

# TES Ground Data System Software

Susan Paradise, Sirvard Akopyan, Richard C. De Baca, Kevin Croft, Kent Fry, Scott Gluck, David Ho, Brooke Koffend, Michael Lampel, James McDuffie, Ruth Monarrez, Hari Nair, Sassaneh Poosti, Doug Shepard, Irina Strickland, Denis Tremblay, Hyejung Yun, and Jia Zong

**Abstract**—Software development in support of the tropospheric emission spectrometer instrument ground data system has undergone many challenges due to the uniqueness of the instrument, complexity of the science, data volume, and performance requirements. In this paper, we describe the software, its development, and the way in which many of these challenges were met.

**Index Terms**—Data processing, EOS-Aura, ground system, Tropospheric Emission Spectrometer (TES).

## I. INTRODUCTION

ON July 15, 2004, NASA's AURA satellite, part of the Earth Observing System (EOS), was launched, carrying the Tropospheric Emission Spectrometer (TES) interferometer. This unique instrument enables through spectroscopy the global measurement of vertical profiles of the Earth's atmospheric constituents from the top-of-atmosphere to the Earth's surface. TES has capabilities of viewing in nadir, off-nadir and limb directions, and using selectable filters spanning the range 650–3050  $\text{cm}^{-1}$ , or roughly 3.3–15.4  $\mu\text{m}$ .

The TES ground data system (GDS) encompasses a multifarious environment of both software and hardware components. The GDS Science Data Processing System (SDPS) software development faced several challenges, starting with the complexity of the algorithms that encompass the transformation of the measured interferogram data into radiometrically calibrated spectra, the physical model of the atmospheric species and behaviors, and the atmospheric retrievals. Next, the immense data volume (47 GB downlinked every two days) and complex configurations and interfaces impose challenges in the building of both hardware and software systems. These challenges are intensified by the performance requirements that call for the ability to process at approximately the rate that data are acquired while also supporting reprocessing loads of up to three times the original data on average throughout the mission. Another challenge comes from the need for visualization and monitoring to present the large volume of instrument and processed data in a form that can be grasped meaningfully by the engineers and scientists.

Manuscript received April 29, 2005; revised November 15, 2005. This work was conducted by the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration.

S. Paradise, S. Akopyan, R. C. De Baca, K. Fry, S. Gluck, B. Koffend, J. McDuffie, R. Monarrez, H. Nair, S. Poosti, D. Shepard, H. Yun, and J. Zong are with the Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA 91109 USA (e-mail: Susan.Paradise@jpl.nasa.gov).

K. Croft, D. Ho, M. Lampel, I. Strickland, and D. Tremblay are with Raytheon ITSS, Pasadena, CA 91101 USA.

Digital Object Identifier 10.1109/TGRS.2006.871205

TABLE I  
TES INSTRUMENT PHYSICAL ATTRIBUTES AND OPERATIONAL MODES

Focal Planes (FP)	1A, 1B, 2A, 2B
Detectors (or pixels)	16 per FP; 64 Total
Optical Filters	1A1, 1A2, 1B1, 1B2, 2A1, 2A2, 2A3, 2A4, 2B1
Viewing Modes	Nadir, Off-Nadir, Limb, Calibration (Cold Space and OBRCS)
Observational Modes	Global Survey, Step-and-Stare, Transect, Stare, Limb Drag

In this paper, we provide an overview of the software functionality and development process, with particular attention to many of the above challenges.

## II. TES INSTRUMENT DATA

TES instrument data are provided from the Aura spacecraft in the Consultative Committee for Space Data Systems (CCSDS) [1] packetized format, with an average data rate of 3.66 Mb/s during global survey operations. Data from each detector in a focal plane (four focal planes with 16 detectors/focal plane) is interleaved during packetization. Data are categorized by an application identification (APID) according to instrument operations type and originating focal plane. This data is termed Level 0 (LO) data, and must be accompanied by ancillary information such as spacecraft attitude and ephemeris to be processed by the ground data system.

For bookkeeping purposes, instrument data are categorized by run, sequence and scan. A scan is a single instrument observation, pointing at a calibration source or a target location. A sequence is a logical grouping of instrument observations. A run is the super-set of observations that is managed and planned by the instrument operations team. The TES ground data system was designed to be adaptive to provide processing for different run types and calibration scenarios as commanded by the instrument operations team. The TES instrument's physical characteristics and operational modes are summarized in Table I.

Nominal operations include runs called global surveys (GSs), which consist of 16 orbits (somewhat more than 26 h) acquired every other day. Each orbit contains 72 sequences of data.

The original GS format contained sequences which included two calibration scans, followed by two nadir scans and three backlooking limb scans for a total of 2304 nadir and 3456 limb observations every other day. The nadir observations were oriented such that they covered approximately the same footprint on the ground track, so that the spectra from all pixels of both these target scenes are averaged to improve signal-to-noise, yielding a total of 1152 target scenes.

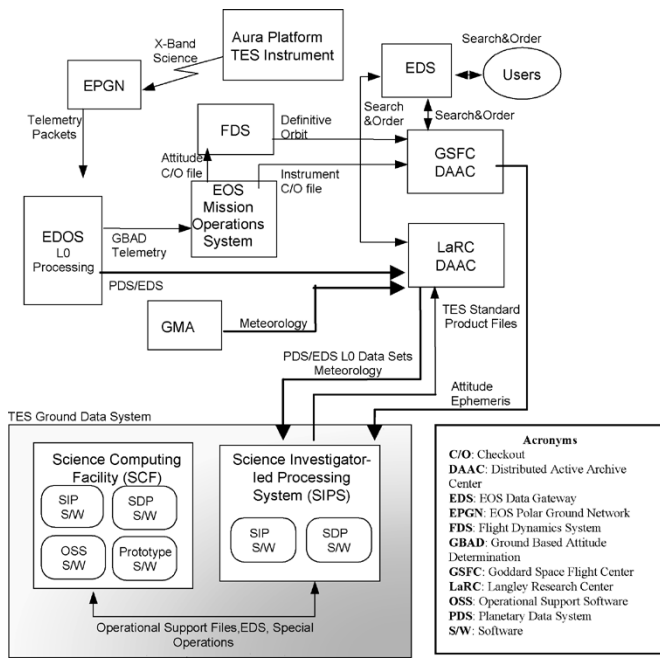


Fig. 1. TES interface to EOSDIS.

