IGC International Growth Centre

Mapping Rural Pakistan: Bottlenecks and Solutions

Sohaib Khan, Amal Aslam, Shehryar Ahmad, Ali Rehan, Ahmad Gul, Usman Alam, Hira Naqvi, Zuhair Bukhari, Imran Iqbal, Khalid Sherdil

This policy brief draws upon the experiences of a major study of mapping demographic and social parameters in District Jhelum, led by Sohaib Khan in partnership with SUPARCO, and the challenges of coordinating relief efforts in Punjab in the aftermath of the 2010 floods.
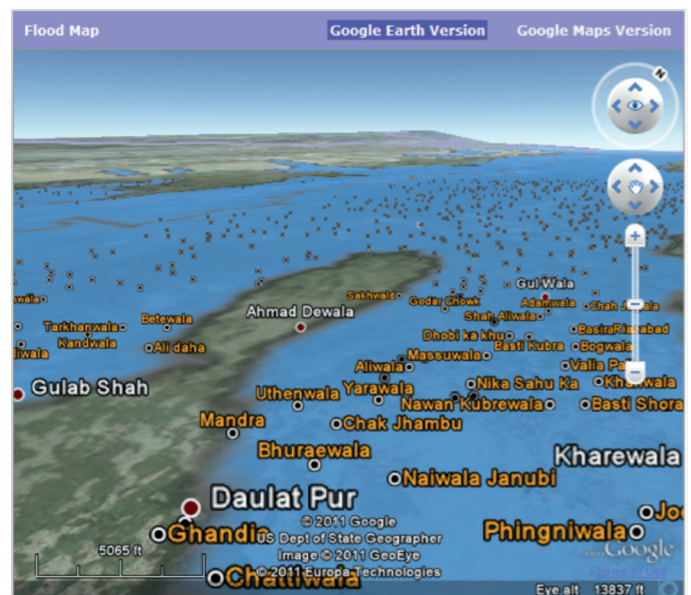

Figure 1: The Floodmaps website, funded by IGC-Pakistan Rapid Response Grant, publically made available the coordinates and names of 9000+ settlements in the flood affected areas, along with the extents of the floods, in an interactive and searchable manner, as an overlay on Google EarthTM.

The Unfulfilled Promise: Bottlenecks in GIS Mapping of Social and Demographic Parameters

Tabular Social and Demographic Datasets

When it comes to generating a visual representation of important social, demographic or economic parameters for rural Pakistan, such as poverty, literacy, crime or access to the internet, the picture is rather blank. Datasets exist in disparate jurisdictions, scattered across many departments, and are often in registers and lists that do not contain geo-coordinates. This discourages evidence-based policy interventions, and severely limits the effectiveness of expensive GIS projects undertaken by many organizations.

As an example, consider the population census, which was last undertaken in 1998, and contains data on many important indicators such as availability of electricity and drinking water, and literacy ratios by gender and age distributions. This dataset is organized for rural areas in a tabular format, indexed by mauza[1]. Despite the criticality of such data, and the fact that more than twelve years have passed since the census, detailed spatial maps do not exist that depict administrative boundaries at the mauza-level, on which these important demographic parameters can be visualized. At best, the information can be seen in some atlases at tehsil or district level, which often is too coarse to be practically useful, and hides the significant non-homogeneities within the large area of a tehsil. This picture is not specific to the

www.theigc.org

Directed and Organised by

LSE · UNIVERSITY OF OXFORD

population census data. It also applies to other datasets, such as the agricultural census, crop reports, and datasets related to schools and health.

While many policy makers and planners would attest to the lack of availability of quality datasets, we believe that whatever datasets exist are heavily underutilized because of the inability to visualize them effectively through spatial mapping. Spatial inferences are impossible in the tabular format that is typical to these datasets, because the neighborhood relations cannot be captured in this format. Spatial planning tasks, for example, where to locate the next school, require creating an overlay of interdependent variables, eg, the population spread, accessibility, existing school locations, and pockets of low literacy, in a wide neighborhood, an exercise in futility if the data is limited to its tabular form

Mapping Needs in 2010 Floods

The recent floods are a stark example of the failings of GIS technologies in Pakistan. While demographic and social mapping is needed for many planning tasks, its criticality is brought to focus against the backdrop of disaster management. It is understood that efficient, streamlined and coordinated disaster response 'starts with a map' (National Research Council 2007), and that mapping challenges are dependent on the socio-economic context. For example, the approach needed for managing the Katrina hurricane must necessarily be different from that required for the Kashmir earthquake (Nourbaksh et al. 2007).

