Design and Validation of the Scientific Data Products for China's Tianwen-1 Mission

Xu Tan[1] · Jianjun Liu[1,2] · Xiaoxia Zhang[1] · Wei Yan[1] · Wangli Chen[1] · Xin Ren[1] ·
Wei Zuo[1] · Chunlai Li[1,2]

Abstract
The Tianwen-1 mission is China's first Mars exploration mission. To complete the scientific objectives of "orbiting, landing, and patrolling" in one mission, thirteen instruments for acquiring multi-disciplinary data are configured. The data products with rich ancillary data and flexible structures will facilitate the data application. To integrate the requirements of experts from disparate disciplines and leverage new information technologies, a complete redesign of data products is promoted in the Tianwen-1 mission. We discuss the whole process of data product generation from the data pipeline design to data validation. We design the various data products according to the application requirements of each data type, including radar echo data, spectrum data, image data, energy spectrum data, magnetic field data, meteorological data, and acoustic data. Since various error sources will exist in the generation chain of the data products, data validation is essential before release to the public; the validation activities are discussed at the end. We deliver these scientific products to the community in a timely manner, with ancillary information and quality information. This paper can provide practical reference for Tianwen-1 data application.

Keywords Tianwen-1 · Data products · Planetary Data System · Data validation

1 Introduction

The Tianwen-1 mission is China's first Mars exploration mission, which will conduct orbiting, landing, and roving in one mission. The spacecraft was launched on July 23, 2020, and will arrive at Mars in February 2021. The main scientific objectives of Tianwen-1 are to

✉ X. Tan
tanx@nao.cas.cn

✉ J. Liu
liujj@nao.cas.cn

[1] Key Laboratory of Lunar and Deep Space Exploration, National Astronomical Observatories, Chinese Academy of Sciences, Beijing, 100101, China

[2] School of Astronomy and Space Science, University of Chinese Academy of Sciences, 100049, Beijing, China

(Wan et al. 2020; Li et al. 2021a): (1) map the morphology and geological structure, (2) investigate the surface soil characteristics and water-ice distribution, (3) analyze the surface material composition, (4) measure the ionosphere and the characteristics of the Martian climate and environment at the surface, and (5) perceive the physical fields (electromagnetic, gravitational) and internal structure of Mars. The spacecraft comprises an orbiter and a lander/rover (Ye et al. 2017), which will acquire a large amount of scientific data to complete the scientific objectives. For the purpose of long-term preservation of these data, a scientific data archive should be developed.

Since the 1960s, many Solar System exploration missions have been performed by several different national and international space agencies (Albee et al. 1998; Saunders et al. 2004; Grotzinger et al. 2012; Banerdt et al. 2020). All the agencies are aware of the need of an archive to preserve scientific data sets. The first data archive system is Planetary Data System (PDS), that is created and maintained by NASA in response to scientists' requests for improved availability of planetary data from NASA missions, with increased scientific involvement and oversight (McMahon 1996). Then, ESA establish Planetary Science Archive (PSA) (Besse et al. 2018; Macfarlane et al. 2018) for its own planetary exploration missions, such as Mars Express(Schmidt 2003). The Japanese Aerospace eXploration Agency (JAXA) also organized its science data sets exploration of the Solar System in an archive (Haruyama et al. 2012; Hoshino et al. 2010), as have the Indian Space Research Organization (ISRO) (Goswami and Annadurai 2009). Nearly all the archive systems share the same set of standards-PDS data standards, but have different implementations.

PDS is a distributed system that has eight discipline nodes (Guinness et al. 1996; Eliason et al. 1996; Grayzeck et al. 1996). Each discipline node is managed by relevant universities or research institutions. This distributed architecture solves the problem of digital preservation of scientific data for long-term use, and the open-source architecture has allowed PDS data standards to gradually become the widely accepted data standards within planetary science community. PDS data standards has two incarnations including PDS version 3 (PDS3) and PDS version 4 (PDS4). PDS3 is the most mature data standards of the PDS archiving system, that has been used for several decades. But now most NASA and ESA planetary mission, such as MAVEN (NASA) (Jakosky et al. 2015), Osiris-REx (NASA) (Lauretta et al. 2017), ExoMars2016 (ESA) (Vago et al. 2015), and BepiColumbo (ESA/JAXA) (Benkhoff et al. 2010), gradually use PDS4 to store and disseminate their science data.

Ground Research and Application System (GRAS) (Liu et al. 2018) is mainly responsible for the acquisition, generation, archiving and releasing of scientific data from China National Space Administration's (CNSA's) planetary mission. To ensure the compatibility with the international planetary mission, in 2007 the first of the CNSA's planetary missions, Chang'e-1 (Ouyang et al. 2010), started to process and deliver data by GRAS using PDS3, as have the Chang'e-2 and Chang'e-3 mission (Li et al. 2015; Zuo et al. 2014; Tan et al. 2014). However, after more than ten years of operation in GRAS, PDS3 began to show some limitations. For example, PDS3 has restrictions on data format, and only a few pieces of software, such as NASA View and ISIS, can be compatible with PDS3. Most special data processing software within the discipline cannot read the PDS3 format directly; the data can only be parsed after format conversion. Moreover, the labels used to describe the scientific data were implemented in the Object Description Language (ODL) (used only by PDS). ODL allows for a human-readable "KEYWORD = VALUE" structure, which is not friendly for software to parse or use efficiently. All the limitations promote a complete redesign of data products in the coming mission, Tianwen-1.

With the development of computer technology and in response to higher requirements for data products, we need more flexible data structure in Tianwen-1 mission. PDS4 (Hughes

et al. 2014; Jet Propulsion Laboratory 2019a,b), the latest incarnation of PDS, is adapted to the design of Tianwen-1 data products for several reasons: First, PDS4 is designed using the principle "Everything is a product". A digital object (e.g., a table of science measurements, a document, an image, etc.), a physical object (e.g., satellites, samples, etc.), and even an idea can be a product. Second, PDS4 just provides four basic structures (Array, Table_Base, Parsable_Byte_Stream, and Encoded_Byte_Stream) to store data objects, and the simple definition of the data format makes the products more flexible and easily available to the scientific community. Third, PDS4 allows copies of data to be archived in supplemental formats, and it can well meet the requirements of the experts from different disciplines. Finally, PDS4 labels replace the use of ODL with eXtensible Markup Language (XML). XML is a widely used international standard that can be read by both humans and software. The use of labels written in XML and constrained by an XML schema and a set of Schematron rules helps ensure that products are thoroughly and consistently documented and that their metadata are available to the wide range of third-party software that reads and writes in XML. On account of above reasons, based on the widely used data standards, we design Tianwen-1 data products and develop our individual data model and validation tool in accordance with Tianwen-1 data pipeline, all the data are validated against the definition in the Tianwen-1 data model and PDS4 information model.

Data generation in GRAS includes several procedures from the prelaunch phase to the in-orbit phase: categorizing science data products, designing data product formats, determination of data processing algorithms via cooperation with the instrument team, validation of data processing algorithms, data acquisition, data processing, validation of data products, and release to the public. In this paper, we present the detailed generation process of Tianwen-1's data products. Thirteen instruments carried in the Tianwen-1 mission are introduced in Sect. 2, and data characteristics of these instruments are summarized. In Sect. 3, the data preprocessing pipeline is described; the multi-antenna data fusion, calibration and indicators of the errors are important parts of the pipeline, and newly developed methods are also introduced in this section. In Sect. 4, the data products' baseline design is presented, we design the Tianwen-1 data model and give the data structures of each data type. The validation processes and tools developed for use on Tianwen-1 are discussed in Sect. 5. This paper can provide practical reference for Tianwen-1 data application.

2 Tianwen-1 Instruments and Data Characteristics

To complete the scientific exploration mission, thirteen instruments are configured (Zou et al. 2021), in which seven instruments are onboard the orbiter, mainly to explore the Martian space environment (including induced magnetosphere, ionosphere, and etc.), gravity field and atmosphere, surface topography, rock composition, soil characteristics, and distribution of water ice and minerals. The instrument configuration and data descriptions of the orbiter are listed in Table 1 (Yu et al. 2020; Kong et al. 2020; Li et al. 2021c; Liu et al. 2020).