Due to life-time concerns arising from instrument motor current monitoring, the nominal GS has been recently modified to remove the limb scans and pre/post calibration runs [2]. Each sequence now contains two calibration scans and three nadir scans, pointing to different locations, so that the pixels for each nadir target scene are averaged separately and the overall nadir coverage is improved to a total of 3456 nadir target scenes for the GS.

TES also supports special observations, such as *Step-and-Stare*: a series of nadir observations along a path of particular geographic interest, for example in coincidence with an aircraft carrying another instrument; *Stare*: pointing at a specific location for up to 4 min; *Transect*: pointing at a set of contiguous locations along the spacecraft flight path; and *Limb Drag*: pointing at the trailing limb (16-s scans) and repeating. Specialized calibration runs are occasionally performed in order to monitor instrument health.

An overview of the TES instrument data and algorithms can be found in [3] and [4].

### III. SUBSYSTEMS AND PROCESSING LEVELS

The TES GDS is a component of the NASA EOS Data and Information System (EOSDIS) [5]. The TES GDS has major interfaces to the EOSDIS Core System (ECS) for the input of TES instrument and ancillary data, and for the output of archived standard data products (Fig. 1).

As for all projects in the EOS Core System (ECS), TES GDS processing is grouped into Level 1A (L1A), Level 1B (L1B), Level 2 (L2), and Level 3 (L3). TES geolocated interferograms are produced at L1A. L1B produces radiometrically and spectrally calibrated spectra with noise equivalent spectral radiance (NESR). L2 produces atmospheric volume mixing ratios (VMRs) and temperature profiles. L3 generates daily and

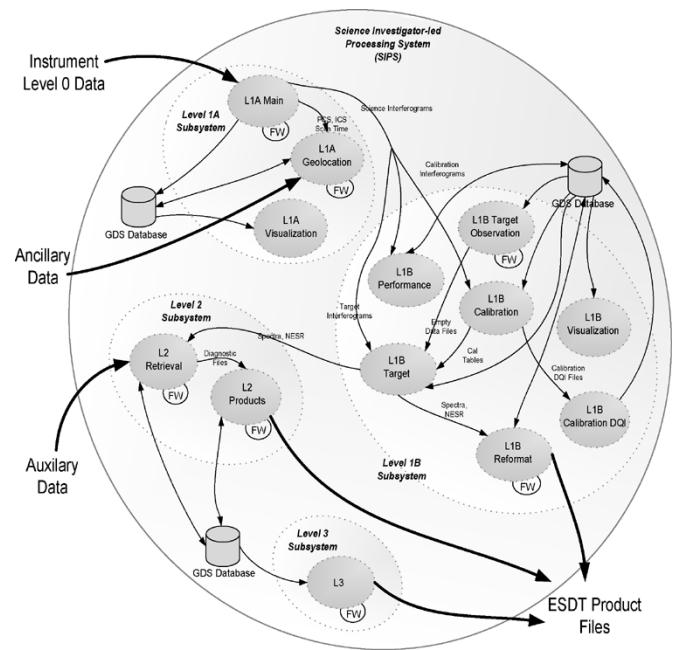


Fig. 2. TES PGE data flow.

monthly gridded plots and monthly zonal means from the L2 data.

One or more product generation executives (PGEs) are executed in support of each of these processing levels. Many PGEs are built upon a shared set of code called Framework/Shared.

The TES PGE interfaces are illustrated in Fig. 2. Ancillary information needed by L1A Geolocation include spacecraft attitude and ephemeris, leap second and UTC pole files from the ECS Science Data Processing (SDP) Toolkit (version 5.2.13 [6]). Auxiliary climatology files from the Global Model and Assimilation Office (GMAO) are required by Level 2. The “late-look” GMAO data are used for standard processing, and are available approximately 15 days after observation date.

Database interfaces are used extensively for data interfaces between PGEs and for long-term trending. Associated with the Level 1 processing is a set of PGEs termed Visualization. These Visualization PGEs use database information processed by a preceding PGE to provide plots and maps to the Science and Instrument Operations team. Summarizing the high data volume provided by the instrument, these plots are critical to day-to-day operations. Most of the Visualization PGEs access the database to obtain trending data that extends over a series of GSs.

#### A. Framework/Shared

In order to maximize software reuse and provide commonality between the PGEs, a collection of class hierarchies implemented as C++ packages was developed to support their development. This collection is known as the *Framework* subsystem. A similar but separate shared set of code is provided under a collection known as *Shared*. The object-oriented design of the Framework and Shared tools provide not only the utilities from which other subsystems benefit such as basic data types, file interfaces, logging, exception handling, database interactions and unit testing, but also a skeleton upon which many PGEs are built. The design and use of the Framework and Shared provide

standards and efficiency in the building and maintenance of the GDS.

The Framework subsystem includes the following packages.