In response to the 2010 floods, remote sensing analysts working at SUPARCO[2] and UNOSAT were able to successfully mark the approximate extents of inundated areas through satellite imagery. However to ensure the establishment of entitlements, coordinate relief and ensure that no village is left behind, it is important that detailed demographic spread is available to the coordinating agency. With this goal in mind, LUMS and SUPARCO, through IGC Rapid Response Funding, put together a website that publically made available the locations and names of more than 9000 settlements in the flood affected area, along with dated updates of the flood extent, search capability and printable map options (http://floodmaps.lums.edu.pk/flood-map). While many NGOs and government agencies made use of this data, with the Punjab Government's disaster management portal making extensive use of it (http://floodrelief.punjab.gov.pk), we also learnt the critical shortcomings of this mapping to derive lessons from more effective use of such exercises in future.

One of the primary customers of such a dataset may be the Punjab Disaster Management Authority (PDMA), which is tasked with coordinating the disaster response. Even in the most technologically advanced scenario, coordination is made complicated by the confusion of an evolving disaster, as was highlighted in the aftermath of Hurricane Katrina. Damaged infrastructure and disrupted local governments add to the pressure of addressing the needs of a population under stress. In Pakistan, where a substantial contribution to the response is undertaken by voluntary philanthropic contributors, the task of ensuring that every villager gets their due share is complicated, with the unorganized relief efforts likely to be biased towards areas which are easily accessible.

From this vantage point, we can view the mapping needs of PDMA, which had to coordinate the response over a large area, spread across 8 districts along the Indus. The government functioning at the rural level is organized at the mauza level, and hence it was natural for PDMA to ask which mauzas were affected by the floods. One reason to ask this question was the distribution of Watan Cards, a Rs 25,000 cash disbursement through debit cards that was undertaken immediately by the government. The PDMA strategized that people in every mauza more than 50 per cent inundated should receive Watan Cards. Given that the flood extents were reasonably marked, all that would be required was an accurate overlay of mauza boundaries to make this decision. This scenario is typical of other flood interventions, such as the agricultural seed subsidy or the damage to building structures. Each of the datasets and the subsequent field-work was organized at the mauza-level.

The rudimentary mauza maps which were digitized from hand-drawn sketches were too approximate for any reasonable decision making by the PDMA. In fact, when they did derive the list of affected mauzas by overlaying the flood extents on such a map, it became apparent that this list contained gross discrepancies from field reports, which led to the scrapping of this 'high-tech' strategy. The PDMA resorted to an extensive field survey of every mauza to establish the list of affected mauzas and to document the severity of damage, with the information collected and organized in tabular form. This slowed down the overall response, which could have been faster, had the first strategy worked.

Most of the work that is needed to reverse this situation and to have an effective disaster management GIS has to be done pre-

Table 1: Administrative Divisions of Pakistan, as defined through the land revenue system.

Mauza / Deh موضع، دی	Revenue village, about 52,000 in Pakistan. They mosaic the entire of settled (non-tribal) Pakistan.
Patwar Halqah پٹوار حلقہ	A combination of a few mauzas, historically under the area of jurisdiction of a single *patwari*, land revenue officer
Qanoongoee (قانونگوئی)	A combination of a few patwar halaqahs, historically the area of jurisdiction of a *naib-tehsildar*
Tehsil / Taluqah (تحصی ل، تعلقہ)	The largest sub-unit of a district, containing a few qanoongoees, historically the area of jurisdiction of a single *tehsildar*.
Zilla / District (ضلع)	The main provincial sub-unit. There are 36 districts in the Punjab, and 105 in Pakistan.
Soubah / Province (صوبہ)	The main sub-national unit. Pakistan has four main provinces.