Additionally, six instruments are onboard the rover to perform high-accuracy and high-resolution investigations of the Mars landing site, mainly including investigation of morphological features, subsurface geological structures, minerals and chemical composition, climate and environmental characteristics, and spatial magnetic field. The instrument configuration and data descriptions of the rover are listed in Table 2 (Peng et al. 2020; Zhou et al. 2020; Du et al. 2020).

It can be seen from the instrument configuration and data description that, to accomplish the scientific objectives, various data types and more complicated data contents are acquired

Table 1 Instrument configuration, data types, data comments and main data applications of the Tianwen-1 orbiter

Instrument	Data comment	Data type	Main data applications
Moderate Resolution Imaging Camera (MoRIC)	RGB color image data of Mars' surface. Resolution is higher than 100 m@400-km orbit altitude. Effective pixel number is 4096×3072, and the pixel size is 5.5 um×5.5 um	Image data	Map global remote sensing images of Mars; explore the topography and its changes on Mars
High Resolution Imaging Camera (HiRIC)	High-resolution image data of key areas on the surface of Mars. Includes gray images and color images (including RGB and Near-infrared band). Resolution of gray images is better than 0.5 m@265-km orbit altitude. Resolution of color images is better than 2.0 m@265-km orbit altitude.	Image data	Perform high-accuracy imaging and detailed survey of typical areas; investigate the characteristics of Mars' topography and its geological structure.
Mars Orbiter Scientific Investigation Radar (MOSIR)	Low-frequency and high-frequency dual-polarized echo data and Very Low Frequency (VLF) electromagnetic data. Center frequency of the low-frequency channel is 15 MHz; center frequency of the high-frequency channel is 40 MHz.	Radar echo data	Explore the subsurface structure and underground water-ice; Interplanetary VLF radio spectrum detection in the Earth-Mars transfer orbit.
Mars Mineralogical Spectrometer (MMS)	High-resolution spectrum data of Mars' surface. Including Visible near infrared band from 0.45 μm to 1.05 μm, Spectral resolution is better than 10 nm, and Near medium wave infrared band from 1.00 μm to 3.40 μm, spectral resolution is better than 12 nm@1.0 μm∼2.0 μm, 25 nm@2.0 μm∼3.4 μm	Spectrum data	Investigate mineral composition and distribution of Mars
Mars Orbiter Magnetometer (MOMAG)	High-accuracy three-axis magnetic field data of Martian space. Accuracy is better than 0.1 nT. Resolution is better than 0.01 nT.	Magnetic field data	Detect the Martian space magnetic environment; in tandem with RoMAG of the rover, acquire the current of the ionospheric dynamo by inversion, and study the conductivity of the Martian ionosphere
Mars Ion and Neutral Particle Analyzer (MINPA)	Energy spectrum of ions (2.8 eV∼25.9 keV) and energetic neutral atoms (ENAs) (50 eV∼3 keV) in near-Mars space environment and Earth-Mars transfer orbit.	Energy spectrum data	Study the upper atmosphere of Mars and interactions with the solar wind
Mars Energetic Particles Analyzer (MEPA)	Energy spectrum of solar energetic electrons (0.1 MeV∼12 MeV), protons(2 MeV∼100 MeV), and heavy ions (25 MeV∼300 MeV) in near-Mars space environment and Earth-Mars transfer orbit.	Energy spectrum data	Measure the energetic charged particles in the Mars space and in the interplanetary space between Earth and Mars; map the spatial distribution of the energetic particle radiation

Table 2 Instrument configuration, data types, data comments and main data applications of the Tianwen-1 rover

Instrument	Data comment	Data type	Main data applications
Navigation and Terrain Camera (NaTeCam)	Terrain data of patrolling area. Color image with RGB bands, effective pixel number is 2048×2048, the pixel size is 5.5 μm×5.5 μm	Image data	Acquire the information of Mars terrain, including slope, undulation, roughness, and other terrain features; high-resolution terrain construction of patrolling areas
Multispectral Camera (MSCam)	High-resolution multispectral image of Mars surface in visible and near infrared bands. Spectral range: 480 nm, 525 nm, 650 nm, 700 nm, 800 nm, 900 nm, 950 nm, 1000 nm	Image data	Acquire material composition and its distribution on the Martian surface
Mars Rover Penetrating Radar (RoPeR)	Ultrawideband full-polarized echo data of the Martian surface and subsurface. Center frequency of low-frequency channel is 55 MHz, center frequency of high-frequency channel is 1300 MHz.	Radar echo data	Detect the subsurface geological structure of roving area
Mars Surface Composition Detector (MarSCoDe)	1. High-resolution spectrum data of target of patrolling area, Spectral range: 240 nm–850 nm. 2. Short-wave infrared spectrum data of target of roving area, Spectral range: 850 nm–2400 nm. 3. Microscopic image of target. Spectral range: 900 nm∼1000 nm.	Spectrum data	Analyze the composition of the chemical elements of Martian surface materials. Perform mineral analysis and rock recognition on the Martian surface.
Mars Rover Magnetometer (RoMAG)	High-accuracy three-axis magnetic field data of the Martian surface. Accuracy is better than 0.1 nT. Resolution is better than 0.01 nT.	Magnetic field data	Detect the magnetic environment of roving area; in tandem with MOMAG of orbiter, acquire the current of ionospheric dynamo by inversion, and study the conductivity of the Martian ionosphere
Mars Climate Station (MCS)	1. Atmosphere temperature, pressure and wind field data of the landing site and patrolling area. 2. Waveform data over a broad frequency band, from low frequency (20 Hz) up to high frequency (20 kHz).	Meteorological data & Acoustic data	Investigate the climate and environmental characteristics, aiming to simultaneously measure the atmosphere temperature, pressure, wind field and sound on Martian ground

in the Tianwen-1 mission. The data characteristics are summarized as follows, and they are considered in the following design of the data pipeline and data products.

1. Very precious data packet due to limited downlink resources

The longest transmission distance is 395 million kilometers, it takes approximately 22 minutes to transmit the data to Earth, the data downlink resources are very limited because of the longer distance, and most data can only be transmitted once, so each data packet is very precious for this mission.

2. Multidisciplinary data

The data detection contents include: atmosphere, ionosphere, magnetosphere, surface and subsurface etc., which covers various discipline (e.g. atmospheres, geosciences, plasma interactions, cartography, and imaging sciences). It is difficult to integrate disparate discipline experts' requirements concerning the data format.

3. Various types of proprietary ancillary data

There are four sources of Tianwen-1 ancillary information: 1) engineering parameters of instruments used to describe the working state, transmitted down in the same package as scientific data; 2) coefficients or parameters obtained by processing, such as geometric information; 3) the inherent parameters related to the detector, such as the technical specifications and installation parameters of the detector; and 4) ephemeris and attitude used to determine the observation geometry. All of these require a data model to describe the relationship between these parameters.

4. The data processing method is new with high uncertainty

The data modeling process is subject to uncertainty due to a lack of knowledge concerning the complicated Martian environment. Although all instruments have been comprehensively calibrated on the ground, differences still exist between measurements made on Mars and those made on the ground. For this purpose, some instruments, such as MMS, MarSCoDe etc., employ onboard calibration systems. Therefore, for the design of data products, the data structure must be flexible enough to support implementation of new data processing methods; for the data validation, the component of validation not only include the inspection of the product itself but also the validation of processing algorithms.

3 Scientific Data Preprocessing Pipeline

To complete the goals of "orbiting, landing and roving" in one mission, multiple functions and various working modes of instruments are implemented, which makes the processing more complicated.

We recognize three broad levels for categorizing scientific data preprocessing, Level 0, Level 1 and Level 2. Level 0 processing is mainly frame/packet decoding according to the CCSDS (Consultative Committee for Space Data Systems) standard and is further divided into Level 0A and Level 0B. Level 1 processing mainly generates original data from an instrument after reformatting, and Level 2 processing mainly provides the calibrated data, which makes values independent of the instrument. Figure 1 illustrates the data flow of the preprocessing pipeline; all the details are provided in the following sections.

The following aspects are considered within the pipeline, and some effects are coupled as argued in Sect. 2.

