NASA/TM—2019-219995

Space Communications and Navigation Validation:
Extracting Data for the Strategic Center for
Networking, Integration, and Communications
Scheduling Algorithms

Noah P. Kontur and Bryan W. Welch
Glenn Research Center, Cleveland, Ohio

NASA STI Program . . . in Profile

Since its founding, NASA has been dedicated to the advancement of aeronautics and space science. The NASA Scientific and Technical Information (STI) Program plays a key part in helping NASA maintain this important role.

The NASA STI Program operates under the auspices of the Agency Chief Information Officer. It collects, organizes, provides for archiving, and disseminates NASA's STI. The NASA STI Program provides access to the NASA Technical Report Server—Registered (NTRS Reg) and NASA Technical Report Server—Public (NTRS) thus providing one of the largest collections of aeronautical and space science STI in the world. Results are published in both non-NASA channels and by NASA in the NASA STI Report Series, which includes the following report types:

- TECHNICAL PUBLICATION. Reports of completed research or a major significant phase of research that present the results of NASA programs and include extensive data or theoretical analysis. Includes compilations of significant scientific and technical data and information deemed to be of continuing reference value. NASA counter-part of peer-reviewed formal professional papers, but has less stringent limitations on manuscript length and extent of graphic presentations.

- TECHNICAL MEMORANDUM. Scientific and technical findings that are preliminary or of specialized interest, e.g., "quick-release" reports, working papers, and bibliographies that contain minimal annotation. Does not contain extensive analysis.

- CONTRACTOR REPORT. Scientific and technical findings by NASA-sponsored contractors and grantees.

- CONFERENCE PUBLICATION. Collected papers from scientific and technical conferences, symposia, seminars, or other meetings sponsored or co-sponsored by NASA.

- SPECIAL PUBLICATION. Scientific, technical, or historical information from NASA programs, projects, and missions, often concerned with subjects having substantial public interest.

- TECHNICAL TRANSLATION. English-language translations of foreign scientific and technical material pertinent to NASA's mission.

For more information about the NASA STI program, see the following:

- Access the NASA STI program home page at http://www.sti.nasa.gov

- E-mail your question to help@sti.nasa.gov

- Fax your question to the NASA STI Information Desk at 757-864-6500

- Telephone the NASA STI Information Desk at 757-864-9658

- Write to:
  NASA STI Program
  Mail Stop 148
  NASA Langley Research Center
  Hampton, VA 23681-2199

NASA/TM—2019-219995

![NASA logo]

Space Communications and Navigation Validation: Extracting Data for the Strategic Center for Networking, Integration, and Communications Scheduling Algorithms

Noah P. Kontur and Bryan W. Welch
Glenn Research Center, Cleveland, Ohio

National Aeronautics and
Space Administration

Glenn Research Center
Cleveland, Ohio 44135

March 2019

Acknowledgments

Level of Review: This material has been technically reviewed by technical management.

Available from

Space Communications and Navigation Validation: Extracting Data for the Strategic Center for Networking, Integration, and Communications Scheduling Algorithms

Noah P. Kontur[*] and Bryan W. Welch
National Aeronautics and Space Administration
Glenn Research Center
Cleveland, Ohio 44135

Summary

Efficiency in communication system architecture performance between Space Communications and Navigation (SCaN) assets and missions is crucial, as space communication is varied, complex, and often not utilized to its full potential. The SCaN Strategic Center for Networking, Integration, and Communications (SCENIC) new scheduling algorithms, which are designed to simulate the allocation of resources between SCaN assets and missions, have the potential to simulate an increase of this efficiency; however, they require real-world data to be validated against. The purpose of this project was to extract said validation data, which details the frequency and duration of utilized contact windows between missions and assets in the Near Earth Network (NEN), Space Network (SN), and Deep Space Network (DSN). Stored as images in daily operations summaries (DOSs), the tabular data existed in a variety of file formats such as .pdf, .docx, and .doc. Since the tables were stored as images, ABBYY® FineReader® (ABBYY Software Ltd.) optical character recognition (OCR) was implemented, which is a proprietary software that reads images from text. The comma separated value (CSV) output was utilized as input to a series of MATLAB® (The MathWorks, Inc.) methods for reformatting, at which point it was ready to be machine read. Finally, the results were converted to a Microsoft Excel format for human readability. Along with being used for validation purposes, the data will also be used to map equipment degradation as a function of time to analyze the reliability of network assets.

