*International Planetary Probe Workshop 11   Caltech  June 2014*

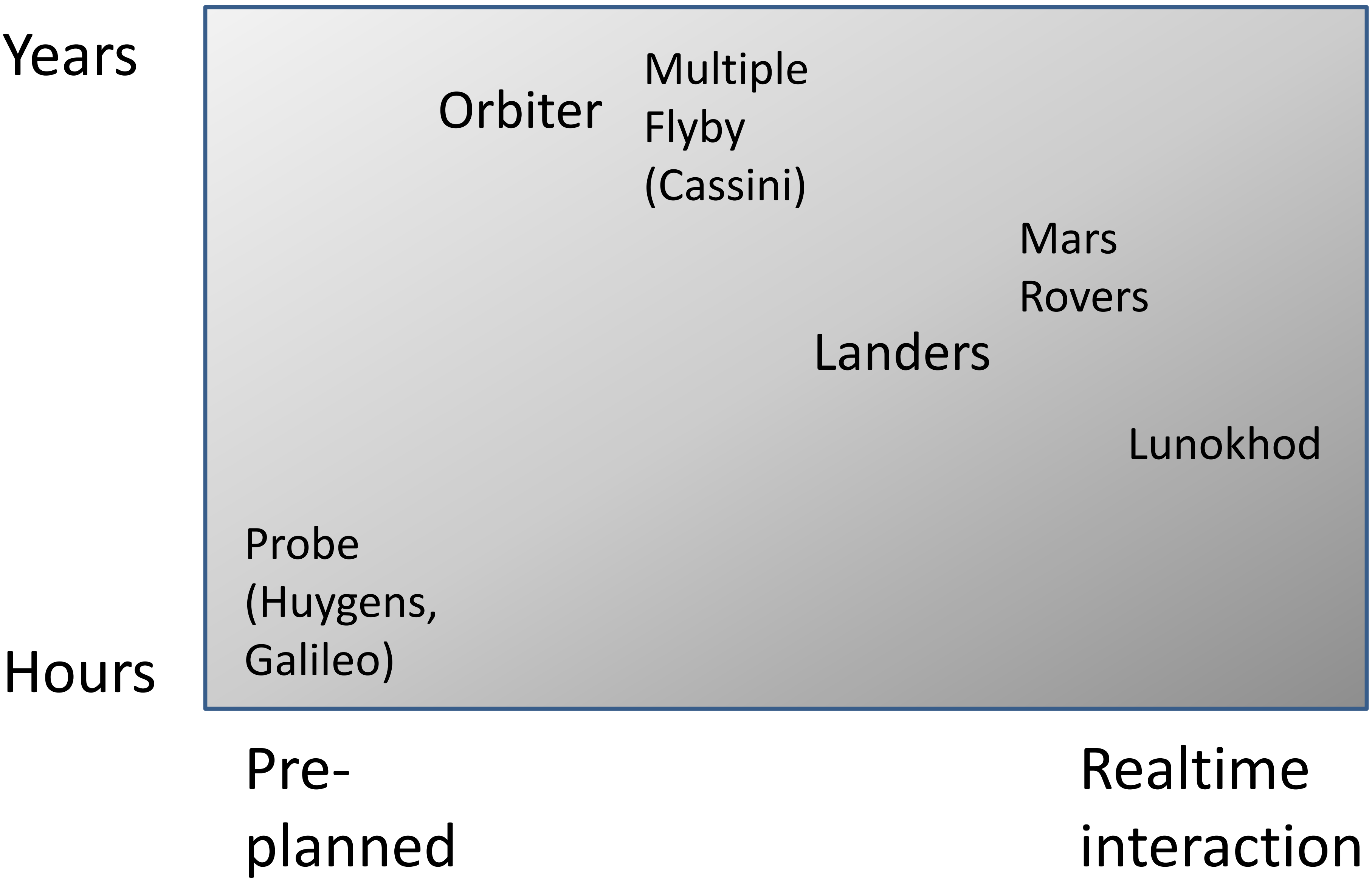
*IPPW-11 Short Course*

# *Flight : Getting, Analyzing, Publishing and Archiving the Data*

***Ralph D. Lorenz***
*Johns Hopkins University Applied Physics Laboratory, Laurel, MD. USA*
*Ralph.lorenz@jhuapl.edu*
*http://www.lpl.arizona.edu/~rlorenz*

JOHNS HOPKINS
APPLIED PHYSICS LABORATORY

Pace, type of a Mission dictates style of operations and science analysis



# NASA Discovery-10 Draft Announcement of Opportunity

## 4.4 Data Policies

### *4.4.1 Data Analysis*

The PI will be responsible for analysis of the mission data (including returned samples) necessary to complete the proposed science objectives and for timely publication of initial scientific results in refereed scientific journals, as part of their mission operations (Phase E) or post-mission (Phase F) activities.

### *4.4.2 Data Rights*

By NASA policy, all science data returned from NASA missions are immediately in the public domain. A short period of exclusive access may be proposed for data calibration and validation, but a compelling justification for it must be demonstrated. Any period of exclusive access should be the minimum that is consistent with optimizing science return from the mission. Barring exceptional circumstances, it may not exceed six months.