1. The whole Tianwen-1 dataset contains five data sources: four ground antenna (the 50 m antenna and 40 m antenna of Miyun Station, 70 m antenna of Wuqing Station, and 40 m antenna of Kunming Station) and one antenna array of GRAS with combined data. Ensuring the integrity of the exploration data can create more opportunities to address the scientific questions. Therefore, the combination and deduplication method of multisource data are designed in Sect. 3.2.
2. Coordinate observations is a highlight of this mission, for example, the RoMAG in rover will combine with the MOMAG in orbiter getting two-site observation data. Hence, the instruments of the same type adopt unified data processing principles. Beside this, for

a reliable conversion from the electromagnetic signal to reflectance, radiance values or other physical quantity, the processing method and parameters will be adapted according to in-flight calibration. Detailed methods of each instrument are given in Sect. 3.4.

3. The scientific exploration of the orbiter will start in the Earth-Mars transfer orbit. The requirements of geometric information include the Earth-Mars transfer orbit, global reconnaissance orbit and rover patrolling route. Different coordinate systems are applied in the pipeline, which is detailed in Sect. 3.5.

4. To oppose ambiguous observations, quality assessment (QA) is an essential process before data release, which is used to provide information on the validity and integrity by flagging data processing. The design of quality assessment information is discussed in Sect. 3.7.

3.1 Virtual Channel Frame Extraction (Level 0A Processing)

After the ground demodulation, frame synchronization, descrambling, and RS/LDPC decoding, the data of each antenna has formed one Raw data file that contains a sequence of fixed-length frames of the orbiter data and lander data. The data frame structures from the orbiter and the rover are different, and the unpacking methods for the two probes are completely different. Therefore, we need to extract the orbiter data and lander data from the Raw data according to the spacecraft ID first. Then, the virtual channel frames are extracted from the data field of the Raw frames according to the virtual channel identifiers (VCID). Each instrument in the orbiter occupies one virtual channel, and therefore the data from a single instrument can be extracted directly after channel frame extraction. However, for the rover instrument, since RoPeR, MarSCoDe, RoMAG and MCS share a virtual channel, this level of processing cannot distinguish the RoPeR, MarSCoDe, RoMAG and MCS data. Then, the quality information of each frame is added to form the Level 0A data products.

3.2 Instrument Packet Extraction (Level 0B Processing)

The instrument packet extraction method is mainly extracting packets from the virtual channel frame data field by the Application Process Identifier (APID). The APID indicates instruments and their working modes; therefore, the Level 0B data products are a sequence of fixed-length packets, and they are generated with several files according to different working modes of each instrument. Particularly, because of the limited data transmission resources, to facilitate data transmission or storage, most instruments' data (e.g., MoRIC, HiRIC, MMS, MarSCoDe, MCS, MINPA, MOSIR, RoPeR, NaTeCam and MSCam) are compressed when transmitted to the ground, and decompression is carried out in this level.

In addition to this, to improve the quality of data, we design the combination and deduplication method of virtual channel frames in this level. We define the default antenna data as the basic data, searching the same frame of other antenna data by time and frame count. The principle of deduplication is to choose the source package with the best quality, and if the quality information is the same, the combined data of the antenna array is selected first. After the above treatment, a packet of Level 0B is formed by a combination of several antennas' virtual channel frames. This will greatly improve the integrity and quality of Level 0B.

3.3 Instrument Data Reformatting and Reorganization (Level 1 Processing)

The purpose of this processing is to form the intuitive original data from an instrument after reformatting and reorganizing according to observation cycles. If temperature, voltage, and
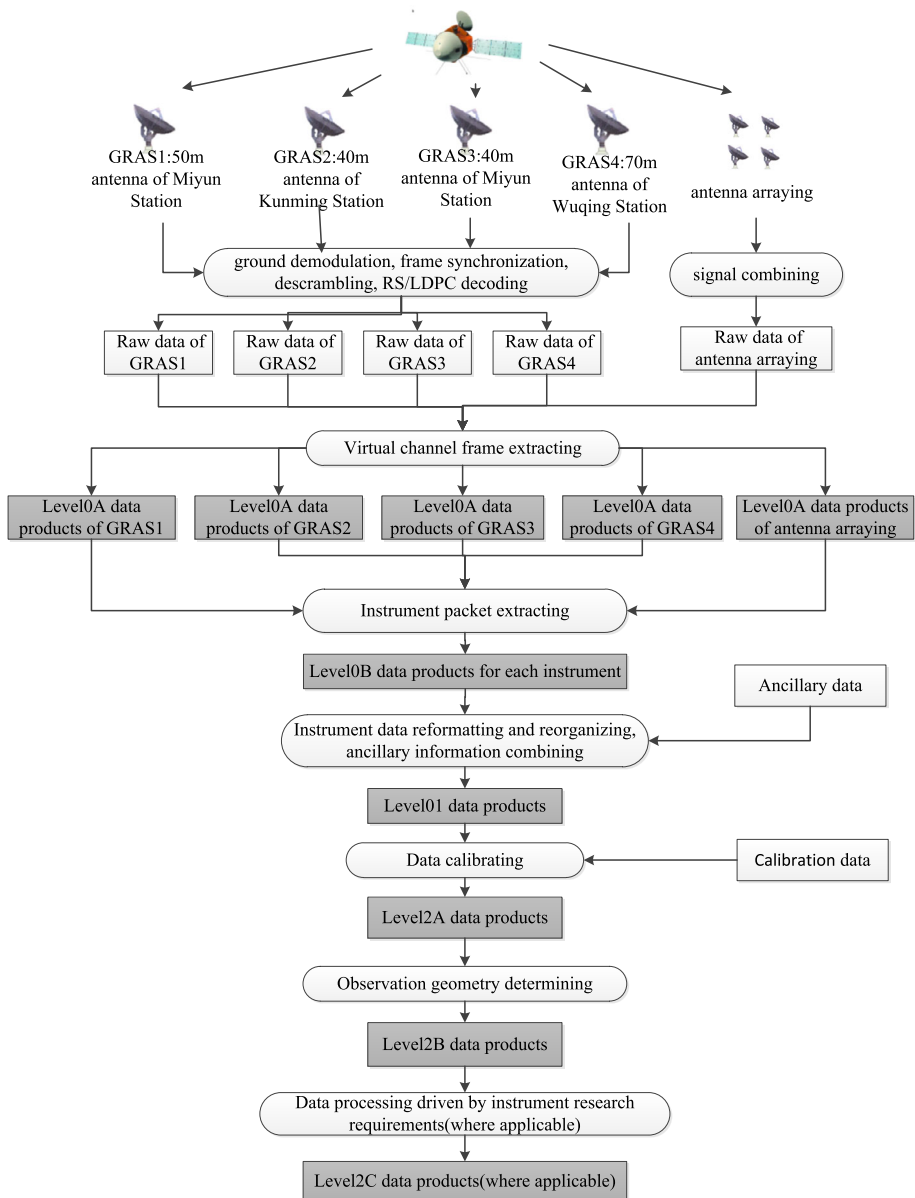
Fig. 1 Data flow in the Tianwen-1 preprocessing pipeline. The data products categorized according to processing level are represented by the boxes with a gray background. The rounded boxes indicate data processing

other instrument engineering parameters are attached in Level 0B data, numerical conversion should be processed in this level.

The observation cycles of the orbiter data and rover data are defined. For the orbiter, the observation cycle is an earth day in the Earth-Mars transfer orbit. In the Mars orbiting phase, the observation cycle is a circle of the orbit around the Mars, started with the apareon. For

Table 3 Detailed Level 2A data processing

Instrument	Data processing
MoRIC	Dark current calibration, radiometric calibration
HiRIC	radiometric calibration
MOSIR	Accumulation removal, Automatic gain control (ACG) correction
MMS	Dark current calibration, temperature calibration, non-uniformity correction, radiometric calibration and image rotation
MOMAG/RoMAG	Zero drift correction, orthogonality correction
MINPA	Conversion from counts to differential particle flux
MEPA	Conversion from sensor ADC value to sensor energy and energy spectrum correction
NaTeCam	Dark current calibration, relative radiometric calibration
MSCam	Dark current calibration, relative radiometric calibration, and calculation of absolute radiation calibration factor
RoPeR	not performed
MarSCoDe	Dark current calibration, absolute radiometric calibration
MCS	Conversion from ADC counts of sensor to physical quantity of sensor including temperature, pressure, and wind field data

the rover, the observation cycle is a SOL, which is the number of complete solar days on Mars landing. The landing day therefore is SOL zero.