*Datatypes*: Encapsulate a common logical data structure which supports the development of compound data type objects. *Environment*: Common interface to the processing environment. *Exception*: Method for handling exceptions throughout the code. *File I/O*: Data access provider, including an interface to the Science Data Processing Toolkit. *Log*: Common logging capabilities, for information such as status, error code, return codes, etc; anything to be saved or archived. *Metadata*: Extensive class library providing tools and mechanisms for all required and optional metadata creation. *Parameters*: Supply run time parameters to the application program. *TES Algorithm Interface*: Implement the algorithm encapsulation; provide some base classes from which all algorithm classes are derived; interface to external operating environment for all science processing algorithms (via Environment). *Oracle DB Access Interface*: Database access utilities; code reused from the EOS Microwave Limb Sounder (MLS) project. *TES Sequencer*: Utility software to support consistent, automated unit testing. *Utilities*: Utility library of common code.

The Shared subsystem includes the following packages.

*Atomic Types*: Common interface to the most basic c++ scalar data types. It ensures that the same data storage unit is used across different operating systems/platforms. *Constants*: Common value to some of the Mathematics, Physical, and TES Hardware constants. *InstantiatedClasses*: Common method to instantiate several RogueWave iterator vector classes. *FileName*: Common interface for all TES subsystems to generate any product file name. *ArrayGenerator*: Common interface to create multidimensional array with contiguous memory. *StringUtility*: Common interface to manipulate c++ strings. *Shared/L1A*: Common interface to some of the Level 1A generic utilities shared across Level 1A subsystem. *Shared/L1B*: Common interface to some of the Level 1B generic utilities shared across Level 1B subsystem. *Shared/L2*: Common interface to some of the Level 2 generic utilities shared across Level 2 subsystem.

## B. Level 1A

TES L1A decommutates the TES science Level 0 datasets and generates geolocated interferograms. L1A consists of the following PGEs.

1) *L1A Main*: This PGE extracts and processes embedded engineering data, decommutates L0 ancillary packet types, and generates diagnostic files for post-processing, visualization and analysis. It provides data quality indicators for use by downstream processing.

The L1A software also responds to anomalies in the instrument. It detects changes in the order that data from individual pixels are packetized and makes an adjustment to the data in response to the shift. These changes are caused by an error in the clocking of the multiplexers that collect data from the detector analog-to-digital converters. It also detects and reports noise bursts (which arise from a number of causes) during interferogram processing. The software monitors engineering data,

reports off tolerance occurrences and places this information in the database tables and files for use by the other PGEs.

2) *L1A Geolocation*: The L1A Geolocation PGE provides target and spacecraft positioning information. Using instrument and spacecraft orientation, this PGE calculates the instrument line of sight (LOS) and LOS intersection with the Earth topographic surface for nadir observations, or LOS intersection of the limb tangent point to the Earth for limb observations. For target observations, the PGE defines the observation footprints with surface information. Doppler effects, sun angles, and related georectification parameters are also derived for both calibration and target observations.

3) *L1A Geolocation Plot*: The L1A Geolocation Plot PGE creates three plots per GS: one showing the limb tangent locations and heights for all three limb scans, and one for both series of nadir scans, showing the location of the footprint and its mean elevation. All plots are in a simple rectangular projection.

4) *L1A ICS Performance Plot*: The L1A\_ICS\_Performance\_Plot PGE plots the motor sine amplitude, motor current, encoder velocity, and fringe velocity from the instrument control system (ICS) diagnostic files produced by L1A\_Main.

5) *L1A PCS Performance Plot*: The L1A\_PCS\_Performance\_Plot PGE plots the track position, error, and motor current from the pointing control system (PCS) diagnostic files produced by L1A\_Main.

6) *L1A Temperature and ICS State Plot*: This PGE creates plots of the instrument temperatures (detectors, optics, interferometer, and black body reference) and the ICS velocity using database values populated by L1A\_Main.

## C. Level 1B

The L1B subsystem performs instrument health checks, transforms interferograms produced by L1A into radiometrically and spectrally calibrated spectra, computes NESR, performs first order cloud and heterogeneous land screening, and writes scientific TES Level 1B data to Hierarchical Data File version 5 (HDF5 [7]) files for public distribution. The following PGEs are part of L1B:

1) *L1B Performance*: The L1B Performance PGE analyzes aspects of the TES instrument performance by using data from the onboard radiometric calibration source (OBRCS). The PGE detects ice buildup on the detectors and co-boresight shear effects (a measure of the optical misalignment), and provides validation of the L1A channel shift correction.

2) *L1B Target Observation*: This PGE creates empty binary observation spectra and NESR files with header information only. These files will later be filled by the L1B Target PGE for subsequent use by the L1B Reformat PGE and Level 2.

3) *L1B Calibration*: The L1B Calibration PGE transforms the OBRCS and Cold Space calibration interferograms into spectra. After sampling phase alignment and averaging of the calibration spectra, the resulting calibration tables are stored for later use by the L1B Target PGE. The software has the flexibility to select between single or double loop phase alignment [8] and to select the Earth's geographical boundaries for calibration. For a given calibration table, the scans have the same set of physical attributes: the run ID, optical filter, resolution, pixel, view, and scan direction. Additionally, this

PGE generates data quality indicator (DQI) files which contain formatted quality assessment data values.

4) *LIB Calibration DQI*: This PGE populates the database with calibration data quality information from the L1B Calibration DQI files.

5) *LIB Target*: The L1B Target PGE comprises several steps. It transforms the target interferograms into spectra. Calibration tables are used to radiometrically calibrate the target spectra, and instrument line shape (ILS) correction is applied. Limb data are corrected for self-apodization. Off-axis compression factor correction is computed using the efficient nonuniform spectral resampling transform (NUSRT) [9] algorithm developed at JPL. NESR is computed for target spectra, which are spectrally calibrated. Finally, target spectra and NESR are interpolated to the L2 frequency grid.

6) *LIB Reformat*: First order cloud and heterogeneous land detection is performed by this PGE through the computation of brightness temperatures (BTs). It also writes TES standard data product HDF5 files and populates the database with L1B Target information.