# NASA Discovery-10 Draft Announcement of Opportunity (cont'd)

### *4.4.3 Delivery of Data to Archive*

Mission data will be made fully available to the public through a NASA-approved data archive (*e.g.,* the Planetary Data System, Atmospheric Data Center, High Energy Astrophysics Science Archive Research Center, etc.), in usable form, in the minimum time necessary but, barring exceptional circumstances, within six months following its collection. The PI will be responsible for collecting the scientific, engineering, and ancillary information necessary to validate and calibrate the data prior to delivery to the archive.

Archival data products will include low-level (raw) data, high-level (processed) data, and derived data products such as maps, ancillary data, calibration data (ground and in flight), documentation, related software, and/or other tools or parameters that are necessary to interpret the data. The PI will be responsible for generating data products that are documented, validated, and calibrated in physical units that are usable by the scientific community at large.

NASA data archives have budgets to support core activities, including the basic ingestion and review of new data. Proposed mission data archiving plans and budgets must be consistent with the policies and practices of the appropriate NASA data archive. Proposals may include funding for up to one year after end-of-operations for the generation and archiving of derived data products. This funding will be included in the capped PI-Managed Mission Cost.

# NASA Discovery-10 Draft Announcement of Opportunity

4. Data Plan.

Requirement B-23. A schedule-based end-to-end data management plan, including approaches for data retrieval, validation, preliminary analysis, and archiving shall be described. The science products (*e.g.*, flight data, ancillary or calibration data, theoretical calculations, higher order analytical or data products, sample returns, witness samples, laboratory data, etc.) shall be identified, including a list of the specific data products and the individual team members responsible for the data products. The plan shall identify the appropriate NASA data archive and the formats and standards to be used. It shall include an estimate of the raw data volume and a schedule for the submission to the data archive of raw and reduced data in physical units accessible to the science community.

No battle plan survives contact with the enemy

- Apocryphal military aphorism

Ev’rybody got a plan, ‘til they get hit.

- Mike Tyson

Sometimes

- operations plans are derailed  (e.g. MGS aerobraking hiatus)
- instrument data cannot be trusted  (e.g. Galileo probe temperature excursion)
- Operations software is not ready (Phoenix arm commanding)
- Science team social structures can have issues

Recommended reading :

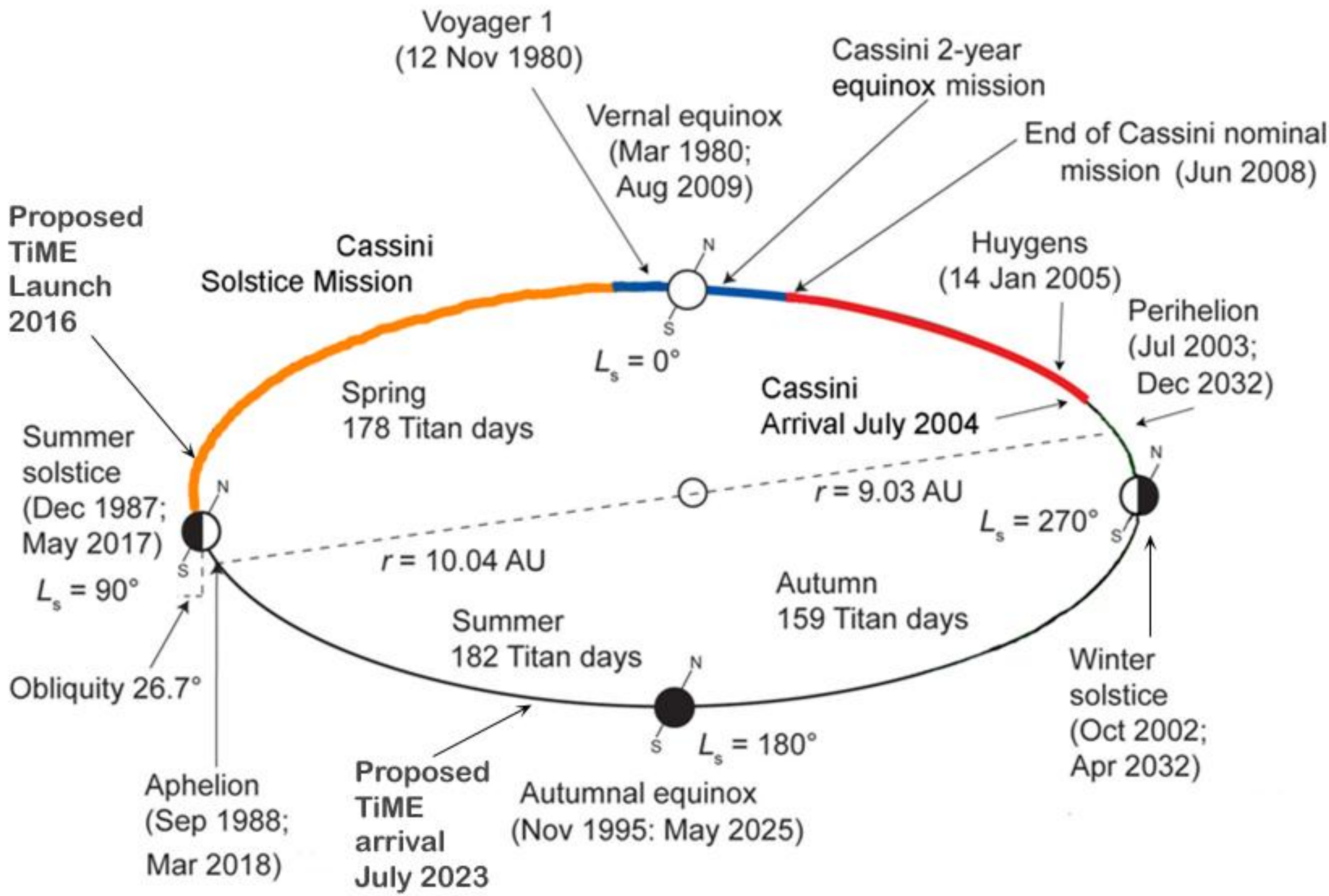
Squyres, Roving Mars : Spirit, Opportunity, and the Exploration of the Red Planet

