



# **Landsat Global Archive Consolidation: Status report to Landsat Science Team**

**19 January 2010**

U.S. Department of the Interior

U.S. Geological Survey

# Background

---

- **Science Team Charge**

- ◆ Acquire data from the global archives
  - Concerned most about data in vulnerable situations
    - ◆ Inactive stations, unmaintained archives, etc.

- **Global Archive Implementation Plan: Spring 2009**

- **Preparation stages:**

- ◆ Contacted all international stations
  - Provided information on Global Archive consolidation effort
  - Requested complete metadata records for their archive
    - ◆ Received Canada, China, Thailand, Pakistan
- ◆ Promoting Global Archive consolidation
  - Cooperators, scientists, and industry
  - GEO Secretariat interested in assisting

# Scope and Concerns

---

- **20 international archives**

- ◆ Each unique in acquisition activity and technical approach
- ◆ Estimated total of 1.3 PB of data to consolidate
  - Approximately the same volume as the US archive

- **Key Technical Risk**

- ◆ All older data will be difficult to ingest (particularly MSS)
  - Science data and ancillary data stored on separate media
  - Old and undocumented formats
  - Obsolete media

# Data in International Archives

<u>Landsat 1-7</u>	<u>Scenes</u>	<u>Size (TB)</u>
Italy	1,545,000	520
Canada	660,000	185
Australia	600,000	125
Brazil	480,000	67
China	415,000	126
Japan (Restec)	403,000	53
Thailand	240,000	43
Argentina	195,000	67
South Africa	153,000	32
Ecuador	80,000	17*
Pakistan	34,000	7*
India	97,000*	20*
Mongolia	9,000*	2*
Saudi Arabia	69,000*	14*
Taiwan	18,000*	4*
<u>Landsat 7 only</u>		
Japan (Hiroshima)	18,000	11
Germany	23,000*	14*
Indonesia	42,000*	26*
Korea	9,000*	6*
Puerto Rico	<u>42,000*</u>	<u>26*</u>
<b>Total</b>	<b>5,132,000</b>	<b>1.3 PB</b>

Numbers are based on information that Landsat has gathered to date from International Cooperator ground receiving stations.

\*Numbers are estimates derived from the length of a station's active data-reception period (Scenes) and the type of sensor data it has received (Size).

