space and time", i.e., gather information that appears seldom in time at one place – say a flood that occurs only every 100 years – by distributing learning processes to a larger number of basins.

This has lot of advantages. Most importantly it allows training the models in regions where lots of data are available and transferring them to regions where almost no data are available. However, it also increases the chance that the model can adapt to climate change. The intuition behind this argument is that the model can use information from one place to compensate for the changes that occur over time. For instance, let's say that the climate conditions in Austria become similar to those of Spain, then – given that the model has seen rivers in Spain – it can simulate the Austrian rivers by treating them as if they were in Spain. Thus, an extrapolation in time would become an interpolation in space. How good this approach works and how far one can go with it remains an open research question.

A potential application for these models (and also for traditional rainfall-runoff models), which is also very relevant for deprived areas since they are often vulnerable to flooding, is the creation of inundation maps. In this context, an important question becomes how to create a good flood map and how to communicate it to people. From experiences of colleagues and other groups, we know that one problem is that deprived areas change very quickly and this also changes the estimation, and the potential impacts of flooding. It is also very difficult to get ground-truth data, e.g., it is difficult to get the real flood extent especially in deprived areas. Furthermore, it can be extremely challenging to find appropriate communication avenues to provide information about the potential extent of floods and the uncertainties of flood forecasts.



Picture 9: from models to communications, in: Klotz, slide 4

Question from the audience: How much data do you need to train the model?

Answer: There is a great paper by colleagues from Canada (Gauch et al. 2019[4]) which examines the question of the amount of needed data. Generally one can say "The less data you use, the worse you get". Roughly speaking, If you use few years of data – I think the actual value was around 3 years – then the traditional methods and other machine learning approaches outperform the LSTM (Long Short-Term Memory). However, once you get more data (in space and time) the LSTM outperforms everything. We can train the model on one river, then we need more time, or we need more spatial extension and less years. With one river, you roughly need 5-10 years to get a good model out of it, if you have more rivers than you need less time. It does however also depend on model specifications, so it might be possible to find very small models that already work fine for small amount of data; or it might be possible to pretrain the models with simulations or such...

4 Gauch, M./Mai, J./Lin, J. (2019): The Proper Care and Feeding of CAMELS: How Limited Training Data Affects Streamflow Prediction. https://arxiv.org/abs/1911.07249

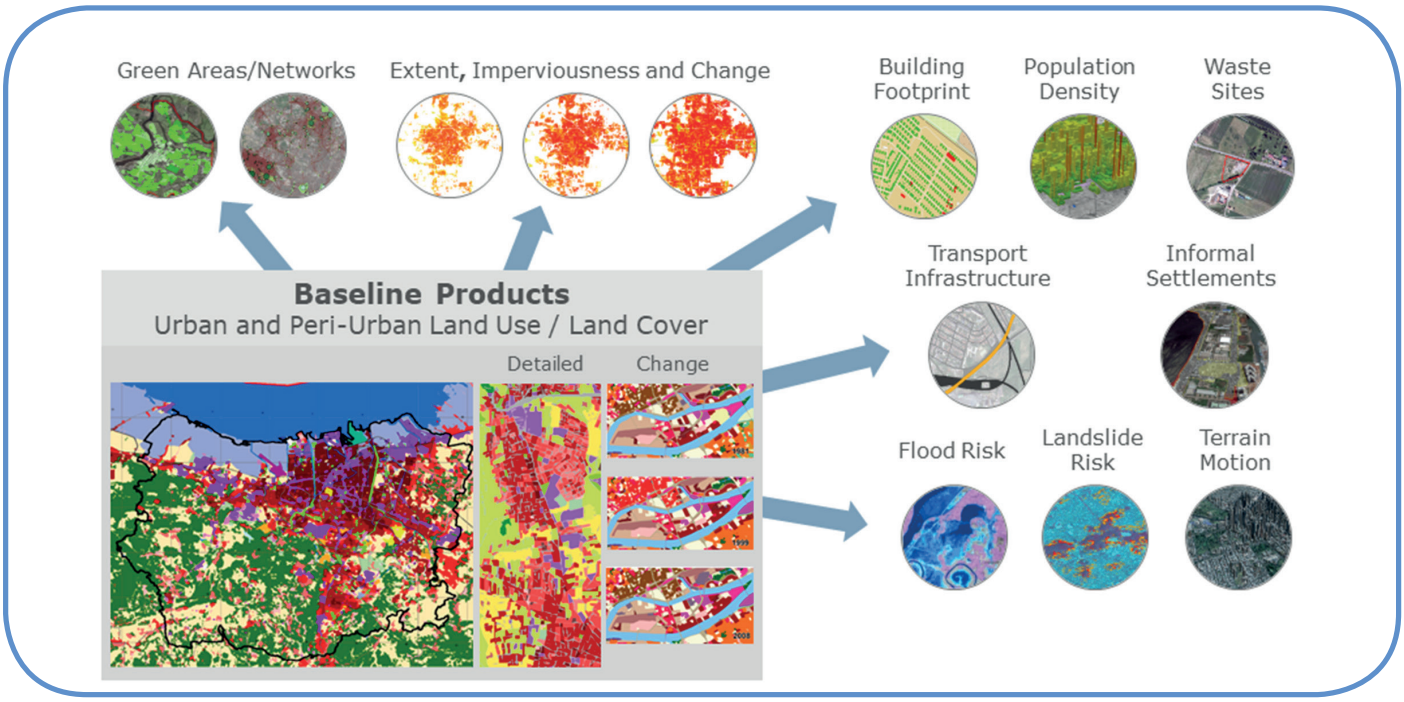
Comment 3: Manuela Hirschmugl, Joanneum Research, Graz: Earth Observation for Sustainable Development

