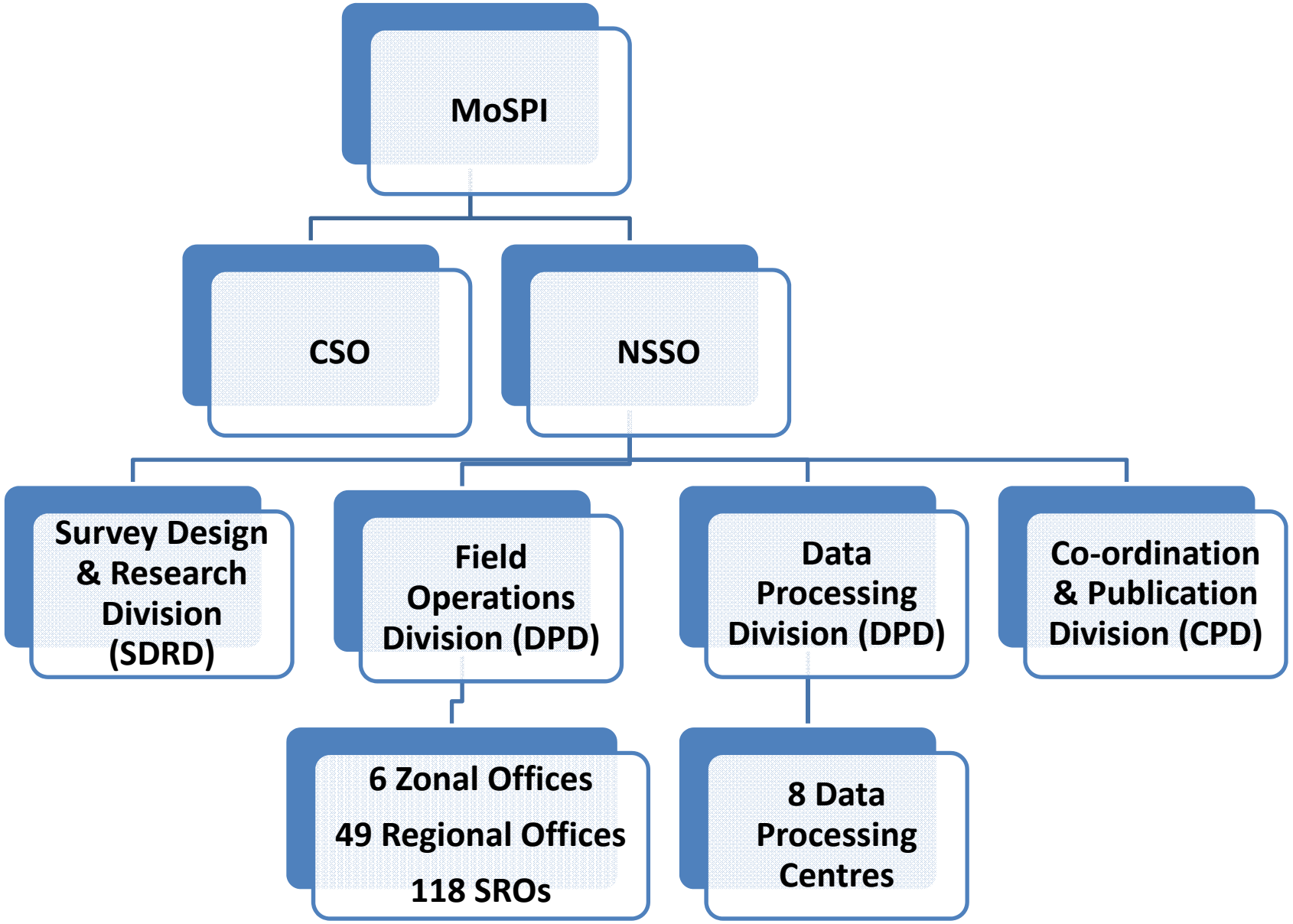
Positioning for the future: Trends in technology, big data, 2020 round of censuses and post 2015 development agenda

Tarak Chandra Patra, ISS
Data Processing Division,
National Sample Survey Office
Ministry of Statistics & Programme Implementation
Government of India