Lorenz and Mitton, Titan Unveiled

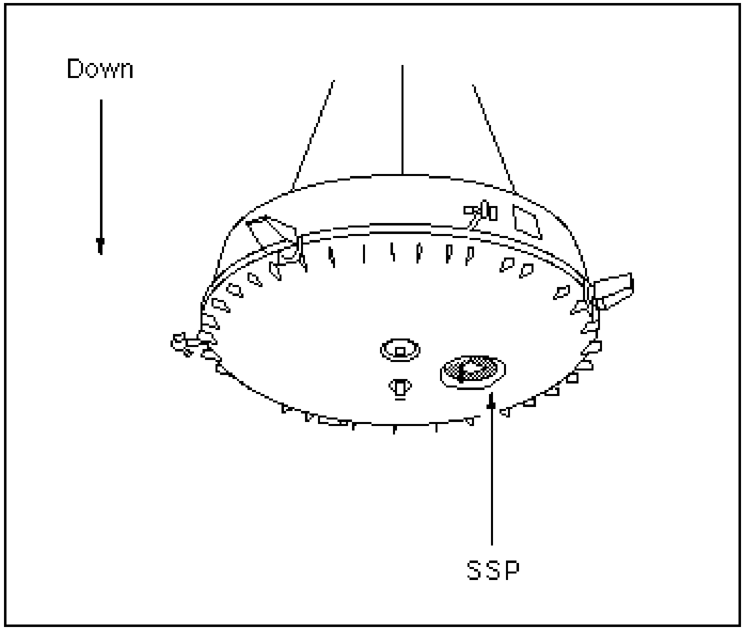
Kessler,   Martian Summer: Robot Arms, Cowboy Spacemen, and My 90 Days with the Phoenix Mars Mission

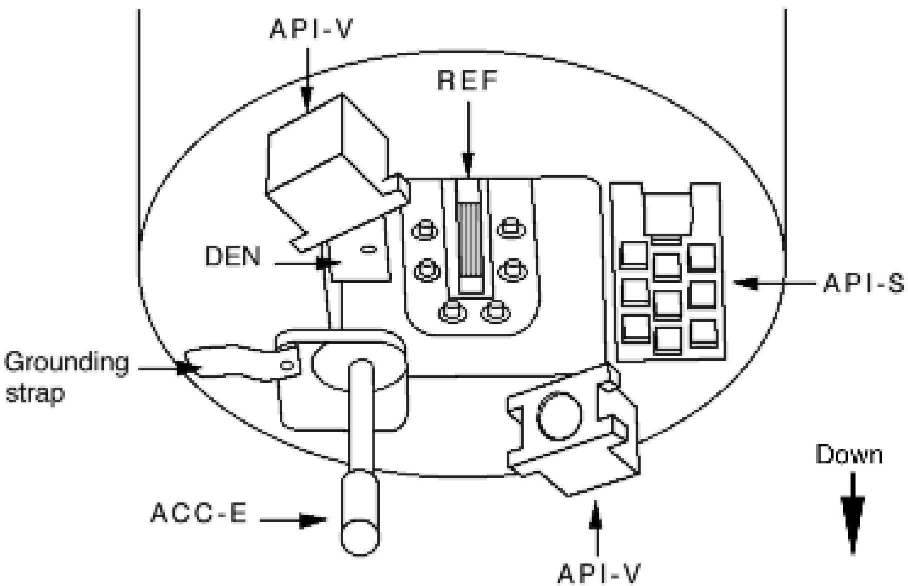
Wall and Ledbetter, Design Of Mission Operations Systems For Scientific Remote Sensing