7) *LIB Performance Visualization*: This PGE plots the instrument  $y$  axis shear and spectrally integrated signal both as a function of time and geographic location for each filter. The spectral magnitude for each filter is also plotted.

8) *LIB Product Visualization*: This PGE plots products from earlier PGEs such as phase alignment, imaginary spectral components, NESR, and BT. Anomalies related to the instrument or the L1B Target (or earlier) algorithms are often seen here first.

#### D. Level 2

In Level 2 the temperature profile and VMR profiles of several atmospheric gases are retrieved and written to HDF-EOS 5 [10] product files (an extension of HDF5) for public distribution.

The following PGEs make up the L2 subsystem.

1) *L2 Retrieval*: This PGE is logically divided into two components: the core L2 Retrieval code is called the Earth Limb and Nadir Operational Retrieval (ELANOR); and Strategy, which prepares the inputs for ELANOR, runs it, and performs error analysis on its output.

Atmospheric retrieval algorithms are computationally expensive and convergence of the retrieval is sensitive to the quality of the input. It is therefore important to provide input that will result in the best atmospheric profile in the smallest number of iterations [11]. Strategy accesses meteorological information to create an initial guess of the atmospheric state. Using this initial guess along with spectroscopic absorption coefficient (ABSCO) tables, ELANOR creates simulated radiances that are compared to the observed radiances. A nonlinear least squares method iteratively retrieves a set of atmospheric species [12]. Multiple iterations are performed at each step and evaluated for convergence. Each subsequent step invokes ELANOR using the updated atmospheric profiles from the prior steps. This technique trades off a simplification of a computational and memory intensive combined retrieval for an increase in complexity of error estimation [13].

The importance of retrievals in the presence of clouds came to light soon after launch [14]. The code was designed for clear-sky

nadir retrievals with the assumption that most of the Earth is clear. Cloudy-sky retrievals were planned but not yet implemented. However, post-launch analysis showed that very little of the Earth is truly “clear,” and instead the assumption must be made that “clouds” (or particulate moisture) exists everywhere in order to achieve reasonable results. A unified approach for dealing with clouds for all optical depths has therefore been developed in the L2 Retrieval software. An initial guess of cloud optical depth is made based on the difference of a predicted and observed BT. The cloud extinction is then included as a retrieved parameter during the gas and temperature retrieval steps.

The error analysis performed after ELANOR in each retrieval step estimates the total error based on smoothing error, cross-state error, measurement error, and systematic error. Smoothing error arises from the effect of both hard and soft constraints; cross-state error arises from the effects of different species on the one being considered; measurement error arises from the NESR; and systematic errors considered may be due to previously retrieved atmospheric trace gas species, atmospheric temperature, surface parameters (surface temperature and emissivity), and spectroscopic line errors.

L2 Retrieval PGE outputs are written to binary files and/or the database.

2) *L2 Products*: The L2 Products PGE reads the binary files and database information provided by the L2 Retrieval PGE for each target scene of a run and populates the TES Level 2 HDF-EOS5 files.

#### E. Level 3

L3, currently under development, will take the constituent profiles retrieved by L2 and generate daily and monthly gridded plots for both nadir and limb retrievals. Monthly zonal means for each species will also be created.

### IV. SOFTWARE METHODOLOGIES AND STANDARDS

#### A. Object-Oriented Design and Languages

Prototypes were coded in object-oriented interactive data language (IDL) [15] for Level 1B, and in C for ELANOR. IDL version 5.6 is currently in use.

In Level 2, the ELANOR C prototype was molded into object-oriented ANSI C++ production code, while writing the new Strategy code in C++ (GNU gcc, version 3.0.4) [16]. This presented a challenge in connecting these two subsystems together as one PGE. Early in the development phase the importance of clean and simple interfaces between them was recognized and as a result, a set of retrieval input classes have been designed. This set of interface classes provides a complete set of input required by ELANOR for retrieval while hiding all the details and complexity in creating those from it. Even though it was not without challenges, this well-defined interface contributed to a successful integration of these two subsystems into one PGE.

#### B. Coding Standards and Error Policy

The coding standards [17] were written by members of the software team in order to promote good coding practices and cohesion in coding style across the project. An error policy was

drawn up to ensure robust operations and error reporting guidelines. Due to the excessive volume of data to be processed, the software must detect and report errors, recover if possible, and otherwise exit gracefully.

## V. DEVELOPMENT AND TESTING

### A. Hardware

The PGEs are compiled, tested and executed within a heterogeneous network of Sparc-Solaris 5.8, i386-Linux Redhat 8, and AMD64-Linux Redhat Enterprise 4 multiprocessor machines. A cross-platform version control system allows seamless development on all platforms.

The multifunction software computing facility (SCF) cluster is the central hub around which software and algorithms are developed and tested. It's composed of 20 dual-processor Opteron compute nodes running Linux, 15 similar nodes outfitted for software development, and 6 Sparc-Solaris nodes. Each machine has roughly 500 GB of local storage and access to more than 40 TB of high-speed network-attached storage.

The Science Processing Cluster (SPC) consists of 20 Athlon and 90 dual-Opteron Linux nodes. This system has been used to process most of TESs publicly presented L2 and L3 science results.

Operational data processing and product generation is performed at the Science Investigator-led Processing System (SIPS) cluster. This cluster consists of four racks of 64-b Opterons, two with 38 nodes, two with 80 nodes, (representing higher density newer technology) for a total of 236 nodes. The SIPS also runs several Sun workstations. More information on the SIPS operating environment is available in [18].

### B. Requirements and Design

A major challenge due to the nature of the project was that the requirements were not fully specified early and that they change over time. The architectural design was developed in the unified modeling language (UML) using Rational Rose [19] first at a high level using the known requirements at that time [20]–[22]. Then, from that conceptual design, as the set of requirements for each software release was defined, high-level and detailed design was built to realize those requirements.