Manuela shared details of the project "Earth Observation for Sustainable Development-Urban" (**EO4SD-Urban**). It is a project of the European Space Agency (ESA) and was initiated in 2016. In this project, ESA provides EO data to support the Sustainable Development Goals (SDGs). The project is implemented in cooperation with the World Bank and the Asian Development Bank. ESA has been cooperating with the International Finance Institutions (IFIs) since 2008 by providing EO information to support their development projects.

The objectives of the project are to improve understanding of EO applications for urban development, and to mainstream EO applications in an operational manner into development programmes.

Background: Urban growth presents many challenges. Such challenges include, infrastructure and access to water. All these issues are embedded in several SDGs. EO-4SD tries to tackle those challenges using EO techniques.

EO4SD works on 38 cities around the globe. For all cities, baseline products were established. These included urban and peri-urban land use and land cover, including changes over time (usually 3 times), and additional products, such as green areas, building footprints, population density, infrastructure, informal settlements, flood risk, and land slide risk.



Picture 10: EO Products for Urban Development, in: Hirschmugl, slide 5

They use the information for urban master plans, infrastructure planning, to improve/green urban neighbourhoods, density estimates, risk assessments for flood risk, or providing data to monitor of urban indicators (e.g., SDG 11). Manuela gave examples about the use of their data in different regions, e.g., for urban planning in Arusha, Tanzania with rapid land-use changes, for risk and disaster assessment in Indonesia, where land is subsiding in many areas, or for delineating slum areas in Kolkata, India.

They have also engaged in capacity building – ESA has a Youtube channel, and conducts live webinars.

One important conclusion that she underlined regards accuracy of data and the importance of people trusting data. She has seen a lot of maps, and with AI there are a lot maps coming. It is critical that people do not lose trust in the data.

In closing, she also noted that crowd sourcing can be important if one gets enough data – but there could also be a psychological bias; people could be interested in making their situation better or worse depending, on what they are asked. This is an open discussion item for her.

C. Discussion: "Basically we are talking about people and not about houses"

In order to lead a fruitful discussion, Ines Omann once again recalled the key questions to which the workshop was devoted:

How can new technologies be designed and regulated to

1. support inclusive urban development,
2. improve living conditions of inhabitants of informal settlements,
3. empower urban citizens to shape their future, and
4. avoid the abuse of data and the input from inhabitants?

"We are not the easiest people to work together with"

In the beginning, the opinion was raised that in order to support inclusive urban development, it is crucial to include municipalities and NGOs in the research process. Experiences from various participants show that collaboration with NGOs (such as MSF, International Committee of the Red Cross (ICRC), Slum Dwellers International, Action Against Hunger (ACF), and OXFAM) works quite well. NGOs seem to be interested in this type of work and at the same time, they already have and are developing many skills in-house. From the perspective of MSF, this assessment was confirmed: So far, cooperation between the organization and the scientific community appeared to be enriching for both sides, even though organizations like MSF might not be the easiest partner to work with. The latter is caused, on the one hand, by NGOs' mandate, which foremost obliges them to their clients, and, on the other hand, by the simple fact that things on the ground tend to change literally overnight. However, NGOs increasingly acknowledge that data and the knowledge about it can positively influence their work and NGO workers on the ground learn to ask for data specifically and to collect it.

"Weird mix of interests at play"

In contrast to NGOs, participants shared the experience that getting municipalities on board, is a lot more challenging. They tend to be quite sceptical about projects, in particular, when it comes to crowd-sourcing activities and citizen's participation. The impression was shared that municipalities often seem to be afraid to empower inhabitants through such activities, which could in turn result in questioning the action of authorities. In this regard, one participant pointed towards the highly political character of informal settlements. The participant reported from field research experiences, in which informal settlements emerged as places with a "weird mix of interests at play". Improving living conditions only in small steps might by a political strategy in the search of votes. The participant concluded by emphasizing the importance of a profound understanding of local government structures in order to decode such power dynamics.