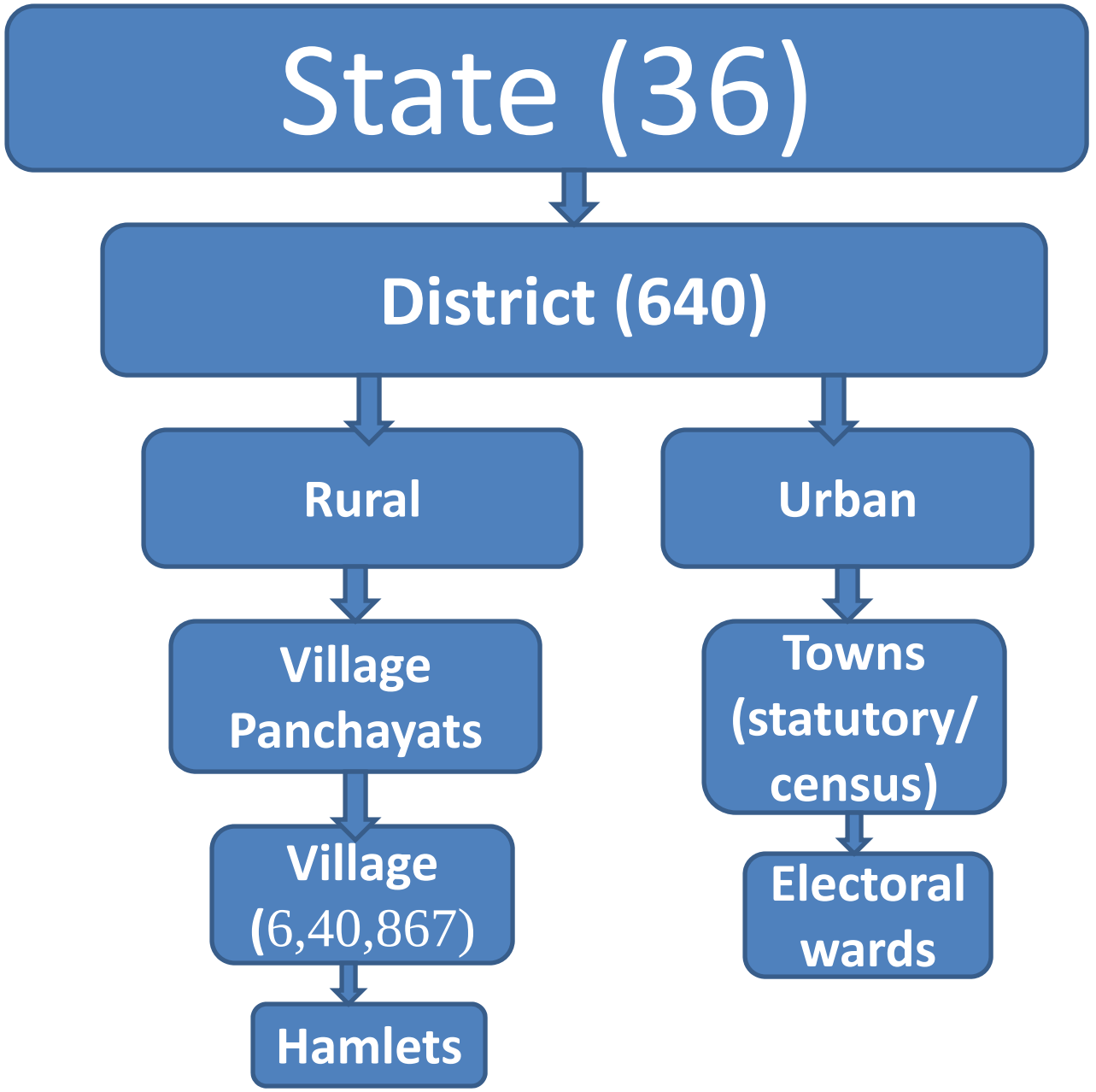
Statistical System in India

- India has a federal structure of Government
- As per the Constitution of India Statistics is a subject in the Concurrent list and it is dealt by both Central Govt. and State Govt.
- The Ministry of Statistics and Programme Implementation (MoSPI) is the nodal agency for planning and facilitating the integrated development of the statistical system in the country
- NSSO under MoSPI is responsible for conducting large scale socio-economic sample surveys of national interest
- NSSO generates estimates which are important for planning & policy making and to measure effects of implementation of different Govt. projects
- DPD under NSSO works as the processing agency as well as developing software for that including tabulation and sharing its software with State Governments

Organisation chart of NSSO



Administrative Units



NSSO experiences

- Data are collected by Field Operations Division (FOD) by canvassing hard copy of schedules by visiting selected sample households
- Sample households are the ultimate stage units and aerial units - Census villages (rural)/ Urban Frame Survey (UFS) blocks (urban) are the First Stage Units
- These villages/ UFS blocks are selected randomly by using in-house developed software by DPD from the list of census villages/ list of UFS blocks
- Maintenance and updation of frame is one of the most important process and it is continuous
- For rural India Census list of Villages forms frame for selection
- In this context Urban Frame Survey plays a crucial role for urban part of the country