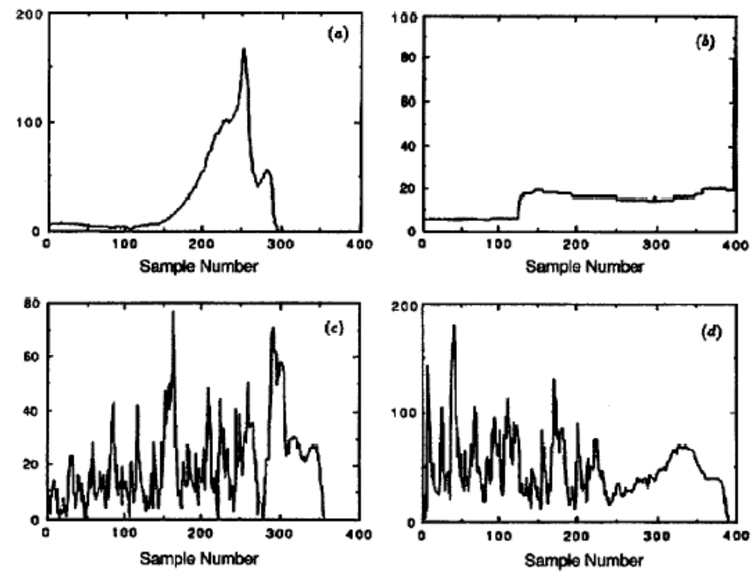
# Seasons on Titan



# The Penetrometer on the Huygens Probe



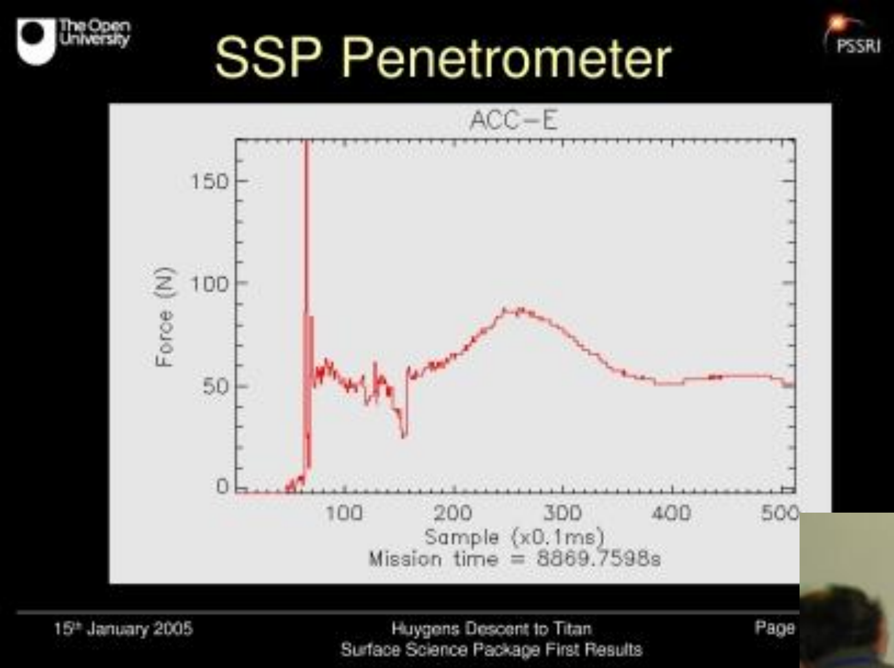



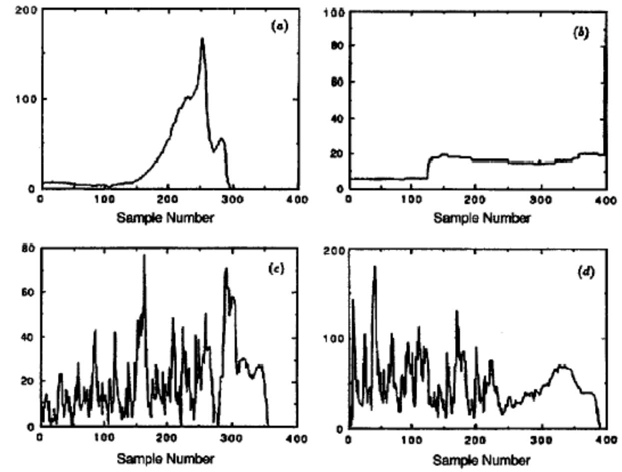

Data taken in the lab in 1994 – (a) dry sand (b) wet clay (c) fine gravel (d) coarse gravel (from *R. D. Lorenz,et al 'An Impact Penetrometer for a Landing Spacecraft', Measurement Science and Technology, vol.5 pp.1033-1041, 1994* also at http://www.lpl.arizona.edu/~rlorenz

I assembled 4 penetrometer sensor heads in August 1994

Best of build – Flight model
Almost as good – Flight Spare

A couple of others for further experimentation, student projects etc.