### C. Implementation

After detailed design is done, code is automatically generated from the design using Rose. During the detailed design phase, C++ features and other setup that code generation requires needs to be set properly in order for code generation to work correctly. Using automatic code generation from design allows for a seamless and easy transition from design to implementation without losing any information from the design and therefore making the code consistent with the design. Formal and informal code reviews are also performed during the implementation phase.

### D. Reuse

We have been fortunate to obtain staff with experience on other EOS missions. This has given us insight into the complexities of a large GDS as well as intimate knowledge of tools that

had been used on previous projects that could be pulled into TES. Further, we had an appreciation for the benefits of reusable software and developed much of the software with this in mind. The Framework and Shared code were built to encourage not only reuse of code between subsystems, but also with the idea in mind that the code could be picked up by later projects for reuse. A white-box testing package designed to explore paths of code that are not easily visited through unit testing was reused from the EOS TERRA Multiangle Imaging SpectroRadiometer (MISR) project [23]. An Oracle 9i database interface package was reused from another EOS AURA instrument, the Microwave Limb Sounder (MLS) [24]. The Cluster-Automated System for Processing Experimental Retrievals (CASPER) utility, which will be discussed in Section VIII-A, also designed with reusability in mind, is adaptable to any system of executables. It has subsequently been adopted by the Orbiting Carbon Observatory (OCO) [25] mission, scheduled for launch in 2007. The TES Database Utility is similarly being developed for reusability (Section VIII-B).

### E. Testing

Early fault detection is important for such a large software system. Testing is therefore performed at all phases of the software development. *Unit tests* are developed for all new code and generally focus on providing code coverage, interface testing, and ensuring reasonable values for inputs and outputs. *White box tests* are included in unit testing to promote maximum code coverage. *Algorithm tests* have been incorporated in many places to ensure proper implementation of the algorithms. *PGE regression tests* are used to demonstrate that PGE software releases provide are consistent with previous releases. *PGE testing* was performed heavily in the development phase of the project by an independent team focused on the documented requirements, but has tapered off as the software matures and is replaced by PGE regression tests. *Chained PGE tests* are performed before every software delivery in order to demonstrate connectivity and impacts of changes on downstream PGEs and on the production environment itself.

*Performance tests* were also done extensively. Level 2 was seen as the major processing bottleneck and was estimated to require as much as 90% of total computational resources due to the intensive forward model calculations and large amounts of data required in the form of the ABSCO tables. Consequently, a special team was created to look into how performance could be improved. Over several years this team addressed multiple issues of code design, memory cache usage, parallelization, instantaneous memory footprint, optimization of disk I/O, and RAM usage. By incorporating multiple incremental design changes, faster processors, and through the use of microwindowing, described below, an overall wall clock improvement of almost 200 times was achieved with a reduction in memory footprint of more than a factor of 10 [18].

Microwindowing involves the selection of limited spectral regions to be used by Level 2 “in order to optimize the quality of a retrieval in part by reducing the effects of known systematic errors on the retrieval as well as by choosing those spectral regions with the best sensitivity to the atmospheric species of

interest. Spectral windows are also selected to reduce computational burden without significantly degrading the quality of a retrieval" [26].

The project supported a series of operational dry runs throughout the development, each with increasing complexity. Starting a few years before launch the *Single Orbit Test* simulated Level 2 ELANOR processing for a single orbit of data. One year before launch the *One Day Test* exercised the entire Level 2 Retrieval PGE for an entire day's worth of simulated data. Finally, a series of *Mission Rehearsals* were employed in the months before and immediately after launch to exercise the entire chain of PGEs, SIPS operations, and the Ground Data System's readiness to respond to issues.

To perform scientific validation of the algorithms and full connectivity through Level 2, simulated L2 radiance was back-produced to L0 data. The L0 data was then forward-produced to L2 radiances that were compared with the original simulated L2 radiance. The backproduction software suite includes instrument effects such as optical bandpass filter, electronic phase, Doppler shift, digitization, and many others. Each forward operation had to be performed with the opposite effect in the reverse direction, and had to be meaningful according to its corresponding geographic location along the data stream, a cumbersome task.

## VI. DATA PRODUCTS

TES generates products at Level 1B and Level 2.

Level 1B HDF5 products are produced for both nadir- limb-viewed data, for every focal plane (1A, 1B, 2A, and 2B) for each of the 16 orbits of a Global Survey. However, limb data are currently unavailable. The primary data stored within the L1B products are atmospheric spectra and NESR. In addition, the standard product files contain geolocation, engineering, production history, and data quality information.

Level 2 products are generated for temperature as well as each of the following atmospheric species in nadir: H<sub>2</sub>O, O<sub>3</sub>, CO, and CH<sub>4</sub>. When they become available, limb products will contain: temperature, H<sub>2</sub>O, O<sub>3</sub>, CH<sub>4</sub>, CO, NO<sub>2</sub> and HNO<sub>3</sub>. Each of these products contains the species or temperature vertical profiles with associated error, quality, geolocation, and other information for all target scenes in a GS. An ancillary product is also produced, containing data common to all species that were not included in the species files. The TES data product description (DPS) document contains detailed descriptions of each of the data products [27].

## VII. OBTAINING TES DATA

L1B nadir beta quality data were made available to the public in February, 2005. L2 nadir beta quality data were released in June, 2005. Data may be obtained from the NASA Langley Research Center (LaRC) Atmospheric Sciences Data Center (ASDC) at <http://eosweb.larc.nasa.gov>.

Products generated from special observations will be available in the early winter of 2005. Nadir methane products, Limb products for both L1B and L2, and L2 summary products (products containing only the primary scientific data fields) are expected to be released in the spring of 2006.