Urban Frame Survey

- UFS covers the whole Indian union comprising of 7933 cities/towns (4041 statutory and 3892 census towns) as per the Census 2011

- Under the UFS, all the towns are physically surveyed and separate UFS blocks (with 80-200 households) and IV units (a group of blocks: 20-50) are carved out for covering the whole geographical area of the town

- UFS Blocks are updated over a five year period, each process of updation is called as phase

- Auxiliary information on type of area such as slum area, residential area, industrial area, education, hospital, bazar(market) area, prohibited area etc. collected during the survey/ updation

- These help in selection of blocks with certain characteristics for the socio economic surveys

- Landmarks such as school, post office, bus stand, clinic etc. are recorded for each block for the purpose of proper identification

- The boundaries of the UFS blocks have been frozen

Urban Frame Survey

- Two notional maps are prepared, one is IV Unit map showing the location of all the blocks included in the Unit

- Other one is the town map depicting the location and relative positioning of IV units in the town

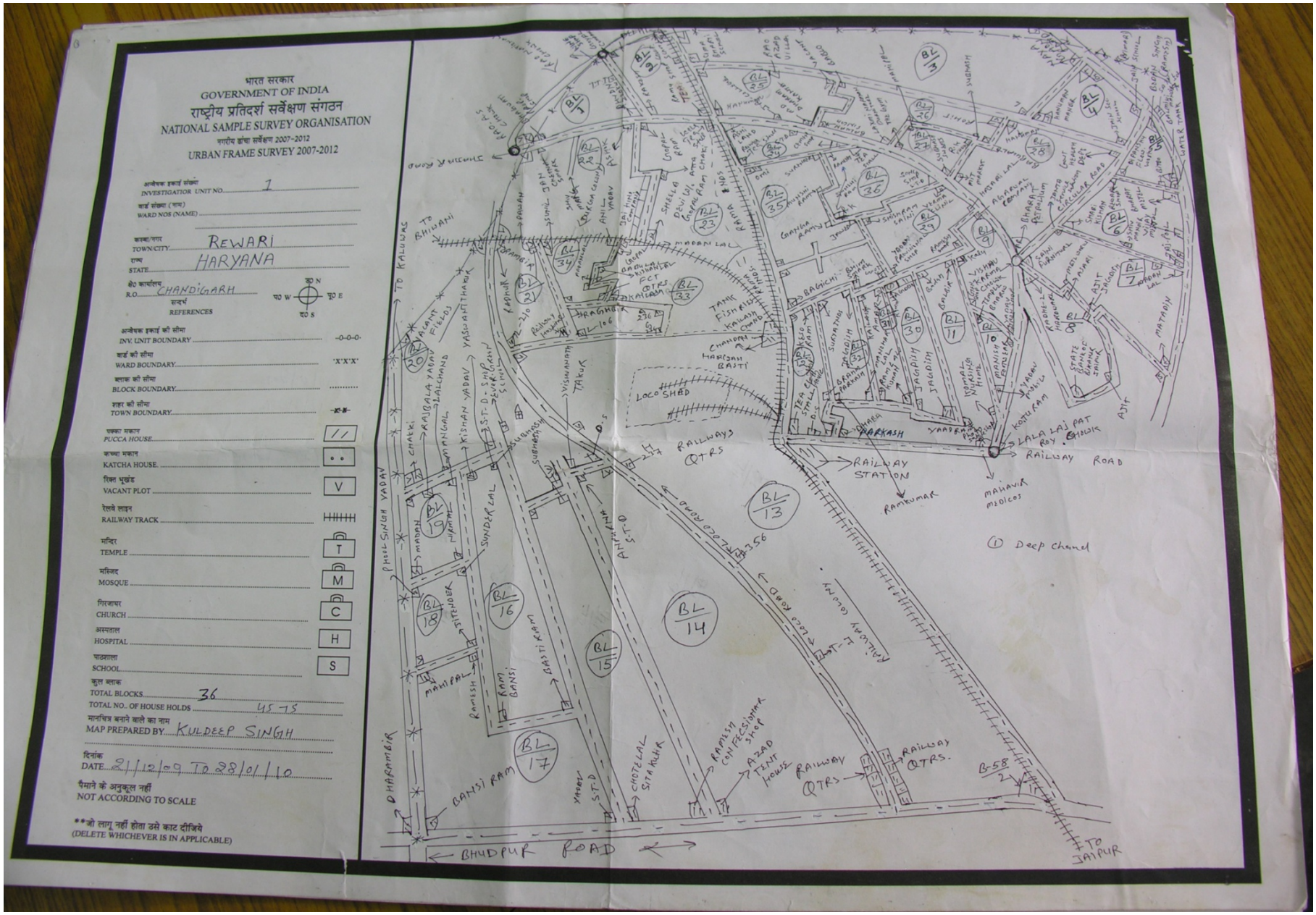
- *NSSO has initiated the conversion of IV unit maps to GIS framework in a uniform scale with the help of* National Informatics Centre (NIC)

