PHO ENIX
workflow tool for automatic science processing

Automatic processing of ESO-VLT science data

Reinhard Hanuschik
ESO Data Management & Operations   rhanusch@eso.org

ESO
European Organisation
for Astronomical
Research in the
Southern Hemisphere

Ground-based science observation

instable but accessible   *really unstable*

science ⊗ instrument ⊗ atmosphere

calibrations:
- bias
- flat
- wave

STD
STD

reduced science observation:

science ⊗ noise, artefacts

- ◆ VLT planned as "science data factory":
  - ◆ any data acquisition is standardized, in templates, observing blocks etc.
  - ◆ true for science, but also for calibrations
  - ◆ calibration plan includes calibrations taken for science reduction and for performance monitoring

- ◆ acquired calibrations used to remove ins+atm signature from science ("reduce")
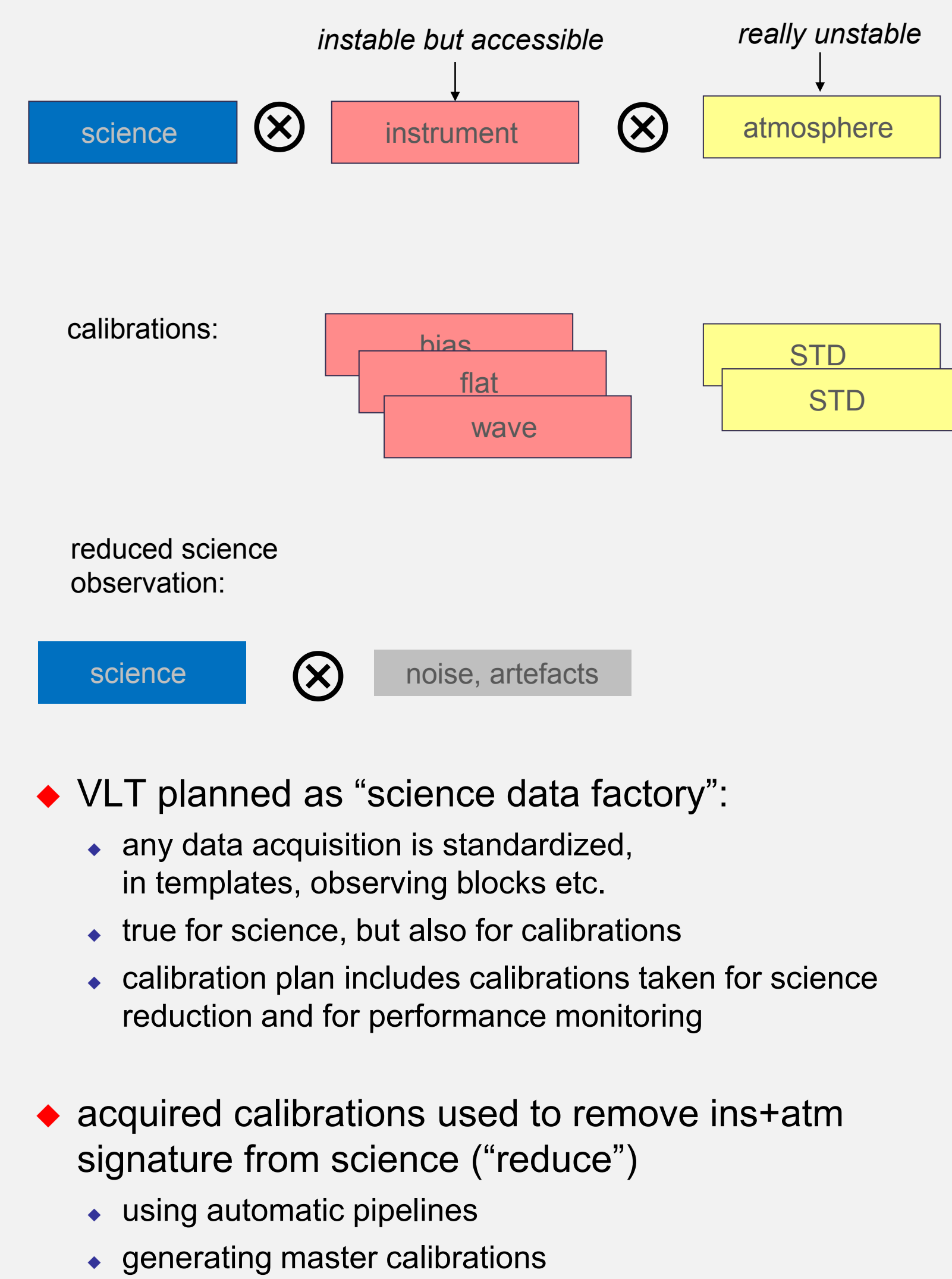 - ◆ using automatic pipelines
  - ◆ generating master calibrations