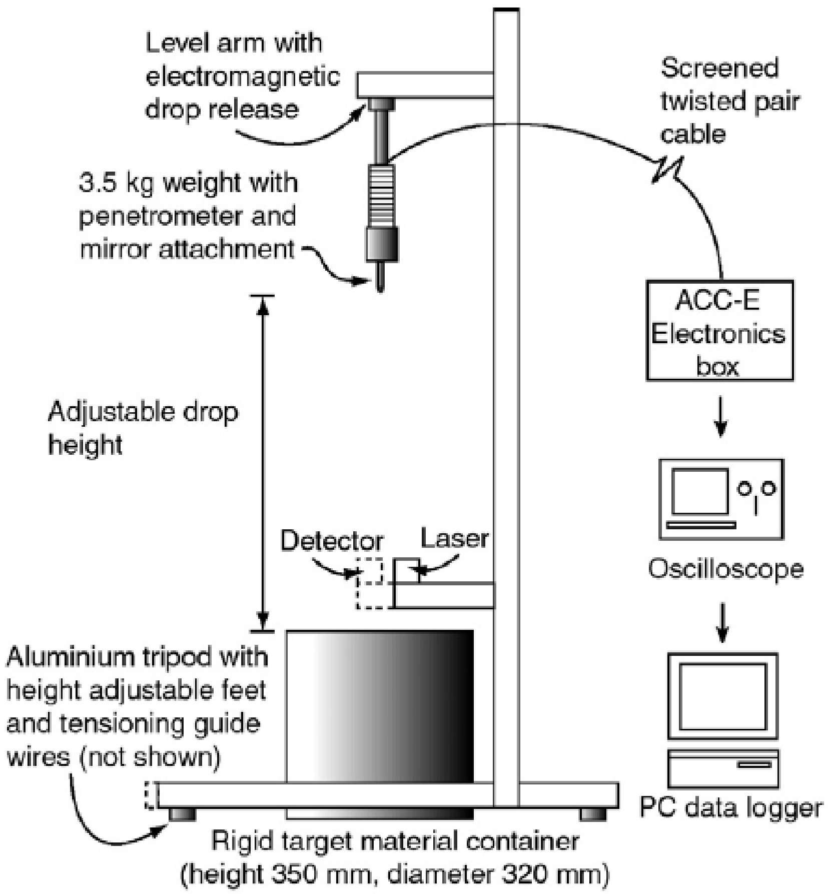


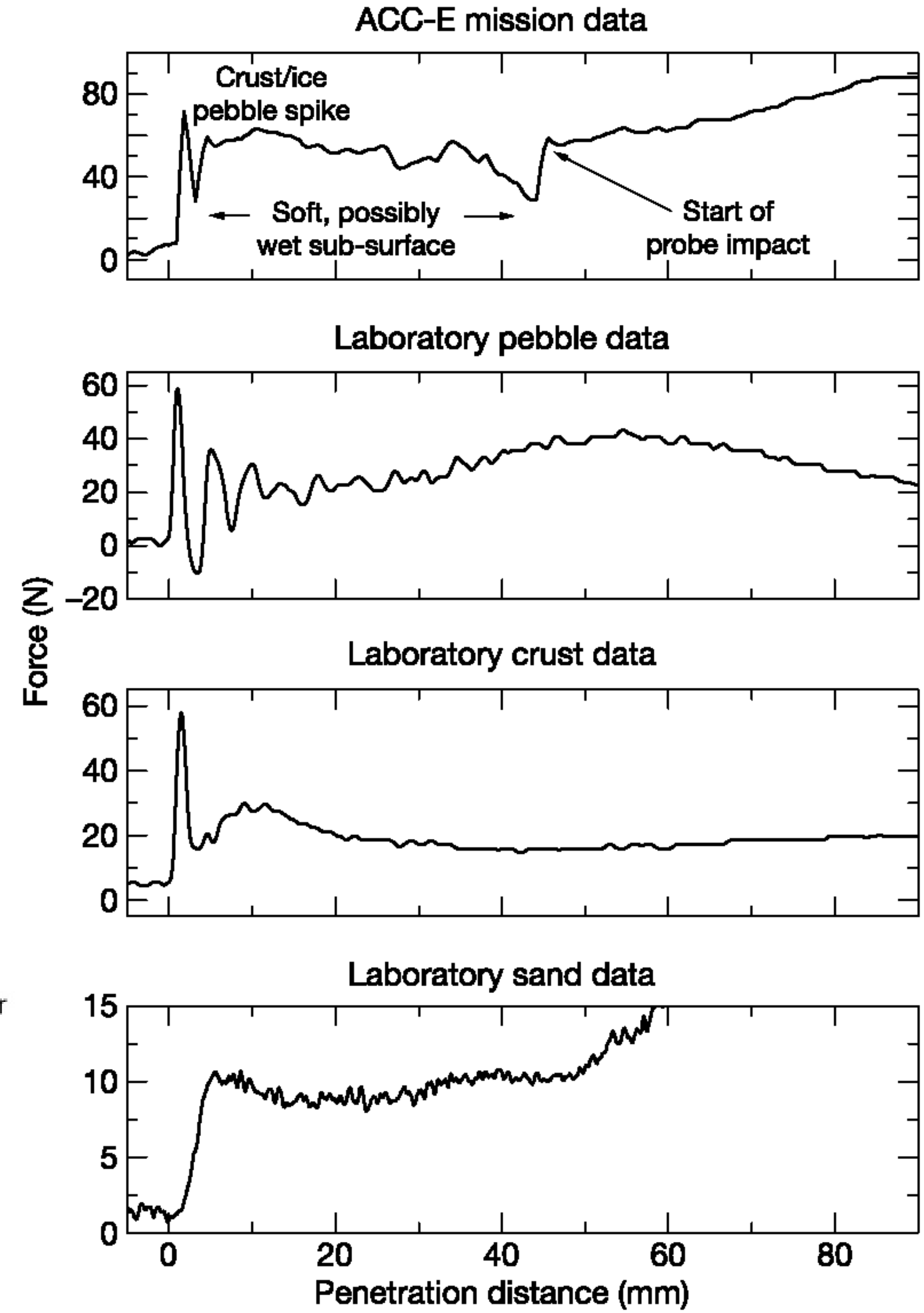




11 years later, flight data look nothing like lab tests! (Interpretation demanded in ~4 hours)

Suggests soft target (packed snow, soft clay, wet sand) plus initial spike (crust? pebble?) - crème brûlée ?!

# Atkinson SSP ACC-E data analysis



512 bytes.  1/20 of a second,

Carefully chosen bits !

Large dynamic range 2-2000N accommodated with logarithmic front-end amplifier.