3.4 Data Calibration (Level 2A Processing)

The data processing methods in Level 2A are determined in pre-flight calibration, also calibration parameters are calculated in the ground calibration. Moreover, in order to get a reliable conversion, most instruments, such as MMS, MarSCoDe, etc., employ onboard calibration unit. Thus the calibration parameters used for processing will be adjusted according to the results of in-flight calibration. The processing procedures of each instrument are given in Table 3. The more detailed information about calibration is published in the overview of each instrument (Tang et al. 2020; Yu et al. 2020; Kong et al. 2020; Liu et al. 2020; Liang et al. 2021; Meng et al. 2021; Fan et al. 2021; He et al. 2021; Du et al. 2020; Peng et al. 2020; Zhou et al. 2020).

3.5 Observation Geometry Determination (Level 2B Processing)

The principle of determining observation geometry is to give the geometry information of the observation point through the transformation between various reference frames, the parameters about ephemeris, attitude and sensor are as input, the rotation and shape of the Mars are also taken into consideration. There are five principal coordinate systems used to represent the data in Tianwen-1 mission. The definition of the coordinate systems are as follows:

- Mars Coordinate System: a body-fixed non-inertial frame associated with Mars.
- Spacecraft Coordinate System: a non-inertial frame associated with the orbiter or rover.
- Landing Station Coordinate System: a topocentric frame located on the surface of Mars, whose center is the landing station, the z-axis is pointed up, the x-axis is East, and the y-axis is North.

- Rover Station Coordinate System: a topocentric frame located on the surface of Mars, whose center is the patrolling station, the z-axis is pointed up, the x-axis is East, and the y-axis is North.
- J2000 Coordinate System: Geocentric mean equator coordinate system for epoch J2000.0, whose center is Mars in the Earth-Mars transfer orbit and the Sun in the global reconnaissance orbit.

Mars orbital data are rendered in the Mars, J2000 and Spacecraft (orbiter) coordinate systems according to application requirements. The in-site data are also provided in the spacecraft (rover), landing station and rover station coordinate system. Cartesian representations are used for all the coordinate systems. Both the orbiter location and rover location are essential information in corresponding products. The rest of the geometric information of each instrument are given in Table 4.

3.6 Data Processing Driven by Instrument Research Requirements (Level 2C Processing)

There is no strict specification to define the processing in this level, and we provide some distilled results from Level 2A and Level 2B data products. For MoRIC and NaTeCam, the color images of Level 2A data products are Bayer patterns, so some commercial software is not compatible. To display the image color information intuitively, color restoration of Bayer pattern images and color correction are implemented in this level.

Because background data will affect the accuracy of measurement data, the remanence of MOMAG and RoMAG is deducted in this level. The deduction method uses the magnetic field data of two probes in the same frequency band, based on cross-correlation spectrum analysis to confirm the background magnetic field signal.

To obtain the meteorological parameters, we build a model for MCS to make a conversion from the physical quantity of the sensor to atmospheric temperature, atmospheric pressure, wind speed and wind direction.

3.7 Design of Quality Assessment Information

This procedure adds quality indicators to the original observation without modifying or removing it. Users can give attention to quality information to obtain the validity and integrity of the record of data products collected from the Tianwen-1 mission. The design of quality information is as follows:

A one-byte quality code is allocated to each frame/packet or record to indicate properties of the original data, the loss of integrity, an exception message sent by the instruments and any errors in data caused by abnormal data processing. Each characteristic element above is represented by the value of the corresponding bit ('0' is normal, '1' is abnormal). The most significant bit represents the validity of data from the previous level through the 'logic-AND' operation. Moreover, different considerations are given for the validity of each level of data.

- Data transmission related error are assessed in Level 0A, including the frame synchronization, bit slip, RS (Reed-Solomon) decoding, and LDPC (Low-density Parity-check) decoding.
- The data integrity in Level 0B, includes the data padding, packet dropout and validity of data decompression.
- Instrument working-related errors, which can affect scientific data, are judged in Level 1 including abnormal temperature, abnormal voltage, raw data out of range, etc.

Table 4 Detailed geometric information of each instrument

Instrument	Type	Coordinate system	Geometric information
MoRIC/ HiRIC	Observation Location	Mars	Longitude and latitude that correspond to the grid points of the image.
	Angle		Solar incidence angle, Solar azimuth angle
MOSIR	Spacecraft Position	J2000	Orbiter location (X, Y, Z)
	Angle		Solar incidence angle, Solar azimuth angle
	Spacecraft Attitude	J2000	Orbiter attitude (roll, pitch, yaw)
MMS	Angle		Solar incidence angle, Solar azimuth angle, Instrument incidence angle, Instrument azimuth angle, Phase angle
MOMAG	Spacecraft Position	Spacecraft (orbiter)	Sensor position (X, Y, Z),
	Spacecraft Attitude	Spacecraft (orbiter)	Sensor attitude (roll, pitch, yaw)
	Angle		Solar incidence angle and solar azimuth angle
MINPA/ MEPA	Spacecraft Position	J2000	Orbiter position in the Earth-Mars transfer orbit and global reconnaissance orbit
		Mars	Orbiter position in the global reconnaissance orbit
	Angle		Solar incidence angle, solar azimuth angle, incidence angle between the symmetry axis of the sensor FOV and the incident illumination.
NaTeCam/ MSCam	Observation Orientation	Rover station	Observation vector of four corner points and center points of the camera
	Spacecraft Attitude	Landing station	Rover attitude (roll, pitch, yaw)
	Angle		Solar incidence angle and solar azimuth angle
RoPeR	Spacecraft Attitude	Landing station	Rover attitude (roll, pitch, yaw)
	Angle		Solar incidence angle and solar azimuth angle
MarSCoDe	Spacecraft Attitude	Landing station	Rover attitude (roll, pitch, yaw)
	Observation Orientation	Rover station	FOV center observation vector, exterior orientation element
	Angle		Solar incidence angle and solar azimuth angle
RoMAG	Spacecraft Position	Spacecraft (rover)	Sensor position (X, Y, Z),
	Spacecraft Attitude	Spacecraft (rover)	Sensor attitude (roll, pitch, yaw)
	Angle		Solar incidence angle and solar azimuth angle
MCS	Angle		Solar incidence angle and solar azimuth angle

Table 5 Comparison between Tianwen-1 data level definition and PDS data level

Tianwen-1 data level	Tianwen-1 data level definition	PDS data level
Level 0A	Virtual channel frame data received at a single ground antenna from an instrument after virtual channel frame extraction.	Telemetry
Level 0B	Packet data from an instrument after instrument packet extraction.	Telemetry
Level 1	Original data from an instrument, after instrument data reformatting and reorganization.	Raw
Level 2A	Calibrated data	Calibrated
Level 2B	Calibrated data with geometric information	Calibrated
Level 2C	Further distilled data	Calibrated/Derived

- The comprehensive evaluation of data validity in Level 2 includes comprehensive analysis of the above factors and processing method-related errors, for example, when the denominator is 0, the data are out of limits, etc.

4 Science Data Products

4.1 General Design

According to the data preprocessing pipeline, the data products are categorized into six levels: level0A, level0B, level1, level2A, level2B and level2C. The data level definitions of Tianwen-1 have the same principle as PDS, but have slightly different implementations. The comparison to PDS data level definition (Jet Propulsion Laboratory 2019c) is shown in Table 5. PDS does not archive telemetry data, however, the Level 0A and Level 0B data are processed and archived in GRAS.

Each Level 0 product consists of one frame file and associated processing report file. The frame file of each instrument is stored in a binary file, separated according to the instrument working mode, detector, compression mode, etc. The processing report file is a text file that provides the source products information, statistical information concerning CCSDS packets, data check information, and exceptions that occurred. The statistical information concerning the CCSDS packets mainly contains the numbers of total packets, packets of each virtual channel and packets of each Application Process Identifier (APID). The data check information includes the check results of Cyclic Redundancy Check (CRC), Reed-Solomon Check and Low-density Parity-check (LDPC). The exception information mainly contains frame numbers of nonsynchronization, frame(/packet) count discontinuity (e.g., packet count jumps from 1 to 10), abnormal frame length (incomplete or too long), and reasons behind the exception of data decompression. With the recorded information of data quality, the data processing report can improve the efficiency of data analysis.