Model of phoenix science processing:

science raw data …

Science-grade data reduction

Two kinds of science data reduction can be defined:
- ◆ advanced
- ◆ science-grade

1. Advanced data products:
- ◆ reduction strategy optimized for science case
- ◆ quality: publication ready

provided by PIs to ESO, as **external data products**

2. Science-grade data products:
automatic pipeline reduction of science data
- ◆ removal of instrumental & atmospheric signature
- ◆ ready for science **analysis**
- ◆ **NO** optimization for science goal
- ◆ quality of reduction defined by calibration plan and pipeline

Why?
- ◆ this can be done *automatically*
- ◆ no knowledge about science goal needed
- ◆ general approach for archive science: this observation could be the answer to a different question than the PI asked
- ◆ called IDPs (internal data products)

provided by ESO, as **internal data products**

grown with care …

Pipeline processing at ESO: step 1 - CALIBs

All VLT instruments, most modes: pipeline-supported

- ◆ pipeline-process all calibrations
- ◆ produce
  - ◆ master calibrations
  - ◆ quality control parameters for trending
  - ◆ scores for quick health checks

- ◆ human supervision for:
  - ◆ certification/rejection of products
  - ◆ identification/analysis of issues

corresponding red scores
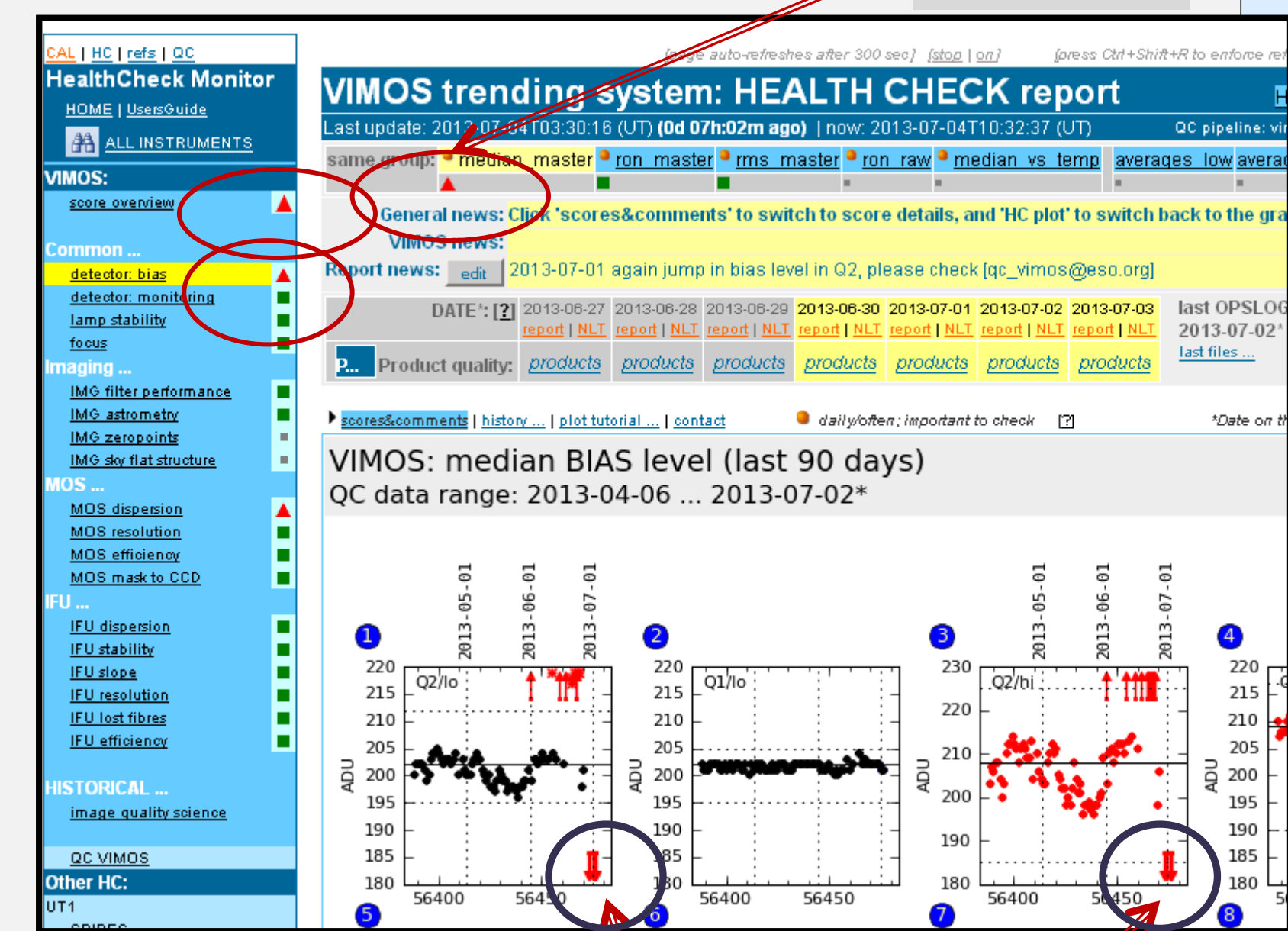
Health Check monitor for VIMOS, displaying & flagging an issue with one of the four CCDs