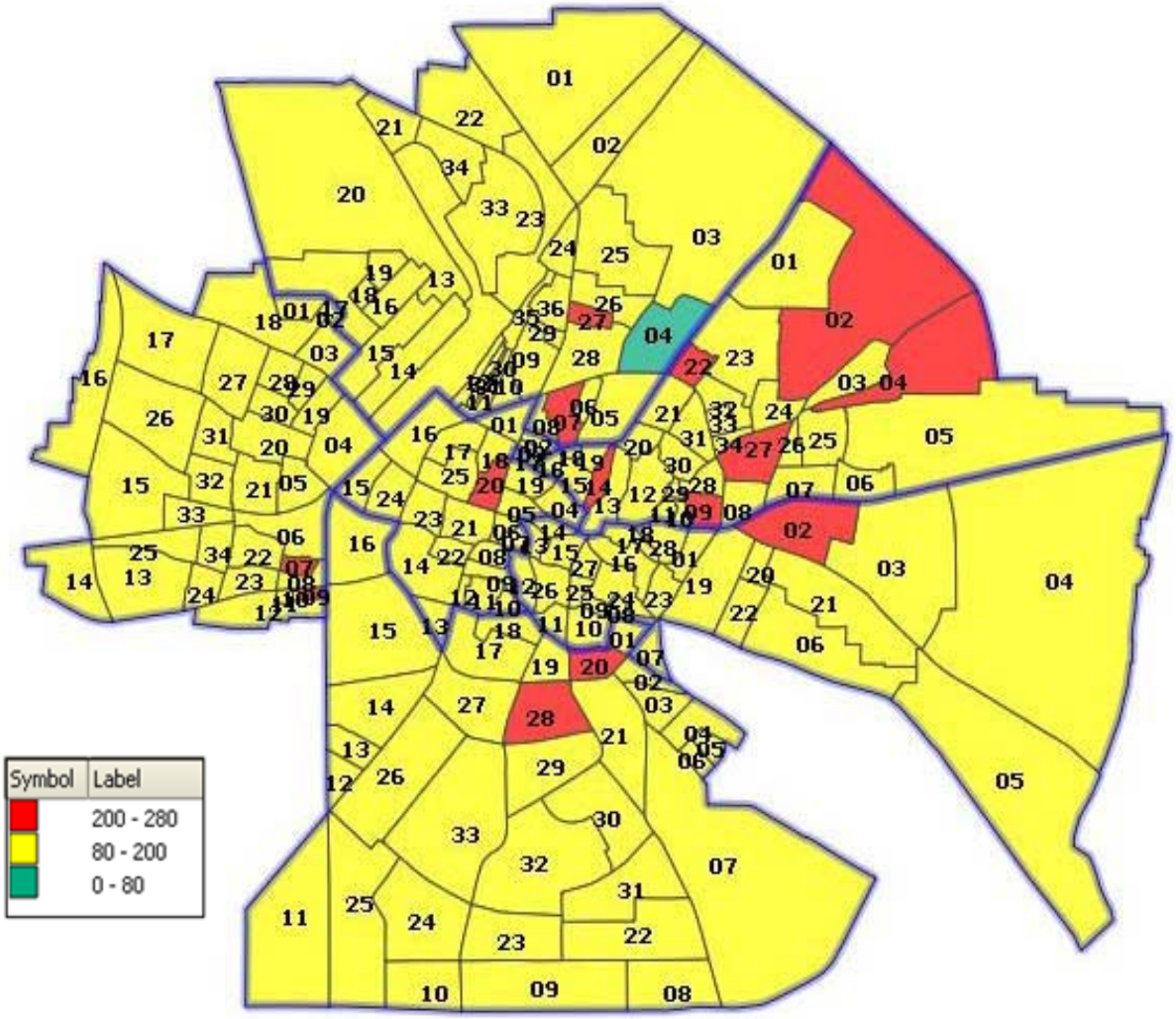
- UFS helps to update both rural and urban frame of survey by incorporating newly declared towns and deleting them from rural part