## VIII. SUPPORT TOOLS

Several tools have been developed in order to support TES software development, testing, and data analysis. A few of these tools are highlighted below.

### A. CASPER

CASPER was developed in order to test the chaining of the PGEs before delivery to the SIPS, allow for rapid testing of software in input file updates, processing special observation data locally, and providing first-look science results.

CASPER allows the user to upgrade the PGEs to official deliveries, pre-deliveries, or packages customized by the user. All interPGE interfaces are provided behind-the-scenes, making its use simple and elegant. Software and input files are versioned automatically by the tool, allowing for easy tracking. Weekly builds, unit tests, regression tests, and chained tests are performed and reported via a web interface, allowing for automated assessment and reporting of software under development. CASPER resolves order-of-execution dependencies between PGEs and executes one or more processes on a cluster to satisfy those dependencies, efficiently load-balancing its processes on a cluster amidst a heterogeneous workload. This tool is used extensively by both the software development team and science team to test the software, interfaces, algorithms, and inputs.

In addition, CASPER is being used extensively for special processing tasks and prototype development which need close supervision of the science team. This currently includes processing step-and-stare observations and GS nadir with science code.

### B. TES Database Browser

The Oracle [28] database (version 9i) is used as the interface between many of the PGEs as well as a repository for information that can be trended. Because of the heavy reliance of the PGEs on the database and the broad audience of both developers and scientists who needed access to it, it became clear that a tool was needed to enable viewing and plotting of this data. A database browser to access the Oracle database was implemented with a HyperText Markup Language (HTML) [29] user interface. Selection criteria are obtained from the user via HTML forms presented. These include the desired database tables and rows to be retrieved under user selected constraints, such as target scene location or observation time. Embedded server-side PHP: Hypertext Preprocessor [30] scripts construct Oracle queries from this input. The PHP Oracle8 Call-Interface (OCI8) invokes these queries to obtain the desired data. Data are displayed in text format. Plots of selected data can also be generated and displayed. Work is in progress to generalize the Database Browser to enable reuse by other projects.

### C. Target Observation Reader (TOR)

A graphical viewer was developed to enable users and developers to view all information contained in the binary L1B target observation files. These files contain the calibrated spectra and

NESR data output by the LIB target PGE. These files also contain headers describing focal plane attributes, observation parameters and additional quality information.

A graphical toolbar facilitates browsing of multiple observation files in the same directory and other submenus allow the user to select and view every header value in the file. Selecting a single spectrum thumbnail plot in the main browser window opens a detailed spectrum plot window with a corresponding data quality table for that pixel.

TOR was developed on Linux in ANSI C using the Gnu Image Manipulation Program (GIMP) Toolkit (GTK+) [31], a popular open source graphical toolkit.

## IX. CONCLUSION

In spite of many challenges in developing and preparing the GDS software for launch operations, including hardware surprises and changing requirements, we have been able to respond to these difficulties by proactively designing easily maintainable and modifiable code, hiring high-quality engineers and leveraging off their experience, encompassing an array of both narrow- and broad-focused testing, and making sound tradeoffs.

It is worth mentioning that the first uncalibrated spectrum was successfully computed two days after receiving the first data from orbit, and the first successful Level 2 retrieval was obtained one week later, well exceeding expectations on the order of several weeks. Further, because of the success of the CASPER utility, the Science Team was able to study and experiment with many entire GSs and special observations, rather than merely a planned set of 22 “golden” target scenes that were identified from the first GS as being cloud-free.

## ACKNOWLEDGMENT

The authors would like to thank the many dedicated members of the TES ground data system and instrument teams.

## REFERENCES

- [1] CSSDS. (2005) The Consultative Committee for Space Data Systems. [Online]. Available: <http://public.ccsds.org/default.aspx>
- [2] NASA Langley Research Center (LaRC) Atmospheric Sciences Data Center (ASDC). (2005) TES Data Sets. [Online]. Available: [http://eosweb.larc.nasa.gov/PRODOCS/tes/table\\_tes.html#data](http://eosweb.larc.nasa.gov/PRODOCS/tes/table_tes.html#data)
- [3] R. Beer, T. Glavich, and D. Rider, “Tropospheric Emission Spectrometer for the Earth Observing system’s Aura satellite,” *Appl. Opt.*, vol. 40, pp. 2356–2367, 2001.
- [4] R. Beer, “TES on the Aura mission: Scientific objectives, measurements, and analysis overview,” *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 5, pp. 1102–1105, May 2006.
- [5] NASA. (2004) Earth Observing Science Data and Information System (EOSDIS) Homepage. [Online]. Available: [http://eosdis-main.gsfc.nasa.gov/eosinfo/EOSDIS\\_Site/index.html](http://eosdis-main.gsfc.nasa.gov/eosinfo/EOSDIS_Site/index.html)
- [6] EOSDIS. (2005) SCF Toolkit 5.2.13 for the EOSDIS Core System (ECS). [Online]. Available: <http://edhs1.gsfc.nasa.gov/wais-data/rel7/pdf/814emd001r3.pdf>
- [7] National Center for Supercomputing Applications (NCSA). (2005) NCSA HDF Home Page. [Online]. Available: <http://hdf.ncsa.uiuc.edu>
- [8] H. Worden, R. Beer, K. Bowman, B. Fisher, M. Luo, E. Sarkissian, D. Tremblay, and J. Zong, “TES level 1 Algorithms: Interferogram processing, geolocation, radiometric and spectral calibration,” *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 5, pp. 1288–1296, May 2006.
- [9] E. Sarkissian and K. W. Bowman, “Application of a nonuniform spectral resampling transform in Fourier-transform spectrometry,” *Appl. Opt.*, vol. 42, pp. 1122–1131, 2003.