The Level 1 and Level 2 product are compliant with PDS4 standards for formatting and labeling files. Each product is the combination of one or more data objects and a data label. Typically, a data object is a file containing the scientific measurements (e.g. a single image or a table etc.); and a label is an XML file, which is the concatenation of metadata, with filename extensions of .01L, .2AL, .2BL or .2CL. The metadata provide the information needed to identify the product within a given data set or mission, as well as details to describe

the content and structure of the product. The metadata and their relationships are abstracted into a series of classes and attributes defined in the PDS4 Information Model (IM). However, the IM just provides a stable common model across its diverse disciplines and missions, the mission-specific accompanying information is not defined in the IM. Therefore, we should develop our individual data models, which will be described in Sect. 4.2.

The data objects classified into observational data and supplemental data are stored in N-dimension array, ASCII table, or binary table. According to the characteristics of the data, seven types of data acquired from the instruments are stored as observational data, including radar echo data, spectrum data, image data, energy spectrum data, magnetic field data, meteorological data, and acoustic data. For the engineering parameter information and geometric information that cannot be directly represented in the data label, engineering files and geometric grid files are created as supplemental data to be carried along with observational data. The detailed data format description of the seven types of data and their associated supplemental data are given in Sect. 4.3–Sect. 4.9.

4.2 Tianwen-1 Data Model

As mentioned in the Sect. 2, the ancillary information plays a key role in using the products. To represent the mission-specific information and their relationships, the Tianwen-1 data models (TwM) are developed as the extension of the PDS IM. Three basic models have been under development, and the classes and attributes are correspondingly defined in the Local Data Dictionaries (LDDs).

4.2.1 Instrument Model

The instrument model is designed to describe the instrument-specific metadata for the thirteen instruments onboard the orbiter and the lander. Two basic classes are provided in this model: the Work_Mode_Parm classe provides the working condition of the instrument, which contains several attributes such as the camera exposure time, gain, working mode, sampling rate etc., the Instrument_Parm classe includes thirteen subclasses, each subclasses represent the specifications of one scientific instrument, in which the information such as band range, frequency range, pixel size, resolution, field of view (FOV), etc. are given in the attributes.

4.2.2 Processing Model

Sometimes the data user wants a partially processed data that have been processed beyond the raw stage but which have not yet completely calibrated, based on such requests from the science community, we design the processing model to represent the typical processing method so that the user can use the typical processing method to do the following process, or they can make their individual calibration model to generate the calibrated data products they want by using the partially data product as input.

The processing model is illustrated using a Unified Modeling Language (UML) class hierarchy diagram in Fig. 2. All data and information required in the practical processing are defined in the Inputs, Outputs and Parameters class, respectively. Method class identifies the mathematical expression and software implementation of the processing method, in which the Algorithm class describes the mathematical formula detailed using MathML language and the Implementation class provides the information in performing the algorithm, includes the method description, the method implementation, and source code of the method.
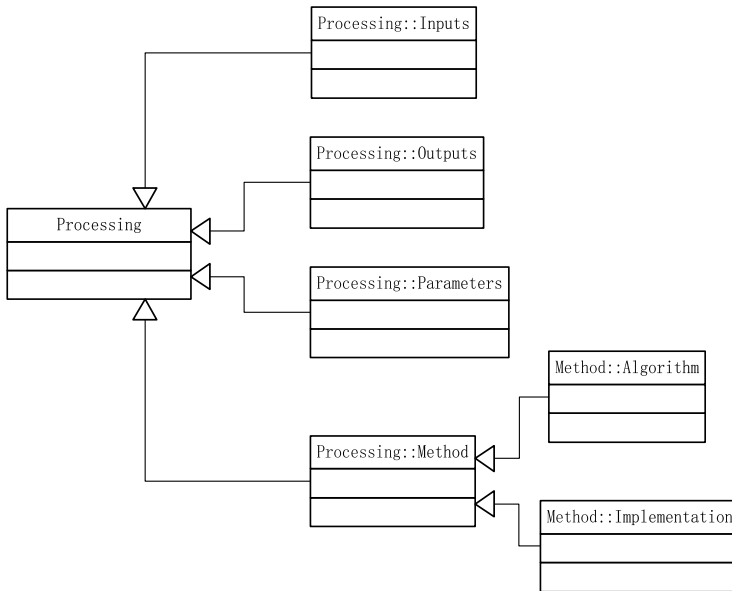
Fig. 2 Processing model of the Tianwen-1 data model

4.2.3 Geometry Model

The geometry metadata is very important ancillary information in Tianwen-1 data archive system, so that users can retrieve location information to obtain the data they want. And well-defined geometry metadata is essential for processing the products into higher level data products, such as DEM, DOM and etc. The geometry model was developed to capture the geometry metadata that processed in Level 2B processing or acquired from the raw observational data. The geometry model defines five classes (see Fig. 3), which support general Mars observations and other observation type. Spacecraft_Position class provide position of the spacecraft (or sensor) relative to Mars surface or landing site in different reference frame. Spacecraft_Attitude class identifies attitude of spacecraft (or sensor) at the time of an observation. Solar_Angle class describes lighting condition at the intercept point and the relationship between the instrument viewing position and solar light. For the orbiter and rover, there are three types of observations, Mars-oriented observation- general observation type in global reconnaissance orbit and rover patrolling route; calibrator observation and calibration field observation (e.g. solar). All three are supported by Observation_Location class and Observation_Orientation class.

4.3 Radar Echo Data

MOSIR and RoPeR are carried on the orbiter and rover, respectively, to detect the subsurface structure of Mars. According to the classification of the electromagnetic spectrum, they belong to radio wave detection. The detection data generated is the echo signal of radar, which is composed of a time stamp (generally 6 bytes), engineering parameters (including the instrument operating mode, gain, operating frequency, etc.) and echo data. In Level 1 and Level 2 radar products, the echo data type is complex, which consists of a 4-byte real
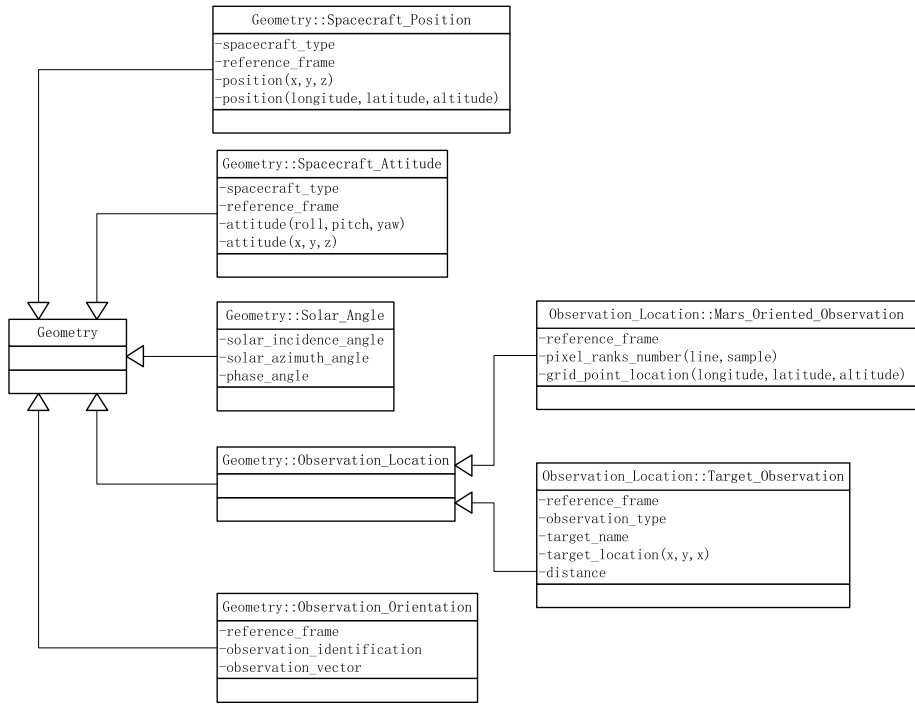
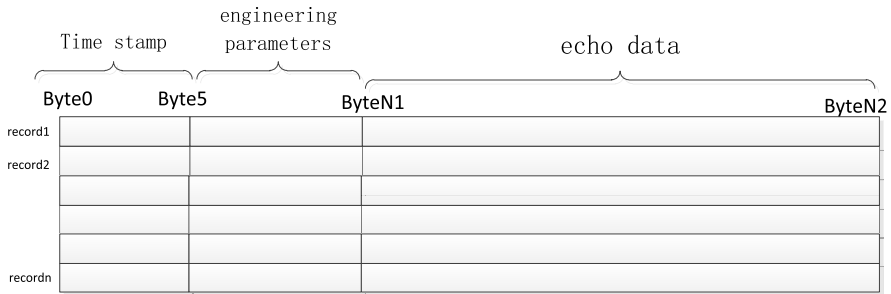**Fig. 3** Geometry model of the Tianwen-1 data model