Triggered at exactly the right moment by rms threshold detector (hardware)

Instrument design published in Meas. Sci. Tech paper, PhD thesis (+workshop proceedings)

Initial intepretation published in main (short) results paper  Zarnecki et al., Nature, 2005

PDS dataset

http://atmos.nmsu.edu/PDS/data/hpssp_0001/DATA/DESCENT/RAW/ACCE/SSP_ACCE_057_1_R_IMPACT.TAB      (see also .LBL and calibration data at data/hpssp_0001/CALIB/SSP_CAL.ASC )

Further analysis, Atkinson et al., Icarus, 2010  (~9 pages) + PhD thesis

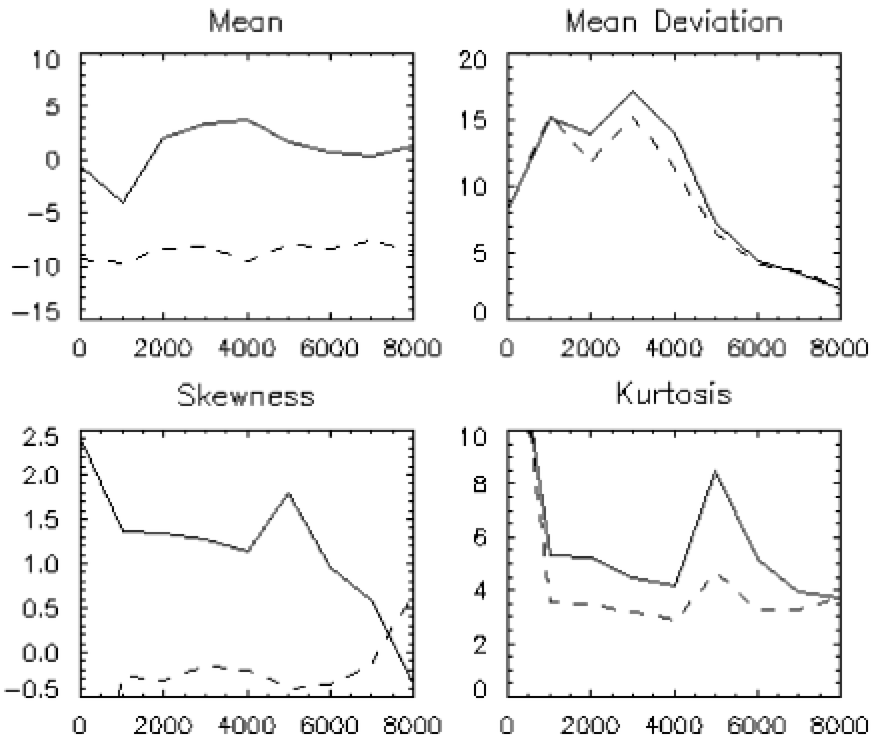
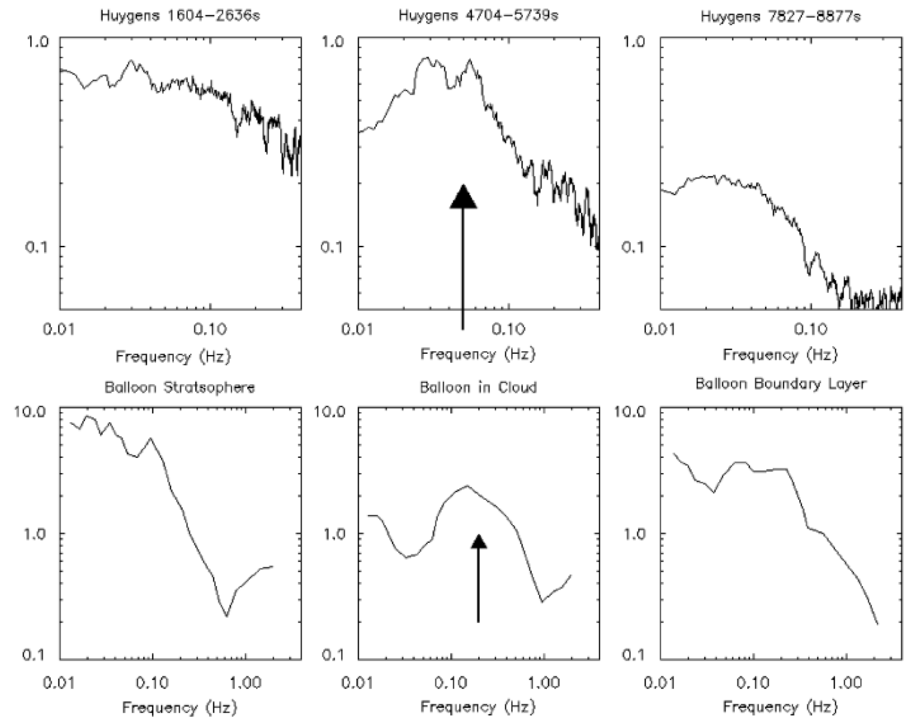
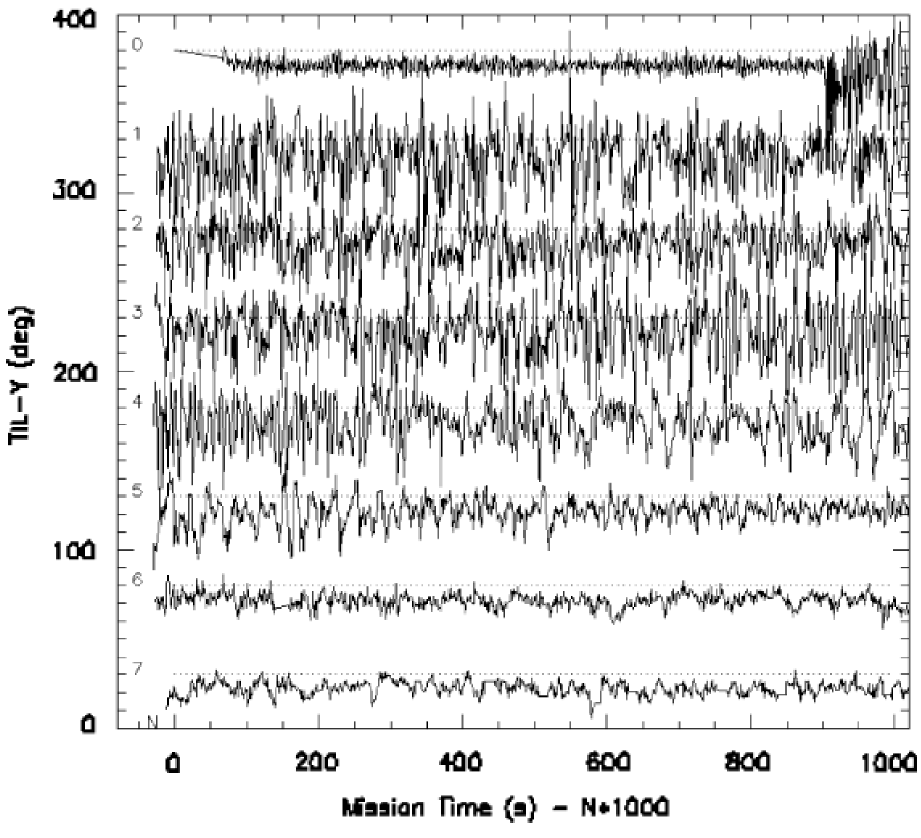