Union Council is an alternate unit of administration, which is a sub-division of the tehsil, and is the primary local government body, headed by a nazim. Its jurisdiction is defined by a collection of mauzas.

disaster. The population demographics can be easily mapped if the locations of mauzas are known, through interpolation of the population census data. Agriculturally productive areas can also be marked in detail on a map by using the crop-census reports, which are again organized at the mauza level. Hence, the mauza is a critical unit in the rural areas of Pakistan, and mapping them will enable the spatial referencing of social, demographic and economic datasets that otherwise do not contain geographic coordinates. The absence of mauza mapping has prevented mapping of socio-economic datasets, and has limited the potential and effectiveness of GIS projects severely.

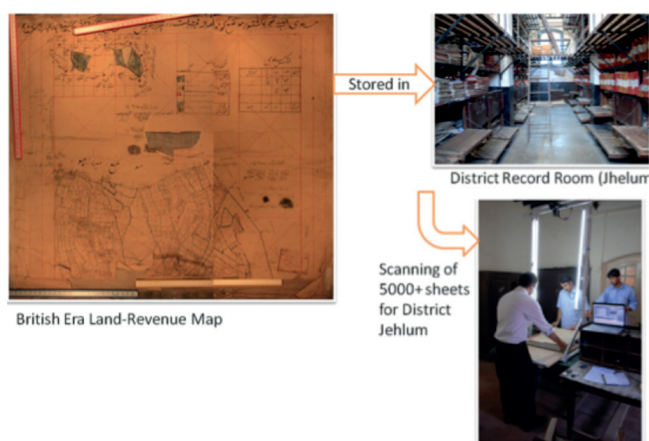## Implementation of Mauza Mapping Solving the Bottleneck

Figure 2: Scanning operation of more than 5000 sheets of British Era Land Revenue Maps.

Why is Mauza a Critical Unit?

From a geo-spatial mapping perspective, the importance of the mauza as a critical spatial identifier derives from its four characteristics:

1. **Small in Size:** There are 26,000 mauzas in Punjab, which implies an average of about 550 households per mauza, making it a good unit of analysis.

2. **Ubiquitously Used Across Government:** At the rural level, the government is organized at mauza level. A mauza may contain more than one settlement, but the government machinery (and datasets) treats it as one integral unit.

3. **Easily Referable by Citizens:** The rural population is typically well aware of their mauza; it is listed on the National ID card, and is frequently referred to in postal addresses. An arbitrary latitude-longitude grid, while useful for mapping tasks, would not be as easily referable as a mauza.

4. **Hierarchical Building Block:** Almost all other relevant boundaries at the rural level, be they Union Councils, election wards (halaqas) or police precincts (thana) are

notified as collections of mauzas. Hence, a mauza map can be used to derive other important maps.

As an example of the utility of a mauza map, consider its application in tracking the spread of dengue virus. If mauza maps were available, the location of each patient reporting to a major tertiary care hospital could be identified from their National ID card (NIC), hence providing a powerful picture of the spread of the disease. Essentially the mauza can then be used as a strong spatial identifier in a GIS system, making it a primary candidate for designing a rural level ZIP-code system for Pakistan. A more pertinent example from the aftermath of the floods is the mapping of damage to household structures. The data on these damages was collected by PDMA at mauza level in tabular form, but could not be mapped because of the non-availability of a mauza map.

We have experimented with two main methodologies to locate mauzas on a map. The primary effort in this regard was to develop a detailed mauza map from land revenue records. Later we also experimented with locating a single point for each mauza through its settlement name. Both strategies are briefly described below.

Mauza Mapping Pilot in District Jhelum

With the help of SUPARCO, we developed what may be termed as the first accurate, integrated and geo-referenced mauza map of a complete district. We chose District Jhelum as the pilot because of its varied geographical features and ease of access.

Approach: Mauza boundaries are defined by the Land Revenue Department. The most recent revenue maps of District Jhelum were prepared in 1940 under the 3rd Settlement of the district. These maps, called masavis, are stored in the District Record Room and form the basis of all administrative boundaries at the rural level. For example, the jurisdiction of a police thana, election halaqas, union councils, patwar circles and qanoongoees are all defined through agglomeration of mauza boundaries. Yet, the maps are largely inaccessible, and some have decayed in the last 70 years to a point where they are unusable. While there could have been other approaches to extract the mauza boundaries, we identified the dire need of preservation of more than 5,000 masavis of District Jhelum as a design constraint, and started preparing a methodology for scanning these valuable documents.