- [10] National Center for Supercomputing Applications (NCSA). (2005) NCSA HDF-EOS Project Home Page. [Online]. Available: <http://hdf.ncsa.uiuc.edu/hdfeos.html>
- [11] S. Poosti, “Providing modifiability and robustness in the science software for atmospheric retrievals,” in *2004 AIAA Meeting—Space 2004 Conf.*, vol. 9.
- [12] S. Kulawik, H. Worden, G. Osterman, M. Luo, R. Beer, J. Worden, K. Bowman, A. Eldering, M. Lampel, T. Steck, and C. Rodgers, “TES atmospheric profile retrieval characterization: An orbit of simulated observations,” *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 5, pp. 1324–1333, May 2006.
- [13] R. Beer, K. W. Bowman, D. Brown, S. A. Clough, A. Goldman, D. J. Jacob, J. A. Logan, M. Luo, F. J. Murcray, D. M. Rider, C. P. Rinsland, C. D. Rodgers, E. A. Ustinov, and H. M. Worden. (1999, Oct.) Tropospheric Emission Spectrometer (TES) Level 2 algorithm theoretical basis document, V. 1.1, JPL D-16474. JPL, Pasadena, CA. [Online]. Available: <http://eosps0.gsfc.nasa.gov/atbd/testables.html>
- [14] S. Kulawik, J. Worden, A. Eldering, K. Bowman, M. Gunson, G. Osterman, L. Zhang, D. Jacob, and R. Beer, “Implementation of cloud retrievals for Tropospheric Emission Spectrometer (TES) atmospheric retrievals: Part I—Description and characterization of errors on trace gas retrievals,” JPL, Pasadena, CA.
- [15] Research Systems Incorporated (RSI). (2005) IDL Homepage. [Online]. Available: <http://www.rsinc.com/idl/index.asp>
- [16] Gnu gcc. (2005) GCC Homepage. [Online]. Available: <http://gcc.gnu.org/>
- [17] S. Watson, S. Larson, and D. Tremblay, “TES science data processing system C++ coding standards,” JPL, Pasadena, CA, JPL D-20315, Jul. 2001.
- [18] E. Chu and D. Tremblay, “TES Science Investigator-led Processing System,” *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 5, pp. 1352–1358, May 2006.
- [19] International Business Machines (IBM). (2005) Rational Rose XDE Developer Plus Homepage. [Online]. Available: <http://www-306.ibm.com/software/awdtools/developer/plus>
- [20] D. Shepard, “TES Level-1A Subsystem Software Requirements, JPL D-17961, Version 2.0,” JPL, Pasadena, CA, Feb. 2004.
- [21] D. Shepard, E. Sarkissian, and S. Gluck, “TES Level 1B Subsystem Software Requirements, JPL D-17962, Version 1.1,” JPL, Pasadena, CA, Aug. 2005.
- [22] S. Paradise, S. Akopyan, A. Eldering, S. Friedman, R. Gonzalez, B. Koffend, M. Lampel, G. Osterman, S. Poosti, E. Sarkissian, S. Sund, H. Worden, and H. Yun, “TES Level-2 Functional Requirements Document, JPL D-23176, Version 1.0,” JPL, Pasadena, CA, Jan. 2005.
- [23] MISR. (2005) Multi-Angle Imaging SpectroRadiometer Homepage. [Online]. Available: <http://www-misr.jpl.nasa.gov>
- [24] MLS. (2005) Microwave Limb Sounder Homepage. [Online]. Available: <http://mls.jpl.nasa.gov>
- [25] OCO. (2005) Orbiting Carbon Observatory Homepage. [Online]. Available: <http://oco.jpl.nasa.gov>
- [26] J. Worden, S. Kulawik, M. Shephard, S. Clough, H. Worden, K. Bowman, and A. Goldman, “Predicted errors of Tropospheric Emission Spectrometer nadir retrievals from spectral window selection,” *J. Geophys. Res.—Atmos.*, vol. 109, no. D9, 2004.
- [27] S. Lewicki, “Science Data Processing Standard and Special Observation Data Product Specifications,” JPL, Pasadena, CA, JPL D-22993. [Online]. Available: <http://eosweb.larc.nasa.gov/PRODOCS/tes/DPS/>
- [28] Oracle. (2005) Oracle Database Home Page. [Online]. Available: <http://www.oracle.com/database/index.html>
- [29] World Wide Web Consortium (W3C). (2005) Hypertext Markup Language (HTML) Home Page. [Online]. Available: <http://www.w3.org/MarkUp/>
- [30] PHP. (2005) PHP: Hypertext Preprocessor. [Online]. Available: <http://www.php.net/>
- [31] GTK+. (2005) The GIMP Toolkit Homepage. [Online]. Available: <http://www.gtk.org>

**Susan Paradise** received the B.S. degree in mathematics and computer science from Purdue University, West Lafayette, IN, and the M.S. degree in mathematics from the University of Southern California, Los Angeles.

She has worked at the Jet Propulsion Laboratory (JPL), Pasadena, CA, as a Software Engineer since 1985, when she joined the Atmospheric Trace MOlecule Spectroscopy (ATMOS) Project, later becoming a Level 2 Cognizant Design Engineer on MISR. She served as the Science Software Project Element Manager on TES.

**Sirvard Akopyan** received the B.S. degree in electronics from Yerevan Polytechnic Institute, Armenia, in 1990, the M.S. degree in computer science from California State University, Northridge, in 2000.

In 1998, during her graduate school years, she supported the TES project, Jet Propulsion Laboratory, Pasadena, CA. After graduating, she continued working on TES project's ground data software as a cognizant design engineer for the TES Level 2 subsystem.

**Richard C. De Baca** received the B.S. degree in computer science from New Mexico State University (NMSU), Las Cruces, in 1995.