Introduction

Efficiency in communication system architectures between Space Communications and Navigation (SCaN) assets consisting of the Near Earth Network (NEN), Space Network (SN), and Deep Space Network (DSN) is crucial, as the networks are varied, complex, and often not utilized to their full potential (Figure 1). The DSN, for instance, suffers from a bevy of constraints, such as transition time between tracks (utilized contact windows between ground stations and missions), an inability of ground stations to communicate with multiple missions at the same time, and nebulous mission requirements (Ref. 1). While some of these constraints are due to hardware limitations, the management and mitigation of these constraints—those that will likely never disappear entirely—is ultimately a scheduling problem. The new SCaN Strategic Center for Networking, Integration, and Communications (SCENIC) scheduling algorithms will employ artificial intelligence to allocate resources between SCaN assets and missions to address this problem by simulating methods that could increase the efficiency across all three networks.

If the SCENIC algorithms are not tested against real-world data, there will be no guarantee that they will operate as intended. Therefore, the algorithms need to be validated before they can be trusted. According to NASA's technical standard for models and simulations (version NASA–STD–7009A), validation is "the process of determining the degree to which a model or a simulation is an accurate representation of the real world from the perspective of the intended uses of the M&S (Ref. 2)." In this particular case, validation is meant to determine whether SCENIC's scheduling algorithms represent the

---

[*]Summer Intern in Lewis' Educational and Research Collaborative Internship Project (LeRCIP).
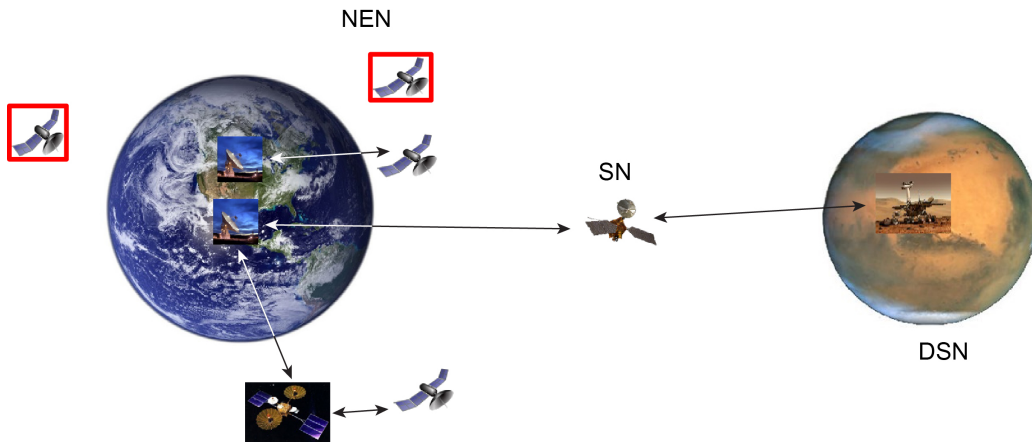
Figure 1.—Near Earth Network (NEN), Deep Space Network (DSN), and Space Network (SN), illustrating often underutilized contact windows between missions and assets.
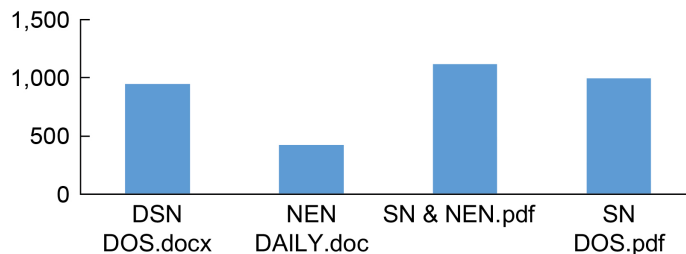
Figure 2.—Number of files for given format. In 2014, Near Earth Network (NEN) and Space Network (SN) daily operations summaries (DOSs) were consolidated into SN_&_NEN format, which is why there are four file formats for three networks.

real-world loading analysis of the missions on the SCaN network assets. In order to make the validation process feasible and time efficient, the data must be made available in a machine-readable format.