Fig. 4 Radar data format: Each radar echo data is stored as a fixed length record, which consists of certain fields to store the time stamp, engineering parameters, etc., and one group to store echo data. The fields of each record begin at fixed locations, and there is no record delimiter between records

part and 4-bye imaginary part after fast Fourier transform (FFT). Since these data do not need to be displayed directly, they can be processed by software to obtain their waveform. To improve the efficiency of software processing and save storage space, the data are represented in binary. The data product format is shown in Fig. 4, which is described using the Table_Binary subclass of Table_Base.
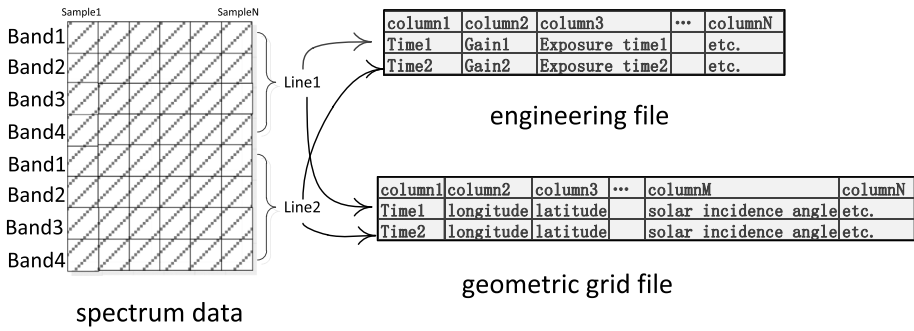
Fig. 5 MMS data format; the spectrum data is stored in form of a 3D array, which is referred to as band interleaved by line (BIL). The engineering file and geometric grid file are stored in ASCII tables

Fig. 6 MarSCoDe data format, the data are stored in a sequence of identically structured records. Each record represents the detection result of one target point, including spectrum data, engineering parameter information, geometric information, and other ancillary data

4.4 Spectrum Data

MMS in the orbiter includes a visible near-infrared module and near medium wave infrared module. Among them, the visible near-infrared module uses a CCD plane array detector to complete the spectrum image acquisition in the range of $450 \sim 1050$ nm; the near and medium wave infrared module uses a Mercury–Cadmium–Telluride (MCT) focal plane array detector to complete the spectrum image acquisition in the range of $1000 \sim 3400$ nm. These two detectors achieve push broom imaging through the motion of the orbiter and finally obtain the three-dimensional spectrum data of the Mars target. The spectrum is stored in form of array_3D_Spectrum with two spatial dimensions and one spectral dimension. In addition to this, the data product also contains an engineering file and geometric grid file, which are used to describe the instrument working parameters and geometric information. The format of data products is shown in Fig. 5.

MarSCoDe in the rover acquires the high-resolution plasma spectrum of the target by using laser-induced breakdown spectroscopy (LIBS). In addition, it also has the function of obtaining the shortwave infrared spectrum in the range of 850 nm~2400 nm. All these data are one-dimensional (1D) spectral data of a single target point. We use Table_Character to store the data of all target points in one observation cycle (see Fig. 6):

4.5 Image Data

Images are the most important and intuitive data in the Mars exploration mission. Several cameras are configured in this mission, such as HiRIC and MoRIC in the orbiter and NaTe-Cam and MSCam in the rover, and all the data of these instruments are image data, which includes linear array images and planar array images.

HiRIC is composed of three TDI CCD detectors and two CMOS detectors. The TDI CCD detectors cover five spectral ranges of panchromatic, blue (B), green (G), red (R) and near-infrared, and each of the R, G, B, panchromatic and near-infrared components are stored in multiple, separate files. Array_2D_Image is used to store linear CCD data; each line of an image is stored sequentially according to the time series. The engineering file and geometric grid file, which are supplementary data objects to be carried along with the image data product, are added at Level 1 and Level 2B, respectively. Panchromatic planar array images of the CMOS detector are stored by the Array_3D_Image structure in Level 1, the image object is a 3D time sequential array, and the three axes are time, line and sample, which have the index storage order (Time, Line, Sample). In Level 2, a single area array CMOS data is stored as a single file, and there is no time axis in the data object, so Array_2D_Image is used to store planar array images of Level 2.

Both Moric and NaTeCam can obtain RGB color images of Mars' surface. For Level 1, Level 2A and Level 2B data, the image object is stored in a Bayer pattern, and in Level 2C, we convert the image from Bayer format to an RGB per-pixel format. The same as HiRIC CMOS data, Array_3D_Image and Array_2D_Image are used in Level 1, Level 2A and Level 2B separately. The Array_3D_Image structure is also used in Level 2C, but the RGB color image objects are stored by a sample interleaved array, which has the index order (Line, Sample, Band).

Image data of MSCam are stored as nine separate Array_2D_Image objects with each of the bands of 480 nm, 525 nm, 650 nm, 700 nm, 800 nm, 900 nm, 950 nm, 1000 nm and panchromatic components being stored separately. The lines or samples of the nine components are not interleaved. The image data object of each band is the same as HiRIC CMOS data.

In addition to this, some engineering parameters, such as the time code, gain, exposure time, temperature, etc., are attached to image data in the same scientific frame and are also generated to form an engineering file as part of the data product. Figure 7 shows a typical data storage structure of multiple panchromatic images. Array_3D_Image and Table_Character are the data storage structure of image data and engineering data.

4.6 Energy Spectrum Data

Tianwen-1 uses MINPA and MEPA to explore the Martian upper atmospheric and ionospheric properties, the solar wind, and solar radiation and particle. The data products are formatted in ASCII tables, the general data structure of which are as follows:

MEPA measures the energy spectrum of light particles, such as protons, helium, and electrons and heavy ions. The data of light particles and heavy ions are separated in two files in Level 1, and variable names are identical for the two types of data. For the light particles data, there is a table file per orbit containing level1 data with time-ordered raw counts in each of the 54 event counters. And level2 data is the energy spectrum after correction. For heavy ions, same formatted table containing level1 data with ADC value of three sensor, i.e. two silicon sensors and one CsI detector and level 2 data with the deposited energy of each sensor.
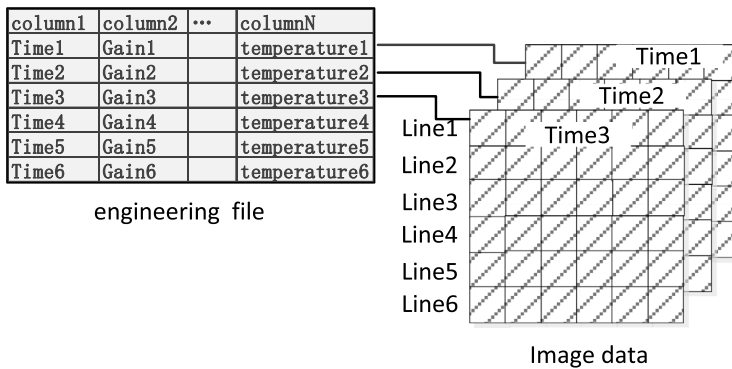
Fig. 7 Multiple panchromatic images' data are stored in a 3D array by time sequence, which has the axis order (Time, Line, Sample), and the engineering files are stored in ASCII tables. Time is the primary parameter by which the relationships are established between image data and engineering files

MINPA is used to study the solar wind/Mars interaction by measuring ions and energetic neutral atoms (ENAs). There is a maximum of 64 energy sweeping steps×16 elevation angle deflection steps×16 azimuth sectors ion data and 8 energy seeping steps ×16 azimuth sectors ENA data for one record. The level1 data contains raw counts of each counter in addition to essential engineering data. The level2 data are in physical units, and include energetic neutral atoms and ion spectra.