- Deurbanised towns are deleted from urban list and included in rural part

- Metadata of UFS are also available

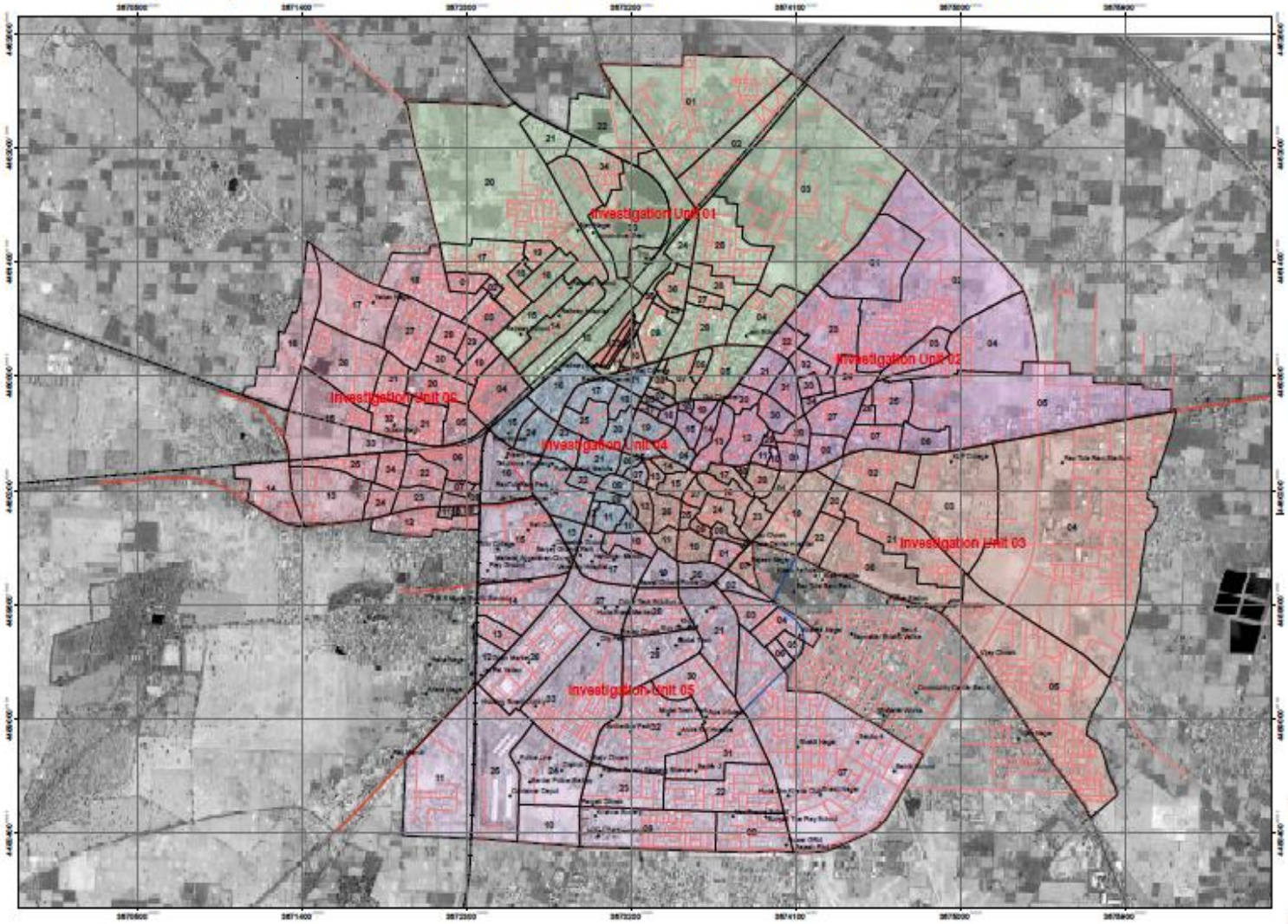
Urban Frame Survey- notional map of a town




Symbol
Label
200 - 280
80 - 200
0 - 80

URBAN FRAME SURVEY : 2007-2012

State : Haryana **District : Rewari** **Town : Rewari**



LEGEND

- Location
- Road
- Bawal Road
- Ghari Boini Road
- N.H-71B
- Nala
- Rail
- Rampura Road (Circular)
- S.H-24
- S.H-26
- Shahjahapur Road