The purpose of this project was to extract and format the data, which will be used to validate the scheduling algorithms. The first phase involved finding a way to extract text from the tabular data. As seen in Figure 2, the data were spread across thousands of files and a variety of formats. Moreover, the information was not confined to one network alone, containing details on the NEN, DSN, and SN, and the files described over 50 assets and over 150 missions. Both the diversity and quantity of the files involved required a whole host of techniques and checks for both an automatic and reliable conversion.

The second phase of the project concerned the extraction of other reliability data stored within each file. In the SN_DOS files alone, for instance, there is detailed information on customer events, tests, table loss corrections, launch forecasts, and daily schedules (see Figure 3). While this auxiliary data will not necessarily be used for validation purposes, the relatively low marginal cost of extracting the additional information qualified it as a secondary objective. This data will likely be used to better understand network loss and equipment degradation.

The third phase of the project involved reformatting and exporting the extracted data to comma separated value (CSV) and Excel formats. CSV format is best for machine readability, and will almost certainly be used for the validation itself; alternately, Excel is better for human readability, a factor that will be important for visualizing and perusing the reliability section.

TDRS-K Test Case 4.12 (NTR-012425)
Objectives: A. Verify proper and timely TDRS K vector processing for Expendable Launch Vehicle (ELV) launch support, including Extended Elliptical Field of View (EEFOV) support. B. Verify proper and timely TDRS K antenna pointing and antenna slewing for ELV launch support including EEFOV support. C. Verify proper and timely TDRS K vector processing for re-entry support. D. Verify Proper and timely TDRS K antenna pointing and antenna slewing for re-entry support.
Results: Objectives met Test Case 4.12 Simultaneous SA Service for launch Support test objectives were completed successfully:

TDRS-K Test Case 4.11 (NTR-012426)
Objectives: Verify that worst case simultaneous slewing of both TDRS K SA antennas in Azimuth or in Elevation will not adversely affect the Space to Ground Link (SGL) antenna pointing system.
Results: Objectives met Test Case A4.11 (Simultaneous SA Antenna Slew) was successfully completed.

ISS/TDRS-K Level 6 SN Target of Opportunity (TOO) Test (NTR-012430)
Objectives: A. Verify TDRS-K support of ISS scheduled on-orbit support configurations. B. Verify operational interfaces that support ISS.
Results: JSC advised that this was an excellent support. No problems were encountered. No GCMRs were sent during this event.

Figure 3.—Space Network test section from SN_DOS file, where EEFOV is Extended Elliptical Field of View, GCMR is Ground Control Message Request, ISS is International Space Station, JSC is Johnson Space Center, NTR is New Technology Report, SA is single access, and TDRS is Tracking and Data Relay Satellite.

The fourth and final phase of the project involved making the process easy for others to use in the future. Admittedly, the main focus of the project was to extract information from existing daily operations summaries (DOSs). However, documentation of the process, along with the MATLAB® code itself, has made it possible for future users to automatically convert new DOSs with relative ease. So long as DOS file formats stay relatively consistent, the SCENIC scheduling algorithms will have an ever-expanding source of real-world data to test against.

Methodology

Software Development and Tools

In the name of cost savings, the first step of the process involved utilizing the MATLAB® Computer Vision System Toolbox™ (The MathWorks, Inc.) (Ref. 3) implementation of Tesseract optical character recognition (OCR) (Ref. 4), an open-source software. The results were initially disappointing, as Tesseract mistook table gridlines as characters; therefore, an attempt was made to implement the MATLAB® line detection software via the Hough transform (Ref. 5). While this approach yielded decent results, it only identified the correct gridlines about 90 percent of the time. Seeing as line detection needed to be perfect to ensure a reasonable output, a method was created that better partitioned a table image according to its gridlines. Rather than using statistical analysis to recognize lines, the method simply found contiguous strings of pixels that stretched across a significant portion of the image. Though far less robust than the previous approach, this method allowed for perfect line detection in this particular case. With the accuracy of the partitions verified, the images were split into subimages where each image represented a cell in the table. The subimages were then fed into Tesseract, and the outputs were reconstructed into a matrix. Although the results were satisfactory in some cases, MATLAB® was unable to natively import vector graphics image formats such as Enhanced MetaFile (EMF)—the standard for images made in Microsoft® Word—and resulted in a necessary format conversion to Portable Network Graphic (PNG) and subsequent image downscaling. This process generally reduced the images' pixel density by anywhere from 2 to 10 times, resulting in garbled OCR results. Moreover, the difficulty of extracting data from Portable Document Format (PDF) documents led to the conclusion that dedicated OCR software that could handle PDFs was necessary.