Methodology: Due to the age and the brittle nature of the paper on which the masavis are made, we did not scan them through a normal roller-scanner. Instead, we manufactured our own camera-based scanner which imaged each masavi in two halves. The scanning operation was completed within two weeks and generated approximately 10,000 images. Advanced image processing techniques were used to stitch the images together, first to combine the two halves of a
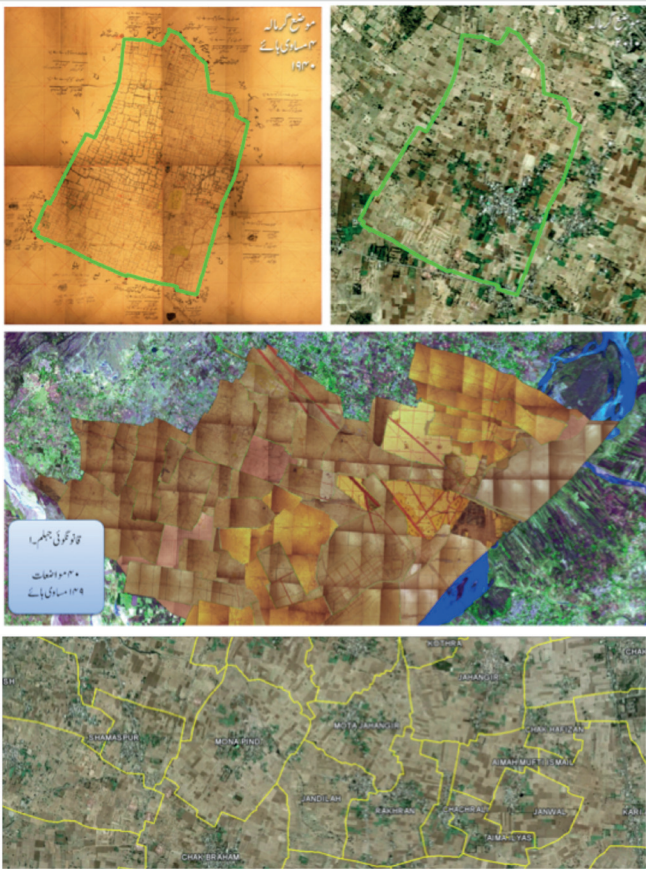
*Figure 3: **Top:** A single mauza map is stitched from four masavis and geo-referenced onto satellite imagery. **Middle:** A mosaic of 40 mauzas, pieced together like a jigsaw puzzle of approximately 150 masavi sheets and shown on satellite image. **Bottom:** Final mauza boundaries extracted from the mosaic and overlaid on Google Earth™.*

masavi into a single image, then to mosaic the masavis of a whole mauza into one image and finally join images of adjacent mauzas together like a giant jigsaw puzzle. These steps required innovation in multiple directions: an understanding of the terminology of the patwari system and their mapping practices were generated, indigenous software was developed to speed up the mosaicing process and protocols for handling missing information, torn sheets and incomplete maps were created. Finally the mosaics of mauzas were geo-referenced on recent satellite imagery. We observed a high degree of alignment between the 70-year old masavis and today's satellite images. Seeing the similarity of features such as roads, fields and streams was a source of great satisfaction for our team; it not only confirmed the high accuracy of the survey procedures employed in the 1930s, but also provided verification that the methodology that we had employed to transfer these maps to digital form was indeed correct.

Alternate Strategy: Mauza Point Mapping

One may argue that instead of mapping the whole boundaries, it may be sufficient for some applications,

and more cost effective, to just locate a single point per mauza on a map. We tested this alternate methodology in the flood-affected district Muzaffargarh of the Punjab province during this research study. This is a time-saving strategy to the problem of jump-starting the critical base layer of mauza boundaries on maps. This is because it does not involve visiting the district record rooms/ scanning the masavis. Instead, it consists of acquiring a list of mauzas and a list of settlements. This strategy is premised on the observation that in Pakistan's rural areas, the name of the mauza always matches the name of one of the settlements contained within it. By matching mauza names to settlement names one is able to point out the locations of the mauzas.

