EPSC Abstracts,
Vol. 2, EPSC2007-A-00053, 2007
European Planetary Science Congress 2007

A new dataset validation system for the Planetary Science Archive

N. Manaud (1), J. Zender (2), D. Heather (2), S. Martinez (1)

(1) European Space Agency, ESAC, Villafranca del Castillo, 28080 Madrid, Spain (nicolas.manaud@sciops.esa.int) (2) European Space Agency, ESTEC, Keplerlaan 1, 2201 AZ Noordwijk, The Netherlands

The Planetary Science Archive is the official archive for the Mars Express mission. It has received its first data by the end of 2004. These data are delivered by the PI teams to the PSA team as datasets, which are formatted conform to the Planetary Data System (PDS). The PI teams are responsible for analyzing and calibrating the instrument data as well as the production of reduced and calibrated data. They are also responsible of the scientific validation of these data. ESA is responsible of the long-term data archiving and distribution to the scientific community and must ensure, in this regard, that all archived products meet quality. To do so, an archive peer-review is used to control the quality of the Mars Express science data archiving process. However a full validation of its content is missing. An independent review board recently recommended that the completeness of the archive as well as the consistency of the delivered data should be validated following well-defined procedures.

A new validation software tool is being developed to complete the overall data quality control system functionality. This new tool aims to improve the quality of data and services provided to the scientific community through the PSA, and shall allow to track anomalies in and to control the completeness of datasets. It shall ensure that the PSA end-users: (1) can rely on the result of their queries, (2) will get data products that are suitable for scientific analysis, (3) can find all science data acquired during a mission.

We defined dataset validation as the verification and assessment process to check the dataset content against pre-defined top-level criteria, which represent the general char-

acteristics of good quality datasets. The dataset content that is checked includes the data and all types of information that are essential in the process of deriving scientific results and those interfacing with the PSA database.

The validation software tool is a multi-mission tool that has been designed to provide the user with the flexibility of defining and implementing various types of validation criteria, to iteratively and incrementally validate datasets, and to generate validation reports.