4.7 Magnetic Field Data

The data acquired by MOMAG of the orbiter and RoMAG of the rover are magnetic field data, which is treated as a 3-dimensional vector. Each product is an ASCII table file containing a time series of magnetic field vectors, in addition to essential engineering paraments (e.g. sampling rate, sensor temperature). The magnetic field data in Level 2A are in physical units (nanotesla, nT) and have been corrected for instrumental and spacecraft effects (calibrated). In Level 2B, the magnetic field vectors have been transformed into spacecraft(orbiter) coordinate system, the sensor position and sensor attitude are also included in the record.

4.8 Meteorological Data

MCS has been designed to record three meteorological parameters: wind speed/direction, pressure, and air temperature. All three are stored in time ordered ASCII data organized as tables of N-columns by M-rows. The columns are separated and are of fixed sizes. Each data file will comprise one SOL of data, therefore, the number of rows in the data products is equivalent to the number of records contained in one SOL. Each row is separated by a carriage-return line-feed pair Level 1 data products provide raw ADC counts recorded by each sensor, and Level 2A data products contains data where ADC counts of sensor have been converted to electrical values using calibration information. Environmental magnitudes of wind speed/direction, pressure, and air temperature are provided in Level 2C.

4.9 Acoustic Data

A Mars microphone as part of MCS obtains the acoustic data on the surface of Mars. The Mars microphone provides two sampling rates at 5 kHz and 40 kHz. The acoustic waveform

is formatted in a 1-D array which records a series 16-bit ADC value of sampling points according to different sampling rates and archived in WAV-formatted data files. The WAV header provide waveform related metadata, including coding format, number of channels, sample rate, bits per sample, etc. Other metadata such as instrument response parameters, timing information are stored in the data label.

5 Data Validation

Since various error sources in the generation chain of the data products exist, only properly validated products can be of value for scientific research. For this reason, all data delivered by Tianwen-1 should undergo three phases of validation activities before release, including 1) Pre-flight data processing algorithm validation, 2) Pre-flight data products validation, and 3) In-flight validation.

5.1 Pre-Flight Data Processing Algorithm Validation

The data processing algorithms are subject to uncertainty. Contributions to uncertainty of algorithms can be based on manifold factors, such as systematic errors on the path from the sensor electromagnetic signal to actual physical value, and retrieval errors consists of components affected by Mars environmental influences and atmospheric absorption and scattering effects. Therefore, each scientific instrument has undergone ground calibration to determine the quantitative relationship between the measurement data and the actual physical value before launch. The data preprocessing algorithms that formed in the ground calibration are delivered to downstream validations. With the support of the GRAS and the instruments team, the ground validation experiments were then carried out under a simulated Mars environment. The GRAS researchers conduct the validation efforts through methods of comparison, accuracy assessment, etc. to evaluate the data processing algorithm. The detailed validation processes are published in scientific validation experiments article of each instrument of the topical collection" The Huoxing-1 (HX-1) / Tianwen-1 (TW-1) mission to Mars". After completion of ground validation, a science peer review meeting, which involves the instrument team, data processing team, and science application experts, will be held under the organization of China National Space Administration (CNSA).

5.2 Pre-Flight Data Products Validation

Testing of algorithm implementation processes ensures the correctness of data products. The validation in this stage has checked the data products from two aspects: 1) assessing the correctness/completeness of the data content to make sure the realization of data processing software is correct; and 2) checking the data format for compliance with PDS4 Standards to ensure that the data can be read and manipulated using standard PDS data-reading tools.

Comparison to reference data is the mainly validation method for assessment of data content. The experiment data obtained in the above ground validation are form of reference data that can be used for comparison. By statistical analysis and typical area comparison, we can know whether the software implementation process is correct. Two major statistical metrics were used to evaluate the accuracies of the data products, including the root mean square error (RMSE), bias and correlation coefficient. Lower RMSE and bias indicate better data processing performance; in addition, the correlation coefficient must approach 1. Typical single sampling points are found using time stamps for comparison. If the comparison results are within the allowable error range, the data are considered to have good consistency.

Checking the data format is another stage of products validation, ensuring that the data products are compliance to the standards and to the archive conventions. We develop a data validation tool to verify a label compliance of PDS4.0 standards, and to validate all syntactic and semantics aspects of the data products structure/content. The data labels are written in eXtensible Markup Language (XML), following its recommendation by the World Wide Web Consortium (W3C). XML is an international standard for information interchange that provides a fixed and well-defined syntax for creating document structures. Therefore, the data validation tool invocates third-party software for the syntax validation of XML files itself. In addition to, the tool can check that all labels correctly followed the PDS4 and Tianwen-1 data Dictionary requirements and the location and format of each data record(item) are complied with the associated data object. The validation tool is also a visualization data browser that provides intuitive graphical data, and it gives a number of analysis plug-ins providing statistical records for validation. For example, drawing a histogram is one of the basic data analysis functions, which can gather statistics about the object, such as the largest value, smallest value, standard deviation, etc. that appear in the stored object. The statistical results in the histogram are just compared with the component of Object_Statistics in the data label. After comparison, the statistical results of the validation tool are fully consistent with the data processing results described in data label, verifying the correctness of the label.

5.3 In-Flight Validation

Once the processing algorithm and processing results are validated, the main purpose of in-flight validation is to ensure that the data pipeline is still valid and that the data format and content still comply with the standards. However, the sensor-prone error sources (e.g., sensor degradation) and processing errors (e.g., data packet dropout) are still introduced in the subsequent processing. To counteract these influences, on one hand, many instruments will carry out onboard calibration regularly, on the other hand, the quality assessment information that mentioned in Sect. 3.7 are used to indicate the data quality. As a result, the shortcoming of processing algorithms can be improved and calibration data may update in this phase.

GRAS and the instruments teams will expect to follow the data release policy proposed by CNSA based on an official 5–6 months' proprietary period that is used by the teams to validate, calibrate, and perform preliminary scientific data exploration (Li et al. 2021a). During this period, the data are released to specific users through offline copying (Li et al. 2021b). After all the procedures are complete, about one year later, the data will be released to the scientific community from the website (http://moon.bao.ac.in/index_en.jsp). Afterward, the data products will be updated periodically and incrementally over a certain amount of time.

6 Conclusions

At present, Tianwen-1 has entered into the Earth-Mars transfer orbit successfully, MINPA and MEPA have been working constantly for the purpose of detecting the interplanetary environments, and other instruments have completed their power-on self-test. The rover will start a 90 sols surface patrolling after the lander is released. The orbiter will perform orbit maneuver and enters one Martian year global reconnaissance orbit when the rover finishes its 90 sols' patrolling. GRAS has completed the prelaunch data validation and has been

generating Level 1 and Level 2 data products of MINPA and MEPA, and the self-check data of other instruments are just processed to Level 0. The data will be released to the public one year after it is available.

There is still much work to do in terms of completing and enhancing the content of the data products. PDS4 standards are newly selected in China's first Mars exploration mission, and a complete validation for all data needs to be performed before they are placed within the scientific archive. For those who have applied China's Lunar Exploration data before, it will take time for the new data product format to be adopted. Accordingly, we should provide more convenient tools and comprehensive and clear documentation and improve the public feedback mechanism to absorb valued feedback from the data users in the future.