Customer	DSCC DSS	GOLDSTONE 14	15	24	25	26	CANBERRA 34	35	43	45	MADRID 54	63	65	Total Passes	Total Time	Service Loss
ACE													1	1	03:30	00:00
CAS									1					1	09:00	00:00
CHDR													2	2	02:00	00:00
DAWN	1						1	1				1		4	24:40	00:00
GTL			1							1			1	3	03:15	00:00
LRO							1			1				2	01:30	00:00
M01O				1		1				1		1	1	5	15:45	00:00
MMS1							1							1	01:00	00:00
MMS2							1							1	01:00	00:00
MMS3							1							1	01:00	00:00
MMS4							1							1	01:00	00:00
MOM				1										1	03:30	00:00
MRO					1			1				1		3	19:00	00:00
MSL													1	1	03:45	00:00
MVN						1								1	09:10	00:00
ROSE									1			1		2	07:50	00:00
SOHO			1							1	1			3	12:50	00:00
STA					1						1			2	08:10	00:01
STB	1													1	02:30	00:00
THB			1											1	01:20	00:00
VGR1	1													1	02:05	00:00
VGR2							1	1						2	13:55	00:00
WIND			1											1	02:35	00:00
Total	3		4	2	2	2	7	3	2	4	2	4	6	41	150:20	00:01

Figure 4.—Example Deep Space Network document being processed by ABBYY® FineReader®. Blue lines show results of its line detection functions, where DSCC is Deep Space Communications Complex and DSS is Deep Space Station.

ABBYY® FineReader® satisfied all these needs, batch converting tabular data into a CSV output. The process was relatively expedient, taking only 2 h for all 3,500 documents. Even so, there were a few hitches in the OCR conversion. For instance, ABBYY® FineReader® requires PDFs as inputs; however, the software has a built-in functionality that converts from .docx to PDF. Another problem concerned the low image quality of the Nen_Daily, which entailed a manual specification of the table regions and vertical text orientation in the table headers. Although, for the most part, ABBYY® FineReader® correctly recognized alphanumeric tabular data (Figure 4).

There were a few formatting issues that were too cumbersome to deal with in ABBYY® FineReader®. For instance, the OCR would occasionally recognize nonexistent columns. Additionally, it was quite common for section titles to be interspersed with garbled characters. Rather than be addressed in ABBYY® FineReader® itself, these mistakes were fixed in the MATLAB® code. For the first problem, a routine was implemented that read the cell-matrix output of the table reader functions. If there was either a row or a column with all indices empty, it was removed. The second problem required a more exhaustive solution. As shown in Figure 5, header titles would often turn out garbled after the OCR conversion, and these mistakes made it difficult for the MATLAB® code to distinguish between sections. If, for instance, the code compared against the phrase NEN EQUIPMENT STATUS to indicate the end of the NEN test section (as might be done in the section in Figure 5), then NEN EQUIPMENT STA TUS_____ would not be recognized. In fact, it often happened in the initial stages of the project that these sections would never compare positively with any phrases in the section's predefined list of end phrases, and the entire rest of the file's text would appear in one section. Two solutions were applied in tandem. One approach was to expand upon the section's list of common phrases (or their garbled relatives). After

NEN TESTS									
None									
NEN EQUIPMENT STA TUS									

Figure 5.—Part of an example SN_&_NEN comma separated value output from ABBYY® FineReader® with garbled section title data, where NEN is Near Earth Network.