Field work in district Muzaffargarh indicated that this strategy has several caveats. We found the settlement name to match to mauza names in less than 50 per cent of instances. For the rest of the mauzas, we met the patwaris in the field and asked them to identify appropriate matches on our settlement map. Moreover, one of the challenges that surrounds the implementation of this strategy, is that because the mauza list and settlement list is acquired from different sources, there is usually a high degree of discrepancy between the spellings of both entities across different departments. For instance, for the mauza 'Bait Hazar Khan', the corresponding settlement is referred to in another dataset as 'Bet Hazar Khan'. This complicates the matching process. This challenge was dealt with, however, by the development of a phonetic matcher by the research team. This mitigated the problem of spelling inconsistencies by generating phonemes and matching names that sounded similar. Given the implementation issues, and the incomplete picture that this strategy generates, it is our recommendation that, on the balance, the mapping of masavis is more useful and complete than the point mapping strategy.

Potential Impact

The mauza-boundaries themselves are a very useful product for district administration, and our mauza maps were termed invaluable for the daily working of the district officials. However, their linkage with socio-economic data opens up the most exciting possibilities. As a demonstration, we developed a few applications of the use of mauza-boundaries for typical government functions. Census data was mapped to these boundaries using the numeric mauza-identifier available in both census and revenue records (hadbast number). To the best of our knowledge, this is the first time in Pakistan that it became possible to visualize census data on a map, hence enabling the inference of spatial trends. Furthermore, we also developed a spatial-analysis visualization to help the government plan the location of a new school or health unit. This was done by analysis of the mauza-level population, the existing location of these facilities and the accessibility of these facilities through roads. An additional significant impact of

the project is the availability of land revenue maps in a geo-referenced format, which will not only helps preserve the decaying original record, but also makes the land-revenue data much more readily accessible and useable.

Policy Recommendations

Building a National Base Map of Mauzas: This project has culminated in a single primary policy imperative: an integrated, geo-referenced and accurate mauza or village-level map of the country does not exist thus far and must be developed. Having such a map available will provide a critical missing layer for easy mapping of socio-economic indicators at rural level. We have tested different strategies in this study as part of broader efforts to determine how best to achieve a useful mapping of rural Pakistan, which will form the necessary base layer for use of GIS systems in the country.

Developing a Standardized Listing of Administrative Units: Our work shows that a key difficulty in integrating datasets from different sources is that the primary identifiers are not picked from a standardized listing. Spelling variations as well as other discrepancies make it difficult to integrate disparate datasets. While our phonetic matching software took care of some of these issues, repetitive manual interventions are always required. Since the source of mauza names are with the land revenue departments, we strongly recommend that standardized name lists, in both English and Urdu, be generated for each district and numbered according to their hadbast number. Wide availability of such lists will help in preparing future datasets for easy integration.

The Need for a Standardized ZIP-Code System: Based on the hadbast number, it is perhaps straightforward to design a ZIP-coding system. While a current system exists with the Postal Service, it is not widely used perhaps because it is not aligned with the administrative units. A coding scheme will allow ZIP based primary identifiers to be used in the datasets, encouraging their integration and use.

Provincial-Level Spatial Data Infrastructure: Our recommendations are in line with the international trend towards spatial data infrastructures, i.e. frameworks that prepare data for multiple uses and interoperability. More than 80 countries are reporting some progress towards developing such infrastructures (Crompovoets et al. 2004, Massers et al. 2005). While governmental boundaries are a critical layer of any SDI, we believe that the route towards building an SDI in Pakistan (or other developing countries) is going to be slightly different and will take our local context into consideration, as evidenced in our work. The concept has already been pitched to the provincial government for further development.
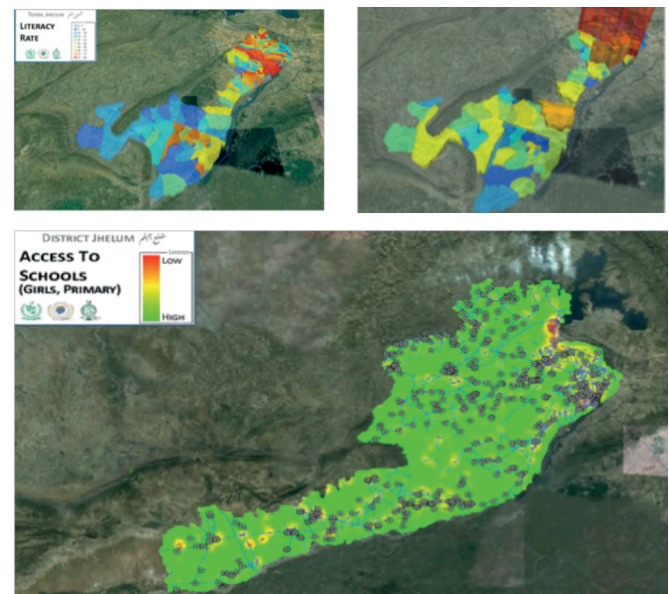