Investigation Unit

- 0001
- 0002
- 0003
- 0004
- 0005
- 0006

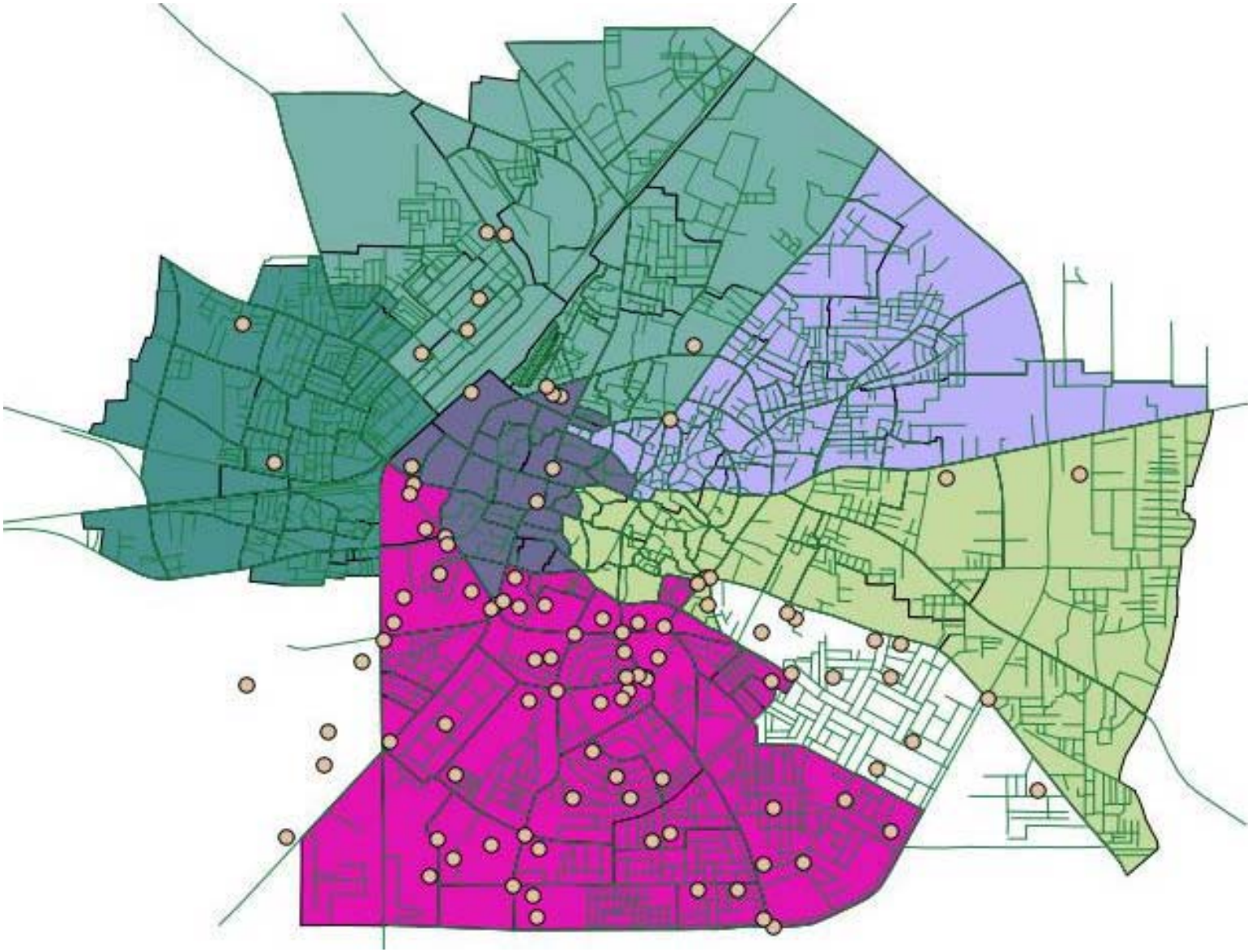
1:10,000

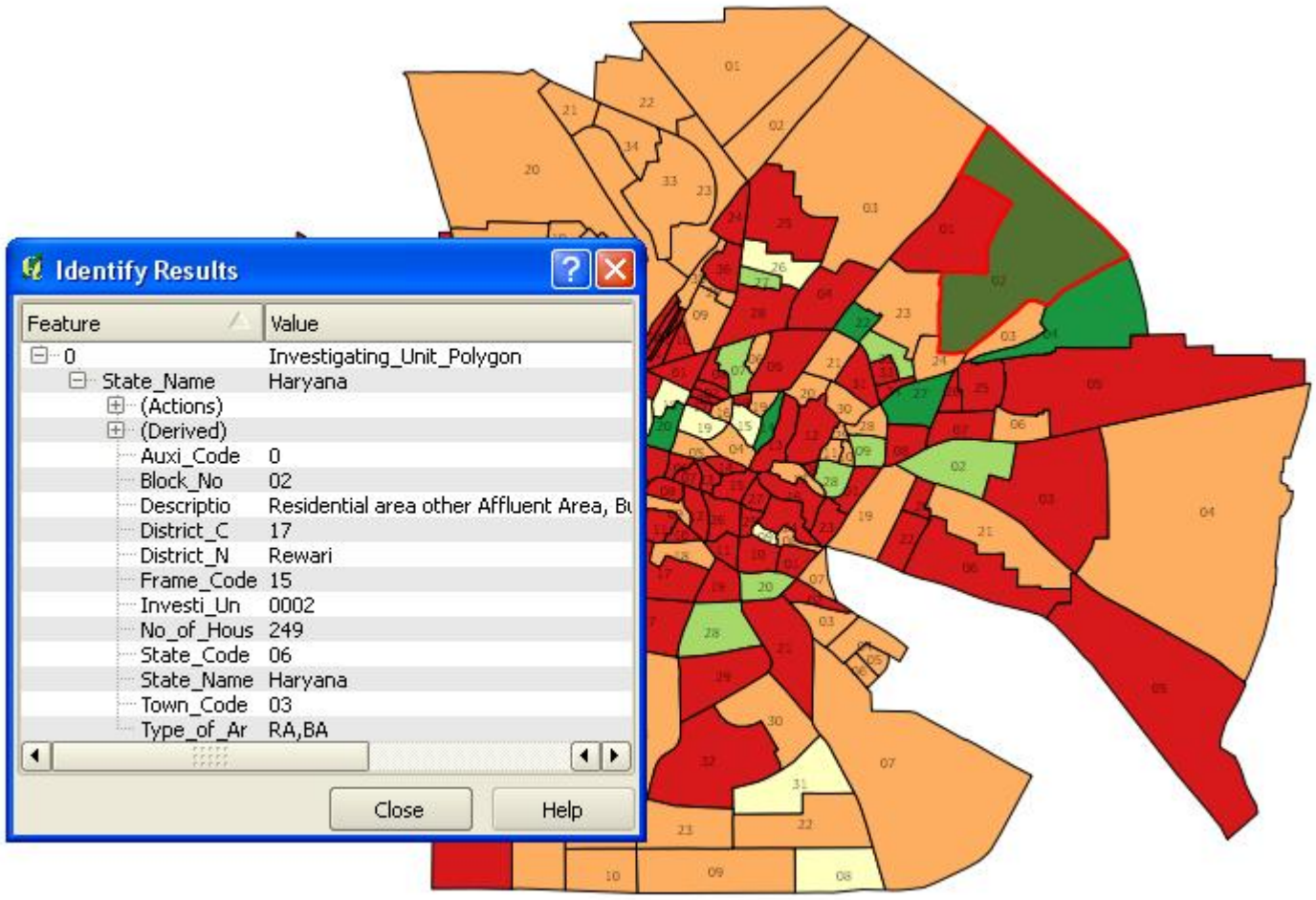
0 375 750 1,500 2,250 3,000 Meters

INPUT SOURCE :

i) Survey Report
ii) QB Satellite Data
iii) Google Map for attribute

Prepared For
NIC (GIS & RS Division) New Delhi
By
Pan India Consultants Pvt. Ltd.
Gurgaon-122015, Haryana




Identify Results
Feature
Value
0
Investigating_Unit_Polygon
State_Name
Haryana
(Actions)
(Derived)
Auxi_Code
0
Block_No
02
Descriptio
Residential area other Affluent Area, Bu
District_C
17
District_N
Rewari
Frame_Code
15
Investi_Un
0002
No_of_Hous
249
State_Code
06
State_Name
Haryana
Town_Code
03
Type_of_Ar
RA,BA
Close
Help

Data Processing Division (DPD)