this approach only achieved sporadic success, a second approach was also used, which consisted of reformatting every line before it was analyzed for end phrases. This method helped resolve some of the idiosyncrasies in the garbled output. Spaces and capitalization might be ignored, for instance, allowing most of the garbled end phrases to be recognized appropriately. Moreover, the method was simpler, as it reduced the size of the string array necessary to recognize the end phrases. However, though this approach might be ostensibly superior to the former, it had to be used with care. If too many changes and simplifications were made for line comparison, there was a chance that the program would identify a false positive. With these risks recognized, spaces and underscores were removed entirely, allowing MATLAB® to correctly separate sections from each other. However, capitalization was not ignored, as many lowercase end phrases are used in textual descriptions of events and tests.

While most formatting challenges revolved around rectifying mistakes in ABBYY® FineReader®, the source data itself posed a few significant challenges. One obstacle was the partitioning of subsections for the reliability data. For instance, while a quick glance at Figure 3 is enough to delineate each of the three tests, many of the geometric attributes that allow this delineation are lost in translation and text alone provided the basis upon which sections could be partitioned. For this particular example, the string New Technology Report (NTR-) served as a suitable signal that a new test was being announced. Even so, there are some cases where this string did not appear in the subsection headers (see Figure 6). Even more problematic, there is the risk that future documents will contain NTR-string within textual descriptions of the tests. Each section then required a rigorous analysis of what words or word combinations might signal the start of a new test, event, or system. These phrases had to be both sufficient to mark separation between subsections and unique in their appearance as a separator. As this sought-after phrase did not always exist for the test sections, regular expressions were also applied. Unlike typical string searches, regular expressions are a more versatile approach, which allow for the identification of patterns in text (Ref. 6). For several of the reliability sections—most notably test, general information, and SN system status—regular expressions were used to identify a pattern that was singular to the subsection header. The test section was separated with the following MATLAB® regular expression: with\s\w*(\s)?(,)?(\s)?(\w*)?(\s)?(\d*)?(\s)?(\w*)?(\s)?(\d*)?(\s)?(.)?\r\nObjectives:. When used as an input into a regular expression function, this particular expression identifies an instance of sandwiched between two words. What makes this expression unique in identifying subsections is the stipulation that the next carriage return and line encountered will be followed by Objectives:. In this way, both the regular expression and the NTR-string identifier were used to recognize test subsections. Regular expressions were used in the general information and system status sections in a similar manner.

One rarer flaw in the data that caused issues was the mislabeling of section titles. In the SN_&_NEN DOS for February 21, 2016, for instance, the NEN Customer Event section was erroneously labeled as the NEN test section. This error has only appeared twice in all of the DOSs combined, but its occurrence has led to a garbled output section in both cases. The rarity of its occurrence, along with the occurrence of other rare errors, has made it impossible to avoid a manual check of the final output. In this case, garbled section data is checked against the source file; if the section is mislabeled, the mistake is recorded in the error logs for each file format (Figure 7).

Throughout the development process, attention has been paid to usability and runtime. Bottlenecks in the code were identified with the MATLAB® profiler, a built-in feature, which displays runtimes for each function. These bottlenecks were then addressed to speed up the code significantly. The checksum routines, for instance, have been modified to reduce loop usage, and variable declarations have been minimized in the main methods. Moreover, the data extraction methods were modified to grab text in larger batches, reducing the bottleneck on the MATLAB® I/O functions. Given the objective of ensuring

ISS / VHF-1 Emergency Comm Checkout with AFRC and WSC					
Objectives: To exercise the following: 1. VHF Scheduling via WOTIS 2. Two Line Element (TLE) retrieval from the FDF CIL, verification and implementation 3. Interface procedures and voice communications between all participants. 4. Verify ISS Tracking with AFRC & WSC VHF-1. 5. Verify VHF-1 Uplink and Downlink communications between ISS and MCC-H MCC-H/International Partners via AFRC and WSC VHF stations					
Results: Objectives partially met Support was evaluated as partially successful. The pass was impacted by ISS Crew availability. It was reported that the crew was not in position to begin VHF communications. At AFRC AOS (14:03:23Z) there was not downlink at AFRC. At 14:10:00Z VTO handed the voice circuits over to WSC, AFRC terminated their uplink and WSC enabled their uplink. At 14:10:19Z WSC received downlink call from ISS who reported receiving uplink calls from JSC. Voice checks were made with MCC-H, MSFC, GSOC, and MCC-M. JSC and ISS continued voice calls until LOS at 14:12:42Z Successful uplink and downlink provided by WSC. AFRC reported receiving good downlink post-handover.					