References

A.L. Albee, F.D. Palluconi, R.E. Arvidson, Mars global surveyor mission: overview and status. Science **279**, 1671–1672 (1998)

W.B. Banerdt, S.E. Smrekar, D. Banfield et al., Initial results from the InSight mission on Mars. Nat. Geosci. **13**, 183–189 (2020)

J. Benkhoff, J. Van Casteren, H. Hayakawa et al., BepiColombo—comprehensive exploration of Mercury: mission overview and science goals. Planet. Space Sci. **58**, 2–20 (2010)

S. Besse, C. Vallat, M. Barthelemy et al., ESA's Planetary Science Archive: preserve and present reliable scientific data sets. J. Geomagn. Geoelectr. **150**, 131–140 (2018)

A.M. Du, Y. Zhang, H.Y. Li et al., The Chinese Mars ROVER fluxgate magnetometers. Space Sci. Rev. **216**, 135 (2020). https://doi.org/10.1007/s11214-020-00766-8

E.M. Eliason, S.K. LaVoie, L.A. Soderblom, The imaging node for the Planetary Data System. Planet. Space Sci. **44**, 23–32 (1996)

M.Y. Fan, P. Lyu, Y. Su et al., The Mars Orbiter Subsurface Investigation Radar (MOSIR) on China's Tianwen-1 mission. Space Sci. Rev. **217**, 8 (2021)

J.N. Goswami, M. Annadurai, Chandrayaan-1: India's first planetary science mission to the Moon. Curr. Sci. **96**, 486–491 (2009)

E.J. Grayzeck Jr., M.F. A'Hearn, A.C. Raugh et al., in *Services of the Small Bodies Node of the Planetary Data System*. Planetary and Space Science, vol. 44 (1996), pp. 47–54

J.P. Grotzinger, J. Crisp, A.R. Vasavada et al., Mars Science Laboratory mission and science investigation. Space Sci. Rev. **170**, 5–56 (2012)

E.A. Guinness, R.E. Arvidson, S. Slavney, The planetary data system geosciences node. Planet. Space Sci. **44**, 13–22 (1996)

J. Haruyama, T. Iwata, H. Ootake et al., *SELENE (Kaguya) Achievements and Perspectives*. EPSC Abstracts, vol. 7 (2012)

Z. He, R. Xu, C.L. Li et al., Mars Mineralogical Spectrometer (MMS) on the Tianwen-1 Mission. Space Sci. Rev. **217**, 27 (2021). https://doi.org/10.1007/s11214-021-00804-z

H. Hoshino, Y. Yamamoto, S. Sobue et al., Data processing at KAGUYA operation and analysis center. Space Sci. Rev. **154**(1–4), 317–342 (2010)

J.S. Hughes, D. Crichton, S. Hardman et al., PDS4: a model-driven planetary science data architecture for long-term preservation, in *2014 IEEE 30th International Conference on Data Engineering Workshops* (IEEE, Chicago, 2014), pp. 134–141

B.M. Jakosky, R.P. Lin, J.M. Grebowsky et al., The Mars atmosphere and volatile evolution (MAVEN) mission. Space Sci. Rev. **195**, 3–48 (2015)

Jet Propulsion Laboratory, Planetary Data System Standards Reference Version 1.12.0, https://pds.nasa.gov/datastandards/documents/sr/v1/StdRef_1.12.0.pdf. Accessed 1 April 2019

Jet Propulsion Laboratory, PDS4 Information Model Specification Version 1.12.0.0, https://pds.nasa.gov/datastandards/documents/im/v1/index_1C00.html. Accessed June, 2019b

Jet Propulsion Laboratory, The PDS4 Data Provider's Handbook Version 1.12.0.0. https://pds.nasa.gov/datastandards/documents/dph/v1/PDS4_DataProvidersHandbook_1.12.0.pdf. Accessed April, 2019c

L.G. Kong, A.B. Zhang, Z. Tian et al., Mars Ion and Neutral Particle Analyzer (MINPA) for Chinese Mars Exploration Mission (Tianwen-1): design and ground calibration. Earth Planet. Phys. **4**, 333–344 (2020)

D.S. Lauretta, S.S. Balram-Knutson, E. Beshore et al., OSIRIS-REx: sample return from asteroid (101955) Bennu. Space Sci. Rev. **212**, 925–984 (2017)

C.L. Li, J.J. Liu, X. Ren et al., The Chang'e 3 mission overview. Space Sci. Rev. **190**, 85–101 (2015). https://doi.org/10.1007/s11214-014-0134-7

C.L. Li, R.Q. Zhang, D.Y. Yu et al., China's Mars exploration mission and science investigation. Space Sci. Rev. **217**, 57 (2021a). https://doi.org/10.1007/s11214-021-00832-9

C.L. Li, W. Zuo, W.B. Wen et al., Overview of the Chang'e-4 mission: opening the frontier of scientific exploration of the Lunar far side. Space Sci. Rev. **217**, 35 (2021b)

C.H. Li, S.W. Tang, X.Y. Hu et al., Design and realization of China Tianwen-1 energetic particle analyzer. Space Sci. Rev. **217**, 26 (2021c). https://doi.org/10.1007/s11214-021-00803-0

X. Liang, W.L. Chen, Z.X. Cao et al., The Navigation and Terrain Cameras on the Tianwen-1 Mars Rover. Space Sci. Rev. **217**, 37 (2021). https://doi.org/10.1007/s11214-021-00813-y

J.J. Liu, Y. Su, W. Zuo et al., Ground research and application system of China's first Mars exploration mission. J. Deep Space Explor. **5**, 414–425 (2018)

K. Liu, X.J. Hao, Y.R. Li et al., Mars Orbiter magnetometer of China's first Mars mission Tianwen-1. Earth Planet. Phys. **4**, 384–389 (2020)

A.J. Macfarlane, R. Docasal, C. Rios et al., Improving accessibility and discovery of ESA planetary data through the new planetary science archive. Planet. Space Sci. **150**, 104–110 (2018)

S.K. McMahon, Overview of the planetary data system. J. Geomagn. Geoelectr. **44**, 3–12 (1996)

Q.Y. Meng, D. Wang, X.D. Wang et al., High Resolution Imaging Camera (HiRIC) on China's first Mars exploration Tianwen-1 mission. Space Sci. Rev. **217**, 42 (2021)

Z.Y. Ouyang, C.L. Li, Y.L. Zou et al., Primary scientific results of Chang'E-1 lunar mission. Sci. China Earth Sci. **53**, 1565–1581 (2010)

Y.Q. Peng, L.B. Zhang, Z.G. Cai et al., Overview of the Mars climate station for Tianwen-1 mission. Earth Planet. Phys. **4**, 371–383 (2020)

R.S. Saunders, R.E. Arvidson, G.D. Badhwar et al., 2001 Mars Odyssey mission summary. Space Sci. Rev. **110**, 1–36 (2004)

R. Schmidt, Mars Express—ESA's first mission to planet Mars. Acta Astronaut. **52**, 197–202 (2003)

X. Tan, J.J. Liu, C.L. Li et al., Scientific data products and the data pre-processing subsystem of the Chang'e-3 mission. Res. Astron. Astrophys. **14**, 1682–1694 (2014)

S.W. Tang, Y. Wang, H.Y. Zhao et al., Calibration of Mars Energetic Particle Analyzer (MEPA). Earth Planet. Phys. **4**, 355–363 (2020)

J. Vago, O. Witasse, H. Svedhem et al., ESA ExoMars program: the next step in exploring Mars. Sol. Syst. Res. **49**, 518–528 (2015)

W.X. Wan, C. Wang, C.L. Li et al., China's first mission to Mars. Nat. Astron. **4**, 721 (2020)

P.J. Ye, Z.Z. Sun, W. Rao et al., Mission overview and key technologies of the first Mars probe of China. Sci. China, Technol. Sci. **60**, 649–657 (2017)

G.B. Yu, E.H. Liu, G.L. Liu et al., Moderate Resolution Imaging Camera (MoRIC) of China's first Mars mission Tianwen-1. Earth Planet. Phys. **4**, 364–370 (2020)

B. Zhou, S.X. Shen, W. Lu et al., The Mars rover subsurface penetrating radar onboard China's Mars 2020 mission. Earth Planet. Phys. **4**, 345–354 (2020)

Y.L. Zou, Y. Zhu, Y.F. Bai et al., Scientific objectives and payloads of Tianwen-1, China's first Mars exploration mission. Adv. Space Res. **67**(2), 812–823 (2021)

W. Zuo, C.L. Li, Z.B. Zhang, Scientific data and their release of Chang'E-1 and Chang'E-2. Chin. J. Geochem. **33**, 24–44 (2014)