archive publication:
archive.eso.org

quality issue

Step 2: Processing SCIENCE with certified calibrations

- ◆ with all ins+atm effects measured & taken into account:
  - ◆ possible to automatically reduce the science data

- ◆ generate IDPs with science-grade pipelines:
  - ◆ selected for impact: spectroscopic instruments (X-Shooter, UVES, GIRAFFE)
  - ◆ also coming: imaging instruments (HAWK-I)
  - ◆ criteria: usefulness, community interest, code and instrument maturity

- ◆ output:
  - ◆ flux-calibrated spectra
  - ◆ standard output format (binary table with multiple columns for wavelength, flux, error etc.)
  - QC parameters (e.g. SNR, FWHM)

- ◆ QC parameters:
  - ◆ stored in a database
  - ◆ can be used to monitor health of the reduction process
  - ◆ e.g.: SNR, FWHM

dedication and skills added…

## Phoenix: workflow tool for automatic science processing

- ◆ IDPs generated with *phoenix*
- ◆ based on previous operational work:
  - ◆ uses certified master calibrations, collected over the years
  - ◆ retrieves previously created & archived associations
  - ◆ processes with latest pipeline version

- ◆ all processing is done in automatic mode

- ◆ monitoring:
  - ◆ health of the process is monitored …
  - ◆ … not the circumstances or the data quality

- ◆ output:
  - ◆ pipeline products
  - ◆ QC reports (for internal storage)
  - ◆ QC1 parameters (to assess quality of reduction process)

Process monitor. *Phoenix UVES reprocessing monitor for month 2008-01. All processing went fine, except for issues on 2008-01-10 where master calibrations were missing (marked "INCL"). Usually these issues are rare and not fixed.*

… delivered to enjoy.

First phoenix results: UVES Echelle IDPs

- ● 91,000 processed UVES spectra
  - ◆ time range: 2000-02-18 … 2013-05-19
  - ◆ POINT sources, no SLICER data (yet)
  - ◆ products are flux-calibrated if possible (for standard setups)

- ◆ the whole UVES-Echelle archive processed in one week
  - ◆ possible to repeat if needed (if data knowledge or algorithms improve)
  - ◆ process & tools ready to be used for science data from other instruments

- ◆ processing platform:
  - ■ Dell M820 using 30 cores
  - ■ condor batch queue system for pipeline and QC jobs
  - ■ processing time: 650,000 sec = 7.5 days (plus overheads for non-pipeline tasks)