Figure 6.—Near Earth Network test subsection, which lacks typical New Technology Report (NTR-) signal string, where AFRC is Armstrong Flight Research Center, AOS is Aircraft Operations Support, CIL is Customer Interface Layer, FDF is Flight Dynamics Facility, GSOC is German Space Operations Center, LOS is loss of signal, MCC-H is Mission Control Center-Houston, MCC-M is Mission Control Center-Moscow, MSFC is Marshall Space Flight Center, ISS is International Space Station, JSC is Johnson Space Center, VHF is very high frequency, VTO is Voice Transfer Officer, WOTIS is Wallops Orbital Tracking Information System, and WSC is White Sands Complex.

LANDSAT-7	042/07:43:37Z	USN-AK	5.4-Meter	Loss 0:05:00	DR# 263489
Project reports low scene quality from review of X-Band playback. SSC/USN investigation of this reported problem ongoing.					

LANDSAT-7	042/09:25:15Z	USN-AK	5.4-Meter	Loss 0:00:30	DR# 263490
Project reported XL data missing first scheduled scene from playback. SSC/USN investigating cause at this time.					

Figure 7.—Near Earth Network (NEN) customer event subsections improperly labeled as NEN test sections, where AK is Alaska, SSC is Swedish Space Corporation, and USN is Universal Space Network, Inc.

human readability, the majority of runtime will still be dedicated to Excel conversion. Batch writing and modifying Excel spreadsheets requires the use of ActiveX® (Microsoft) servers, which in turn open an instance of Excel for every modified sheet. Presently, there is no way to create and modify Excel spreadsheets in batch writing without opening a new Excel instance, so code efficiency efforts were here constrained to limiting Excel instances per file to one. After all these efforts to streamline the code were implemented, the SN_&_NEN converter, the program that converts and writes ABBYY SN_&_NEN output to CSV and Excel, still took around 2 h to run for all 1,100 files. Although, future DOSs will likely be processed in batches not exceeding 100, so current runtimes are most definitely within the realm of feasibility. In addition to modifying the code for efficiency, work has been done to ensure future usability. Comments have been judiciously placed to make the code more readable. The code is also broken up into header sections, and information about the author, function, and other required functions is included at the top of each function file. User input has been reduced to specifying destinations for the outputs and toggling an optional debug tool. Warnings are displayed when the input data is garbled or unexpected. Moreover, a user manual was written, which details how to use both ABBYY® FineReader® and the MATLAB® code. The manual also describes some of the more common errors in the ABBYY® FineReader® conversion, which require manual user intervention.

Verification and Validation Testing

While validation tests simulations against real-world systems, software verification is the process of checking that software performs its intended purpose. As stated before, the purpose of the software employed—whether MATLAB® or ABBYY® FineReader®—is to convert DOS data to a machine- and human-readable format. Admittedly, verification is a difficult process for proprietary software on account of the software's immutability. If ABBYY® FineReader® fails verification, the errors in the software that lead to the failure cannot be fixed. Ultimately, the CSVs could only be altered after the fact, either

manually or in MATLAB®. Nevertheless, both the ABBYY® FineReader® and MATLAB® software were verified, even if a failure in the former could only entail modification to the latter.

One method of performing verification and validation was done by checking the numeric data and seeing if the totals were correct. If data in the numeric section of the table showed up as nonnumeric, then ABBYY® FineReader® had made a mistake in the conversion to CSV and the output was modified. If, however, the error was present in the original source document, then the document was noted to have failed validation, and all validation errors were recorded in text files. Though the checksum errors were generally infrequent, some DOSs were littered with them (Figure 8). In one SN DOS from March 2014, for instance, over 15 sum errors were identified and recorded, and when the scheduling algorithms are validated, this kind of unreliable data will be discarded.