# Examples of stations useful to large efforts

---

## Forest Carbon Tracking

Ecuador  
Brazil  
Thailand  
India  
Indonesia  
China

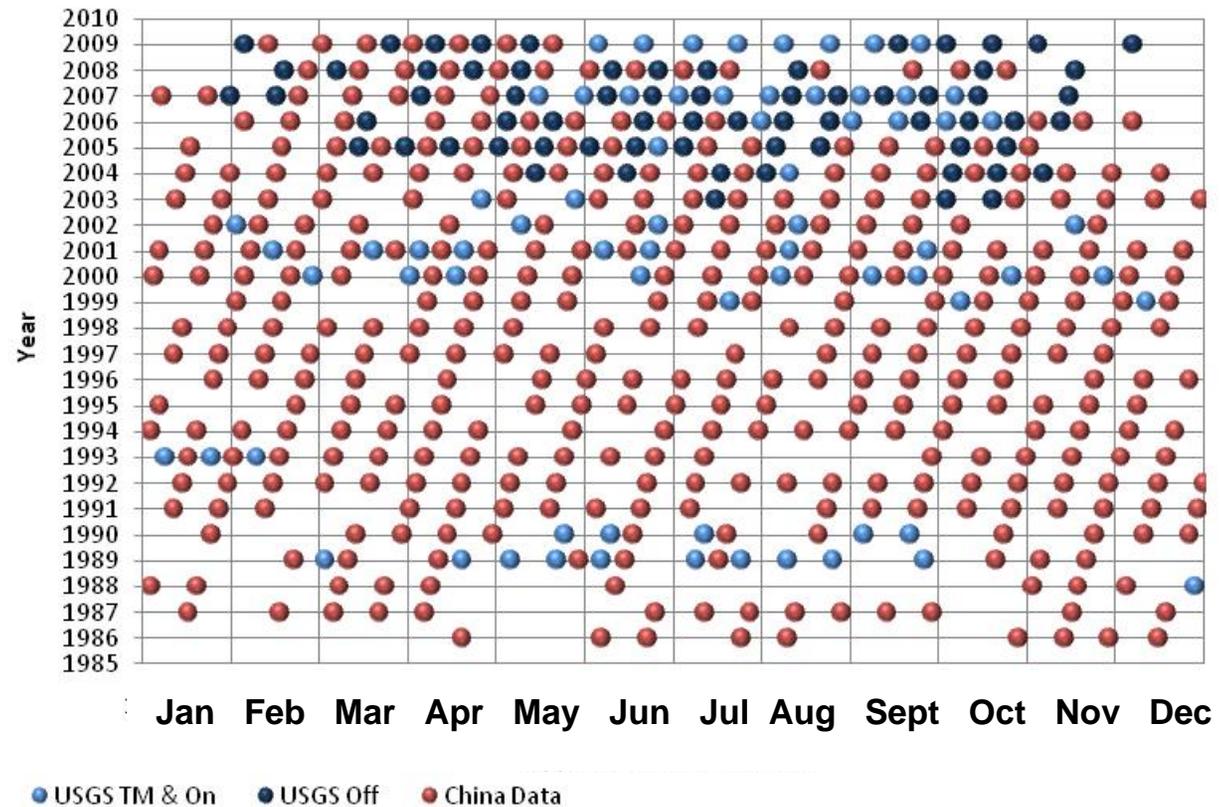
National Security  
interests  
others

## Boreal/Arctic

China  
Mongolia  
Kiruna, Sweden  
Canada  
Korea

# Large IC archive; remote location

Northeast of Ulaanbaatar, Mongolia: p128r025



# Scenes: p128r25

USGS archive

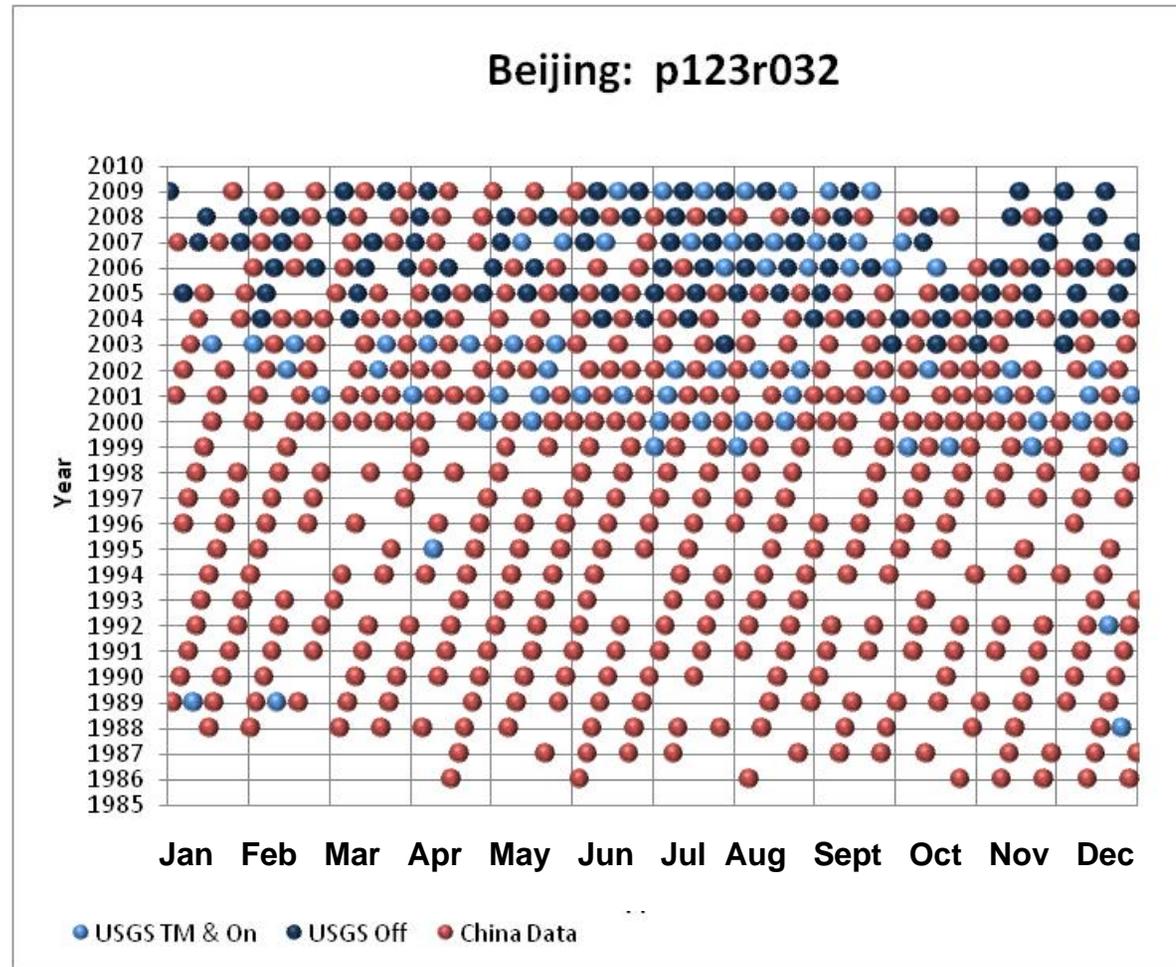
TM&On: 71

Off: 68

China archive: 379



# Large IC archive; central location



# Scenes: p123r32

USGS archive

TM & On: 72

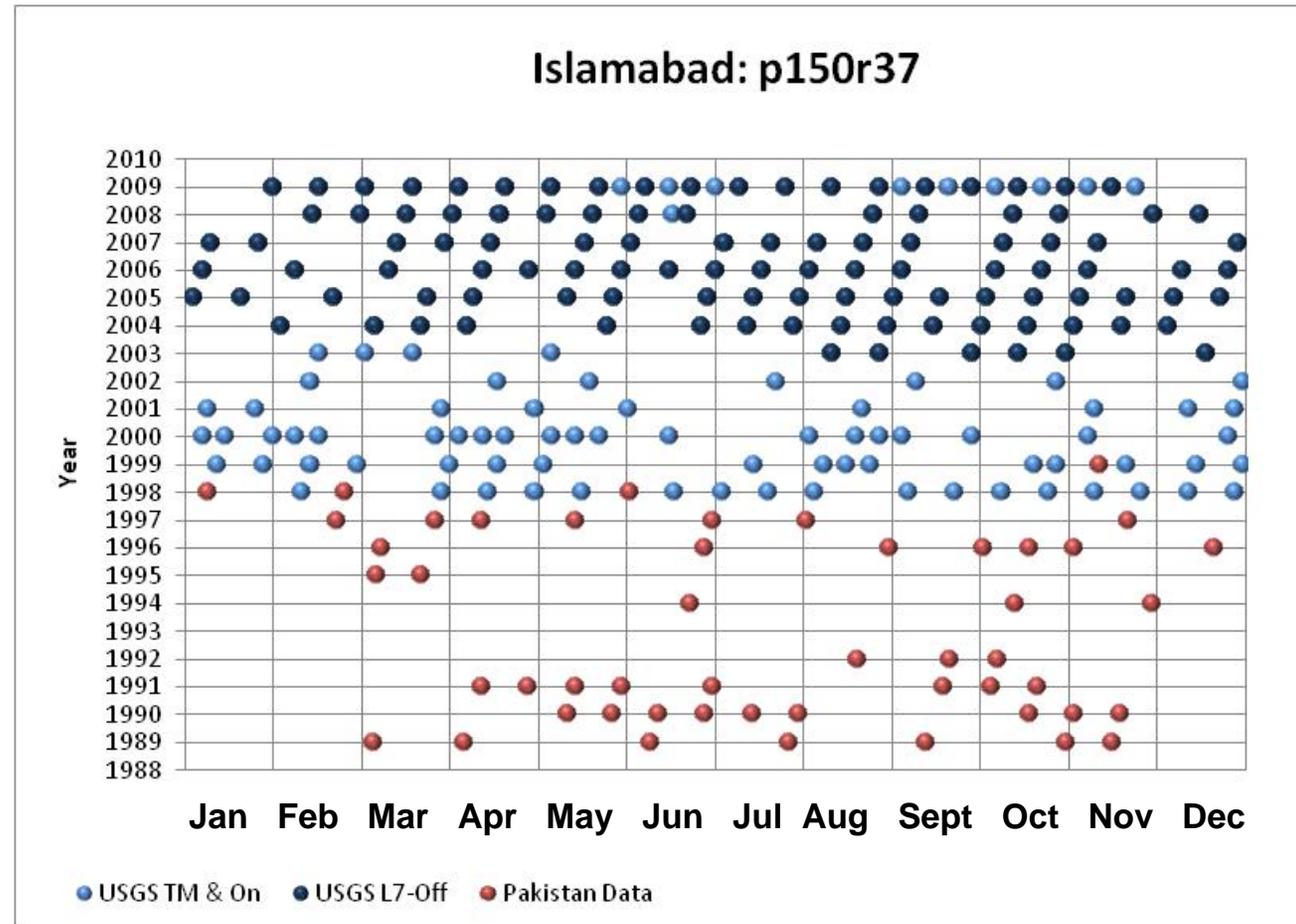
Off: 100

China archive: 450



# Small IC archive; central location

Extends the archive over 10 years.



# Scenes: p150r37

USGS archive

TM&On: 83

Off: 109

Pakistan archive: 50



# Required Steps

---

## Preparation

- **Promote Global Archive Consolidation to international community**
- **Acquire and reconcile detailed archive metadata from each station**
  - ◆ Current plan is to ask for everything they have
- **Verify data exchange format and media capability**
- **“Agreement” with each station – does not have to be formal**

## Implementation

- **Reception and organization of media**
  - ◆ Large shipments will take time to be worked through, so we will need some procedure to manage them in the meanwhile, particularly if we have multiple stations arriving simultaneously.
- **Ingest of tapes (FTP may be an option for small archives)**
- **Database management for overlapping path/rows, mismatched metadata, data management, and reconciliation of archive listings.**

# Where we are

---

- **Preparation**

- ◆ Promote Global Archive consolidation
- ◆ Contacted all international stations
  - Requested archive metadata
    - ◆ China, Pakistan, Ecuador, Canada received & analyzed

- **Implementation – Pakistan & Ecuador**

- ◆ Pakistan

- Landsat requested their archive metadata.
- Pakistan sent metadata and asked for help reading old tapes.
  - ◆ Should receive sample tapes in mid-January

- ◆ Ecuador

- Communication with Ecuador originated through LRS.
- We have submitted a USGS data rescue proposal.