- ◆ spectra visible in archive soon (September 2013)
  - ◆ future data to be added as a stream (once a month)
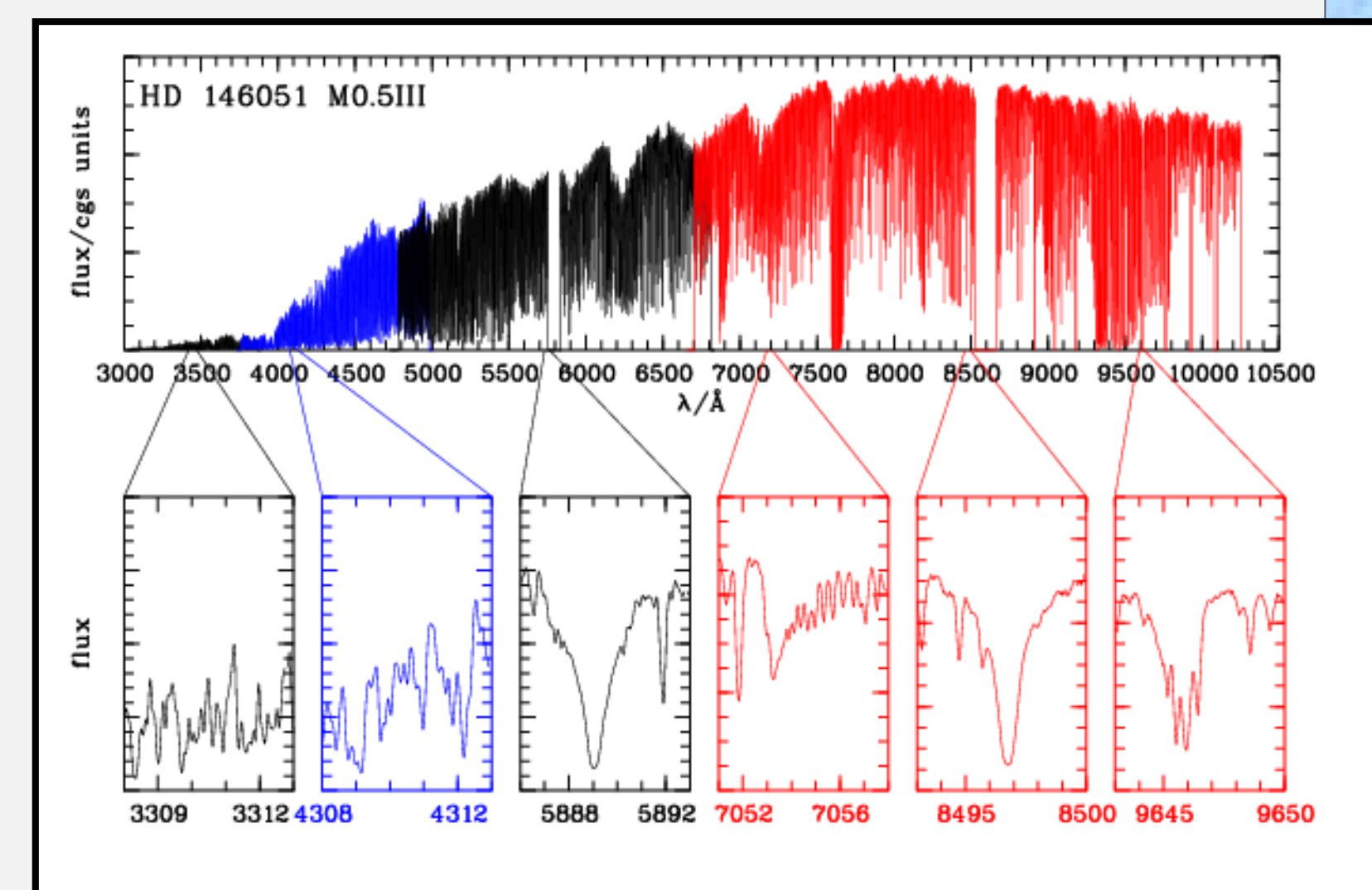
Example of phoenix science processing. *Fluxed UVES IDP spectra of M giant, composed of 4 individual UVES spectra ranging from 300nm to 1 micron (coded black-blue-black-red from left), with close-ups of major stellar features (SNR is about 100-200, the "noise" is actually stellar spectral structure). These spectra have been reduced **entirely automatically**.*