Fig. 15. Statistical moments of the tilt signals. Clockwise from top left are mean (°), square root of the variance (°), skewness and kurtosis, all as a function of mission time in seconds. In each panel, $X$ data are plotted with a solid line, $Y$ data with a dashed line. Notable features are the mean offset in $Y$, the substantial skewness of the $X$ data, and the bump in kurtosis at around 5000 s.



Data often defy expectations

Simple pendulum motion expected - SSP tilt sensor data from Titan showed much stronger high-frequency variations. Interpretation helped by comparison with balloon data on Earth. (experiment preparation would have been greatly enhanced by balloon flights – descent is not the same as the lab!)



Data Rights

Project-by-Project “Rules of the Road”

Strike a balance between timely delivery of data to the wider scientific community (and, indeed, the taxpaying public), and the need to “validate” the data.

“validate” is code for two things.

1. The need to properly assess data quality, determine/document obfuscating conditions, apply calibrations, reconstruct trajectory etc.
2. Give investigators who have invested years/decades of their careers ‘first crack’ at the data generated by the instruments or observations they conceived. (reference used to be made to ‘proprietary period’)

In the old days investigators sat on ‘their’ data indefinitely. Present posture is typically to deliver data to archive 3-6 months after acquisition ; one-off probe missions typically have required longer due to various complications.

Different rules have emerged for image data (since Spirit/Opportunity) – unvalidated images released more or less instantly, sometimes with caveats to discourage ‘scientific analysis”.   Has created  occasional complications for scientific publication.

(Huygens image data inadvertently released)

Engineering data typically not archived publically (a few exceptions).
I think this is bad.

Data sharing dynamics can be challenging – not entirely successful on Huygens ?     DTWG approach worked well (learning Galileo lesson) but inter-instrument comparisons hampered by proprietary considerations.

Some science results even now being derived and published, 10 years after mission !

## Data Restrictions can reduce Science Quality

Phoenix lander MET package included pressure sensors from Finnish Meteorological Institute. Pressure transducers have significant temperature sensitivity.

Pressure sensor was mounted near radio transmitter with significant power dissipation (heating). Details of the thermal architecture, nor the time history of transmitter dissipation, were not communicated to science team. (ITAR considerations cited).

Planetary Data System

Archive of raw and reduced science data products and documentation from planetary missions, paid for by NASA. Implemented at several Discipline Nodes (Rings, Geosciences, Atmospheres etc), at a number of institutions (e.g. Atmospheres Node is hosted at NMSU, Las Cruces).

Originally data distributed as hard copy/tape etc., then as CD-ROMs in the 1990s. Now essentially all online.

Data maintained in nonproprietary formats (text tables, .csv, some binary formats), platform-independent. Sometimes tools offered. Documentation typically
Systematic structure (label files, documentation)
Moving (slowly) towards a new standard – PDS4. XML-based.

Archive can (should!) include ground calibration data e.g. Aerodynamic database for trajectory/atmosphere reconstruction (see next talk)

ESA has similar ‘PSA’ (Planetary Science Archive). Uses same standards as PDS. Some data (e.g. Huygens) is mirrored on PSA and PDS.

Archiving is a nontrivial exercise.   Takes significant effort that must be budgeted.   Archive Plan, Software Interface Specification for archive products .
Delivery schedule, including peer review.

NASA mission and instrument proposals require archive plan.
PDS provides a cost estimation tool to ballpark FTE requirements for a given instrument.

Archive can (should) accommodate documentation

Data formats have traditionally been ASCII text-based  (independent of platform, proprietary software.  Some tools etc. – spreadsheets, IDL code etc. do get archived as better the tool that was used than nothing)    More voluminous data such as images in binary format.