As a student he worked at NMSU's Physical Science Laboratory developing software for ground-based microwave radiometers and joined the Jet Propulsion Laboratory, Pasadena, CA, full-time in 1996. Since then, he has worked in the development of ground data systems software for the MISR and TES projects.

**Kevin Croft** received the Geomatics Engineering degree from the University of Calgary, Calgary, AB, Canada.

He contributed to the Level 2 software optimization effort and developed the CASPER special-observation processing system for TES, Raytheon ITSS, Pasadena, CA.

**Kent Fry** came to the TES Project, Jet Propulsion Laboratory, Pasadena, CA, from the NASA Scatterometer (NSCAT), Quick Scatterometer (QuikSCAT) and SeaWinds Projects. He is currently assigned the task of Cognizant Design Engineer of the L1A subsystems and the CASPER test and development aid tool.

**Scott Gluck** received the B.S. degree from the University of California at Santa Barbara and the M.S. degree from the University of Southern California, Los Angeles, in 1985 and 1994, respectively, both in computer science.

He has worked on numerous missions at the Jet Propulsion Laboratory, Pasadena, CA, since joining in 1990. These include the ground data systems for MISR and the All Source Analysis System query support processing. He is currently the cognizant software engineer for the TES Level 1B subsystem processing.

**David Ho** was born in Hong Kong. He received the B.S. degree from the University of California, Los Angeles, in 1988.

He has been working in the aerospace industry since his graduation, on projects from military airplanes such as the F-14 Tomcat and the F-22 Raptor to various commercial and NASA satellites, including the XM-radio, DirecTV, NASA TDRS, and NASA EOS Aura. He is currently with Raytheon ITSS, Pasadena, CA.

**Brooke Koffend** received the Ph.D. degree in physical chemistry from the Massachusetts Institute of Technology, Cambridge.

He was a Visiting Professor of physics at the University of Lyon, France. He then was Scientist at the Aerospace Corporation where his research involved chemical lasers, molecular spectroscopy, and chemical kinetics. He has been a Software Engineer following his work at Aerospace, involved in numerous scientific and defense related work, including support for the TES Level 2 subsystem. He is currently with the Jet Propulsion Laboratory, Pasadena, CA.

**Michael Lampel** is with Raytheon ITSS as Group Manager, Atmospheric and Ocean Sciences, Pasadena, CA.

He has been working on TES for seven years in aspects of algorithm development, data verification and validation, system performance and architecture. He has worked in multiple fields after receiving his doctorate in plasma and accelerator physics, including medical X-ray imaging, free electron laser technology development and systems design, physics design tools, and most recently atmospheric remote sensing.

**James McDuffie** received the B.S. degree in computer science in 2002 from the Georgia Institute of Technology, Atlanta, and the MBA degree from the University of Redlands, Redlands, CA, in 2002.

Initially a student employee at Jet Propulsion Laboratory, Pasadena, CA, he worked on a web-based Mars data image retrieval system. Since 2002, he has contributed to development of the TES L1B, L2 and Visualization ground data processing software subsystems.

**Ruth Monarrez** received the B.S. degree from the University of California at Los Angeles and the M.S. degree from Claremont Graduate School, Claremont, CA, both in applied mathematics.

While earning her degrees, she worked at the Jet Propulsion Laboratory, Pasadena, CA, as a Data Engineer for the Planetary Data System (PDS) project. Upon receiving her master's degree she worked for the MISR and TES projects developing ground data system software.

**Hari Nair** received the B.S. degree in mathematics and physics from Gannon University, Erie, PA, and the M.S. and Ph.D. degrees in planetary science from the California Institute of Technology, Pasadena.

He served on the TES L1A and L2 software teams from 2000 to 2005. He is presently a member of the Orbiting Carbon Observatory (OCO) L2 software team, Jet Propulsion Laboratory, Pasadena.

**Sassaneh Poosti** received the B.S. degree in mechanical engineering and the M.S. degree in computer science from California State University, Northridge.

She has been working on the TES project at the Jet Propulsion Laboratory, Pasadena, CA, since 1998.

**Doug Shepard** is the System Engineer on the Ground Data System of TES and has been employed at the Jet Propulsion Laboratory, Pasadena, CA, in different capacities for 20 years.

**Irina Strickland** received the B.A. degree in mathematics from Occidental College, Los Angeles, CA, in 1997.

She has been working as a Software Engineer for the TES Ground Data System, Raytheon ITSS, Pasadena, CA.

**Denis Tremblay** received the Ph.D. degree in aerospace engineering from the University of Colorado, Boulder, in 1995.

He has been a member of the TES Ground Data System, Raytheon ITSS, Pasadena, CA, since 1998. He contributed to the software development with emphasis on calibration (Level 1B), geolocation, and science validation.



**Hyejung Yun** received the B.S. degree in mathematics and computer science from the University of California, Los Angeles, in 1989. After working in industry, she went on to continue graduate study and received the M.S. degree in computer science at the Florida Institute of Technology, Melbourne, in 1991.

Since then, she has worked in the defense/aerospace industry developing large-scale, complex software systems until she joined Raytheon ITSS, Pasadena, CA, in 1998, supporting Jet Propulsion Laboratory (JPL), Pasadena, CA, JPL/NASA projects. Employed at JPL since 2002, she has worked on Shuttle Radar Topography Mission (SRTM), Navigation software, and TES.

**Jia Zong** received the M.S. degree in physics from the Southern Illinois University, Carbondale, in 1989 and the M.S. degree in photogrammetry from The Ohio State University, Columbus, in 1993.

She served as the Cognizant Design Engineer for the Tropospheric Emission Spectrometer Level 1A Geolocation subsystem. She joined the Jet Propulsion Laboratory, Pasadena, CA, in 1994.