Figure 4: Example Results Derived from Mauza Mapping. **Top Left:** *Literacy Ratio at mauza level for Tehsil Jhelum, from population census data.* **Top Right:** *color coded population of each mauza. Bottom Population, school locations and roads are combined to form a 'heat map' of accessibility to girls primary schools in district Jhelum.*

Partnerships between multiple quarters at both the national and the local level: The federal and provincial governments along with bodies such as the Land Revenue Department, SUPARCO (Space and Upper Atmosphere Research Commission) and PITB (Punjab Information Technology Board) are amongst key players who must collaborate and share resources and technical expertise with one other in the mapping process. There also needs to be a clear division of responsibilities across these bodies. A sustained presence of political will is necessary to ensure that the whole country is mapped. A positive outcome of this study has been the formation of a consortium by the above mentioned stakeholders that will pilot the mauza boundary mapping exercise in 18 more districts of the Punjab.

The Need for Urban Zoning: The advantages of the mauza as a primary rural mapping unit has been demonstrated in this study. However, what an equivalent urban unit that carries the same benefits as those of a mauza at a rural level is not clear. Urban space is characterized by overlapping jurisdictions and many departments seem to use different spatial determinants. An urban unit which is small, common to many government departments, easily referable by citizens and hierarchical in nature will have to be developed.

Preservation of Revenue Maps: Finally, we wish to highlight that the preservation of land revenue maps is also needed as they are in a crumbling state. As we have demonstrated, these maps are not only useful for land-revenue system, but are also critical for making the base layer of any social, demographic or economic mapping.

---

[1] Mauza is a revenue village, the smallest unit in the administrative hierarchy. Also termed deh in some areas.

[2] SUPARCO is Pakistan's national space agency, and a primary provider of GIS and remote sensing services to the government.

About the International Growth Centre

The IGC offers independent advice on economic growth to governments of developing countries. Based at the London School of Economics and in partnership with Oxford University, the IGC is initiated and funded by the UK Department for International Development (DFID).

The IGC has active country programmes in Bangladesh, Ethiopia, Ghana, India, Pakistan, Sierra Leone, Tanzania, Mozambique, Zambia and Rwanda and supports over seventy individual research projects on issues of governance, human capital, agriculture, infrastructure, trade, firm capability, state capacity, macroeconomics and political economy.

The IGC is directed by a Steering Group consisting of an Executive Director (Gobind Nankani) in collaboration with a Deputy Executive Director (Mark Henstridge) and two Academic Directors, one from LSE (Robin Burgess) and one from Oxford University (Paul Collier). The Steering Group also includes Chang-Tai Hsieh from the University of Chicago, Timothy Besley at LSE and Stefan Dercon at Oxford University.

The organisational structure of the IGC spans a London hub, country offices in partner countries, a group of 10 research programmes with participation from academics in world-class institutions, a network of policy stakeholders in the developing world and a range of public, civil society and private sector partners.

Contact us
International Growth Centre
The London School of Economics and Political Science
4th Floor, Tower Two
Houghton Street
London WC2A 2AE
United Kingdom

General Office Tel: +44 (0)20 7955 6144

For enquiries about the IGC, please contact Adam Green:
a.r.green@lse.ac.uk
+44 (0)20 7955 3665

Directed and Organised by

Initiated and Funded by