Another method used for verification and validation was a name check of mission and asset names against a known list. If there was an unknown name, it would be searched for manually in a provided mission commitment database. Once the name was found, its existence was recorded and future instances of the name were converted according to the most recent name on record. If the name could not be found, which was often the case when ABBYY® FineReader® mistakenly recognized it, a quick perusal of the original document would provide the correct name, which would be applied manually to the input.

Another check that was performed involved duplicate DOSs. When the NEN_Daily format was transitioning to the SN_&_NEN format in late 2014, some NEN DOSs were included in both formats. To prevent validation data from being redundant, dates were extracted from the file names for the NEN_Daily's and SN_&_NEN files and compared by using PowerShell (Microsoft). Dates with duplicates were found and output to a text file for perusal.

Finally, a manual check of the output in totum was performed at the end of project. Ideally, this check would be unnecessary, as the sheer amount of output files precludes an expedient, exhaustive check of each and every one for exotic errors. Though this final check was by no means perfect, it helped to rectify some of the more glaring errors in the output, like an incorrect partitioning of sections and garbled text in the customer event sections.

ANTENNA	S-South	W Central	S2	W-North	10 Meter	S-North	G-South	G Central	DAS	TOTAL
ACRIMSAT	2	1		3		2		6		8
AIM	3	4		5		2		5		20
AQUA	4	9		11		5		2		34
AURA	2	9		9		4				26
DRAGON		3								3
FGST	2	3		3		4		11	45	57
GPM	5	9		4		8				37
HST		3				12		2		15
ISS	14	2		15		1				34
LSAT-7				1				2		1
NEOWISE	2			1		3				8
NUSTAR	2									2
SCANTB HARRIS	2									2

Figure 8.—Part of SN_DOS_041414 scheduled pass table. Notice how missions' passes with individual Tracking and Data Relay Satellite assets seldom sum to total daily passes listed in final column shown.

```
,TDRS-9,TDRS-3,TDRS-6 S-Band ,TDRS-5,TDRS-10,TDRS-8
Low ,TDRS-7,,,,,
LOCATION,40.9W,49W,62.4W,167.5W,174W,271W,275W,,,,,
SGLT,3,5,S-STTC,4,1,7,6,,,,,
DESIGNATOR,TDE,TDS/DASE,,171/DASW,TDW,,TDZ/DASZ,,,,,
ANTENNA,S-South,W-Central,10 Meter,W-North,S-North,G-South,G-
Central,DAS,TOTAL,SCHEDULED,LOSS,
AIM,6,4,0,6,4,0,6,0,26,21:38:44,0:00:00
AQUA,5,13,0,13,2,0,3,0,36,29:52:45,0:00:00
AURA,0,15,0,3,15,0,4,0,37,27:10:15,0:00:00
CNOFS,0,0,0,0,0,0,0,0,42,42,41:12:00,0:00:00
Fermi Gamma-ray Space
Telescope,4,0,0,2,4,0,0,40,50,43:39:08,0:00:00
HST,0,8,0,0,8,0,0,0,16,24:56:38,0:00:00
ISS,0,13,0,12,0,0,8,0,33,78:53:24,0:00:00
LDBP-BLAST,0,11,0,0,0,0,16,5,32,21:07:57,0:00:00
LDBP-EBEX,0,0,0,6,2,0,5,5,18,41:24:14,0:00:00
LDBP-SUPERTIGER,0,0,0,0,0,0,0,3,3,22:31:00,0:00:00
LANDSAT-5,2,0,0,0,3,0,0,0,5,2:53:34,0:00:00
LANDSAT-7,0,0,0,1,0,0,0,0,1,0:29:30,0:00:00
NUSTAR,0,0,0,2,0,0,0,0,2,1:20:00,0:00:00
SORCE,2,3,0,2,1,0,4,0,12,4:00:00,0:00:00
Special Projects and
Missions,0,0,0,0,0,0,0,0,42,94:14:28,0:00:00
SPTR-2,0,0,0,2,0,0,0,0,2,20:22:30,0:00:00
SWIFT,0,0,0,0,0,0,0,42,42,23:49:58,0:00:00
TERRA,11,0,0,12,6,0,2,0,31,26:22:30,0:00:00
TRMM,6,4,0,2,5,0,0,0,17,9:59:30,0:00:00
TOTALS,0,0,0,0,0,0,0,0,447,535:58:05,0:00:00
Bilateration Ranging Transponder System,12,14,0,14,12,0,17,0,0,,
```

Figure 9.—Example Space Network table in comma
separated value format output from MATLAB®.