"Get to know their visions – it is not only about budget"

The participants showed a broad consensus that, if research is truly aimed at empowering people and communities in deprived urban areas, it is crucial to know and understand their needs and visions. In connection with new technologies, this additionally means getting an idea of locals' worries and concerns about the use of those technologies, as one participant noted. Only in this way, can researchers – by showing later how they dealt with these concerns – ensure that people trust the data and the results. In this context, participants agreed that community empowerment is only possible through direct engagement with communities and that it must be part of a research project from the very beginning.

The question of how best to implement stakeholder/community engagement remains open. However, it is clear that money is needed for this. Research projects on local topics should always provide a budget for community engagement. This is something that is not being done sufficiently at present and is due, among other things, to the structuring of research funds, which still place too little emphasis on participatory approaches. Nevertheless, some experiences have been exchanged with projects that included participatory and interdisciplinary elements such as Sustainable Living Labs and training with partners. In the opinion of one participant, training could and should have two objectives: first, to ensure that people see changes and positive effects of the use of new technologies for themselves, and second, to help these people gain the necessary knowledge about new technology to be able to negotiate their use vis-à-vis municipalities and local authorities. However, a concern was also raised,

which called into question the representativeness of stakeholder engagement processes. The participant argued for a careful combination of top-down and bottom-up approaches. Another participant added that s/he can say from her/his experience that the machine learning community is highly interested in questions concerning the common good, climate and ethics, but that the community struggles with framing the problems right and that interdisciplinary exchange is appreciated.

"Do no harm and if you think there might be harm done, just don't do it"

Finally, the fourth aspect of the workshop's key question – how to make sure that the abuse of data to the disadvantage of inhabitants is avoided – led to a somehow controversial debate about cost-benefit considerations: How to deal with the fact that high resolution data can be very beneficial on the one hand, but also be easily misused on the other? And who is in the position to judge, what kind of data is used or not? In this regard, one participant pointed to the discussion concerning satellite images. About ten years ago, there was a big controversy about who should have access to these images. Nowadays, nearly everybody can access high-resolution satellite information via Google Maps – for the good and the bad. Other participants, however, agreed that this does not mean that there is no need to reflect on these topics again and again. One suggested that it is necessary to evaluate such considerations from case to case and to not only look at the amount of costs and benefits, but on the severity of possible violations. There might be many benefits, but as long as there is a fundamental right, which is violated, one should refrain from any action. The concerns of another participant were not so much directed towards the availability of high-resolution data, but to the labelling process conducted by researchers. S/he pointed to the (unintended) consequences, with which the labelling of an area as a slum is intertwined. Such consequences span from the social stigmatization of inhabitants to devaluation processes of surrounding areas. There was agreement among participants that such labelling processes – if necessary at all – have to be conducted carefully and that they tend to undermine the temporal process of urban development. On the other hand, the opinion was raised that sometimes it is not possible to avoid labels (e.g., flood risk), even though errors might occur. Labels might be applied by using degrees and in terms of likelihoods.

During the overall discussion, participants articulated two specific needs with regard to their research. First, they called for better coordination among research efforts. A lot of work is done in parallel and greater transparency and harmonization of data would be highly appreciated. Second, participants appeared to be very interested in best practice examples of projects, which managed to design their research towards the end of local communities. Such best practices would not only inform participants' research. They would also serve as low-threshold tools in future projects to communicate the project's potential gain to communities.

"Basically we are talking about people and not about houses"

When participants then eventually shared their workshop's key takeaways in a final round, their wish to connect technical with social and ethical questions were once again clearly articulated. This is well reflected in the statement that, even though earth observation and GIS are used to map physical things, this process is never disconnected from social questions. In the face of unintended consequences of labelling processes, for example, it is not enough to make assumptions about these connections and then extrapolate poverty levels. But it is necessary to really prove and scrutinize the linkages between the physical/technical and social realms. That this endeavour is complex, is obvious. The call to let research be guided by the needs of people on the ground seems to reduce this complexity somehow. However, it should not be overlooked that those needs and interests are heterogeneous, perhaps in very important and even contradictory ways.

Imprint: ÖFSE, Sensengasse 3, 1090 Vienna, Austria. Design: Alexandra Erös, ÖFSE

This work was conducted in part as a component of a planning grant of the initiative "Artificial Intelligence and the Society of the Future" of the Volkswagen foundation and of the Belmont Forum Sustainable Urbanisation Global Initiative (SUGI) Food-Water-Energy Nexus theme for which coordination was supported by the US National Science Foundation under grant ICER/EAR-1829999 to Stanford University. Any opinions, findings, and conclusions or recommendations expressed in this material do not necessarily reflect the views of the funding organizations.