PDS slowly migrating to PDS4 standard (xml based)

PRODUCER_ID = "HASI_TEAM"
PRODUCER_FULL_NAME = "ABOUDAN ALESSIO"
PRODUCER_INSTITUTION_NAME = "CISAS-UPD"
TARGET_NAME = "TITAN"

/* INSTRUMENT DESCRIPTION */
INSTRUMENT_ID = "HASI"
INSTRUMENT_NAME = "HUYGENS ATMOSPHERIC STRUCTURE INSTRUMENT"
INSTRUMENT_TYPE = "THERMOMETER"

DATA_QUALITY_ID = 1
DATA_QUALITY_DESC = "1 = high quality ... 5 = low quality"

INSTRUMENT_MODE_ID = "DESCENT"

/* DATA OBJECT DEFINITION */
OBJECT = TABLE
 INTERCHANGE_FORMAT = ASCII
 ROWS = 1811
 COLUMNS = 3
 ROW_BYTES = 36
 DESCRIPTION = "TEM sensor head 2 thermometer fine data during DESCENT mission phase"

PDS label file defines origin and format of data file

Example here is Dataset 4 from the analysis exercise

```
PDS_VERSION_ID                    = PDS3

/* FILE CHARACTERISTICS DATA ELEMENTS */
RECORD_TYPE                       = FIXED_LENGTH
RECORD_BYTES                      = 36
FILE_RECORDS                      = 1811

/* DATA OBJECT POINTER IDENTIFICATION ELEMENTS */
^TABLE                            = "HASI_L3_TEMD_FINE2.TAB"

/* INSTRUMENT AND DETECTOR DESCRIPTIVE DATA ELEMENTS */
FILE_NAME                         = "HASI_L3_TEMD_FINE2.TAB"
DATA_SET_ID                       = "HP-SSA-HASI-2-3-4-MISSION-V1.1"
DATA_SET_NAME                     = "HUYGENS HASI MISSION RAW AND
CALIBRATED DATA V1.1"
PRODUCT_ID                        = "HASI_L3_TEMD_FINE2"
PRODUCT_NAME                      = "HASI_L3_TEMD_FINE2.TAB"
MISSION_NAME                      = "CASSINI-HUYGENS"
INSTRUMENT_HOST_NAME              = "HUYGENS PROBE"
INSTRUMENT_HOST_ID                = HP
MISSION_PHASE_NAME                = "DESCENT"
PRODUCT_TYPE                      = EDR
START_TIME                        = 2005-01-14T09:10:33.453
STOP_TIME                         = 2005-01-14T11:38:09.703
SPACECRAFT_CLOCK_START_COUNT      = " 00:00:12.625"    /* DDB
time in HH:MM:SS.MS  ('-' for preT0) */
SPACECRAFT_CLOCK_STOP_COUNT       = " 02:27:48.875"
NATIVE_START_TIME                 =      12625    /* Elapsed time from T0
in milliseconds ('-' for preT0) */
NATIVE_STOP_TIME                  =    8868875
PRODUCT_CREATION_TIME             = 2006-07-05T09:09:22.000

PRODUCER_ID                       = "HASI_TEAM"
PRODUCER_FULL_NAME                = "ABOUDAN ALESSIO"
PRODUCER_INSTITUTION_NAME         = "CISAS-UPD"
TARGET_NAME                       = "TITAN"

/* INSTRUMENT DESCRIPTION */
INSTRUMENT_ID                     = "HASI"
INSTRUMENT_NAME                   = "HUYGENS ATMOSPHERIC
STRUCTURE INSTRUMENT"
INSTRUMENT_TYPE                   = "THERMOMETER”
```

```
/* DATA OBJECT DEFINITION */
OBJECT                     = TABLE
 INTERCHANGE_FORMAT        = ASCII
 ROWS                      = 1811
 COLUMNS                   = 3
 ROW_BYTES                 = 36
 DESCRIPTION               = "TEM sensor head 2 thermometer fine
data during DESCENT mission phase"

 OBJECT                    = COLUMN
  COLUMN_NUMBER            = 1
  NAME                     = "Time"
  UNIT                     = "milliseconds"
  DATA_TYPE                = ASCII_INTEGER
  START_BYTE               = 1
  BYTES                    = 8
  FORMAT                   = "I8"
  DESCRIPTION              = "NULL"
 END_OBJECT                = COLUMN


 OBJECT                    = COLUMN
  COLUMN_NUMBER            = 2
  NAME                     = "Resistance"
  UNIT                     = "Ohm"
  DATA_TYPE                = ASCII_REAL
  START_BYTE               = 10
  BYTES                    = 12
  FORMAT                   = "F12.5"
  DESCRIPTION              = "NULL"
 END_OBJECT                = COLUMN

 OBJECT                    = COLUMN
  COLUMN_NUMBER            = 3
  NAME                     = "Temperature"
  UNIT                     = "Kelvin"
  DATA_TYPE                = ASCII_REAL
  START_BYTE               = 23
  BYTES                    = 12
  FORMAT                   = "F12.5"
  DESCRIPTION              = "NULL"
 END_OBJECT                = COLUMN

END_OBJECT                 = TABLE
```

# Closing Thoughts

The data are the legacy of the mission.  Be diligent in defining your legacy.   Just because science analysis and archiving happens at the end doesn't mean it is a cost reserve to be raided by other elements which get into trouble first.

Treat your flight spare and EM well – you may need them!
Analog tests are great for rehearsing data handling and thinking.

Plan for the long haul.  Throw nothing away !
Huygens project start 1990.
Launch 1997
(need for instrument flight software updates ~2002)
Arrival 2005

Engineering model hardware needs to be maintained.  Compilers get updated – retain legacy versions?      People retire!