Results and Discussion

After the code and the ABBYY® FineReader® output were verified and validated, reliability data were appropriately partitioned and organized according to test, event, and section. Then, each file was written piecemeal to CSV. Under the destination directory, a subdirectory was made for each file, where CSVs for each section were then stored.

Though the output from MATLAB® is sufficient for machine readability, Figure 9 demonstrates how the data might be hard for a human to swiftly peruse. For the reliability data especially, there was a need to ensure human readability, as that data consists of textual descriptions of tests and events. By using the MATLAB® xlswrite function and applying ActiveX® servers to batch modify and create Excel files, the code transformed the CSVs into a consolidated Excel format, which could easily be understood by humans. Each Excel file represents 1 month of DOSs for each format, with a different sheet for each day—in this way, a reader can visualize the data without necessarily resorting to more arcane techniques.

Conclusion

In addition to the uses already enumerated, the data extracted could be utilized in a variety of circumstances. For instance, it could be made available in the Strategic Center for Networking, Integration, and Communications (SCENIC) interface, so that users would have access to past data on equipment failures and scheduled passes. It could also be used to visualize network loss and utilized network capacity. It is likely that the data, along with the code that brought it from image to machine-readable text, will be useful for years to come.

Appendix—Nomenclature

AFRC	Armstrong Flight Research Center
AK	Alaska
AOS	Aircraft Operations Support
CIL	Customer Interface Layer
CSV	comma separated value
DOS	daily operations summary
DSCC	Deep Space Communications Complex
DSN	Deep Space Network
DSS	Deep Space Station
EEFOV	Extended Elliptical Field of View
ELV	Expendable Launch Vehicle
EMF	Enhanced MetaFile
FDF	Flight Dynamics Facility
GCMR	Ground Control Message Request
GSOC	German Space Operations Center
ISS	International Space Station
JSC	Johnson Space Center
LOS	loss of signal
M&S	models and simulations
MCC-H	Mission Control Center-Houston
MCC-M	Mission Control Center-Moscow
MSFC	Marshall Space Flight Center
NEN	Near Earth Network
NTR	New Technology Report
OCR	Optical Character Recognition
PDF	Portable Document Format
PNG	Portable Network Graphic
SA	single access
SCaN	Space Communications and Navigation
SCENIC	Strategic Center for Networking, Integration, and Communications
SGL	Space to Ground Link
SN	Space Network
SSC	Swedish Space Corporation
TDRS	Tracking and Data Relay Satellite
TLE	Two Line Element
TOO	Target of Opportunity
USN	Universal Space Network, Inc.
VHF	very high frequency
VTO	Voice Transfer Officer
WOTIS	Wallops Orbital Tracking Information System
WSC	White Sands Complex

References

1. Clement, Bradley J.; and Johnston, Mark D.: The Deep Space Network Scheduling Problem. IAAI–05, Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA, 2005.
2. NASA–STD–7009A: Standard for Models and Simulations. 2016.
3. Design and Simulate Computer Vision and Video Processing Systems. MathWorks, 2019. https://www.mathworks.com/products/computer-vision.html Accessed Feb. 8, 2019.
4. Tesseract OCR. 2017. https://opensource.google.com/projects/tesseract Accessed Feb. 8, 2019.
5. Hough Transform. MathWorks, 2019. https://www.mathworks.com/help/images/hough-transform.html Accessed Feb. 8, 2019.
6. Regular Expressions. MathWorks, 2019. https://www.mathworks.com/help/matlab/matlab_prog/regular-expressions.html Accessed Feb. 8, 2019.