- Data Processing Division (DPD) is responsible for processing of data collected by FOD in paper schedules
- Around 1.71 crores of records are processed by DPD in a Quinquennial round for consumption expenditure survey
- DPD handles data of different subjects every year. So, no standard set of software. However, the data processing system remains more or less the same
- Data are entered & verified by using in-house developed software of DPD in RDBMS platform using ORACLE as back end database
- This enables on line validation and checking of data, process has become much easier to control and manage
- Time frame for finalisation of survey results of unit level data of a round is one year after the end of field work of data collection
- Various complicated (multivariate tables ranges from 600 or more) are generated by DPD using its own tabulation package “DOSTAB”

NSSO Initiative

- In order to reduce the time frame further to release the final unit level data to the users and policy makers and researchers NSSO has been pursuing different experiments so that the basic objective of reducing time can be achieved without compromising its standard quality
- One of them is the experiment of decentralisation of data entry at FOD
- It is also being thought of data collection by electronic media using data collection software instead of canvassing paper schedule
- Main constraint in this regard is length of the NSS schedules and their complexities
- Experiment of shortening and simplification of schedules of various subjects has been taken up by NSSO

NSSO Initiative

- As Statistics is in the Concurrent list of Constitution of India, State Governments also participate in NSS surveys
- One important objective of a sample survey is to generate estimates at lower level of administrative unit such as district or sub-district level for planning and policy making
- Sample size of NSSO is not enough to generate robust estimate below the state level
- In order to generate district level estimate pooling of central sample data (done by NSSO) and state sample data (done by various state Governments) is required
- To meet this requirement DPD has been sharing all its software and technical inputs for processing of their data whereas other branches of NSSO also share their survey instruments with state governments
- DPD has developed software to test poolability of some important parameters so that estimate can be built for these parameters at district level

Population Census experiences

- In the 2010-2011 round Census of India implemented three different approaches – a traditional paper-schedule-based approach for the Housing and Population Census, a biometric database for the National Population Register, and direct data collection on an electronic platform for the Socio-Economic and Caste Census

- For use in Census 2011, information on changes in the jurisdiction of the administrative boundaries of 35 States/Union Territories, 640 districts, 5,924 sub-districts, 7,935 towns and 6,40,867 villages were meticulously collected along with official notifications and maps

- Digitised maps were prepared using latest GIS software. Detailed digital maps of 33 capital cities of the country based on satellite imagery were also prepared

- These maps show detailed layout of buildings, houses, other structures, road network and important landmarks and were used in Census 2011

- Household level information not available

Population Census experiences

- The Census Schedules were scanned using high speed duplex scanners and information read using ICR technology
- Use of the ICR technology not only saved time for data capture and data tabulation thus ultimately making it available to the users early, but also was very cost effective saving public money
- In the National Population Register the digital database that contain certain basic demographic characteristics of every usual resident of India along with 3 biometric attributes - 10 fingerprints, 2 iris prints and photograph
- It is envisaged that the register would be a dynamic one with linkages to the Civil Registration System (CRS) to ensure that every birth & death would be properly reflected in the electronic database
- In the Socio Economic and Caste Census (SECC) tablet PCs were used for collection of data where database created during NPR was preloaded
- *The methodology of direct data capture has exciting implications for the future of census in India*

MDG framework in India

❖India's MDGs framework is based on the 2003 United Nations Development Goals (UNDGs) guidelines on concepts, definition and methodology of MDGs indicators

❖ 53 indicators (48 basic and 5 alternatives) recognized by this framework has been contextualized for India

❖ All the 8 Millennium Development Goals, 12 of the 18 targets (target 1 to target 11 and target 18) are relevant for India

❖ These targets and indicators under the 8 Goals constitute the instrument for statistical tracking of the MDGs in India

❖ MoSPI is engaged in the task of statistically tracking the MDGs on the basis of data sets generated by different Ministries/ Departments and information gathered from periodic national surveys and censuses

❖ Some of the indicators achieved ahead of given dead line and some are expected to reach closer

Note for the future

❖ MDGs have helped in bringing a much needed focus and pressure on basic development issues and in turn led the governments at national and sub national levels to do better planning and implement more intensive policies and programmes

❖ In India the various development programmes / schems are formulated and implemented under the Five year Plans (FYP)

❖ The 12th FYP (2012-2017) goal is to achieve " Faster, More Inclusive and Sustainable Growth" which is in conformity with the MDGs

❖ The 12th Plan has identified 25 core indicators which reflect the vision of rapid, sustainable and more inclusive growth

❖ The 12th Plan also emphasizes to encourage research and innovation for more use of geospatial information

❖ An in-house geospatial software solution for crop production forecasting (FASAL) is already operational using information from different sources like remote sensing, meteorology and land observations

THANKS
for
ATTENTION